Supplementary Material for "Nonequivalence of Two Least-Absolute-Deviation Estimators for Mediation Effects"

WenWu Wang[a,b,*], Ping Yu[b]

[a] *School of Statistics and Data Science, Qufu Normal University, Shandong, China*
[b] *HKU Business School, University of Hong Kong, Hong Kong*

Abstract

This supplementary material contains the proofs of Proposition 1, Theorems 1-3, Corollary 1, and simplifications in Case three and simulations.

Keywords: LAD mediation, Model consistency, Difference estimator, Product estimator, Second-order asymptotic,
2020 MSC: : 62D20, 62E20, 62F12

---

*Corresponding author
Email addresses: wengewsh@sina.com (WenWu Wang), pingyu@hku.hk (Ping Yu)

Appendix A: Proof of Proposition 1

Lemma 1. *If ϵ_{2i} and ϵ_{3i} are independent and symmetrically distributed about* 0*, then $\epsilon_i = b\epsilon_{2i} + \epsilon_{3i}$ for any b is symmetrically distributed about* 0*.*

Proof. Denote the pdfs of ϵ_{2i} and ϵ_{3i} as $f_{\epsilon_2}(\cdot)$ and $f_{\epsilon_3}(\cdot)$, and the cdfs as $F_{\epsilon_2}(\cdot)$ and $F_{\epsilon_3}(\cdot)$. For any δ, the equality $\mathrm{pr}(\epsilon_i \le \delta) = \mathrm{pr}(\epsilon_i \ge -\delta)$ implies our conclusion. Now,

$$
\begin{aligned}
pr(\epsilon_i \le \delta) &= \int_{-\infty}^{\infty} f_{\epsilon_2}(\epsilon_{2i}) \int_{-\infty}^{\delta - b\epsilon_{2i}} f_{\epsilon_3}(\epsilon_{3i}) \mathrm{d}\epsilon_{2i} \mathrm{d}\epsilon_{3i} \\
&= \int_{-\infty}^{\infty} f_{\epsilon_2}(\epsilon_{2i}) F_{\epsilon_3}(\delta - b\epsilon_{2i}) \mathrm{d}\epsilon_{2i} = \int_{\infty}^{-\infty} f_{\epsilon_2}(-\epsilon_{2i}) F_{\epsilon_3}(\delta - b(-\epsilon_{2i})) \mathrm{d}(-\epsilon_{2i}) \\
&= \int_{-\infty}^{\infty} f_{\epsilon_2}(\epsilon_{2i}) F_{\epsilon_3}(\delta + b\epsilon_{2i}) \mathrm{d}\epsilon_{2i} = \int_{-\infty}^{\infty} f_{\epsilon_2}(\epsilon_{2i}) \{1 - F_{\epsilon_3}(-\delta - b\epsilon_{2i})\} \mathrm{d}\epsilon_{2i} \\
&= \int_{-\infty}^{\infty} f_{\epsilon_2}(\epsilon_{2i}) \int_{-\delta - b\epsilon_{2i}}^{\infty} f_{\epsilon_3}(\epsilon_{3i}) \mathrm{d}\epsilon_{3i} \mathrm{d}\epsilon_{2i} = \mathrm{pr}(\epsilon_i \ge -\delta)
\end{aligned}
$$

This completes the proof of Lemma 1.

Proof. For case (i), we condition on X_i everywhere. Since ϵ_{2i} and ϵ_{3i} have symmetric densities about 0, ϵ_i has a symmetric density by Lemma 1 with $\mathrm{Med}(\epsilon_i \mid X_i) = 0$. Taking median of Equations (1) and (4), we have

$$
\begin{aligned}
\mathrm{Med}(Y_i \mid X_i) &= \beta_1 + cX_i, \\
\mathrm{Med}(Y_i \mid X_i) &= (\beta_3 + b\beta_2) + (c' + ab)X_i.
\end{aligned}
$$

Thus, $ab = c - c'$. In addition, we have $\beta_1 = \beta_3 + b\beta_2$ and $\epsilon_i = \epsilon_{1i}$.

In case (ii), $\mathrm{Med}(\epsilon_i \mid X_i) = \mathrm{Med}(\epsilon_i) \equiv d$. So taking median of Equations (1) and (4) , we have

$$
\begin{aligned}
\mathrm{Med}(Y_i \mid X_i) &= \beta_1 + cX_i, \\
\mathrm{Med}(Y_i \mid X_i) &= (\beta_3 + b\beta_2 + d) + (c' + ab)\, X_i.
\end{aligned}
$$

Although $\beta_1 \ne \beta_3 + b\beta_2$, we still have $c - c' = ab$. Note also that $\epsilon_{1i} = \epsilon_i - d$.

Appendix B: Proof of Theorem 1

PROOF. First, note that Condition 1(iv) implies $E\left[|M|^2\right] < \infty$ by Cauchy-Schwarz inequality. As a result, $E\left[\|\mathbf{x}_2\|^2\right] < \infty$, and the conditions required for the first-order expansion are met, see, e.g., Pollard (1991) and Knight (1998b). Specifically, if we define $\mathbf{x}_1^{\mathrm{T}} = (1, X)$, $\mathbf{x}_2^{\mathrm{T}} = (1, X, M)$, and $s(\epsilon_{ki}) = \{1/2 - 1(\epsilon_{ki} \le 0)\}/f_{\epsilon_k}(0)$ for $k = 1, 2, 3$, we have

$$
\begin{aligned}
\sqrt{n}\begin{pmatrix} \widehat{\beta}_1 - \beta_1 \\ \widehat{c} - c \end{pmatrix} &\approx E\left[f_{\epsilon_1|X}(0 \mid X)\,\mathbf{x}_1\mathbf{x}_1^{\mathrm{T}}\right]^{-1} n^{-1/2}\sum_{i=1}^{n}\mathbf{x}_{1i}\left\{\frac{1}{2} - 1(\epsilon_{1i} \le 0)\right\}, \\
&= n^{-1/2}\sum_{i=1}^{n} E\left[\mathbf{x}_1\mathbf{x}_1^{\mathrm{T}}\right]^{-1}\mathbf{x}_{1i}s(\epsilon_{1i}), \\
n^{1/2}\begin{pmatrix} \widehat{\beta}_2 - \beta_2 \\ \widehat{a} - a \end{pmatrix} &\approx E\left[f_{\epsilon_2|X}(0 \mid X)\,\mathbf{x}_1\mathbf{x}_1^{\mathrm{T}}\right]^{-1} n^{-1/2}\sum_{i=1}^{n}\mathbf{x}_{1i}\left\{\frac{1}{2} - 1(\epsilon_{2i} \le 0)\right\} \\
&= n^{-1/2}\sum_{i=1}^{n} E\left[\mathbf{x}_1\mathbf{x}_1^{\mathrm{T}}\right]^{-1}\mathbf{x}_{1i}s(\epsilon_{2i}), \\
n^{1/2}\begin{pmatrix} \widehat{\beta}_3 - \beta_3 \\ \widehat{c'} - c' \\ \widehat{b} - b \end{pmatrix} &\approx E\left[f_{\epsilon_3|X,M}(0 \mid X, M)\,\mathbf{x}_2\mathbf{x}_2^{\mathrm{T}}\right]^{-1} n^{-1/2}\sum_{i=1}^{n}\mathbf{x}_{2i}\left\{\frac{1}{2} - 1(\epsilon_{3i} \le 0)\right\} \\
&= n^{-1/2}\sum_{i=1}^{n} E\left[\mathbf{x}_2\mathbf{x}_2^{\mathrm{T}}\right]^{-1}\mathbf{x}_{2i}s(\epsilon_{3i}),
\end{aligned}
$$

where $\epsilon_1 = b\epsilon_2 + \epsilon_3 - d$ if $d \neq 0$. We are interested in $\widehat{c}, \widehat{c'}, \widehat{a}$ and $\widehat{b}$, so the key is to express $E\left[\mathbf{x}_1\mathbf{x}_1^{\mathrm{T}}\right]^{-1}\mathbf{x}_1$ and $E\left[\mathbf{x}_2\mathbf{x}_2^{\mathrm{T}}\right]^{-1}\mathbf{x}_2$ in a convenient way. For this purpose, we conduct the following transformations on $\mathbf{x}_1$ and $\mathbf{x}_2$:

$$
\begin{aligned}
\mathbf{C}_1\mathbf{x}_1 &= \begin{pmatrix} 1 & 0 \\ -\mu_X & 1 \end{pmatrix}\begin{pmatrix} 1 \\ X \end{pmatrix} = \begin{pmatrix} 1 \\ \widetilde{X} \end{pmatrix} \equiv \widetilde{\mathbf{x}}_1, \\
\mathbf{C}_2\mathbf{x}_2 &= \begin{pmatrix} 1 & 0 & 0 \\ -\mu_X & 1 & 0 \\ -\beta_2 - \mu_{\epsilon_2} & -a & 1 \end{pmatrix}\begin{pmatrix} 1 \\ X \\ M \end{pmatrix} = \begin{pmatrix} 1 \\ \widetilde{X} \\ \widetilde{\epsilon}_2 \end{pmatrix} \equiv \widetilde{\mathbf{x}}_2;
\end{aligned}
$$

where the "tilde" variable is the demeaned version of the original variable, and then

$$\begin{aligned}
E\left[\mathbf{x}_1\mathbf{x}_1^{\mathrm{T}}\right]^{-1}\mathbf{x}_1 &= \mathbf{C}_1^{\mathrm{T}}E\left[\mathbf{C}_1\mathbf{x}_1\mathbf{x}_1^{\mathrm{T}}\mathbf{C}_1^{\mathrm{T}}\right]^{-1}\mathbf{C}_1\mathbf{x}_1 \\
&= \begin{pmatrix} 1 & -\mu_X \\ 0 & 1 \end{pmatrix} E\left[\widetilde{\mathbf{x}}_1\widetilde{\mathbf{x}}_1^{\mathrm{T}}\right]^{-1}\widetilde{\mathbf{x}}_1 \\
&= \begin{pmatrix} 1 & -\mu_X \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \sigma_X^2 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ \widetilde{X} \end{pmatrix} \\
&= \begin{pmatrix} 1 & -\mu_X \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \widetilde{X}/\sigma_X^2 \end{pmatrix} = \begin{pmatrix} 1 - \frac{\mu_X}{\sigma_X^2}\widetilde{X} \\ \frac{\widetilde{X}}{\sigma_X^2} \end{pmatrix},
\end{aligned}$$

and similarly,

$$E\left[\mathbf{x}_2\mathbf{x}_2^{\mathrm{T}}\right]^{-1}\mathbf{x}_2 = \begin{pmatrix} 1 & -\mu_X & -\beta_2-\mu_{\epsilon_2} \\ 0 & 1 & -a \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \widetilde{X}/\sigma_X^2 \\ \widetilde{\epsilon}_2/\sigma_2^2 \end{pmatrix} = \begin{pmatrix} 1 - \frac{\mu_X}{\sigma_X^2}\widetilde{X} - \frac{\beta_2+\mu_{\epsilon_2}}{\sigma_2^2}\widetilde{\epsilon}_2 \\ \widetilde{X}/\sigma_X^2 - a\widetilde{\epsilon}_2/\sigma_2^2 \\ \widetilde{\epsilon}_2/\sigma_2^2 \end{pmatrix}.$$

As a result, we have the following first-order expansions:

$$\begin{aligned}
n^{1/2}\left(\widehat{c}-c\right) &\approx n^{-1/2}\sum_{i=1}^{n}\frac{\widetilde{X}_i}{\sigma_X^2}s\left(\epsilon_{1i}\right), \\
n^{1/2}\left(\widehat{a}-a\right) &\approx n^{-1/2}\sum_{i=1}^{n}\frac{\widetilde{X}_i}{\sigma_X^2}s\left(\epsilon_{2i}\right), \\
n^{1/2}(\widehat{c'}-c') &\approx n^{-1/2}\sum_{i=1}^{n}\left(\frac{\widetilde{X}}{\sigma_X^2}-a\frac{\widetilde{\epsilon}_2}{\sigma_2^2}\right)s\left(\epsilon_{3i}\right), \\
n^{1/2}(\widehat{b}-b) &\approx n^{-1/2}\sum_{i=1}^{n}\frac{\widetilde{\epsilon}_2}{\sigma_2^2}s\left(\epsilon_{3i}\right).
\end{aligned}$$

It follows that

$$n^{1/2}\{\widehat{c}-\widehat{c'}-(c-c')\} \approx \frac{1}{n^{1/2}}\sum_{i=1}^{n}\left\{\frac{\widetilde{X}_i s\left(\epsilon_{1i}\right)}{\sigma_X^2} - \left(\frac{\widetilde{X}}{\sigma_X^2}-a\frac{\widetilde{\epsilon}_2}{\sigma_2^2}\right)s\left(\epsilon_{3i}\right)\right\},$$

and

$$n^{1/2}(\widehat{a}\widehat{b}-ab) = n^{1/2}\widehat{b}(\widehat{a}-a)+n^{1/2}a(\widehat{b}-b) \approx \frac{1}{n^{1/2}}\sum_{i=1}^{n}\left\{\frac{b\widetilde{X}_i s\left(\epsilon_{2i}\right)}{\sigma_X^2}+a\frac{\widetilde{\epsilon}_2}{\sigma_2^2}s\left(\epsilon_{3i}\right)\right\}.$$

So

$$n^{1/2}(\widehat{c}-\widehat{c'}-\widehat{ab}) = \frac{1}{n^{1/2}}\frac{1}{\sigma_X^2}\sum_{i=1}^{n}\widetilde{X}_i\left[s\left(\epsilon_{1i}\right)-s\left(\epsilon_{3i}\right)-bs\left(\epsilon_{2i}\right)\right].$$

Next, we calculate the asymptotic variances of $\widehat{c}-\widehat{c'}$, $\widehat{ab}$ and $\widehat{c}-\widehat{c'}-\widehat{ab}$. The two terms in the first-order expansion are correlated in $\widehat{c}-\widehat{c'}$, but are not in $\widehat{ab}$. Specifically,

$$\begin{aligned}\text{Avar}(\widehat{c}-\widehat{c'}) &= \frac{1}{4f_{\epsilon_1}\left(0\right)^2\sigma_X^2}+\frac{1/\sigma_X^2+a^2/\sigma_2^2}{4f_{\epsilon_3}\left(0\right)^2}-\frac{2E\left[s\left(\epsilon_1\right)s\left(\epsilon_3\right)\right]}{\sigma_X^2},\\ \text{Avar}(\widehat{ab}) &= \frac{b^2}{4f_{\epsilon_2}\left(0\right)^2\sigma_X^2}+\frac{a^2}{4f_{\epsilon_3}\left(0\right)^2\sigma_2^2}\end{aligned}$$

where $\text{Avar}(X_n)$ is the asymptotic variance of a generic sequence of random variables X_n, and in $E\left[s\left(\epsilon_1\right)s\left(\epsilon_3\right)\right]$,

$$\begin{aligned}E\left[\left\{\frac{1}{2}-1\left(\epsilon_1\leq 0\right)\right\}\left\{\frac{1}{2}-1\left(\epsilon_3\leq 0\right)\right\}\right] &= \frac{1}{4}-\frac{1}{4}-\frac{1}{4}+E\left\{1\left(\epsilon_1\leq 0\right)1\left(\epsilon_3\leq 0\right)\right\}\\ &= E\left\{1\left(b\epsilon_2+\epsilon_3\leq d\right)1\left(\epsilon_3\leq 0\right)\right\}-\frac{1}{4},\end{aligned}$$

with

$$\begin{aligned}&E\left[1\left(b\epsilon_2+\epsilon_3\leq d\right)1\left(\epsilon_3\leq 0\right)\right]\\ =\ &E\left[1\left(\epsilon_3\leq -b\epsilon_2+d\right)1\left(\epsilon_3\leq 0\right)1\left(\epsilon_2<\frac{d}{b}\right)\right]+E\left[1\left(\epsilon_{3i}\leq -b\epsilon_2+d\right)1\left(\epsilon_3\leq 0\right)1\left(\epsilon_2\geq\frac{d}{b}\right)\right]\\ =\ &\begin{cases}\frac{1}{2}F_{\epsilon_2}\left(\frac{d}{b}\right)+\int_{d/b}^{\infty}F_{\epsilon_3}\left(-b\epsilon_2+d\right)\mathrm{d}F_{\epsilon_2}\left(\epsilon_2\right), & \text{if } b>0,\\ \frac{1}{2}\left(1-F_{\epsilon_2}\left(\frac{d}{b}\right)\right)+\int_{-\infty}^{d/b}F_{\epsilon_3}\left(-b\epsilon_2+d\right)\mathrm{d}F_{\epsilon_2}\left(\epsilon_2\right), & \text{if } b<0.\end{cases}\end{aligned}$$

Finally,

$$\begin{aligned}&\text{AVar}(\widehat{c}-\widehat{c'}-\widehat{ab})\\ =\ &\frac{1}{\sigma_X^4}\left[\frac{\sigma_X^2}{4f_{\epsilon_1}\left(0\right)^2}+\frac{\sigma_X^2}{4f_{\epsilon_3}\left(0\right)^2}+\frac{b^2\sigma_X^2}{4f_{\epsilon_2}\left(0\right)^2}-2\frac{\sigma_X^2\left\{E\left[1\left(\epsilon_1\leq 0\right)1\left(\epsilon_3\leq 0\right)\right]-1/4\right\}}{f_{\epsilon_1}\left(0\right)f_{\epsilon_3}\left(0\right)}\right.\\ &\left.-2\frac{b\sigma_X^2\left\{E\left[1\left(\epsilon_1\leq 0\right)1\left(\epsilon_2\leq 0\right)\right]-1/4\right\}}{f_{\epsilon_1}\left(0\right)f_{\epsilon_2}\left(0\right)}\right]\\ =\ &\frac{1}{\sigma_X^2}\left[\frac{1}{4f_{\epsilon_1}\left(0\right)^2}+\frac{1}{4f_{\epsilon_3}\left(0\right)^2}+\frac{b^2}{4f_{\epsilon_2}\left(0\right)^2}-2\frac{E\left[1\left(\epsilon_1\leq 0\right)1\left(\epsilon_3\leq 0\right)\right]-1/4}{f_{\epsilon_1}\left(0\right)f_{\epsilon_3}\left(0\right)}\right.\\ &\left.-2\frac{b\left\{E\left[1\left(\epsilon_1\leq 0\right)1\left(\epsilon_2\leq 0\right)\right]-1/4\right\}}{f_{\epsilon_1}\left(0\right)f_{\epsilon_2}\left(0\right)}\right],\end{aligned}$$

where

$$
\begin{aligned}
E\left[\widetilde{X}s(\epsilon_1)\,\widetilde{X}s(\epsilon_3)\right] &= \frac{\sigma_X^2\left\{E\left[1(\epsilon_1\le 0)\,1(\epsilon_3\le 0)\right]-1/4\right\}}{f_{\epsilon_1}(0)\,f_{\epsilon_3}(0)}, \\
E\left\{\widetilde{X}s(\epsilon_1)\,b\widetilde{X}s(\epsilon_2)\right\} &= \frac{b\sigma_X^2\left\{E\left[1(\epsilon_1\le 0)\,1(\epsilon_2\le 0)\right]-1/4\right\}}{f_{\epsilon_1}(0)\,f_{\epsilon_2}(0)}, \\
E\left[\widetilde{X}s(\epsilon_3)\,b\widetilde{X}s(\epsilon_2)\right] &= 0
\end{aligned}
$$

with

$$
\begin{aligned}
E\left[1(\epsilon_1\le 0)\,1(\epsilon_2\le 0)\right] &= E\left[1(\epsilon_3\le -b\epsilon_2+d)\,1(\epsilon_2\le 0)\right] \\
&= \int_{-\infty}^{0} F_{\epsilon_3}(-b\epsilon_2+d)\,\mathrm{d}F_{\epsilon_2}(\epsilon_2).
\end{aligned}
$$

Appendix C: Proof of Corollary 1

Proof. When ϵ_2 and ϵ_3 both follow $N(0,1)$, $d=0$, so the difference of the asymptotic variances of $\widehat{c}-\widehat{c'}$ and $\widehat{ab}$ is $1/\sigma_X^2$ times

$$
\begin{aligned}
&\frac{1}{4f_{\epsilon_1}(0)^2}+\frac{1}{4f_{\epsilon_3}(0)^2}-\frac{b^2}{4f_{\epsilon_2}(0)^2}-\frac{2\int_0^\infty F_{\epsilon_3}(-|b|\,\epsilon_2)\,\mathrm{d}F_{\epsilon_2}(\epsilon_2)}{f_{\epsilon_1}(0)\,f_{\epsilon_3}(0)} \\
=\;& \frac{b^2+1}{4/(2\pi)}+\frac{1}{4/(2\pi)}-\frac{b^2}{4/(2\pi)}-\frac{\sqrt{b^2+1}}{1/(2\pi)}\mathrm{pr}\left(\epsilon_3\le -|b|\,\epsilon_2\mid\epsilon_2>0\right) \\
=\;& \pi-2\pi\sqrt{b^2+1}P\left(\epsilon_3\le -|b|\,\epsilon_2\mid\epsilon_2>0\right) \\
=\;& \pi-2\pi\sqrt{b^2+1}\left(\frac{1}{2}-\frac{1}{\pi}\arctan|b|\right) \\
=\;& \pi-\frac{\sqrt{b^2+1}}{|b|}\frac{\pi-2\arctan|b|}{1/|b|},
\end{aligned}
$$

and by L'Hôpital's rule,

$$
\lim_{|b|\to\infty}\pi-\frac{\sqrt{b^2+1}}{|b|}\frac{\pi-2\arctan|b|}{1/|b|}=\lim_{|b|\to\infty}\pi-\frac{-2/(1+b^2)}{-1/b^2}=\pi-2,
$$

where

$$\begin{aligned}\frac{\mathrm{d}\int_0^\infty F_{\epsilon_3}\left(-\left|b\right|\epsilon_2\right)\mathrm{d}F_{\epsilon_2}\left(\epsilon_2\right)}{\mathrm{d}\left|b\right|} &= \frac{\mathrm{d}\int_0^\infty \Phi\left(-\left|b\right|\epsilon_2\right)\mathrm{d}\Phi\left(\epsilon_2\right)}{\mathrm{d}\left|b\right|} = -\int_0^\infty \epsilon_2\phi_{\epsilon_3}\left(-\left|b\right|\epsilon_2\right)\phi_{\epsilon_2}\left(\epsilon_2\right)\mathrm{d}\epsilon_2 \\ &= -\frac{1}{\sqrt{2\pi\left(b^2+1\right)}}\frac{1}{\sqrt{2\pi/\left(b^2+1\right)}}\int_0^\infty \epsilon_2\exp\left(-\frac{b^2+1}{2}\epsilon_2^2\right)\mathrm{d}\epsilon_2 \\ &= -\frac{1}{\sqrt{2\pi\left(b^2+1\right)}}\frac{1}{\sqrt{b^2+1}}\phi\left(0\right) = -\frac{1}{2\pi\left(b^2+1\right)}\end{aligned}$$

with $\phi\left(\cdot\right)$ an $\Phi\left(\cdot\right)$ being the pdf and cdf of the standard normal distribution, respectively, so

$$\begin{aligned}\int_0^\infty F_{\epsilon_3}\left(-\left|b\right|\epsilon_2\right)\mathrm{d}F_{\epsilon_2}\left(\epsilon_2\right) &= -\int_0^{\left|b\right|}\frac{1}{2\pi\left(b^2+1\right)}\mathrm{d}b + \int_0^\infty \Phi\left(0\right)\phi_{\epsilon_2}\left(\epsilon_2\right)\mathrm{d}\epsilon_2 \\ &= \frac{1}{4} - \frac{1}{2\pi}\arctan\left|b\right|.\end{aligned}$$

Appendix D: Proof of Theorem 2

Proof. The asymptotic distributions of $\widehat{c}-\widehat{c'}$ and $\widehat{a}\widehat{b}$ when $b=0$ but $a\neq 0$ are implied by Theorem 1. Specifically, the asymptotic variance of $\widehat{c}-\widehat{c'}$ is

$$\frac{1}{4f_{\epsilon_3}\left(0\right)^2\sigma_X^2} + \frac{1/\sigma_X^2 + a^2/\sigma_2^2}{4f_{\epsilon_3}\left(0\right)^2} - \frac{1}{2f_{\epsilon_3}\left(0\right)^2\sigma_X^2} = \frac{a^2}{4f_{\epsilon_3}\left(0\right)^2\sigma_2^2},$$

which is the same as that of $\widehat{a}\widehat{b}$. As the asymptotic variance of $\widehat{c}-\widehat{c'}-\widehat{a}\widehat{b}$ degenerates to zero, we next refine its asymptotic distribution.

First of all, for a general LAD regression, $y_i = \mathbf{x}_i^{\mathrm{T}}\beta+\epsilon_i$, where $\mathbf{x}_i\in\mathbb{R}^k$, ϵ_i is independent of $\mathbf{x}_i$ with $\mathrm{Med}\left(\epsilon_i\right)=0$, $E(\left[\left|\mathbf{x}_i\right|^3\right]<\infty$ and f_ϵ are differentiable at 0, Theorem 3 of Knight (1997) (see also Knight (1998a)) shows that

$$\begin{aligned}& n^{1/4}\left\{n^{1/2}\left(\widehat{\beta}-\beta\right) - E\left[\mathbf{x}\mathbf{x}^{\mathrm{T}}\right]^{-1}\frac{1}{n^{1/2}}\sum_{i=1}^n \mathbf{x}_i s\left(\epsilon_i\right)\right\} \\ = \quad & -f_\epsilon\left(0\right)^{-1}E\left[\mathbf{x}\mathbf{x}^{\mathrm{T}}\right]^{-1}n^{-1/4}\sum_{i=1}^n \mathbf{x}_i\left\{1(\epsilon_i\leq Z^{\mathrm{T}}\mathbf{x}_i/n^{1/2}) - 1(\epsilon_i\leq 0) - f_\epsilon\left(0\right)Z^{\mathrm{T}}\mathbf{x}_i/n^{1/2}\right\} + o_p(1) \\ \Longrightarrow \quad & -f_\epsilon\left(0\right)^{-1}E\left[\mathbf{x}\mathbf{x}^{\mathrm{T}}\right]^{-1}D(Z),\end{aligned}$$

where $D(\mathbf{u})$ is a zero-mean Gaussian process (independent of Z) with $D(\mathbf{0}) = \mathbf{0}$ and

$$E\left[\left(D(\mathbf{u}) - D(\mathbf{v})\right)\left(D(\mathbf{u}) - D(\mathbf{v})\right)^{\mathrm{T}}\right] = f_{\epsilon}(0)\, E\left[\mathbf{x}\mathbf{x}^{\mathrm{T}} \left|(\mathbf{u}-\mathbf{v})^{\mathrm{T}}\mathbf{x}\right|\right]$$

for $\mathbf{u}, \mathbf{v} \in \mathbb{R}^k$, and $Z \sim N(\mathbf{0}, (4f_{\epsilon}(0)^2)^{-1} E\left[\mathbf{x}\mathbf{x}^{\mathrm{T}}\right]^{-1})$. We apply this general result to our case. Note that Condition 2(iv) implies $E(\left[|M|^3\right] < \infty$ by Hölder's inequality, so $E\left[\|\mathbf{x}_2\|^3\right] < \infty$ holds. Also, $\epsilon_1 = \epsilon_3$ when $b = 0$, so we require only the differentiability of ϵ_3.

First,

$$\begin{aligned}
& n^{1/4}\left[n^{1/2}\left\{\widehat{c} - \widehat{c'} - (c - c')\right\} - \frac{1}{n^{1/2}}\sum_{i=1}^{n}\frac{a\widetilde{\epsilon}_{2i}}{\sigma_2^2}s\left(\epsilon_{3i}\right)\right] \\
= \quad & -\frac{(0,1)}{f_{\epsilon_3}(0)}n^{-1/4}\sum_{i=1}^{n} E\left[\mathbf{x}_1\mathbf{x}_1^{\mathrm{T}}\right]^{-1}\mathbf{x}_{1i}\left\{1(\epsilon_{3i} \le Z_1^{\mathrm{T}}\mathbf{x}_{1i}/n^{1/2}) - 1(\epsilon_{3i} \le 0) - f_{\epsilon_3}(0)\, Z_1^{\mathrm{T}}\mathbf{x}_{1i}/n^{1/2}\right\} \\
& + \frac{(0,1,0)}{f_{\epsilon_3}(0)}n^{-1/4}\sum_{i=1}^{n} E\left[\mathbf{x}_2\mathbf{x}_2^{\mathrm{T}}\right]^{-1}\mathbf{x}_{2i}\left\{1(\epsilon_{3i} \le Z_2^{\mathrm{T}}\mathbf{x}_{2i}/n^{1/2}) - 1(\epsilon_{3i} \le 0) - f_{\epsilon_3}(0)\, Z_2^{\mathrm{T}}\mathbf{x}_{2i}/n^{1/2}\right\} \\
\Longrightarrow \quad & \widetilde{D}_2\left(Z_2\right) - \widetilde{D}_1\left(Z_1\right),
\end{aligned}$$

which cannot be simplified if $a \neq 0$, where the covariance kernel of $\widetilde{D}_1(\cdot)$ is

$$\frac{1}{f_{\epsilon_3}(0)}E\left[\left(\frac{\widetilde{X}}{\sigma_X^2}\right)^2 \left|\mathbf{x}_1^{\mathrm{T}}(\mathbf{u}-\mathbf{v})\right|\right],$$

the covariance kernel of $\widetilde{D}_2(\cdot)$ is

$$\frac{1}{f_{\epsilon_3}(0)}E\left[\left(\frac{\widetilde{X}}{\sigma_X^2} - a\frac{\widetilde{\epsilon}_2}{\sigma_2^2}\right)^2 \left|\mathbf{x}_2^{\mathrm{T}}(\mathbf{u}-\mathbf{v})\right|\right],$$

$\widetilde{D}_1(\cdot)$ and $\widetilde{D}_2(\cdot)$ are correlated with the covariance kernel equal to

$$\begin{aligned}
& \lim_{n\to\infty} n^{1/2}\frac{1}{f_{\epsilon_3}(0)^2}E\Big[(0,1)E\left[\mathbf{x}_1\mathbf{x}_1^{\mathrm{T}}\right]^{-1}\mathbf{x}_{1i}\mathbf{x}_{2i}^{\mathrm{T}}E\left[\mathbf{x}_2\mathbf{x}_2^{\mathrm{T}}\right]^{-1}(0,1,0) \\
& \left\{1(\epsilon_{3i} \le \mathbf{u}^{\mathrm{T}}\mathbf{x}_{1i}/n^{1/2}) - 1(\epsilon_{3i} \le 0) - f_{\epsilon_3}(0)\,\mathbf{u}^{\mathrm{T}}\mathbf{x}_{1i}/n^{1/2}\right\} \\
& \left\{1(\epsilon_{3i} \le \mathbf{v}^{\mathrm{T}}\mathbf{x}_{2i}/n^{1/2}) - 1(\epsilon_{3i} \le 0) - f_{\epsilon_3}(0)\,\mathbf{v}^{\mathrm{T}}\mathbf{x}_{2i}/n^{1/2}\right\}\Big] \\
= \quad & \frac{1}{f_{\epsilon_3}(0)}E\left[\frac{\widetilde{X}}{\sigma_X^2}\left(\frac{\widetilde{X}}{\sigma_X^2} - a\frac{\widetilde{\epsilon}_2}{\sigma_2^2}\right)\left(|\mathbf{u}^{\mathrm{T}}\mathbf{x}_1| \wedge |\mathbf{v}^{\mathrm{T}}\mathbf{x}_2|\right)1(\mathbf{u}^{\mathrm{T}}\mathbf{x}_1\mathbf{x}_2^{\mathrm{T}}\mathbf{v} > 0)\right],
\end{aligned}$$

both $\widetilde{D}_1(\cdot)$ and $\widetilde{D}_2(\cdot)$ are independent of Z_1 and Z_2,

$$
\begin{aligned}
Z_1 &\sim \{2f_{\epsilon_3}(0)\}^{-1} N\left(\mathbf{0}, E\left[\mathbf{x}_1\mathbf{x}_1^{\mathrm{T}}\right]^{-1}\right) \equiv \frac{1}{2f_{\epsilon_3}(0)} N(\mathbf{0}, \Sigma_1), \\
Z_2 &\sim (2f_{\epsilon_3}(0))^{-1} N\left(\mathbf{0}, E\left[\mathbf{x}_2\mathbf{x}_2^{\mathrm{T}}\right]^{-1}\right) \equiv \frac{1}{2f_{\epsilon_3}(0)} N(\mathbf{0}, \Sigma_2), \\
E\left[Z_1 Z_2^{\mathrm{T}}\right] &= E\left[\mathbf{x}_1\mathbf{x}_1^{\mathrm{T}}\right]^{-1} E\left[\mathbf{x}_1\mathbf{x}_2^{\mathrm{T}} s(\epsilon_1) s(\epsilon_3)\right] E\left[\mathbf{x}_2\mathbf{x}_2^{\mathrm{T}}\right]^{-1} \\
&= \frac{1}{4f_{\epsilon_3}(0)^2} E\left[\mathbf{x}_1\mathbf{x}_1^{\mathrm{T}}\right]^{-1} E\left[\mathbf{x}_1\mathbf{x}_2^{\mathrm{T}}\right] E\left[\mathbf{x}_2\mathbf{x}_2^{\mathrm{T}}\right]^{-1} \equiv \frac{1}{4f_{\epsilon_3}(0)^2} \Sigma_{12},
\end{aligned}
$$

with

$$
\begin{aligned}
\Sigma_1 &= \mathbf{C}_1^{\mathrm{T}} E\left[\mathbf{C}_1\mathbf{x}_1\mathbf{x}_1^{\mathrm{T}}\mathbf{C}_1^{\mathrm{T}}\right]^{-1} \mathbf{C}_1 \\
&= \begin{pmatrix} 1 & -\mu_X \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \sigma_X^2 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ -\mu_X & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 + \frac{\mu_X^2}{\sigma_X^2} & -\frac{\mu_X}{\sigma_X^2} \\ -\frac{\mu_X}{\sigma_X^2} & \frac{1}{\sigma_X^2} \end{pmatrix},
\end{aligned}
$$

$$
\begin{aligned}
\Sigma_2 &= \mathbf{C}_2^{\mathrm{T}} E\left[\mathbf{C}_2\mathbf{x}_2\mathbf{x}_2^{\mathrm{T}}\mathbf{C}_2^{\mathrm{T}}\right]^{-1} \mathbf{C}_2 \\
&= \begin{pmatrix} 1 & -\mu_X & -\beta_2 - \mu_{\epsilon_2} \\ 0 & 1 & -a \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sigma_X^2} & 0 \\ 0 & 0 & \frac{1}{\sigma_2^2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -\mu_X & 1 & 0 \\ -\beta_2 - \mu_{\epsilon_2} & -a & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 + \frac{\mu_X^2}{\sigma_X^2} + \frac{\left(\beta_2 + \mu_{\epsilon_2}\right)^2}{\sigma_2^2} & a\frac{\beta_2 + \mu_{\epsilon_2}}{\sigma_2^2} - \frac{\mu_X}{\sigma_X^2} & -\frac{\beta_2 + \mu_{\epsilon_2}}{\sigma_2^2} \\ a\frac{\beta_2 + \mu_{\epsilon_2}}{\sigma_2^2} - \frac{\mu_X}{\sigma_X^2} & \frac{1}{\sigma_X^2} + \frac{a^2}{\sigma_2^2} & -\frac{a}{\sigma_2^2} \\ -\frac{\beta_2 + \mu_{\epsilon_2}}{\sigma_2^2} & -\frac{a}{\sigma_2^2} & \frac{1}{\sigma_2^2} \end{pmatrix}
\end{aligned}
$$

and

$$
\begin{aligned}
\Sigma_{12} &= E\left[\begin{pmatrix} 1 - \frac{\mu_X}{\sigma_X^2}\widetilde{X} \\ \frac{\widetilde{X}}{\sigma_X^2} \end{pmatrix} \begin{pmatrix} 1 - \frac{\mu_X}{\sigma_X^2}\widetilde{X} - \frac{\beta_2 + \mu_{\epsilon_2}}{\sigma_2^2}\widetilde{\epsilon}_2 \\ \widetilde{X}/\sigma_X^2 - a\widetilde{\epsilon}_2/\sigma_2^2 \\ \widetilde{\epsilon}_2/\sigma_2^2 \end{pmatrix} \right] \\
&= \begin{pmatrix} 1 + \frac{\mu_X^2}{\sigma_X^2} & -\frac{\mu_X}{\sigma_X^2} & 0 \\ -\frac{\mu_X}{\sigma_X^2} & \frac{1}{\sigma_X^2} & 0 \end{pmatrix}.
\end{aligned}
$$

Second, because

$$\begin{aligned}
n^{1/2}(\widehat{ab}-ab) &= n^{1/2}\widehat{b}(\widehat{a}-a)+n^{1/2}a(\widehat{b}-b) \\
&= n^{1/2}O_p(n^{-1/2})\{O_p(n^{-1/2})+O_p(n^{-3/4})\}+n^{1/2}a\{O_p(n^{-1/2})+O_p(n^{-3/4})\} \\
&= aO_p(1)+aO_p(n^{-1/4})+o_p(n^{-1/4}),
\end{aligned}$$

we have

$$\begin{aligned}
& n^{1/4}\left\{n^{1/2}\left(\widehat{ab}-ab\right)-\frac{1}{n^{1/2}}\sum_{i=1}^{n}\frac{a\widetilde{\epsilon}_{2i}}{\sigma_2^2}s\left(\epsilon_{3i}\right)\right\} \\
= \;& a\frac{(0,0,1)}{f_{\epsilon_3}(0)}n^{-1/4}\sum_{i=1}^{n}\Big[E\left[\mathbf{x}_2\mathbf{x}_2^{\mathrm{T}}\right]^{-1}\mathbf{x}_{2i}\{1(\epsilon_{3i}\le Z_2^{\mathrm{T}}\mathbf{x}_{2i}/n^{1/2}) \\
& -1(\epsilon_{3i}\le 0)-f_{\epsilon_3}(0)\,Z_2^{\mathrm{T}}\mathbf{x}_{2i}/n^{1/2}\}\Big]+o_p(1) \\
\Longrightarrow \;& a\overline{D}_2(\cdot),
\end{aligned}$$

where the covariance kernel of $\overline{D}_2(\cdot)$ is

$$\frac{1}{f_{\epsilon_3}(0)}E\left[\left(\frac{\widetilde{\epsilon}_2}{\sigma_2^2}\right)^2\left(|\mathbf{u}^{\mathrm{T}}\mathbf{x}_2|\wedge|\mathbf{v}^{\mathrm{T}}\mathbf{x}_2|\right)1(\mathbf{u}^{\mathrm{T}}\mathbf{x}_2\mathbf{x}_2^{\mathrm{T}}\mathbf{v}>0)\right],$$

the covariance kernel between $\widetilde{D}_1(\cdot)$ and $\overline{D}_2(\cdot)$ is

$$\frac{1}{f_{\epsilon_3}(0)}E\left[\frac{\widetilde{X}}{\sigma_X^2}\frac{\widetilde{\epsilon}_2}{\sigma_2^2}\left(|\mathbf{u}^{\mathrm{T}}\mathbf{x}_1|\wedge|\mathbf{v}^{\mathrm{T}}\mathbf{x}_2|\right)1(\mathbf{u}^{\mathrm{T}}\mathbf{x}_1\mathbf{x}_2^{\mathrm{T}}\mathbf{v}>0)\right],$$

and between $\widetilde{D}_2(\cdot)$ and $\overline{D}_2(\cdot)$ is

$$\frac{1}{f_{\epsilon_3}(0)}E\left[\left(\frac{\widetilde{X}}{\sigma_X^2}-a\frac{\widetilde{\epsilon}_2}{\sigma_2^2}\right)\frac{\widetilde{\epsilon}_2}{\sigma_2^2}\left|\mathbf{x}_2^{\mathrm{T}}(\mathbf{u}-\mathbf{v})\right|\right],$$

In summary,

$$n^{3/4}\left(\widehat{c}-\widehat{c'}-\widehat{ab}\right)\Longrightarrow\widetilde{D}_2(Z_2)-\widetilde{D}_1(Z_1)-a\overline{D}_2(Z_2),$$

where $\widetilde{D}_2(\cdot),\widetilde{D}_1(\cdot)$ and $\overline{D}_2(\cdot)$ are correlated zero-mean Gaussian processes.

Appendix E: Proof of Theorem 3

Proof. From Theorem 1, we have

$$\begin{aligned} n\widehat{a}\widehat{b} &= n^{1/2}\left(\widehat{a}-a\right)n^{1/2}\left(\widehat{b}-b\right) \\ &\Longrightarrow N\left(0,\frac{1}{4f_{\epsilon_2}\left(0\right)^2\sigma_X^2}\right)N\left(0,\frac{1}{4f_{\epsilon_3}\left(0\right)^2\sigma_2^2}\right)=\frac{1}{4f_{\epsilon_2}\left(0\right)f_{\epsilon_3}\left(0\right)\sigma_X\sigma_2}z_1z_2. \end{aligned}$$

From Theorem 2,

$$n^{3/4}\left(\widehat{c}-\widehat{c'}\right)=n^{3/4}\left\{\widehat{c}-\widehat{c'}-(c-c')\right\}\Longrightarrow \widetilde{D}_2\left(Z_2\right)-\widetilde{D}_1\left(Z_1\right),$$

where the covariance kernel of $\widetilde{D}_1\left(\cdot\right)$ remains as

$$\frac{1}{f_{\epsilon_3}\left(0\right)}E\left[\left(\frac{\widetilde{X}}{\sigma_X^2}\right)^2\left|\mathbf{x}_1^{\mathrm{T}}(\mathbf{u}-\mathbf{v})\right|\right],$$

the covariance kernel of $\widetilde{D}_2\left(\cdot\right)$ reduces to

$$\frac{1}{f_{\epsilon_3}\left(0\right)}E\left[\left(\frac{\widetilde{X}}{\sigma_X^2}\right)^2\left|\mathbf{x}_2^{\mathrm{T}}(\mathbf{u}-\mathbf{v})\right|\right],$$

and the covariance kernel between $\widetilde{D}_1\left(\cdot\right)$ and $\widetilde{D}_2\left(\cdot\right)$ reduces to

$$=\frac{1}{f_{\epsilon_3}\left(0\right)}E\left[\left(\frac{\widetilde{X}}{\sigma_X^2}\right)^2\left(|\mathbf{u}^{\mathrm{T}}\mathbf{x}_1|\wedge|\mathbf{v}^{\mathrm{T}}\mathbf{x}_2|\right)1(\mathbf{u}^{\mathrm{T}}\mathbf{x}_1\mathbf{x}_2^{\mathrm{T}}\mathbf{v}>0)\right].$$

Finally,

$$n^{3/4}(\widehat{c}-\widehat{c'}-\widehat{a}\widehat{b})=n^{3/4}(\widehat{c}-\widehat{c'})-n^{-1/4}n\widehat{a}\widehat{b}=n^{3/4}(\widehat{c}-\widehat{c'})+o_p(1),$$

so $n^{3/4}(\widehat{c}-\widehat{c'}-\widehat{a}\widehat{b})$ has the same asymptotic distribution as $n^{3/4}(\widehat{c}-\widehat{c'})$.

When $a=0$, Σ_1 and Σ_{12} cannot be simplified, but Σ_2 reduces to

$$\begin{pmatrix} 1+\frac{\mu_X^2}{\sigma_X^2}+\frac{\left(\beta_2+\mu_{\epsilon_2}\right)^2}{\sigma_2^2} & -\frac{\mu_X}{\sigma_X^2} & -\frac{\beta_2+\mu_{\epsilon_2}}{\sigma_2^2} \\ -\frac{\mu_X}{\sigma_X^2} & \frac{1}{\sigma_X^2} & 0 \\ -\frac{\beta_2+\mu_{\epsilon_2}}{\sigma_2^2} & 0 & \frac{1}{\sigma_2^2} \end{pmatrix}.$$

Appendix F: Simplifications and Simulations in Case Three

In Case three, $a = b = 0$, and Σ_1, Σ_{12} and Σ_2 in Theorem 2 can be simplified as

$$\Sigma_1 = \begin{pmatrix} 1 + \frac{\mu_X^2}{\sigma_X^2} & -\frac{\mu_X}{\sigma_X^2} \\ -\frac{\mu_X}{\sigma_X^2} & \frac{1}{\sigma_X^2} \end{pmatrix}, \quad \Sigma_{12} = \begin{pmatrix} 1 + \frac{\mu_X^2}{\sigma_X^2} & -\frac{\mu_X}{\sigma_X^2} & 0 \\ -\frac{\mu_X}{\sigma_X^2} & \frac{1}{\sigma_X^2} & 0 \end{pmatrix},$$

$$\Sigma_2 = \begin{pmatrix} 1 + \frac{\mu_X^2}{\sigma_X^2} + \frac{\left(\beta_2+\mu_{\epsilon_2}\right)^2}{\sigma_2^2} & -\frac{\mu_X}{\sigma_X^2} & -\frac{\beta_2+\mu_{\epsilon_2}}{\sigma_2^2} \\ -\frac{\mu_X}{\sigma_X^2} & \frac{1}{\sigma_X^2} & 0 \\ -\frac{\beta_2+\mu_{\epsilon_2}}{\sigma_2^2} & 0 & \frac{1}{\sigma_2^2} \end{pmatrix}.$$

As a result, we can write $Z_2^{\mathrm{T}} = \left(Z_1^{\mathrm{T}} + (z_3, 0), z_2\right)$ with (z_2, z_3) independent of Z_1, $Z_1 \sim \frac{1}{2f_{\epsilon_3}(0)} N(\mathbf{0}, \Sigma_1)$, and

$$(z_2, z_3)^{\mathrm{T}} \sim \frac{1}{2f_{\epsilon_3}(0)\,\sigma_2} N\left(0, \begin{pmatrix} 1 & -\mu_M \\ -\mu_M & \mu_M^2 \end{pmatrix}\right),$$

where $\mu_M = \beta_2 + \mu_{\epsilon_2}$. In consequence,

$$\frac{1}{4f_{\epsilon_3}(0)^2}\Sigma = E\begin{pmatrix} Z_1Z_1^{\mathrm{T}} & Z_1Z_1^{\mathrm{T}} & \mathbf{0} \\ Z_1Z_1^{\mathrm{T}} & Z_1Z_1^{\mathrm{T}} + \begin{pmatrix} z_3^2 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} z_2z_3 \\ 0 \end{pmatrix} \\ \mathbf{0} & (z_2z_3, 0) & z_2^2 \end{pmatrix}$$

with $|\Sigma| = 0$. Also, when $\mu_X = 0$, the variance of $\widetilde{D}_2(Z_2) - \widetilde{D}_1(Z_1)$ reduces to

$$\{\sigma_X^2 f_{\epsilon_3}(0)\}^{-1} \int \Big\{ E\left[\overline{X}^2 |\mathbf{x}_1^{\mathrm{T}}\mathbf{z}_1|\right] + E\left[\overline{X}^2 |\mathbf{x}_2^{\mathrm{T}}\mathbf{z}_2|\right]$$
$$-2E\left[\overline{X}^2(|\mathbf{x}_1^{\mathrm{T}}\mathbf{z}_1| \wedge |\mathbf{x}_2^{\mathrm{T}}\mathbf{z}_2|)1(\mathbf{z}_1^{\mathrm{T}}\mathbf{x}_1\mathbf{x}_2^{\mathrm{T}}\mathbf{z}_2 > 0)\right]\Big\} f\left(\mathbf{z}_1, \mathbf{z}_2 \mid \mathbf{0}, \frac{1}{4f_{\epsilon_3}(0)^2}\Sigma\right) d\mathbf{z}_1 d\mathbf{z}_2,$$

where $\overline{X} = X/\sigma_X$ has mean zero and variance 1, and Σ can be further simplified. Note that assuming $\mu_X = 0$ does not lose generality, e.g., in Equation (1), $\beta_1 + cX = (\beta_1 + c\mu_X) + c\widetilde{X}$.

In our simulations, $\sigma_X^2 = 1$, $\sigma_2^2 = 1$, $\mu_X = 0$ and $\mu_M = 0$, so

$$\Sigma_1 = I_2, \quad \Sigma_2 = I_3, \quad \Sigma_{12} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

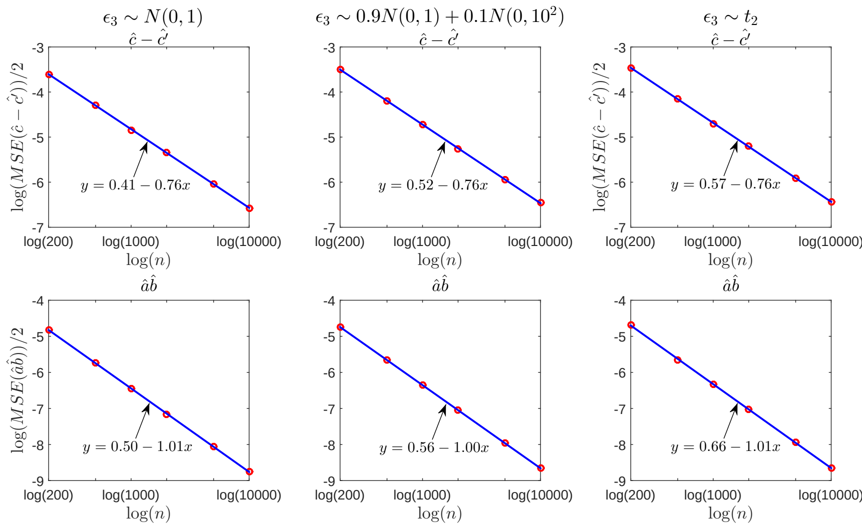

Figure 1: $\log\{MSE(\widehat{c}-\widehat{c'})\}/2$ and $\log\{MSE(\widehat{ab})\}/2$ against $\log n$ for $a=b=0$ and three ϵ_3 distributions.

As a result, the variance of $\widetilde{D}_2(Z_2)-\widetilde{D}_1(Z_1)$ in Case three is

$$\frac{E\left[X^2\left|\mathbf{x}_1^{\mathrm{T}}Z_1\right|\right]+E\left[X^2\left|\mathbf{x}_2^{\mathrm{T}}Z_2\right|\right]-2E\left[X^2\left(\left|\mathbf{x}_1^{\mathrm{T}}Z_1\right|\wedge\left|\mathbf{x}_2^{\mathrm{T}}Z_2\right|\right)1(Z_1^{\mathrm{T}}\mathbf{x}_1\mathbf{x}_2^{\mathrm{T}}Z_2>0)\right]}{2f_{\epsilon_3}(0)^2},$$

where $\mathbf{x}_1=(1,X)^{\mathrm{T}}$ and $\mathbf{x}_2=(1,X,M)^{\mathrm{T}}$ with $(X,M)\sim N(\mathbf{0},I_2)$, $Z_2^{\mathrm{T}}=(Z_1^{\mathrm{T}},z_2)$ with z_2 independent of Z_1, $Z_1\sim N(\mathbf{0},I_2)$, and $z_2\sim N(0,1)$, and Z_2 is independent of $\mathbf{x}_2$. By simulation,

$$E\left[X^2\left|\mathbf{x}_1^{\mathrm{T}}Z_1\right|\right]+E\left[X^2\left|\mathbf{x}_2^{\mathrm{T}}Z_2\right|\right]-2E\left[X^2\left(\left|\mathbf{x}_1^{\mathrm{T}}Z_1\right|\wedge\left|\mathbf{x}_2^{\mathrm{T}}Z_2\right|\right)1(Z_1^{\mathrm{T}}\mathbf{x}_1\mathbf{x}_2^{\mathrm{T}}Z_2>0)\right]=0.63.$$

We finally analyze the convergence rates of the two LAD estimates in Case three. Fig. 1 shows $\log\{MSE(\widehat{c}-\widehat{c'})\}/2$ and $\log\{MSE(\widehat{ab})\}/2$ against $\log n$ when $a=b=0$. Different from the common convergence rate $n^{1/2}$ of $\widehat{c}-\widehat{c'}$ and $\widehat{ab}$ in Cases one and two, the convergence rate of $\widehat{c}-\widehat{c'}$ is $n^{3/4}$ and that of $\widehat{ab}$ is n, which are clearly shown in Table 3. Also, the asymptotic variances of both $\widehat{c}-\widehat{c'}$ and $\widehat{ab}$ increase with the heaviness of ϵ_3's tail. Comparing Fig. 1 and Fig. 4, we can see that MSE$(\widehat{c}-\widehat{c'}-\widehat{ab})$ and MSE$(\widehat{c}-\widehat{c'})$ are close, which is because $\widehat{ab}$ has a faster convergence rate so its MSE in MSE$(\widehat{c}-\widehat{c'}-\widehat{ab})$ is neglectable.

Table 1: MSE ($\times 10^{-3}$) for LS and two LAD estimates with non-normal ϵ_2 error

		$a = b = 0.14$			$a = b = 1.4$		
ϵ_3	n	MSE_{LS}	MSE_P	MSE_D	MSE_{LS}	MSE_P	MSE_D
(I)	200	5.63E5	0.26	4.07	5.63E7	26.14	41.01
	1000	1.62E4	0.05	0.78	1.62E6	4.91	8.03
(II)	200	1.78E4	0.27	4.63	1.78E6	27.05	48.97
	1000	6.18E5	0.05	0.89	6.18E6	4.89	9.15
(III)	200	1.73E5	0.27	4.94	1.73E7	26.76	50.81
	1000	6.86E6	0.05	0.91	6.86E8	4.95	9.71
		$b = 0, a = 0.14$			$b = 0, a = 1.4$		
(I)	200	0.03	0.02E-1	0.36	0.10	0.11	0.44
	1000	0.01E-1	0.05E-4	0.02	0.04E-1	0.04E-1	0.02
(II)	200	0.27	0.00	0.43	1.04	0.21	0.65
	1000	0.01	0.00	0.02	0.05	0.01	0.03
(III)	200	0.28	0.00	0.50	1.69	0.22	0.71
	1000	0.01	0.00	0.02	0.07	0.01	0.03
		$a = b = 0$					
(I)	200	2.51	0.08	33.29			
	1000	0.10	0.00	1.72			
(II)	200	26.73	0.15	43.22			
	1000	1.10	0.00	2.31			
(III)	200	25.10	0.14	49.64			
	1000	1.24	0.00	2.39			

For each ϵ_3 distribution and n, we only list the finite-sample MSE based on 10000 replications, but not the MSE predicted by the asymptotic theory .

Appendix G: Simulation Results with Non-normal ϵ_2 Errors

Following the reviewer's comments, we set $\epsilon_2 \sim Laplace(0, 1)$ (non-normal distribution), and ϵ_3 follows $N(0; 1)$, $0.9N(0; 1) + 0.1N(0; 10^2)$, or t_2. The other settings are the same as those in Section 4.1 of the paper. Simulation results are provided in Table 1, which show that our conclusions still hold in these cases.

Acknowledgements

We thank Professor Keith Knight for providing us his papers on second-order asymptotics of quantile regression. WenWu Wang was supported by National Natural Science Foundation of China (No.12071248) and National Statistical Science Research Foundation of China (No.2020LZ26). Yu acknowledges support from the GRF of Hong Kong Government under Grant No. 106200228.