Manuscript (2020), **xx**, x, *pp.* 1–9

Supplementary material for "Nonequivalence of Two Least-Absolute-Deviation Estimators for Mediation Effect"

BY WENWU WANG

School of Statistics, Qufu Normal University, Jingxuan West Road, Qufu 273165, China

wengewsh@sina.com

PING YU

HKU Business School, University of Hong Kong, Pokfulam Road, Hong Kong

pingyu@hku.hk

SUMMARY

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

APPENDIX A: PROOF OF PROPOSITION 1

LEMMA S1. *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 $\mathrm{pr}(\epsilon_i \leq \delta) = \mathrm{pr}(\epsilon_i \geq -\delta)$ implies our conclusion. Now,

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

This completes the proof of Lemma S1. □

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 S1 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

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

Although $\beta_1 \neq \beta_3 + b\beta_2$, 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 $E(|M|^2) < \infty$ by Cauchy-Schwarz inequality. As a result, $E(\|\mathbf{x}_2\|^2) < \infty$, 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 $s\,(\epsilon_{ki}) = \{1/2 - 1\,(\epsilon_{1i} \leq 0)\}/f_{\epsilon_k}\,(0)$ for $k = 1, 2, 3$, we have

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

where for generic random variables X and Y, $\sigma_{XY} = \mathrm{cov}\,(X, Y)$, $\epsilon_1 = b\epsilon_2 + \epsilon_3 - d$ if $d \neq 0$, and

$$E(\mathbf{x}_2\mathbf{x}_2^{\mathrm{T}})^{-1} = (\sigma_X^2\sigma_2^2)^{-1}\begin{pmatrix} \mu_{X^2}\mu_{M^2} - \mu_{XM}^2 & \mu_M\mu_{XM} - \mu_X\mu_{M^2} & \mu_X\mu_{XM} - \mu_M\mu_{X^2} \\ \mu_M\mu_{XM} - \mu_X\mu_{M^2} & \sigma_M^2 & -\sigma_{XM} \\ \mu_X\mu_{XM} - \mu_M\mu_{X^2} & -\sigma_{XM} & \sigma_X^2 \end{pmatrix}.$$

From these first-order expansions, we have

$$\begin{aligned}
n^{1/2}\,(\widehat{c} - c) &\approx \frac{1}{\sigma_X^2} n^{-1/2}\sum_{i=1}^{n}\widetilde{X}_i s\,(\epsilon_{1i}), \\
n^{1/2}\,(\widehat{a} - a) &\approx \frac{1}{\sigma_X^2} n^{-1/2}\sum_{i=1}^{n}\widetilde{X}_i s\,(\epsilon_{2i}), \\
n^{1/2}(\widehat{c'} - c') &\approx \frac{1}{\sigma_X^2\sigma_2^2} n^{-1/2}\sum_{i=1}^{n}\left(\sigma_M^2 X_i - \sigma_{XM}M_i + \mu_M\mu_{XM} - \mu_X\mu_{M^2}\right) s\,(\epsilon_{3i}),
\end{aligned}$$

$$n^{1/2}(\widehat{b}-b) \approx \frac{1}{\sigma_X^2\sigma_2^2} n^{-1/2} \sum_{i=1}^{n} \left(\sigma_X^2 M_i - \sigma_{XM} X_i + \mu_X \mu_{XM} - \mu_M \mu_{X^2}\right) s\left(\epsilon_{3i}\right),$$

where $\widetilde{X}_i = X_i - \mu_X$. As a result,

$$\begin{aligned}
& n^{1/2}\{\widehat{c}-\widehat{c'}-\left(c-c'\right)\} \\
\approx\ & \frac{1}{n^{1/2}} \sum_{i=1}^{n} \left\{ \frac{\widetilde{X}_i s\left(\epsilon_{1i}\right)}{\sigma_X^2} - \frac{\left(\sigma_M^2 X_i - \sigma_{XM} M_i + \mu_M \mu_{XM} - \mu_X \mu_{M^2}\right) s\left(\epsilon_{3i}\right)}{\sigma_X^2 \sigma_2^2} \right\} \\
=\ & \frac{1}{n^{1/2}} \sum_{i=1}^{n} \left\{ \frac{\widetilde{X}_i s\left(\epsilon_{1i}\right)}{\sigma_X^2} - \frac{\left(\sigma_M^2 - a\sigma_{XM}\right) X_i - \sigma_{XM}\epsilon_{2i} + \left(\mu_M \mu_{XM} - \mu_X \mu_{M^2} - \sigma_{XM}\beta_2\right)}{f_{\epsilon_3}\left(0\right) \sigma_X^2 \sigma_2^2} s\left(\epsilon_{3i}\right) \right\} \\
=\ & \frac{1}{n^{1/2}} \sum_{i=1}^{n} \left\{ \frac{\widetilde{X}_i s\left(\epsilon_{1i}\right)}{\sigma_X^2} - \frac{\sigma_2^2 \widetilde{X}_i - a\sigma_X^2 \widetilde{\epsilon}_{2i}}{\sigma_X^2 \sigma_2^2} s\left(\epsilon_{3i}\right) \right\} = \frac{1}{n^{1/2}} \sum_{i=1}^{n} \left\{ \frac{\widetilde{X}_i}{\sigma_X^2} s\left(\epsilon_{1i}\right) - \left(\frac{\widetilde{X}_i}{\sigma_X^2} - \frac{a}{\sigma_2^2}\widetilde{\epsilon}_{2i} \right) s\left(\epsilon_{3i}\right) \right\}
\end{aligned}$$

and

$$\begin{aligned}
& n^{1/2}(\widehat{a}\widehat{b} - ab) = n^{1/2}\widehat{b}(\widehat{a}-a) + n^{1/2}a(\widehat{b}-b) \\
\approx\ & \frac{1}{n^{1/2}} \sum_{i=1}^{n} \left\{ \frac{b\widetilde{X}_i s\left(\epsilon_{2i}\right)}{\sigma_X^2} + \frac{a\left(\sigma_X^2 M_i - \sigma_{XM} X_i + \mu_X \mu_{XM} - \mu_M \mu_{X^2}\right) s\left(\epsilon_{3i}\right)}{\sigma_X^2 \sigma_2^2} \right\} \\
=\ & \frac{1}{n^{1/2}} \sum_{i=1}^{n} \left\{ \frac{b\widetilde{X}_i s\left(\epsilon_{2i}\right)}{\sigma_X^2} + a \frac{\left(a\sigma_X^2 - \sigma_{XM}\right) X_i + \sigma_X^2 \epsilon_{2i} + \left(\mu_X \mu_{XM} - \mu_M \mu_{X^2} + \sigma_X^2 \beta_2\right)}{\sigma_X^2 \sigma_2^2} s\left(\epsilon_{3i}\right) \right\} \\
=\ & \frac{1}{n^{1/2}} \sum_{i=1}^{n} \left\{ \frac{b\widetilde{X}_i s\left(\epsilon_{2i}\right)}{\sigma_X^2} + \frac{a\sigma_X^2 \widetilde{\epsilon}_{2i} s\left(\epsilon_{3i}\right)}{\sigma_X^2 \sigma_2^2} \right\} = \frac{1}{n^{1/2}} \sum_{i=1}^{n} \left\{ \frac{b\widetilde{X}_i}{\sigma_X^2} s\left(\epsilon_{2i}\right) + \frac{a\widetilde{\epsilon}_{2i}}{\sigma_2^2} s\left(\epsilon_{3i}\right) \right\},
\end{aligned}$$

where $\widetilde{\epsilon}_{2i} = \epsilon_{2i} - \mu_{\epsilon_2}$, and

$$\begin{aligned}
\mu_M \mu_{XM} - \mu_X \mu_{M^2} - \sigma_{XM}\beta_2 &= \mu_M \sigma_{XM} - \mu_X \sigma_M^2 - \sigma_{XM}\beta_2 \\
&= \left(\beta_2 + a\mu_X + \mu_{\epsilon_2}\right) a\sigma_X^2 - \mu_X \left(a^2\sigma_X^2 + \sigma_2^2\right) - a\sigma_X^2 \beta_2 \\
&= \mu_X \sigma_2^2 + a\mu_{\epsilon_2}\sigma_X^2, \\
\mu_X \mu_{XM} - \mu_M \mu_{X^2} + \sigma_X^2 \beta_2 &= \mu_X \sigma_{XM} - \mu_M \sigma_X^2 + \sigma_X^2 \beta_2 \\
&= \mu_X a\sigma_X^2 - \left(\beta_2 + a\mu_X + \mu_{\epsilon_2}\right) \sigma_X^2 + \sigma_X^2 \beta_2 = -\mu_{\epsilon_2}\sigma_X^2.
\end{aligned}$$

So

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

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

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

where $\mathrm{Avar}(X_n)$ is the asymptotic variance of a generic sequence of random variable 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,

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

where

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

with

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

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

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

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

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

□

APPENDIX D: PROOF OF THEOREM 2

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

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

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

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

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

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

$$
E\left\{\left(D(\mathbf{u})-D(\mathbf{v})\right)\left(D(\mathbf{u})-D(\mathbf{v})\right)'\right\}=f_\epsilon\left(0\right)E\left\{\mathbf{x}\mathbf{x}'\left|(\mathbf{u}-\mathbf{v})'\mathbf{x}\right|\right\}
$$

for $\mathbf{u},\mathbf{v}\in\mathbb{R}^k$, and $Z\sim N(\mathbf{0},(4f_\epsilon(0)^2)^{-1}E(\mathbf{x}\mathbf{x}')^{-1})$. We apply this general result to our case. Note that Assumption II(iv) implies $E(|M|^3)<\infty$ by Hölder's inequality, so $E(\|\mathbf{x}_2\|^3)<\infty$ holds. Also, $\epsilon_1=\epsilon_3$ when $b=0$, we require only the differentiability of ϵ_3.

First,

$$
\begin{aligned}
& n^{1/4}\left[n^{1/2}\left\{\widehat{c}-\widehat{c'}-\left(c-c'\right)\right\}-\frac{1}{n^{1/2}}\sum_{i=1}^{n}\frac{a\widetilde{\epsilon}_{2i}}{\sigma_2^2}s\left(\epsilon_{3i}\right)\right] \\
= & -\frac{(-\mu_X,1)}{f_{\epsilon_1}\left(0\right)\sigma_X^2}n^{-1/4}\sum_{i=1}^{n}\mathbf{x}_{1i}\left\{1(\epsilon_{3i}\leq Z_1'\mathbf{x}_{1i}/n^{1/2})-1(\epsilon_{3i}\leq 0)-f_{\epsilon_3}\left(0\right)Z_1'\mathbf{x}_{1i}/n^{1/2}\right\} \\
& +\frac{\left(\mu_M\mu_{XM}-\mu_X\mu_{M^2},\sigma_M^2,-\sigma_{XM}\right)}{f_{\epsilon_3}\left(0\right)\sigma_X^2\sigma_2^2}n^{-1/4}\sum_{i=1}^{n}\mathbf{x}_{2i}\left\{1(\epsilon_{3i}\leq Z_2'\mathbf{x}_{2i}/n^{1/2})-1(\epsilon_{3i}\leq 0)-f_{\epsilon_3}\left(0\right)Z_2'\mathbf{x}_{2i}/n^{1/2}\right\} \\
\Longrightarrow & -\frac{(-\mu_X,1)}{f_{\epsilon_3}\left(0\right)\sigma_X^2}D_1\left(Z_1\right)+\frac{\left(\mu_M\mu_{XM}-\mu_X\mu_{M^2},\sigma_M^2,-\sigma_{XM}\right)}{f_{\epsilon_3}\left(0\right)\sigma_X^2\sigma_2^2}\mathbf{D}_2\left(Z_2\right),
\end{aligned}
$$

which cannot be simplified if $a\neq 0$. The covariance kernel of $D_1(\cdot)$ is $f_{\epsilon_1}(0)E\{\mathbf{x}_{1i}\mathbf{x}_{1i}'\,|\mathbf{x}_{1i}'(\mathbf{u}-\mathbf{v})|\}=f_{\epsilon_3}(0)E\{\mathbf{x}_{1i}\mathbf{x}_{1i}'\,|\mathbf{x}_{1i}'(\mathbf{u}-\mathbf{v})|\}$, the covariance kernel of $\mathbf{D}_2(\cdot)$ is $f_{\epsilon_3}(0)E\{\mathbf{x}_{2i}\mathbf{x}_{2i}'\,|\mathbf{x}_{2i}'(\mathbf{u}-\mathbf{v})|\}$, $D_1(\cdot)$ and $\mathbf{D}_2(\cdot)$ are correlated with the covariance kernel equal to

$$
\begin{aligned}
& \lim_{n\to\infty}n^{1/2}E\left[\mathbf{x}_{1i}\mathbf{x}_{2i}'\left\{1(\epsilon_{3i}\leq\mathbf{u}'\mathbf{x}_{1i}/n^{1/2})-1(\epsilon_{3i}\leq 0)-f_{\epsilon_3}\left(0\right)\mathbf{u}'\mathbf{x}_{1i}/n^{1/2}\right\}\right. \\
& \left.\left\{1(\epsilon_{3i}\leq\mathbf{v}'\mathbf{x}_{2i}/n^{1/2})-1(\epsilon_{3i}\leq 0)-f_{\epsilon_3}\left(0\right)\mathbf{v}'\mathbf{x}_{2i}/n^{1/2}\right\}\right] \\
= & E\left\{\mathbf{x}_{1i}\mathbf{x}_{2i}'f_{\epsilon_3}\left(0\right)\left(\left|\mathbf{u}'\mathbf{x}_{1i}\right|\wedge\left|\mathbf{v}'\mathbf{x}_{2i}\right|\right)1(\mathbf{u}'\mathbf{x}_{1i}\mathbf{x}_{2i}'\mathbf{v}>0)\right\},
\end{aligned}
$$

both are independent of Z_1 and Z_2,

$$
\begin{aligned}
Z_1 &\sim \{2f_{\epsilon_3}\left(0\right)\}^{-1}N\left(\mathbf{0},E\left(\mathbf{x}_{1i}\mathbf{x}_{1i}'\right)^{-1}\right)\equiv N\left(\mathbf{0},\Sigma_1\right), \\
Z_2 &\sim \left(2f_{\epsilon_3}\left(0\right)\right)^{-1}N\left(\mathbf{0},E\left[\mathbf{x}_{2i}\mathbf{x}_{2i}'\right]^{-1}\right)\equiv N\left(\mathbf{0},\Sigma_2\right), \\
E\left(Z_1Z_2'\right) &= E\left(\mathbf{x}_{1i}\mathbf{x}_{1i}'\right)^{-1}E\left\{\mathbf{x}_{1i}\mathbf{x}_{2i}'s\left(\epsilon_{1i}\right)s\left(\epsilon_{3i}\right)\right\}E\left(\mathbf{x}_{2i}\mathbf{x}_{2i}'\right)^{-1} \\
&= \frac{1}{4f_{\epsilon_3}\left(0\right)^2}E\left(\mathbf{x}_{1i}\mathbf{x}_{1i}'\right)^{-1}E\left(\mathbf{x}_{1i}\mathbf{x}_{2i}'\right)E\left(\mathbf{x}_{2i}\mathbf{x}_{2i}'\right)^{-1}\equiv\Sigma_{12}.
\end{aligned}
$$

Second, because

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

we have

$$\begin{aligned} & n^{1/4}\left\{n^{1/2}\left(\widehat{a}\widehat{b}-ab\right)-\frac{1}{n^{1/2}}\sum_{i=1}^{n}\frac{a\widetilde{\epsilon}_{2i}}{\sigma_2^2}s\left(\epsilon_{3i}\right)\right\} \\ = \;& a\frac{\left(\mu_X\mu_{XM}-\mu_M\mu_{X^2},-\sigma_{X,M},\sigma_X^2\right)}{f_{\epsilon_3}(0)\,\sigma_X^2\sigma_2^2}n^{-1/4}\sum_{i=1}^{n}\Big[\mathbf{x}_{2i}\{1(\epsilon_{3i}\le Z_2'\mathbf{x}_{2i}/n^{1/2}) \\ & -1(\epsilon_{3i}\le 0)-f_{\epsilon_3}(0)\,Z_2'\mathbf{x}_{2i}/n^{1/2}\}\Big] \\ \Longrightarrow \;& a\frac{\left(\mu_X\mu_{XM}-\mu_M\mu_{X^2},-\sigma_{X,M},\sigma_X^2\right)}{f_{\epsilon_3}(0)\,\sigma_X^2\sigma_2^2}\mathbf{D}_2\left(Z_2\right). \end{aligned}$$

In summary,

$$\begin{aligned} n^{3/4}\left(\widehat{c}-\widehat{c'}-\widehat{a}\widehat{b}\right) \Longrightarrow & -\frac{(-\mu_X,1)}{f_{\epsilon_3}(0)\,\sigma_X^2}D_1\left(Z_1\right)+\frac{\left(\mu_M\mu_{XM}-\mu_X\mu_{M^2},\sigma_M^2,-\sigma_{XM}\right)}{f_{\epsilon_3}(0)\,\sigma_X^2\sigma_2^2}\mathbf{D}_2\left(Z_2\right) \\ & -\frac{a\left(\mu_X\mu_{XM}-\mu_M\mu_{X^2},-\sigma_{X,M},\sigma_X^2\right)}{f_{\epsilon_3}(0)\,\sigma_X^2\sigma_2^2}\mathbf{D}_2\left(Z_2\right), \end{aligned}$$

where

$$\begin{aligned} & \left(\mu_M\mu_{XM}-\mu_X\mu_{M^2},\sigma_M^2,-\sigma_{XM}\right)-a\left(\mu_X\mu_{XM}-\mu_M\mu_{X^2},-\sigma_{X,M},\sigma_X^2\right) \\ = & \left[\mu_{XM}\left(\mu_M-a\mu_X\right)-\left(\mu_X\mu_{M^2}-a\mu_M\mu_{X^2}\right),\sigma_M^2+a\sigma_{X,M},-\left(\sigma_{XM}+a\sigma_X^2\right)\right] \\ = & \left[2a\sigma_X^2\left(\mu_M-a\mu_X\right)-\mu_X\sigma_2^2,2a^2\sigma_X^2+\sigma_2^2,-2a\sigma_X^2\right]. \end{aligned}$$ □

Appendix E: Proof of Theorem 3

Proof. From Theorem 1, we have

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

From Theorem 2,

$$\begin{aligned} n^{3/4}\left(\widehat{c}-\widehat{c'}\right) &= n^{3/4}\left\{\widehat{c}-\widehat{c'}-\left(c-c'\right)\right\} \\ &\Longrightarrow -\frac{(-\mu_X,1)}{f_{\epsilon_3}(0)\,\sigma_X^2}D_1\left(Z_1\right)+\frac{\left(\mu_M\mu_{XM}-\mu_X\mu_{M^2},\sigma_M^2,-\sigma_{XM}\right)}{f_{\epsilon_3}(0)\,\sigma_X^2\sigma_2^2}\mathbf{D}_2\left(Z_2\right) \end{aligned}$$

$$= -\frac{(-\mu_X, 1)}{f_{\epsilon_3}(0)\,\sigma_X^2} D_1(Z_1) + \frac{\left(-\sigma_2^2\mu_X, \sigma_2^2, 0\right)}{f_{\epsilon_3}(0)\,\sigma_X^2\sigma_2^2} \mathbf{D}_2(Z_2)$$
$$= \frac{(-\mu_X, 1)}{f_{\epsilon_3}(0)\,\sigma_X^2} \{D_2(Z_2) - D_1(Z_1)\},$$

where $D_2(\cdot)$ is the first two element of $\mathbf{D}_2(\cdot)$, so the covariance kernel of $D_2(\cdot)$ is $f_{\epsilon_3}(0)E\{\mathbf{x}_{1i}\mathbf{x}_{1i}' \, |\mathbf{x}_{2i}'(\mathbf{u}-\mathbf{v})|\}$ and it covariance kernel with $D_1(\cdot)$ is $f_{\epsilon_3}(0)E\{\mathbf{x}_{1i}\mathbf{x}_{1i}'(|\mathbf{u}'\mathbf{x}_{1i}| \wedge |\mathbf{v}'\mathbf{x}_{2i}|)1(\mathbf{u}'\mathbf{x}_{1i}\mathbf{x}_{2i}'\mathbf{v} > 0)\}$. 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 $n^{3/4}(\widehat{c} - \widehat{c'})$. □

APPENDIX F: SIMPLIFICATIONS IN CASE THREE AND SIMULATIONS

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

$$\Sigma_1 = \frac{1}{4f_{\epsilon_3}(0)^2}\frac{1}{\sigma_X^2}\begin{pmatrix} \mu_{X^2} & -\mu_X \\ -\mu_X & 1 \end{pmatrix}, \quad \Sigma_{12} = \frac{1}{4f_{\epsilon_3}(0)^2}\frac{1}{\sigma_X^2}\begin{pmatrix} \mu_{X^2} & -\mu_X & 0 \\ -\mu_X & 1 & 0 \end{pmatrix},$$
$$\Sigma_2 = \frac{1}{4f_{\epsilon_3}(0)^2}\frac{1}{\sigma_2^2\sigma_X^2}\begin{pmatrix} \mu_{X^2}\sigma_2^2 + \mu_M^2\sigma_X^2 & -\mu_X\sigma_2^2 & -\mu_M\sigma_X^2 \\ -\mu_X\sigma_2^2 & \sigma_2^2 & 0 \\ -\mu_M\sigma_X^2 & 0 & \sigma_X^2 \end{pmatrix}.$$

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

$$(z_2, z_3)^{\mathrm{T}} \sim N\left(0, \frac{1}{4f_{\epsilon_3}(0)^2\sigma_2^2}\begin{pmatrix} 1 & -\mu_M \\ -\mu_M & \mu_M^2 \end{pmatrix}\right).$$

In consequence,

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

with $|\Sigma| = 0$. Also, when $\mu_X = 0$, the variance of $(-\mu_X, 1)\{D_2(Z_2) - D_1(Z_1)\}/\{\sigma_X^2 f_{\epsilon_3}(0)\}$ reduces to

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

where $\widetilde{X} = X/\sigma_X$ 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

$$\Sigma_1 = \frac{I_2}{4f_{\epsilon_3}(0)^2}, \quad \Sigma_2 = \frac{I_3}{4f_{\epsilon_3}(0)^2}, \quad \Sigma_{12} = \frac{1}{4f_{\epsilon_3}(0)^2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

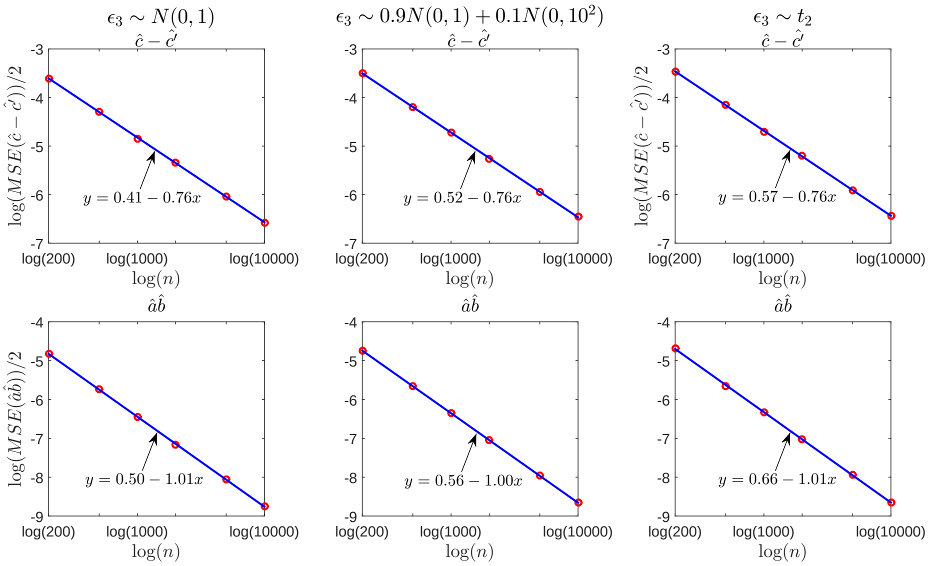

Fig. S1. $\log\{MSE(\widehat{c}-\widehat{c'})\}/2$ and $\log\{MSE(\widehat{ab})\}/2$ against $\log n$ for $a=b=0$ and three ϵ_3 distributions

As a result, the variance of $(-\mu_X, 1)\{D_2(Z_2) - D_1(Z_1)\}/\{\sigma_X^2 f_{\epsilon_3}(0)\}$ in Case three is

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

where $\mathbf{x}_1=(1,X)'$ and $\mathbf{x}_2=(1,X,M)'$ with $(X,M)\sim N(\mathbf{0},I_2)$, $Z_2'=(Z_1',z_2)$ 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,

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

We finally analyze the convergence rates of the two LAD estimates in Case three. Figure S1 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. S1 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.

REFERENCES

KNIGHT, K. (1997). Asymototics for L_1 regression estimators under general conditions. Technical Report 9716, Department of Statistics, University of Toronto.

KNIGHT, K. (1998a). A "delta method" approach to bahadur-kiefer theorems. *Scand. J. Stat.* **25**, 555–568.

KNIGHT, K. (1998b). Limiting distributions for L_1 regression estimators under general conditions. *Ann. Statist.* **26**, 755–770.

POLLARD, D. (1991). Asymptotics for least absolute deviation regression estimators. *Econ. Theory* **7**, 186–199.