To appear in the *Journal of Nonparametric Statistics*
Vol. 00, No. 00, Month 20XX, 1–27

An Alternative Local Polynomial Estimator for the Error-in-Variables Problem
Supplementary Materials

Xianzheng Huang[a*] and Haiming Zhou[b]

[a]*Department of Statistics, University of South Carolina, Columbia, South Carolina, U.S.A.;*
[b]*Division of Statistics, Northern Illinois University, DeKalb, Illinois, U.S.A.*

(*v4.0 released June 2015*)

Appendix A: Asymptotic bias of $\hat{m}_{\rm HZ}(x)$

A.1. Dominating asymptotic bias of $\hat{m}_{\rm HZ}(x)$

Define $\mathbb{W}=(W_1,\ldots,W_n)$. Under the regularity conditions that guarantees the consistency of $\hat{f}_X(x)$, the dominating terms in $E\{\hat{m}_{\rm HZ}(x)-m(x)|\mathbb{W}\}$ are the same as those in

$$E[\{\hat{m}_{\rm HZ}(x)-m(x)\}\hat{f}_X(x)/f_X(x)|\mathbb{W}]=[E\{\mathcal{B}(x)|\mathbb{W}\}-m(x)\hat{f}_X(x)]/f_X(x). \tag{A.1}$$

The majority of the following derivation is to elaborate $E\{\mathcal{B}(x)|\mathbb{W}\}$.

By the relationship between $\mathcal{B}(x)$ and $\mathcal{A}(w)$ indicated by equation (8) in the main article, we immediate have $E\{\mathcal{B}(x)|\mathbb{W}\}=\{E(\mathcal{A}|\mathbb{W})*D\}(x)$. This motivates us to first look into $E\{\mathcal{A}(w)|\mathbb{W}\}$, that is,

$$E\left\{\hat{m}^*(w)\hat{f}_W(w)|\mathbb{W}\right\}=E\left\{\hat{m}^*(w)|\mathbb{W}\right\}\hat{f}_W(w).$$

The following two results for kernel-based estimators in the absence of measurement error can be used to derive $E\{\mathcal{A}(w)|\mathbb{W}\}$. The first result is about the local polynomial estimator of the regression function $m^*(w)$. In particular, by Theorem 3.1 in Fan and Gijbels (1996), if $f'_W(\cdot)$ and $m^{*(p+2)}(\cdot)$ are continuous in a neighborhood of w and $nh\to\infty$, then

$$E\left\{\hat{m}^*(w)|\mathbb{W}\right\}=m^*(w)+\boldsymbol{e}_1^{\rm T}\mathbf{S}^{-1}\boldsymbol{c}_p\frac{1}{(p+1)!}m^{*(p+1)}(w)h^{p+1}+o_P(h^{p+1}) \tag{A.2}$$

*Corresponding author. Email: huang@stat.sc.edu

when p is odd, and

$$\begin{aligned} E\{\hat{m}^*(w)|\mathbb{W}\} &= m^*(w) + \boldsymbol{e}_1^{\mathrm{T}}\mathbf{S}^{-1}\tilde{\boldsymbol{c}}_p \frac{1}{(p+2)!}\Big\{m^{*(p+2)}(w) \\ &\quad + (p+2)m^{*(p+1)}(w)\frac{f'_W(w)}{f_W(w)}\Big\}h^{p+2} + o_P(h^{p+2}) \end{aligned} \quad (A.3)$$

when p is even, where $\mathbf{S} = (\mu_{\ell_1+\ell_2})_{0\le\ell_1,\ell_2\le p}$, $\boldsymbol{c}_p = (\mu_{p+1},\ldots,\mu_{2p+1})^{\mathrm{T}}$, $\tilde{\boldsymbol{c}}_p = (\mu_{p+2},\ldots,\mu_{2p+2})^{\mathrm{T}}$, $\mu_\ell = \int u^\ell K(u)du$, and $\boldsymbol{e}_1$ is the $(p+1)\times 1$ vector with the first entry being 1 and the remaining p entries being 0. The second result is about the kernel-based density estimator, $\hat{f}_W(w)$. By the definition of $\hat{f}_W(w)$ and using Taylor expansion around $h=0$, one has

$$E\left\{\hat{f}_W(w)\right\} = f_W(w) + \sum_{\ell=1}^{p+2} f_W^{(\ell)}(w)\mu_\ell h^\ell/\ell! + o(h^{p+2}).$$

Furthermore, $\mathrm{Var}\{\hat{f}_W(w)\} = O\{1/(nh)\}$. It follows that

$$\begin{aligned} \hat{f}_W(w) &= E\left\{\hat{f}_W(w)\right\} + O_P\left[\sqrt{\mathrm{Var}\{\hat{f}_W(w)\}}\right] \\ &= f_W(w) + \sum_{\ell=1}^{p+2} f_W^{(\ell)}(w)\mu_\ell h^\ell/\ell! + o_P(h^{p+2}) + O_P(1/\sqrt{nh}). \end{aligned} \quad (A.4)$$

Using the two sets of results in (A.2)–(A.4), provided that $nh^{2p+3}\to\infty$ when p is odd and $nh^{2p+5}\to\infty$ when p is even, we have

$$E\{\mathcal{A}(w)|\mathbb{W}\} = \begin{cases} A(w) + N_p(w) + M_p(w)h^{p+1}/(p+1)! + o_P(h^{p+1}), & \text{if } p \text{ is odd,} \\ A(w) + N_p(w) + M_p(w)h^{p+2}/(p+2)! + o_P(h^{p+2}), & \text{if } p \text{ is even,} \end{cases}$$

where $N_p(w) = m^*(w)\sum_{\ell=1}^p f_W^{(\ell)}(w)\mu_\ell h^\ell/\ell!$ and

$$M_p(w) = m^*(w) f_W^{(p+1)}(w)\mu_{p+1} + m^{*(p+1)}(w) f_W(w)\boldsymbol{e}_1^{\mathrm{T}}\mathbf{S}^{-1}\boldsymbol{c}_p,$$

if p is odd, and

$$M_p(w) = m^*(w) f_W^{(p+2)}(w)\mu_{p+2} + \left\{m^{*(p+2)}(w) f_W(w) + (p+2)m^{*(p+1)}(w) f'_W(w)\right\}\boldsymbol{e}_1^{\mathrm{T}}\mathbf{S}^{-1}\tilde{\boldsymbol{c}}_p,$$

if p is even.

With $E\{\mathcal{A}(w)|\mathbb{W}\}$ derived, we have

$$\begin{aligned} & E\{\mathcal{B}(x)|\mathbb{W}\} = \{E(\mathcal{A}|\mathbb{W}) * D\}(x) \\ = & \begin{cases} B(x) + (N_p * D)(x) + (M_p * D)(x)h^{p+1}/(p+1)! + o_P(h^{p+1}), & \text{if } p \text{ is odd,} \\ B(x) + (N_p * D)(x) + (M_p * D)(x)h^{p+2}/(p+2)! + o_P(h^{p+2}), & \text{if } p \text{ is even.} \end{cases} \end{aligned} \quad (A.5)$$

For the deconvolution kernel density estimator $\hat{f}_X(x)$, by equation (1.9) in Stefanski and Carroll (1990), one has

$$E\left\{\hat{f}_X(x)\right\} = f_X(x) + \sum_{\ell=1}^{p+2} f_X^{(\ell)}(x)\mu_\ell h^\ell/\ell! + o(h^{p+2}).$$

Under the conditions in Theorem 2.1 in Stefanski and Carroll (1990), $\text{Var}\{\hat{f}_X(x)\}$ is bounded from above by $(nh)^{-1}\sup_x f_X(x)\int K^2(t)dt$. It follows that

$$\begin{aligned}\hat{f}_X(x) &= E\left\{\hat{f}_X(x)\right\} + O_P\left[\sqrt{\text{Var}\{\hat{f}_X(x)\}}\right] \\ &= f_X(x) + \sum_{\ell=1}^{p+2} f_X^{(\ell)}(x)\mu_\ell h^\ell/\ell! + o_P(h^{p+2}) + O_P(1/\sqrt{nh}).\end{aligned}$$

Therefore,

$$m(x)\hat{f}_X(x) = B(x) + m(x)\sum_{\ell=1}^{p+2} f_X^{(\ell)}(x)\mu_\ell h^\ell/\ell! + o_P(h^{p+2}) + O_P(1/\sqrt{nh}). \tag{A.6}$$

By (A.5) and (A.6), provided that $nh^{2p+5}\to\infty$ when p is even and $nh^{2p+3}\to\infty$ when p is odd, (A.1) is equal to $\{f_X(x)\}^{-1}$ times

$$\begin{aligned}&(N_p * D)(x) + \left\{(M_p * D)(x) - m(x)f_X^{(p+1)}(x)\mu_{p+1}\right\}h^{p+1}/(p+1)! \\ - \; & m(x)\sum_{\ell=1}^{p} f_X^{(\ell)}(x)\mu_\ell h^\ell/\ell! + o_P(h^{p+1})\end{aligned}$$

when p is odd, and

$$\begin{aligned}&(N_p * D)(x) + \left\{(M_p * D)(x) - m(x)f_X^{(p+2)}(x)\mu_{p+2}\right\}h^{p+2}/(p+2)! \\ - \; & m(x)\sum_{\ell=1}^{p} f_X^{(\ell)}(x)\mu_\ell h^\ell/\ell! + o_P(h^{p+2})\end{aligned}$$

when p is even. This gives the dominating bias for $\hat{m}_{\text{HZ}}(x)$ of order h^{p+1} when p is odd and that of order h^{p+2} when p is even. It is worth noting that, although the derivation of the asymptotic bias of $\hat{m}_{\text{HZ}}(x)$ is conditional on $\mathbb{W}$, the leading terms of the asymptotic bias do not depend on $\mathbb{W}$, and thus these leading terms can be interpreted as the unconditional dominating asymptotic bias. This is in line with the remarks in Ruppert and Wand (1994, Remark 1 on page 1351) regarding their asymptotic bias and variance of the nonparametric estimator of $m(x)$ with $\mathbb{X} = (X_1,\ldots,X_n)$ observed.

A.2. Dominating asymptotic bias of $\hat{m}_{\mathrm{HZ}}(x)$ when $m(x)$ is a polynomial

Under the assumptions that $m(x) = \sum_{k=0}^{r} \beta_k x^k$, where $r \geq 2$, $X \sim N(0, 1)$, $U \sim N(0, \sigma_u^2)$, and $X \perp U$, we show in this section that

$$f_X^{-1}(x) \left\{ (M * D)(x) - m(x) f_X^{(2)}(x) \right\}, \tag{A.7}$$

is a polynomial of order r. For notational brevity, $\mathrm{Low}_k(t)$ is used in the sequel to stand for a generic polynomial in t of order lower than k, for $k > 0$.

With $f_X(x) = \exp(-x^2/2)/\sqrt{2\pi}$, straightforward induction reveals that,

$$f_X^{(k)}(x) = f_X(x) \left\{ (-1)^k x^k + (-1)^{k-1} \binom{k}{2} x^{k-2} + \mathrm{Low}_{k-2}(x) \right\}. \tag{A.8}$$

It follows that

$$m(x) f_X^{(2)}(x) = f_X(x) \left\{ \beta_r x^{r+2} + \beta_{r-1} x^{r+1} + (\beta_{r-2} - \beta_r) x^r + \mathrm{Low}_r(x) \right\}.$$

This solves half of the "mystery" in (A.7), that is, we have

$$f_X^{(-1)}(x) m(x) f_X^{(2)}(x) = \beta_r x^{r+2} + \beta_{r-1} x^{r+1} + (\beta_{r-2} - \beta_r) x^r + \mathrm{Low}_r(x). \tag{A.9}$$

The other half of the mystery is about $f_X^{-1}(x)(M * D)(x)$ in (A.7). In what follows, we will show that this half is equal to

$$\beta_r x^{r+2} + \beta_{r-1} x^{r+1} + [\beta_{r-2} + \beta_r \{2r(\lambda - 1) - 1\}] x^r + \mathrm{Low}_r(x), \tag{A.10}$$

where $\lambda = 1/(1 + \sigma_u^2)$. Once this is established, subtracting (A.9) from (A.10) reveals that (A.7) is equal to $2r(\lambda - 1)\beta_r x^r + \mathrm{Low}_r(x)$, i.e., a polynomial of order r as long as $\lambda \neq 1$.

Recall that $M(w) = m^*(w) f_W^{(2)}(w) + m^{*(2)}(w) f_W(w)$, which involves $f_W(w)$ and $m^*(w)$. Because $X \sim N(0, 1)$ is independent of $U \sim N(0, \sigma_u^2)$, one has $f_W(w) = (1 + \sigma_u^2)^{-1/2} \phi(w/\sqrt{1 + \sigma_u^2})$, and thus $f_W^{(2)}(w) = \lambda f_W(w)(\lambda w^2 - 1)$, where $\phi(\cdot)$ denotes the pdf of the standard normal. Given the current $m(x)$, the naive regression function

$m^*(w)$ is equal to

$$\begin{aligned} E(Y|W=w) &= E\{E(Y|X)|W=w\} \\ &= f_W^{-1}(w)\int m(x)f_X(x)f_U(w-x)dx \\ &= f_W^{-1}(w)\sum_{k=0}^{r}\beta_k\int x^k\phi(x)\sigma_u^{-1}\phi\{(w-x)/\sigma_u\}dx \\ &= f_W^{-1}(w)f_W(w)\sum_{k=0}^{r}\beta_k\int x^k\frac{\sqrt{1+\sigma_u^2}}{\sigma_u}\phi\left\{\frac{\sqrt{1+\sigma_u^2}}{\sigma_u}\left(x-\frac{w}{1+\sigma_u^2}\right)\right\}dx \\ &= \sum_{k=0}^{r}\beta_k \times \text{the } k\text{th moment of } N(\lambda w,\, 1-\lambda) \\ &= \sum_{k=0}^{r}\beta_k\lambda^k\sum_{\ell=0}^{\lfloor k/2\rfloor}\binom{k}{2\ell}(2\ell-1)!!(1-\lambda)^\ell\lambda^{-2\ell}w^{k-2\ell}, \end{aligned} \tag{A.11}$$

where !! is the double factorial symbol, with $(-1)!!$ defined to be 1. It follows that

$$m^{*(2)}(w)=\sum_{k=0}^{r}\beta_k\lambda^k\sum_{\ell=0}^{\lfloor k/2-1\rfloor}\binom{k}{2\ell}(2\ell-1)!!(1-\lambda)^\ell\lambda^{-2\ell}(k-2\ell)(k-2\ell-1)w^{k-2\ell-2}.$$

Putting $f_W(w)$, $f_W^{(2)}(w)$, $m^*(w)$, and $m^{*(2)}(w)$ back in $M(w)$, one can see that $M(w)$ is equal to $f_W(w)$ times a polynomial in w of order $r+2$. Hence, the key to deriving

$$(M*D)(x)=\frac{1}{2\pi}\int e^{-itx}\frac{\phi_M(t)}{\phi_U(t)}dt, \tag{A.12}$$

is to understand, for $k\geq 0$, $(2\pi)^{-1}\int e^{-itx}\phi_{w^k f_W(w)}(t)/\phi_U(t)dt$.

It is straightforward to show that $\phi_{w^k f_W}(t)=i^{-k}\phi_W^{(k)}(t)$, where $\phi_W(t)$ is the characteristic function of W. With a normal W, using induction one can show that

$$\phi_W^{(k)}(t)=\{(-1)^k\lambda^{-k}t^k+\text{Low}_{k-1}(t)\}\phi_W(t). \tag{A.13}$$

Noting that $\phi_W(t)=\phi_X(t)\phi_U(t)$, and using the result that, for $k=0,1,\ldots$,

$$(-i)^k(2\pi)^{-1}\int e^{-itx}t^k\phi_X(t)dt=f_X^{(k)}(x), \tag{A.14}$$

we now have

$$
\begin{aligned}
& (2\pi)^{-1}\int e^{-itx}\phi_{w^k f_W(w)}(t)/\phi_U(t)dt \\
= \;& i^{-k}(2\pi)^{-1}\int e^{-itx}\{(-1)^k\lambda^{-k}t^k + \mathrm{Low}_{k-1}(t)\}\phi_W(t)/\phi_U(t)dt \\
= \;& i^{-k}(-1)^k\lambda^{-k}(-i)^{-k}(-i)^k(2\pi)^{-1}\int e^{-itx}t^k\phi_X(t)dt + \int e^{-itx}\mathrm{Low}_{k-1}(t)\phi_X(t)dt \\
= \;& (-1)^k\lambda^{-k}f_X^{(k)}(x) + \text{ (some coefficient free of } x) \times f_X^{(k-2)}(x), \text{ by (A.14)} \\
= \;& (-1)^k\lambda^{-k}(-1)^k x^k f_X(x) + \mathrm{Low}_{k-1}(x)f_X(x), \text{ by (A.8)}, \\
= \;& \lambda^{-k}x^k f_X(x) + \mathrm{Low}_{k-1}(x)f_X(x). \qquad (A.15)
\end{aligned}
$$

We next focus on the terms in $M(w)$ with the two highest powers of w. Tracing the coefficients of w^{r+2} and w^{r+1} in $M(w)$ and applying (A.15) for $k = r+2$ and $r+1$, one can see that (A.12) is equal to

$$
\begin{aligned}
& \lambda^{r+2}\beta_r\lambda^{-(r+2)}x^{r+2}f_X(x) + \lambda^{r+1}\beta_{r-1}\lambda^{-(r+1)}x^{r+1}f_X(x) + \mathrm{Low}_{r+1}(x)f_X(x) \\
= \;& \{\beta_r x^{r+2} + \beta_{r-1}x^{r+1} + \mathrm{Low}_{r+1}(x)\}f_X(x),
\end{aligned}
$$

which proves the first two terms in (A.10).

To prove the third term in (A.10), we first find the term in $M(w)$ of the third highest order in w, i.e., w^r, because this term leads to a term with x^r in $(M*D)(x)$ according to (A.15). One such term shows up in $m^*(w)f_W^{(2)}$ is $\lambda^r\{\beta_{r-2} - \beta_r\lambda + \beta_r(1-\lambda)\binom{r}{2}\}w^r f_W(w)$. Convoluting this term with $D(\cdot)$ yields

$$\left\{\beta_{r-2} - \beta_r\lambda + \beta_r(1-\lambda)\binom{r}{2}\right\} x^r f_X(x). \qquad (A.16)$$

Secondly, note that $m^*(w)f_W^{(2)}$ contains $\lambda^{r+2}\beta_r w^{r+2}f_W(w)$, which, after convoluting with $D(\cdot)$ produces $f_X^{(r+2)}(x)$, which itself contributes x^r according to (A.8). More specifically, this term is

$$-\beta_r\binom{r+2}{2}x^r f_X(x). \qquad (A.17)$$

Thirdly, also due to the involvement of $\lambda^{r+2}\beta_r w^{r+2}f_W(w)$ in $M(w)$, which, when convoluting with $D(\cdot)$, yields $\phi_W^{(r+2)}(t)$, which contains a terms with t^r according to (A.13). This eventually translates to

$$\lambda\beta_r x^r\binom{r+2}{2}f_X(x). \qquad (A.18)$$

Summing (A.16), (A.17), and (A.18) together gives the third term in (A.10).

Now (A.10) is established, and combining it with (A.9), we show that (A.7) is equal to a polynomial in x of order r.

In the case with $r = 2$, we have

$$\begin{aligned}\phi_{m^* f_W^{(2)}}(t) &= \{\beta_2 t^4 - i\beta_1 t^3 - (\beta_0 + \beta_2 + 4\lambda\beta_2)t^2 + i2\lambda\beta_1 + 2\lambda^2\beta_2\}\phi_W(t), \\ \phi_{m^{*(2)} f_W}(t) &= 2\lambda^2\beta_2\phi_W(t).\end{aligned}$$

It follows that (A.12) is equal to the inverse Fourier transform of $\{\beta_2 t^4 - i\beta_1 t^3 - (\beta_0 + \beta_2 + 4\lambda\beta_2)t^2 + i2\lambda\beta_1 + 4\lambda^2\beta_2\}\phi_X(t)$. Using (A.14), one can show that this inverse Fourier transform is equal to

$$\begin{aligned}& \beta_2 f_X^{(4)}(x) - \beta_1 f_X^{(3)}(x) + (\beta_0 + \beta_2 + 4\lambda\beta_2) f_X^{(2)}(x) - 2\lambda\beta_1 f_X'(x) + 4\lambda^2\beta_2 f_X(x) \\ = & \{\beta_2 x^4 + \beta_1 x^3 + (\beta_0 - 5\beta_2 + 4\lambda\beta_2)x^2 + (2\lambda - 3)\beta_1 x + \\ & 4\lambda^2\beta_2 - 4\lambda\beta_2 - \beta_0 + 2\beta_2\} f_X(x). \qquad \text{(A.19)}\end{aligned}$$

Subtracting $m(x) f_X^{(2)}(x) = f_X(x)(x^2 - 1)(\beta_0 + \beta_1 x + \beta_2 x^2)$ from (A.19) reveals that (A.7) reduces to $2\{(\lambda - 1)(\beta_1 + 2\beta_2 x)x + \beta_2(2\lambda^2 - 2\lambda + 1)\}$.

Appendix B: Detailed derivations for Var$\{\mathcal{B}(x)|\mathbb{W}\}$

The five steps of the road map outlined in Section 4.1 in the main article are elaborated in this section.

B.1. Step 1: Relating Var$\{\mathcal{B}(x)|\mathbb{W}\}$ to Cov$\{\mathcal{A}(w_1), \mathcal{A}(w_1)|\mathbb{W}\}$

Assuming interchangeability of expectation and integration, one has

$$E\{\mathcal{B}^2(x)|\mathbb{W}\} = E\left\{\int D(v_1)\mathcal{A}(x - v_1)\mathcal{B}(x)dv_1|\mathbb{W}\right\} = \int D(v_1)E\left\{\mathcal{A}(x - v_1)\mathcal{B}(x)|\mathbb{W}\right\}dv_1,$$

where

$$\begin{aligned}E\left\{\mathcal{A}(x - v_1)\mathcal{B}(x)|\mathbb{W}\right\} &= E\left[\mathcal{A}(x - v_1)\int \mathcal{A}(x - v_2)D(v_2)dv_2|\mathbb{W}\right] \\ &= \int D(v_2)E\left\{\mathcal{A}(x - v_1)\mathcal{A}(x - v_2)|\mathbb{W}\right\}dv_2.\end{aligned}$$

Thus

$$E\{\mathcal{B}^2(x)|\mathbb{W}\} = \int D(v_1)\int D(v_2)E\left\{\mathcal{A}(x - v_1)\mathcal{A}(x - v_2)|\mathbb{W}\right\}dv_2 dv_1. \qquad \text{(B.1)}$$

In addition,

$$\begin{aligned}[E\{\mathcal{B}(x)|\mathbb{W}\}]^2 &= \int D(v_1)E\{\mathcal{A}(x - v_1)|\mathbb{W}\}dv_1 \int D(v_2)E\{\mathcal{A}(x - v_2)|\mathbb{W}\}dv_2 \\ &= \int D(v_1)\int D(v_2)E\{\mathcal{A}(x - v_1)|\mathbb{W}\}E\{\mathcal{A}(x - v_2)|\mathbb{W}\}dv_2 dv_1 \text{(B.2)}\end{aligned}$$

Subtracting (B.2) from (B.1) gives

$$\begin{aligned}\mathrm{Var}\{\mathcal{B}(x)|\mathbb{W}\} &= \int D(v_1)\int D(v_2)\mathrm{Cov}\left\{\mathcal{A}(x-v_1),\mathcal{A}(x-v_2)|\mathbb{W}\right\}dv_2dv_1\\ &= \int D(x-w_1)\int D(x-w_2)\mathrm{Cov}\left\{\mathcal{A}(w_1),\mathcal{A}(w_2)|\mathbb{W}\right\}dw_2dw_1,\end{aligned} \tag{B.3}$$

where

$$\mathrm{Cov}\left\{\mathcal{A}(w_1),\,\mathcal{A}(w_2)|\mathbb{W}\right\} = \mathrm{Cov}\{\hat{m}^*(w_1),\,\hat{m}^*(w_2)|\mathbb{W}\}f_W(w_1)f_W(w_2)\{1+o_P(1)\}. \tag{B.4}$$

The next two steps are devoted to deriving $\mathrm{Cov}\{\hat{m}^*(w_1),\,\hat{m}^*(w_2)|\mathbb{W}\}$.

B.2. Step2: Approximating $\mathrm{Cov}\{\hat{m}^*(w_1),\,\hat{m}^*(w_2)|\mathbb{W}\}$

Naive estimation of $m(x)$ based on error-contaminated data, $\{(Y_j,W_j)\}_{j=1}^n$, entails implementing the weighted least squares estimation in Fan and Gijbels (1996, Section 3.1) with X_j and x there replaced by W_j and w, respectively, for $j=1,\ldots,n$. In particular, one may consider the naive regression, $Y_j = m^*(W_j)+\nu(W_j)\epsilon_j^*$, where $E(\epsilon^*)=0$, $\mathrm{Var}(\epsilon^*)=1$, and W_j and ϵ_j^* are independent. Then a set of estimators of $m^{*(\ell)}(w_k)$, for $k=1,2$, and $\ell=0,1,\ldots,p$, can be obtained by minimizing the following weighted sum of squares,

$$\sum_{j=1}^n\left\{Y_j-\sum_{\ell=0}^p\beta_{\ell k}^*(W_j-w_k)^\ell\right\}^2K_h(W_j-w_k), \tag{B.5}$$

where $\beta_{\ell k}^* = m^{*(\ell)}(w_k)/\ell!$, for $\ell=0,\ldots,p$, and $K_h(t)=K(t/h)/h$. Denote by $\boldsymbol{\beta}_k^* = (\beta_{0k}^*,\beta_{1k}^*,\ldots,\beta_{pk}^*)^{\mathrm{T}}$ and by $\hat{\boldsymbol{\beta}}_k^*$ the minimizer of (B.5), for $k=1,2$. Then $\mathrm{Cov}\{\hat{m}^*(w_1),\,\hat{m}^*(w_2)|\mathbb{W}\}$ is the $[1,1]$ element of the $(p+1)\times(p+1)$ variance-covariance matrix $\mathrm{Cov}(\hat{\boldsymbol{\beta}}_1^*,\hat{\boldsymbol{\beta}}_2^*|\mathbb{W})$.

As in equation (3.5) in Fan and Gijbels (1996), the minimizer of (B.5) is, for $k=1,2$, $\hat{\boldsymbol{\beta}}_k^* = (\mathbf{G}_k^{\mathrm{T}}\mathbf{W}_k\mathbf{G}_k)^{-1}\mathbf{G}_k^{\mathrm{T}}\mathbf{W}_k\mathbf{Y}$, where $\mathbf{W}_k=\mathrm{diag}\{K_h(W_1-w_k),\ldots,K_h(W_n-w_k)\}$ and

$$\mathbf{G}_k = \begin{bmatrix}1 & (W_1-w_k) & \ldots & (W_1-w_k)^p\\ \vdots & \vdots & \ldots & \vdots\\ 1 & (W_n-w_k) & \ldots & (W_n-w_k)^p\end{bmatrix}.$$

It follows that, since $\mathrm{Var}(Y|W=w)=\nu^2(w)$ under the naive regression,

$$\mathrm{Cov}(\hat{\boldsymbol{\beta}}_1^*,\,\hat{\boldsymbol{\beta}}_2^*|\mathbb{W}) = \left\{\mathbf{S}_{nW}^{(1)}\right\}^{-1}\mathbf{S}_{nW}^*\left\{\mathbf{S}_{nW}^{(2)}\right\}^{-1}. \tag{B.6}$$

where $\mathbf{S}_{nW}^{(k)} = \mathbf{G}_k^{\mathrm{T}}\mathbf{W}_k\mathbf{G}_k = (S_{nW,\ell_1+\ell_2}^{(k)})_{0\le\ell_1,\ell_2\le p}$, for $k=1,2$, in which, for $\ell=0,1,\ldots,2p$, $S_{nW,\ell}^{(k)} = \sum_{j=1}^n K_h(W_j-w_k)(W_j-w_k)^\ell$; and

$$\mathbf{S}_{nW}^* = \mathbf{G}_1^{\mathrm{T}}\boldsymbol{\Sigma}_{12}\mathbf{G}_2 = (S_{nW,\ell_1,\ell_2}^*)_{0\le\ell_1,\ell_2\le p}, \tag{B.7}$$

in which, $\mathbf{\Sigma}_{12} = \text{diag}\{K_h(W_1 - w_1)K_h(W_1 - w_2)\nu^2(W_1), \ldots, K_h(W_n - w_1)K_h(W_n - w_2)\nu^2(W_n)\}$, and, for $\ell_1, \ell_2 = 0, 1, \ldots, p$,

$$S^*_{nW,\ell_1,\ell_2} = \sum_{j=1}^{n}(W_j - w_1)^{\ell_1}(W_j - w_2)^{\ell_2}K_h(W_j - w_1)K_h(W_j - w_2)\nu^2(W_j). \tag{B.8}$$

Finally, extracting the $[1, 1]$ element of (B.6) gives $\text{Cov}\{\hat{m}^*(w_1), \hat{m}^*(w_2)|\mathbb{W}\}$.

Now, to derive a large-sample approximation of $\text{Cov}\{\hat{m}^*(w_1), \hat{m}^*(w_2)|\mathbb{W}\}$, we need to approximate $\mathbf{S}^{(k)}_{nW}$, for $k = 1, 2$, and $\mathbf{S}^*_{nW}$. Both approximations follow the same spirit as those in Fan and Gijbels (1996, page 101) that lead to their (3.54) and (3.55).

B.2.1. *Approximate* $\mathbf{S}^{(k)}_{nW}$

For $k = 1, 2$ and $\ell = 0, 1, \ldots, 2p$,

$$\begin{aligned}
S^{(k)}_{nW,\ell} &= E\left\{S^{(k)}_{nW,\ell}\right\} + O_P\left[\sqrt{\text{Var}\left\{S^{(k)}_{nW,\ell}\right\}}\right] \\
&= n\int K_h(w - w_k)(w - w_k)^{\ell} f_W(w)dw + O_P\left[\sqrt{n\text{Var}\left\{K_h(W_1 - w_k)(W_1 - w_k)^{\ell}\right\}}\right] \\
&= n\int K(u)h^{\ell}u^{\ell} f_W(hu + w_k)du + O_P\left[\sqrt{nE\left\{K_h^2(W_1 - w_k)(W_1 - w_k)^{2\ell}\right\}}\right] \\
&= nh^{\ell}\left\{f_W(w_k) + o_P(1)\right\}\int K(u)u^{\ell}du + nh^{\ell}O_P(1/\sqrt{nh}) \\
&= nh^{\ell} f_W(w_k)\mu_{\ell}\{1 + o_P(1)\}.
\end{aligned}$$

Hence,

$$\mathbf{S}^{(k)}_{nW} = nf_W(w_k)\mathbf{HSH}\{1 + o_P(1)\}, \quad \text{for } k = 1, 2, \tag{B.9}$$

where $\mathbf{H} = \text{diag}(1, h, \ldots, h^p)$ and $\mathbf{S} = (\mu_{\ell_1+\ell_2})_{0\le\ell_1,\ell_2\le p}$.

B.2.2. *Approximate* $\mathbf{S}^*_{nW}$

For $\ell_1, \ell_2 = 0, 1, \ldots, p$, $S^*_{nW,\ell_1,\ell_2} = E\left(S^*_{nW,\ell_1,\ell_2}\right) + O_P\left\{\sqrt{\mathrm{Var}\left(S^*_{nW,\ell_1,\ell_2}\right)}\right\}$, where

$$\begin{aligned}
& E\left(S^*_{nW,\ell_1,\ell_2}\right) \\
= \; & n\int (w-w_1)^{\ell_1}(w-w_2)^{\ell_2} K_h(w-w_1)K_h(w-w_2)\nu^2(w) f_W(w)dw \\
= \; & nh^{\ell_1+\ell_2-1}\int\left(u-\frac{w_1-w_2}{2h}\right)^{\ell_1}\left(u+\frac{w_1-w_2}{2h}\right)^{\ell_2} K\left(u-\frac{w_1-w_2}{2h}\right)K\left(u+\frac{w_1-w_2}{2h}\right)\times \\
& \nu^2\left(hu+\frac{w_1+w_2}{2}\right) f_W\left(hu+\frac{w_1+w_2}{2}\right)du \\
= \; & nh^{\ell_1+\ell_2-1}\{\nu^2_W\left(\frac{w_1+w_2}{2}\right) f_W\left(\frac{w_1+w_2}{2}\right)\times \\
& \int\left(u-\frac{w_1-w_2}{2h}\right)^{\ell_1}\left(u+\frac{w_1-w_2}{2h}\right)^{\ell_2} K\left(u-\frac{w_1-w_2}{2h}\right)K\left(u+\frac{w_1-w_2}{2h}\right)du + o(1)\} \\
= \; & nh^{\ell_1+\ell_2-1}\left\{\nu^2\left(\frac{w_1+w_2}{2}\right) f_W\left(\frac{w_1+w_2}{2}\right)\xi_{\ell_1,\ell_2}\left(\frac{w_1-w_2}{2}, h\right) + o(1)\right\}, \qquad \text{(B.10)}
\end{aligned}$$

in which

$$\xi_{\ell_1,\ell_2}(w,h) = \int (u-w/h)^{\ell_1}(u+w/h)^{\ell_2}K(u-w/h)K(u+w/h)du; \qquad \text{(B.11)}$$

and

$$\begin{aligned}
& \mathrm{Var}\left(S^*_{nW,\ell_1,\ell_2}\right) \\
= \; & nE\left\{(W_1-w_1)^{2\ell_1}(W_1-w_2)^{2\ell_2}K_h^2(W_1-w_1)K_h^2(W_1-w_2)\nu^4(W_1)\right\} - \\
& n\left[E\left\{(W_1-w_1)^{\ell_1}(W_1-w_2)^{\ell_2}K_h(W_1-w_1)K_h(W_1-w_2)\nu^2(W_1)\right\}\right]^2 \\
= \; & n\int (w-w_1)^{2\ell_1}(w-w_2)^{2\ell_2}K_h^2(w-w_1)K_h^2(w-w_2)\nu^4(w)f_W(w)dw - \\
& n\left[h^{\ell_1+\ell_2-1}\left\{\nu^2\left(\frac{w_1+w_2}{2}\right) f_W\left(\frac{w_1+w_2}{2}\right)\xi^{(h)}_{\ell_1,\ell_2}\left(\frac{w_1-w_2}{2}\right) + o(1)\right\}\right]^2 \\
= \; & nh^{2(\ell_1+\ell_2-1)-1}\int\left(u-\frac{w_1-w_2}{2h}\right)^{2\ell_1}\left(u+\frac{w_1-w_2}{2h}\right)^{2\ell_2}K^2\left(u-\frac{w_1-w_2}{2h}\right)\times \\
& K^2\left(u+\frac{w_1-w_2}{2h}\right)\nu^4\left(hu+\frac{w_1+w_2}{2}\right) f_W\left(hu+\frac{w_1+w_2}{2}\right)du \\
& +o\left\{nh^{2(\ell_1+\ell_2-1)-1}\right\} \qquad \text{(B.12)} \\
= \; & nh^{2(\ell_1+\ell_2-1)-1}\zeta_{\ell_1,\ell_2}\left(\frac{w_1-w_2}{2}\right)\left\{\nu^4\left(\frac{w_1+w_2}{2}\right) f_W\left(\frac{w_1+w_2}{2}\right)\right. \\
& \left.+o(1)\right\} + o\left\{nh^{2(\ell_1+\ell_2-1)-1}\right\},
\end{aligned}$$

in which

$$\zeta_{\ell_1,\ell_2}(w,h) = \int (u-w/h)^{2\ell_1}(u+w/h)^{2\ell_2}K^2(u-w/h)K^2(u+w/h)dw. \tag{B.13}$$

Note that (B.12) is reached under the assumptions that $\nu^2(w)$ is bounded, and $\xi_{\ell_1,\ell_2}(w,h)$ is bounded for all w, $h>0$, and $0\le \ell_1,\ell_2\le p$.

Now we see that, if $\zeta_{\ell_1,\ell_2}(w,h)$ is bounded for all w, $h>0$, and $0\le \ell_1,\ell_2\le p$,

$$O_P\left\{\sqrt{\mathrm{Var}\left(S^*_{nW,\ell_1,\ell_2}\right)}\right\} = nh^{\ell_1+\ell_2-1}O_P(1/\sqrt{nh}). \tag{B.14}$$

Combining (B.10) and (B.14), we have

$$S^*_{nW,\ell_1,\ell_2} = nh^{\ell_1+\ell_2-1}f_W\left(\frac{w_1+w_2}{2}\right)\nu^2\left(\frac{w_1+w_2}{2}\right)\xi_{\ell_1,\ell_2}\left(\frac{w_1-w_2}{2},h\right)\{1+o_P(1)\}, \tag{B.15}$$

which is similar to (3.56) in Fan and Gijbels (1996) although they have $\nu_{\ell_1+\ell_2}$ in the place of $\xi_{\ell_1,\ell_2}\{(w_1-w_2)/2,h\}$ above, where $\nu_\ell = \int u^\ell K^2(u)du$. We shall point out that their ν_ℓ is free of h, whereas our $\xi_{\ell_1,\ell_2}\{(w_1-w_2)/2,h\}$ depends on h. In fact, the dependence of $\xi_{\ell_1,\ell_2}\{(w_1-w_2)/2,h\}$ on h is crucial in the follow-up derivations.

Putting (B.15) inside the matrix in (B.7), we have

$$\mathbf{S}^*_{nW} = nh^{-1}f_W\left(\frac{w_1+w_2}{2}\right)\nu^2\left(\frac{w_1+w_2}{2}\right)\mathbf{H}\mathbf{S}^*_{W,h}\mathbf{H}\{1+o_P(1)\}, \tag{B.16}$$

where

$$\mathbf{S}^*_{W,h} = \left(\xi_{\ell_1,\ell_2}\left(\frac{w_1-w_2}{2},h\right)\right)_{0\le\ell_1,\ell_2\le p}. \tag{B.17}$$

The result in (B.16) is the counterpart of (3.57) in Fan and Gijbels (1996)

B.3. Step 3: Go from Cov $\{\mathcal{A}(w_1),\,\mathcal{A}(w_2)|\mathbb{W}\}$ to Var $\{\mathcal{B}(x)|\mathbb{W}\}$

Substituting (B.9) and (B.16) in (B.6) yields

$$\mathrm{Cov}(\hat{\boldsymbol\beta}^*_1,\,\hat{\boldsymbol\beta}^*_2|\mathbb{W}) = \frac{\nu^2\{(w_1+w_2)/2\}}{nh}\frac{f_W\{(w_1+w_2)/2\}}{f_W(w_1)f_W(w_2)}\mathbf{H}^{-1}\mathbf{S}^{-1}\mathbf{S}^*_{W,h}\mathbf{S}^{-1}\mathbf{H}^{-1}\{1+o_P(1)\}, \tag{B.18}$$

which is the counterpart of (3.58) in Fan and Gijbels (1996). Hence

$$\mathrm{Cov}\{\hat m^*(w_1),\,\hat m^*(w_2)|\mathbb{W}\} = \frac{\nu^2\{(w_1+w_2)/2\}}{nh}\frac{f_W\{(w_1+w_2)/2\}}{f_W(w_1)f_W(w_2)}\boldsymbol{e}_1^{\mathrm{T}}\mathbf{S}^{-1}\mathbf{S}^*_{W,h}\mathbf{S}^{-1}\boldsymbol{e}_1+o_P\left(\frac{1}{nh}\right), \tag{B.19}$$

as a counterpart of (3.7) in Fan and Gijbels (1996). Finally, by (B.4), we have

$$\mathrm{Cov}\left\{\mathcal{A}(w_1),\,\mathcal{A}(w_2)|\mathbb{W}\right\} = \frac{\gamma\{(w_1+w_2)/2\}}{nh}\boldsymbol{e}_1^{\mathrm{T}}\mathbf{S}^{-1}\mathbf{S}^*_{W,h}\mathbf{S}^{-1}\boldsymbol{e}_1+o_P\left(\frac{1}{nh}\right), \tag{B.20}$$

where $\gamma(w) = \nu^2(w) f_W(w)$.

Plugging (B.20) in (B.3) gives

$$\begin{aligned} \text{Var}\{\mathcal{B}(x)|\mathbb{W}\} &= \int D(x-w_1)\int D(x-w_2)\Big[\frac{\gamma\{(w_1+w_2)/2\}}{nh}\boldsymbol{e}_1^{\mathrm{T}}\mathbf{S}^{-1}\mathbf{S}^*_{W,h}\mathbf{S}^{-1}\boldsymbol{e}_1 \\ &\quad + o_P\left(\frac{1}{nh}\right)\Big]dw_2dw_1. \end{aligned} \tag{B.21}$$

Given the definition of $\mathbf{S}^*_{W,h}$ in (B.17) and the definition of its entries in (B.11), we shall elaborate the following integral,

$$\int D(x-w_1)\int D(x-w_2)\xi_{\ell_1,\ell_2}\left(\frac{w_1-w_2}{2},h\right)\gamma\left(\frac{w_1+w_2}{2}\right)dw_2dw_1. \tag{B.22}$$

The next step tackles this integral in detail.

B.4. Step 4: Elaborate (B.22)

First, substituting $\xi_{\ell_1,\ell_2}(\cdot,\cdot)$ (B.22) with its definition in (B.11) yields

$$\begin{aligned} &\int D(x-w_1)\int D(x-w_2)\gamma\left(\frac{w_1+w_2}{2}\right)\int\left(u-\frac{w_1-w_2}{2h}\right)^{\ell_1}\times \\ &\left(u+\frac{w_1-w_2}{2h}\right)^{\ell_2}K\left(u-\frac{w_1-w_2}{2h}\right)K\left(u+\frac{w_1-w_2}{2h}\right)dudw_2dw_1. \end{aligned} \tag{B.23}$$

Using multivariate change-of-variable and letting $s_1 = u-(w_1-w_2)/(2h)$ and $s_2 = u+(w_1-w_2)/(2h)$, (B.23) becomes

$$\begin{aligned} & h\int D(x-w_2)\int s_1^{\ell_1}K(s_1)\int D\{x-w_2-h(s_2-s_1)\}\gamma\{w_2+h(s_2-s_1)/2\}\times \\ & s_2^{\ell_2}K(s_2)ds_2ds_1dw_2 \\ =\; & h\int D(x-w_2)\{\gamma(w_2)+O(h)\}\int s_1^{\ell_1}K(s_1)\int D\{x-w_2-h(s_2-s_1)\}\times \\ & s_2^{\ell_2}K(s_2)ds_2ds_1dw_2. \end{aligned} \tag{B.24}$$

Second, zooming on the inner integral (with respect to s_2) in (B.24), we have

$$\begin{aligned} &\int D\{x-w_2-h(s_2-s_1)\}s_2^{\ell_2}K(s_2)ds_2 \\ =\; & (2\pi)^{-1}\int e^{-it\{x-w_2-h(s_2-s_1)\}}\{\phi_U(t)\}^{-1}\int e^{iths_2}s_2^{\ell_2}K(s_2)ds_2dt. \end{aligned}$$

Using the fact that $\phi_K^{(\ell)}(t) = i^\ell \int e^{itv} v^\ell K(v)dv$, the preceding expression is equal to

$$\begin{aligned} & i^{-\ell_2}(2\pi)^{-1}\int e^{-it(x-w_2+hs_1)}\frac{\phi_K^{(\ell_2)}(th)}{\phi_U(t)}dt \\ = & h^{-1}i^{-\ell_2}(2\pi)^{-1}\int e^{-is(x-w_2+hs_1)/h}\frac{\phi_K^{(\ell_2)}(s)}{\phi_U(s/h)}ds \\ = & h^{-1}K_{U,\ell_2}\{(x-w_2)/h+s_1\}, \end{aligned} \quad \text{(B.25)}$$

where we use equation (5) in the main article to introduce the "transformed kernel" in Delaigle, Fan, and Carroll (2009), $K_{U,\ell}(x)$.

Third, putting (B.25) back in (B.24) to deal with the remaining two-dimensional integral (with respect to s_1 and w_2), we have

$$\int D(x-w_2)\{\gamma(w_2)+O(h)\}\int s_1^{\ell_1}K(s_1)K_{U,\ell_2}\left(\frac{x-w_2}{h}+s_1\right)ds_1dw_2.$$

Letting $v = (s-w_2)/h + s_1$, the above integral is equal to

$$\begin{aligned} & h\int K_{\mathrm{U},\ell_2}(v)\int D\{(v-s_1)h\}[\gamma\{x-(v-s_1)h\}+O(h)]s_1^{\ell_1}K(s_1)ds_1dv \\ = & h\{\gamma(x)+O(h)\}\int K_{\mathrm{U},\ell_2}(v)\int(2\pi)^{-1}e^{-ithv}\frac{1}{\phi_U(t)}\int e^{iths_1}s_1^{\ell_1}K(s_1)ds_1dtdv \\ = & \{\gamma(x)+O(h)\}\int K_{\mathrm{U},\ell_2}(v)i^{-\ell_1}(2\pi)^{-1}\int e^{-isv}\frac{\phi_K^{(\ell_1)}(s)}{\phi_U(s/h)}dsdv \\ = & \{\gamma(x)+O(h)\}\int K_{\mathrm{U},\ell_1}(v)K_{\mathrm{U},\ell_2}(v)dv. \end{aligned} \quad \text{(B.26)}$$

B.5. Step 5: Lemmas needed for elaborating (B.26)

To elaborate (B.26), as related in Section 4.1 in the main article, we use directly Lemma B.4, Lemma B.6 (for ordinary smooth U) and Lemma B.9 (for super smooth U) in Delaigle et al. (2009). For completeness, these lemmas are restated next.

Lemma B.4: Assume that, for $\ell = \ell_1, \ell_2$, $\|\phi_K^{(\ell)}\|_\infty < \infty$, $\|\phi_K^{(\ell+1)}\|_\infty < \infty$, $\|\phi_U'\|_\infty < \infty$, $\int(|t|^b+|t|^{b-1})\{|\phi_K^{(\ell)}|+|\phi_K^{(\ell+1)}|\}dt < \infty$, and $\int |t|^b|\phi_K^{(\ell)}|dt < \infty$, then, for a bounded function g,

$$\begin{aligned} & \lim_{n\to\infty} h^{2b}\int K_{U,\ell_1}(v)K_{U,\ell_2}(v)g(x-hv)dv \\ = & i^{-\ell_1-\ell_2}(-1)^{-\ell_2}\frac{g(x)}{c^2}\frac{1}{2\pi}\int |t|^{2b}\phi_K^{(\ell_1)}(t)\phi_K^{(\ell_2)}(t)dt. \end{aligned}$$

Lemma B.6: Suppose, for $\ell = \ell_1, \ell_2$, $\|\phi_K^{(\ell)}(t)\|_\infty < \infty$ and $\int |t|^{2b}|\phi_K^{(\ell)}(t)|^2dt < \infty$. Then $|\int_{-\infty}^{\infty} K_{U,\ell_1}(v)K_{U,\ell_2}(v)dv| \le Ch^{-2b}$ for some finite positive constant C.

Lemma B.9: Suppose that $\phi_K(t)$ is supported on $[-1,1]$, and, for $\ell = \ell_1$ and ℓ_2, $\|\phi_K^{(\ell)}(t)\|_\infty < \infty$. Then $|\int_{-\infty}^{\infty} K_{U,\ell_1}(v)K_{U,\ell_2}(v)dv| \le Ch^{2b_2}\exp(2h^{-b}/d_2)$, where

$b_2 = b_0 I(b_0 < 1/2)$.

The conditions required in Lemma B.6 and Lemma B.9 are included in or implied by **Condition O** (for ordinary smooth U) and **Condition S** (for super smooth U), respectively.

B.6. Elaboration of $\gamma(\cdot)$

Define $\tau^2(x) = \text{Var}(Y|X = x)$, then

$$\begin{aligned}
\nu^2(w) &= \text{Var}(Y|W = w) \\
&= E\{\text{Var}(Y|X)|W = w\} + \text{Var}\{E(Y|X)|W = w\} \\
&= E\left\{\tau^2(X)|W = w\right\} + \text{Var}\{m(X)|W = w\} \\
&= \{f_W(w)\}^{-1}\int \tau^2(x)f_X(x)f_U(w-x)dx + \\
&\quad E\left\{m^2(X)|W = w\right\} - [E\{m(X)|W = w\}]^2 \\
&= \frac{\int\left\{\tau^2(x) + m^2(x)\right\}f_X(x)f_U(w-x)dx}{f_W(w)} - \frac{\left\{\int m(x)f_X(x)f_U(w-x)dx\right\}^2}{\{f_W(w)\}^2} \\
&= \frac{\left\{\left(\tau^2 + m^2\right)f_X\right\} * f_U(w)}{f_W(w)} - \frac{\{(mf_X) * f_U(w)\}^2}{\{f_W(w)\}^2}.
\end{aligned}$$

In the above elaboration, we use the following identity according to Billingsley (1979, Theorem 34.4),

$$E\{g(Y)|W\} = E\left[E\{g(Y)|X, W\}|W\right],$$

where $g(\cdot)$ a generic function such that the relevant expectations exist. Under the assumption of nondifferential measurement error, the right-hand side of this identity is equal to $E[E\{g(Y)|X\}|W]$.

It follows that

$$\begin{aligned}
\gamma(w) &= \left\{\left(\tau^2 + m^2\right)f_X\right\} * f_U(w) - \{f_W(w)\}^{-1}\{(mf_X) * f_U(w)\}^2 \\
&= \left\{\left(\tau^2 + m^2\right)f_X\right\} * f_U(w) - \frac{\{(mf_X) * f_U(w)\}^2}{(f_X * f_U)(w)}.
\end{aligned}$$

Appendix C: Asymptotic normality of $\hat{m}_{\text{HZ}}(x)$

C.1. A sufficient condition for asymptotic normality

To show $\hat{m}_{\text{HZ}}(x)$ converges in distribution to a normal distribution as $n \to \infty$, by Slutsky's Theorem and the fact that $\hat{f}_X(x) \xrightarrow{p} f(x)$ (Stefanski and Carroll 1990, Theorem 2.1), it suffices to show the asymptotic normality for the difference $\mathcal{B}(x) - B(x)$, that is,

$$\hat{m}_{\mathrm{HZ}}(x)\hat{f}_X(x) - m(x)f_X(x) = \frac{1}{2\pi}\int e^{-itx}\frac{\phi_{\hat{m}^*\hat{f}_W - m^*f_W}(t)}{\phi_U(t)}dt = \{(\mathcal{A} - A) * D\}(x). \quad \text{(C.1)}$$

To show the asymptotic normality of (C.1), we first show that (C.1) can be approximated by an average, $n^{-1}\sum_{j=1}^n \tilde{U}_{n,j}(x)$, for some independent and identically distributed (i.i.d) random variables (at each fixed x) $\{\tilde{U}_{n,j}(x)\}_{j=1}^n$, each of which depends on n via its dependence on h. Then we show that, for some positive constant η,

$$\lim_{n\to\infty}\frac{E|\tilde{U}_{n,1}|^{2+\eta}}{n^{\eta/2}\{E(\tilde{U}_{n,1}^2)\}^{(2+\eta)/2}} = 0, \quad \text{(C.2)}$$

which is a sufficient condition for

$$\frac{\sum_{j=1}^n \tilde{U}_{n,j} - nE(\tilde{U}_{n,j})}{\sqrt{n\mathrm{Var}(\tilde{U}_{n,j})}} \xrightarrow{\mathcal{L}} N(0,1).$$

Because convolution is a linear operator, approximating (C.1) via an average of i.i.d random variables can be realized by approximating $\mathcal{A}(w) - A(w)$ via an average of another set of n i.i.d random variables at a fixed w in the support of W. We achieve this goal following four steps described next.

C.2. Step 1: Re-express $\mathcal{A}(w) - A(w)$ as a summation

Assuming $\hat{m}^*(w)$ bounded, we have

$$\begin{aligned}\mathcal{A}(w) - A(w) &= \hat{m}^*(w)\hat{f}_W(w) - m^*(w)f_W(w) \\ &= \left\{\sum_{\ell=0}^{p} S_{nW}^{0,\ell}(w)T_{nW,\ell}(w) - m^*(w)\right\} f_W(w) + o_P(1), \end{aligned} \quad \text{(C.3)}$$

where $T_{nW,\ell} = \sum_{j=1}^n Y_j(W_j - w)^\ell K_h(W_j - w)$, for $\ell = 0, 1, \ldots, p$, and $S_{nW}^{0,\ell}(w)$ is the $[1, \ell+1]$ element of $\mathbf{S}_{nW}^{-1}(w)$, with $\mathbf{S}_{nW}(w) = (S_{nW,\ell_1+\ell_2}(w))_{0\le\ell_1,\ell_2\le p}$ and

$$S_{nW,\ell} = \sum_{j=1}^n (W_j - w)^\ell K_h(W_j - w), \text{ for } \ell = 0, 1, \ldots, 2p.$$

Because $\sum_{\ell=0}^p S_{nW}^{0,\ell}(w)S_{nW,\ell+\ell'}(w) = I(\ell' = 0)$, where $I(\cdot)$ is the indicator function, inside the curly brackets in (C.3) we have

$$\begin{aligned}
&\sum_{\ell=0}^{p} S_{nW}^{0,\ell}(w) T_{nW,\ell}(w) - m^*(w) \\
=\ & \sum_{\ell=0}^{p} S_{nW}^{0,\ell}(w) T_{nW,\ell}(w) - \sum_{\ell'=0}^{p} h^{\ell'} \frac{m^{*(\ell')}(w)}{\ell'!} \sum_{\ell=0}^{p} S_{nW}^{0,\ell}(w) S_{nW,\ell+\ell'}(w) \\
=\ & \sum_{\ell=0}^{p} S_{nW}^{0,\ell}(w) \left\{ T_{nW,\ell}(w) - \sum_{\ell'=0}^{p} h^{\ell'} \frac{m^{*(\ell')}(w)}{\ell'!} S_{nW,\ell+\ell'}(w) \right\} \\
=\ & \sum_{\ell=0}^{p} S_{nW}^{0,\ell}(w) T^*_{nW,\ell}(w), \qquad \text{(C.4)}
\end{aligned}$$

where $T^*_{nW,\ell}(w) = T_{nW,\ell}(w) - \sum_{\ell'=0}^{p} h^{\ell'}\{m^{*(\ell')}(w)/\ell'!\} S_{nW,\ell+\ell'}(w)$, for $\ell = 0, 1, \ldots, p$.

In what follows, we show that (C.4) is equivalent to

$$\begin{aligned}
&\sum_{\ell=0}^{p} [\text{a nonrandom function of } w \text{ as an approximation of } nS_{nW}^{0,\ell}(w)] \\
\times\ & \left\{ n^{-1} T^*_{nW,\ell}(w) \right\} + O_P(h^{\text{some positive power}}).
\end{aligned}$$

This is accomplished in two steps. First, studying $n^{-1}T^*_{nW,\ell}(w)$ to understand its order in h. Second, approximating $nS_{nW}^{0,\ell}(w)$.

C.3. Step 2: The order of $n^{-1}T^*_{nW,\ell}(w)$

Because $n^{-1}T^*_{nW,\ell}(w) = E\{n^{-1}T^*_{nW,\ell}(w)\} + O_P[\sqrt{\text{Var}\{n^{-1}T^*_{nW,\ell}(w)\}}]$, we study the order of the expectation and variance separately in this section.

For the expectation, we have $E\{n^{-1}T^*_{nW,\ell}(w)\}$ equal to

$$\begin{aligned}
& E\left\{n^{-1} T_{nW,\ell}(w)\right\} - \sum_{\ell'=0}^{p} h^{\ell'} \frac{m^{*(\ell')}(w)}{\ell'!} E\left\{n^{-1} S_{nW,\ell+\ell'}(w)\right\} \\
=\ & E\left\{Y(W-w)^{\ell} K_h(W-w)\right\} - \sum_{\ell'=0}^{p} h^{\ell'} \frac{m^{*(\ell')}(w)}{\ell'!} E\left\{(W-w)^{\ell+\ell'} K_h(W-w)\right\} \\
=\ & \int m(v_2) \int h^{\ell} v_1^{\ell} K(v_1) f_U(hv_1 - v_2 + w) dv_1 f_X(v_2) dv_2 - \\
& \sum_{\ell'=0}^{p} h^{\ell'} \frac{m^{*(\ell')}(w)}{\ell'!} \int\int h^{\ell+\ell'} v_1^{\ell+\ell'} K(v_1) f_U(hv_1 - v_2 + w) dv_1 f_X(v_2) dv_2.
\end{aligned}$$

Using the first-order Taylor expansion of $f_U(hv_1 - v_2 + w)$ around $h = 0$, i.e.,

$$f_U(hv_1 - v_2 + w) = f_U(w - v_2) + h f_U'(w - v_2) v_1 + O(h^2), \qquad \text{(C.5)}$$

in the above integral, we have $E\{n^{-1}T^*_{nW,\ell}(w)\}$ equal to

$$\begin{aligned}
& h^\ell\Big\{\mu_\ell \int m(v_2)f_U(w-v_2)f_X(v_2)dv_2 + h\mu_{\ell+1}\int m(v_2)f'_U(w-v_2)f_X(v_2)dv_2 \\
& +O(h^2)\Big\} - \sum_{\ell'=0}^{p} h^{\ell+2\ell'}\frac{m^{*(\ell')}(w)}{\ell'!}\Big\{\mu_{\ell+\ell'}f_W(w)+ \\
& h\mu_{\ell+\ell'+1}\int f'_U(w-v_2)f_X(v_2)dv_2 + O(h^2)\Big\} \\
= & \; h^\ell\mu_\ell f_W(w)m^*(w) + h^{\ell+1}\mu_{\ell+1}\int m(v_2)f'_U(w-v_2)f_X(v_2)dv_2 - h^\ell\mu_\ell f_W(w)m^*(w) \\
& -h^{\ell+2}m^{*'}(w)\mu_{\ell+1}f_W(w) - h^{\ell+1}m^*(w)\mu_{\ell+1}\int f'_U(w-v_2)f_X(v_2)dv_2 + O(h^{\ell+2}) \\
= & \; h^{\ell+1}\mu_{\ell+1}\int \{m(v_2)-m^*(w)\}\, f'_U(w-v_2)f_X(v_2)dv_2 - h^{\ell+2}m^{*'}(w)\mu_{\ell+1}f_W(w) \\
& +O(h^{\ell+2}) \\
= & \left\{\begin{array}{ll} O(h^{\ell+1}) & \text{if } \ell \text{ is odd} \\ O(h^{\ell+2}) & \text{if } \ell \text{ is even}\end{array}\right. . \qquad \text{(C.6)}
\end{aligned}$$

For the variance, we have $\text{Var}\{n^{-1}T^*_{nW,\ell}(w)\}$ equal to

$$\begin{aligned}
& \text{Var}\left\{n^{-1}T_{nW,\ell}(w) - n^{-1}\sum_{\ell'=0}^{p} h^{\ell'}\frac{m^{*(\ell')}(w)}{\ell'!}S_{nW,\ell+\ell'}(w)\right\} \\
= & \; \text{Var}\left\{n^{-1}T_{nW,\ell}(w) - n^{-1}m^*(w)S_{nW,\ell}(w) - n^{-1}\sum_{\ell'=1}^{p} h^{\ell'}\frac{m^{*(\ell')}(w)}{\ell'!}S_{nW,\ell+\ell'}(w)\right\} \\
= & \; O\left[\text{Var}\left\{n^{-1}T_{nW,\ell}(w) - n^{-1}m^*(w)S_{nW,\ell}(w)\right\}\right] + O\left[\sum_{\ell'=1}^{p}\text{Var}\left\{n^{-1}S_{nW,\ell+\ell'}(w)\right\}\right]. \qquad \text{(C.7)}
\end{aligned}$$

Looking into the first variance in (C.7), we have

$$\begin{aligned}
& \text{Var}\left\{n^{-1}T_{nW,\ell}(w) - n^{-1}m^*(w)S_{nW,\ell}(w)\right\} \\
= & \; \text{Var}\left[n^{-1}\sum_{j=1}^{n}\{Y_j - m^*(w)\}(W_j-w)^\ell K_h(W_j-w)\right] \\
\le & \; \frac{1}{nh^2}E\left[\{Y-m^*(w)\}^2(W-w)^{2\ell}K^2\left(\frac{W-w}{h}\right)\right] \\
= & \; \frac{1}{nh^2}\int E\left[\{Y-m^*(w)\}^2|X=v_2\right]\int h^{2\ell+1}v_1^{2\ell}K^2(v_1)f_U(hv_1-v_2+w)dv_1 f_X(v_2)dv_2.
\end{aligned}$$

Define $\tilde{\kappa}(w,X) = E\left[\{Y-m^*(w)\}^2|X\right]$ and $\delta_\ell = \int v^\ell K^2(v)dv$ for $\ell = 0, 1, \ldots, 2p$, using

(C.5), the preceding expression becomes

$$\begin{aligned}
& \frac{h^{2\ell}}{nh}\int \tilde{\kappa}(w,v_2)\left\{\delta_{2\ell}f_U(w-v_2)+h\delta_{2\ell+1}f_U'(w-v_2)+O(h^2)\right\}f_X(v_2)dv_2 \\
= & \frac{h^{2\ell}}{nh}\delta_{2\ell}E\left\{\tilde{\kappa}(w,X)f_U(w-X)\right\}+\frac{h^{2\ell+1}}{nh}\delta_{2\ell+1}E\left\{\tilde{\kappa}(w,X)f_U'(w-X)\right\}+\frac{1}{nh}O(h^{2\ell+2}) \\
= & \frac{h^{2\ell}}{nh}\delta_{2\ell}E\left\{\tilde{\kappa}(w,X)f_U(w-X)\right\}+\frac{1}{nh}O(h^{2\ell+2}),\ \text{as } \delta_k=0 \text{ when } k \text{ is odd.} \\
= & O\left(\frac{h^{2\ell}}{nh}\right),\ \text{assuming } \delta_{2\ell}E\left\{\tilde{\kappa}(w,X)f_U(w-X)\right\} \text{ bounded and nonzero.}
\end{aligned}$$

Hence, $\text{Var}\left\{n^{-1}T_{nW,\ell}(w)-n^{-1}m^*(w)S_{nW,\ell}(w)\right\}$ is bounded from above by some nonrandom quantity of order $(nh)^{-1}O\left(h^{2\ell}\right)$.

Similarly, for the second variance in (C.7), we show that

$$\text{Var}\{n^{-1}S_{nW,\ell+\ell'}(w)\}\leq h^{2(\ell+\ell')}/(nh)\delta_{2k}f_W(w)+(nh)^{-1}O(h^{2k+2}).$$

Because $\ell+\ell'\geq\ell+1$ in (C.7), assuming $\delta_{2k}f_W(w)$ bounded and nonzero, we see that $\sum_{\ell'=1}^{p}\text{Var}\left\{n^{-1}S_{nW,\ell+\ell'}(w)\right\}$ is bounded from above by some nonrandom quantity of order $(nh)^{-1}O(h^{2\ell+2})$, which converges to 0 faster than the first variance in (C.7). Therefore, $\text{Var}\{n^{-1}T^*_{nW,\ell}(w)\}\leq Ch^{2\ell}/(nh)$, for some positive constant C. If $1/\sqrt{nh}=O(h^2)$, i.e., $h=O(n^{-1/5})$, then $\sqrt{\text{Var}\{n^{-1}T^*_{nW,\ell}(w)\}}\leq\sqrt{C}h^{\ell+2}$, which tends to 0 at least as fast as $E\{n^{-1}T^*_{nW,\ell}(w)\}$ according to (C.6).

In conclusion, we establish that

$$n^{-1}T^*_{nW,\ell}(w)=\left\{\begin{array}{ll} h^{\ell}\left\{Ch+O_P(h^2)\right\} & \text{if } \ell \text{ is odd} \\ h^{\ell}\left\{C'h^2+O_P(h^3)\right\} & \text{if } \ell \text{ is even}\end{array}\right., \tag{C.8}$$

for some finite nonzero nonrandom quantities C and C' that depend on w (but not on n). This completes the first task stated in Section C.2.

C.4. Step 3: Approximate $nS^{0,\ell}_{nW}(w)$

Since $S^{0,\ell}_{nW}(w)$ is an element of $\mathbf{S}^{-1}_{nW}(w)$, we may first study the elements in $\mathbf{S}_{nW}(w)$, namely $S_{nW,\ell}(w)$. Following similar strategies used in Section C.3, we begin with $n^{-1}S_{nW,\ell}(w)=E\{n^{-1}S_{nW,\ell}(w)\}+O_P[\sqrt{\text{Var}\{n^{-1}S_{nW,\ell}(w)\}}]$

For the expectation above, we have $E\{n^{-1}S_{nW,\ell}(w)\}$ equal to

$$
\begin{aligned}
& E\left\{(W-w)^{\ell}K_h(W-w)\right\} \\
= & \int\int (u+v_2-w)^{\ell}h^{-1}K\left(\frac{u+v_2-w}{h}\right) f_U(u)du f_X(v_2)dv_2 \\
= & h^{\ell}\int\int v_1^{\ell}K(v_1)f_U(hv_1-v_2+w)dv_1 f_X(v_2)dv_2 \\
= & h^{\ell}\int\int v_1^{\ell}K(v_1)\left\{f_U(w-v_2)+hf_U'(w-v_2)v_1+O(h^2)\right\}dv_1 f_X(v_2)dv_2 \\
= & h^{\ell}\left[\mu_{\ell}f_W(w)+h\mu_{\ell+1}E\left\{f_U'(w-X)\right\}+O(h^2)\right]. \qquad (C.9)
\end{aligned}
$$

For the variance, we have $\text{Var}\{n^{-1}S_{nW,\ell}(w)\}$ equal to

$$
\begin{aligned}
& n^{-1}\text{Var}\left\{(W-w)^{\ell}K_h(W-w)\right\} \\
\leq & \frac{1}{nh^2}E\left\{(W-w)^{2\ell}K^2\left(\frac{W-w}{h}\right)\right\} \\
= & \frac{1}{nh^2}\int\int h^{2\ell+1}v_1^{2\ell}K^2(v_1)f_U(hv_1-v_2+w)dv_1 f_X(v_2)dv_2 \\
= & \frac{1}{nh^2}\int\int h^{2\ell+1}v_1^{2\ell}K^2(v_1)\left\{f_U(w-v_2)+hf_U'(w-v_2)v_1+O(h^2)\right\}dv_1 f_X(v_2)dv_2 \\
= & \frac{1}{nh^2}h^{2\ell+1}\left[\delta_{2\ell}f_W(w)+h\delta_{2\ell+1}E\left\{f_U'(w-X)\right\}+O(h^2)\right] \\
= & \frac{h^{2\ell}}{nh}\left\{\delta_{2\ell}f_W(w)+O(h^2)\right\}.
\end{aligned}
$$

Therefore, $\sqrt{\text{Var}\{n^{-1}S_{nW,\ell}(w)\}} \leq Ch^{\ell}/\sqrt{nh}$, for some positive constant C that depends on w but not on n. And if $h = O(n^{-1/5})$, $\sqrt{\text{Var}\left\{n^{-1}S_{nW,\ell}(w)\right\}} \leq C'h^{\ell+2}$, for some positive constant C' that depends on w but not on n. Hence, the dominating terms in $n^{-1}S_{nW,\ell}(w)$ are in the expectation elaborated in (C.9).

Define $\widetilde{\mathbf{S}} = (\mu_{\ell_1+\ell_2+1})_{0\leq\ell_1,\ell_2\leq p}$. By (C.9), we now have

$$
n^{-1}\mathbf{S}_{nW}(w) = \mathbf{H}\left[f_W(w)\mathbf{S}+hE\left\{f_U'(w-X)\right\}\widetilde{\mathbf{S}}+O_P(h^2)\right]\mathbf{H}.
$$

It follows that

$$
\mathbf{H}\left\{n\mathbf{S}_{nW}^{-1}(w)\right\}\mathbf{H} = \left[\mathbf{I}_{p+1}+hE\left\{f_U'(w-X)\right\}f_W^{-1}(w)\mathbf{S}^{-1}\widetilde{\mathbf{S}}\right]^{-1}\mathbf{S}^{-1}f_W^{-1}(w)+O_P(h^2).
$$

Using the first order Taylor expansion of $\left[\mathbf{I}_{p+1}+hE\left\{f_U'(w-X)\right\}f_W^{-1}(w)\mathbf{S}^{-1}\widetilde{\mathbf{S}}\right]^{-1}$ around $h=0$, the above expression is equal to

$$
\begin{aligned}
& \left[\mathbf{I}_{p+1}-hE\left\{f_U'(w-X)\right\}f_W^{-1}(w)\mathbf{S}^{-1}\widetilde{\mathbf{S}}+O(h^2)\right]\mathbf{S}^{-1}f_W^{-1}(w)+O_P(h^2) \\
= & \mathbf{S}^{-1}f_W^{-1}(w)-hE\left\{f_U'(w-X)\right\}f_W^{-2}(w)\mathbf{S}^{-1}\widetilde{\mathbf{S}}\mathbf{S}^{-1}+O_P(h^2).
\end{aligned}
$$

Hence,

$$n\mathbf{S}_{nW}^{-1}(w) = \mathbf{H}^{-1}\left[\mathbf{S}^{-1}f_W^{-1}(w) - hE\left\{f_U'(w-X)\right\}f_W^{-2}(w)\mathbf{S}^{-1}\widetilde{\mathbf{S}}\mathbf{S}^{-1} + O_P(h^2)\right]\mathbf{H}^{-1}. \tag{C.10}$$

Denote by $S_Z^{0,\ell}(w)$ the $[1,\ell+1]$ element of the matrix $\mathbf{S}^{-1}f_W^{-1}(w)$, and by $\breve{S}_W^{0,\ell}(w)$ the $[1,\ell+1]$ element of the matrix $E\{f_U'(w-X)\}f_W^{-2}(w)\mathbf{S}^{-1}\widetilde{\mathbf{S}}\mathbf{S}^{-1}$, for $\ell = 1,\ldots,p$. Then (C.10) indicates that

$$nS_{nW}^{0,\ell}(w) = h^{-\ell}\left\{S_Z^{0,\ell}(w) - h\breve{S}_W^{0,\ell}(w) + O_P(h^2)\right\} = h^{-\ell}\left\{R_W^{0,\ell}(w) + O_P(h^2)\right\}, \tag{C.11}$$

where $R_W^{0,\ell}(w) = -h\breve{S}_W^{0,\ell}(w)$ if ℓ is odd, and $R_W^{0,\ell}(w) = S_Z^{0,\ell}(w)$ if ℓ is even. The definition of $R_W^{0,\ell}(w)$ comes from the observation that the locations in $\mathbf{S}$ where the elements are 0 remain to be 0 in the same locations in $\mathbf{S}^{-1}$, and the locations in $\widetilde{\mathbf{S}}$ where the elements are 0 remain to be 0 in the same locations in $\mathbf{S}^{-1}\widetilde{\mathbf{S}}\mathbf{S}^{-1}$. More specifically, with an even kernel $K(v)$, for $\ell = 0,1,\ldots,p$, the $[1,\ell+1]$ element of $\mathbf{S}$, μ_ℓ, is equal to 0 if ℓ is odd, and thus the $[1,\ell+1]$ element of $\mathbf{S}^{-1}f_W^{-1}(w)$, $S_Z^{0,\ell}(w)$, is also 0 if ℓ is odd. Similarly, the $[1,\ell+1]$ element of $\widetilde{\mathbf{S}}$, $\mu_{\ell+1}$, is equal to 0 if ℓ is even, and thus the $[1,\ell+1]$ element of $E\left\{f_U'(w-X)\right\}f_W^{-2}(w)\mathbf{S}^{-1}\widetilde{\mathbf{S}}\mathbf{S}^{-1}$, $\breve{S}_W^{0,\ell}(w)$, is also 0 if ℓ is even. This completes the second task stated in Section C.2.

C.5. Step 4: Reexpress (C.4)

By (C.11), the summand in (C.4) is equal to

$$\begin{aligned}
& \left\{nS_{nW}^{0,\ell}(w)\right\}\left\{n^{-1}T_{nW,\ell}^*(w)\right\} \\
= \; & h^{-\ell}\left\{R_W^{0,\ell}(w) + O_P(h^2)\right\}\left\{n^{-1}T_{nW,\ell}^*(w)\right\} \\
= \; & h^{-\ell}R_W^{0,\ell}(w)\left\{n^{-1}T_{nW,\ell}^*(w)\right\} + h^{-\ell}O_P(h^2)\left\{n^{-1}T_{nW,\ell}^*(w)\right\},
\end{aligned}$$

where the second term above is $O_P(h^3)$ if ℓ is odd and $O_P(h^4)$ if ℓ is even according to (C.8). Assuming $p \geq 1$ (so that ℓ is odd at least once in (C.4)), we deduce that (C.4) is equal to

$$
\begin{aligned}
&\sum_{\ell=0}^{p} h^{-\ell} R_W^{0,\ell}(w)\left\{n^{-1} T^*_{nW,\ell}(w)\right\} + O_P(h^3) \\
= &\sum_{\ell=0}^{p} h^{-\ell} R_W^{0,\ell}(w) n^{-1} \Bigg\{ \sum_{j=1}^{n} Y_j (W_j - w)^\ell K_h(W_j - w) - \\
&\sum_{\ell'=0}^{p} h^{\ell'} \frac{m^{*(\ell')}(w)}{\ell'!} \sum_{j=1}^{n} (W_j - w)^{\ell+\ell'} K_h(W_j - w) \Bigg\} + O_P(h^3) \\
= &\; n^{-1} \sum_{j=1}^{n} \Bigg[\sum_{\ell=0}^{p} h^{-\ell} R_W^{0,\ell}(w) Y_j (W_j - w)^\ell K_h(W_j - w) - \\
&\sum_{\ell=0}^{p} \sum_{\ell'=0}^{p} h^{\ell'-\ell} R_W^{0,\ell}(w) \frac{m^{*(\ell')}(w)}{\ell'!} (W_j - w)^{\ell+\ell'} K_h(W_j - w) \Bigg] + O_P(h^3) \\
= &\; n^{-1} \sum_{j=1}^{n} \Bigg[\sum_{\ell=0}^{p} h^{-\ell} R_W^{0,\ell}(w) \{Y_j - m^*(w)\} (W_j - w)^\ell K_h(W_j - w) \\
&- \sum_{\ell=0}^{p} \sum_{\ell'=1}^{p} h^{\ell'-\ell} R_W^{0,\ell}(w) \frac{m^{*(\ell')}(w)}{\ell'!} (W_j - w)^{\ell+\ell'} K_h(W_j - w) \Bigg] + O_P(h^3),
\end{aligned}
\tag{C.12}
$$

which is finally in the form of an average of n i.i.d. random variables for a fixed w plus $O_P(h^3)$. Denote by $U_{W,j}(w)$ the summand inside the square brackets in (C.12) and decompose it as $U_{W,j}(w) = P_{W,j}(w) + Q_{W,j}(w)$, where

$$
\begin{aligned}
P_{W,j}(w) &= \sum_{\ell=0}^{p} h^{-\ell} R_W^{0,\ell}(w) \{Y_j - m^*(w)\} (W_j - w)^\ell K_h(W_j - w), \\
Q_{W,j}(w) &= -\sum_{\ell=0}^{p} \sum_{\ell'=1}^{p} h^{\ell'-\ell} R_W^{0,\ell}(w) \frac{m^{*(\ell')}(w)}{\ell'!} (W_j - w)^{\ell+\ell'} K_h(W_j - w).
\end{aligned}
$$

Now (C.3) reduces to

$$
\hat{m}^*(w)\hat{f}_W(w) - m^*(w) f_W(w) = n^{-1} \sum_{j=1}^{n} U_{W,j}(w) f_W(w) \{1 + O_P(1)\}. \tag{C.13}
$$

After repeating the exercise already seen in Sections C.3 and C.4, by looking into $\hat{f}_W(w) = E\{\hat{f}_W(w)\} + O_P[\sqrt{\mathrm{Var}\{\hat{f}_W(w)\}}]$, we show that $\hat{f}_W(w) = f_W(w) + O_P(h^2)$. So the $O_P(h^3)$ in (C.12) is dominated by (or absorbed in) this $O_P(h^2)$. Hence, we actually have

$$
\hat{m}^*(w)\hat{f}_W(w) - m^*(w) f_W(w) = n^{-1} \sum_{j=1}^{n} U_{W,j}(w) f_W(w) + O_P(h^2)
$$

before concluding (C.13). We reach (C.13) by showing that $n^{-1}\sum_{j=1}^{n} U_{W,j}(w) f_W(w)$ is of order $O_P(h^2)$ or tends to zero at a slower rate than h^2.

Finally, plugging (C.13) in (C.1), we obtain the following desired form,

$$\hat{m}_{\mathrm{HZ}}(x)\hat{f}_X(x) - m(x)f_X(x) = \left\{ n^{-1}\sum_{j=1}^{n} \frac{1}{2\pi}\int e^{-itx}\frac{\phi_{U_{W,j}f_W}(t)}{\phi_U(t)}dt \right\}\{1+O_P(1)\}$$
$$= \left\{ n^{-1}\sum_{j=1}^{n} \tilde{U}_{n,j}(x) \right\}\{1+O_P(1)\}, \qquad (C.14)$$

where, for $j = 1, \ldots, n$, $\tilde{U}_{n,j}(x) = \tilde{P}_{n,j}(x) + \tilde{Q}_{n,j}(x)$, with

$$\tilde{P}_{n,j}(x) = \frac{1}{2\pi}\int e^{-itx}\frac{\phi_{P_{W,j}f_W}(t)}{\phi_U(t)}dt, \quad \tilde{Q}_{n,j}(x) = \frac{1}{2\pi}\int e^{-itx}\frac{\phi_{Q_{W,j}f_W}(t)}{\phi_U(t)}dt.$$

C.6. The order (in h) of $E|\tilde{P}_{n,j}(x)|^{2+\eta}$, $E|\tilde{Q}_{n,j}(x)|^{2+\eta}$, and $E|\tilde{U}_{n,j}(x)|^2$

In order to show (C.2), we need to study the orders (in h) of $E|\tilde{P}_{n,j}(x)|^{2+\eta}$, $E|\tilde{Q}_{n,j}(x)|^{2+\eta}$, and $E|\tilde{U}_{n,j}(x)|^2$. The orders of these quantities mainly depend on two facts. First, the orders of $E|\phi_{P_{W,j}f_W}(t)|^{2+\eta}$ and $E|\phi_{Q_{W,j}f_W}(t)|^{2+\eta}$; second, the smoothness of U. We first look into the first factor in the upcoming subsection, which leads to the intermediate results needed for showing asymptotic normality. In this section, we use $s \asymp t$ to indicate that s and t are of the same order in h as $n \to \infty$.

C.6.1. *Intermediate results*

First, by the definition of $P_{W,j}(w)$,

$$P_{W,j}(w)f_W(w) = f_W(w)\sum_{\ell=0}^{p} h^{-\ell}R_W^{0,\ell}(w)\{Y_j - m^*(w)\}(W_j - w)^\ell K_h(W_j - w)$$
$$= \sum_{\ell=0}^{p} h^{-\ell}\tilde{R}_W^{0,\ell}(w)\{Y_j - m^*(w)\}(W_j - w)^\ell K_h(W_j - w),$$

where $\tilde{R}_W^{0,\ell}(w) = f_W(w)R_W^{0,\ell}(w)$. By the definition of $R_W^{0,\ell}(w)$ given after (C.11), when ℓ is odd, $\tilde{R}_W^{0,\ell}(w)$ is equal to $-hE\{f_U(w-X)\}f_W^{-1}(w)$ times the $[1, \ell+1]$ entry of $\mathbf{S}^{-1}\widetilde{S}\mathbf{S}^{-1}$, and, when ℓ is even, $\tilde{R}_W^{0,\ell}(w)$ is equal to the $[1, \ell+1]$ entry of $\mathbf{S}^{-1}$.

Define $\kappa_W(w, W) = E\{|Y - m^*(w)|^{2+\eta}|W\}$. Assuming $\tilde{R}_W^{0,\ell}(w)$ and $\|\kappa_W(w, W)\|_\infty$

bounded, we have, for some positive finite constants C and C',

$$\begin{aligned}
& E|\phi_{P_{W,j}f_W}(t)|^{2+\eta} \\
= \; & h^{-2-\eta}E\left|\int e^{-itw}\{Y_j - m^*(w)\}\sum_{\ell=0}^{p}\tilde{R}_W^{0,\ell}(w)\left(\frac{W_j-w}{h}\right)^{\ell}K\left(\frac{W_j-w}{h}\right)dw\right|^{2+\eta} \\
\leq \; & Ch^{-2-\eta}E\left|\sum_{\ell=0}^{p}\int e^{-itw}\{Y - m^*(w)\}\left(\frac{W_j-w}{h}\right)^{\ell}K\left(\frac{W_j-w}{h}\right)dw\right|^{2+\eta} \\
\asymp \; & Ch^{-2-\eta}\sum_{\ell=0}^{p}E\left|\int e^{-itw}\{Y - m^*(w)\}\left(\frac{W_j-w}{h}\right)^{\ell}K\left(\frac{W_j-w}{h}\right)dw\right|^{2+\eta} \\
\leq \; & C'h^{-2-\eta}\sum_{\ell=0}^{p}E\left|\int e^{-itw}\left(\frac{W_j-w}{h}\right)^{\ell}K\left(\frac{W_j-w}{h}\right)dw\right|^{2+\eta} \\
= \; & C'h^{-2-\eta}\sum_{\ell=0}^{p}\int\int\left|e^{-itw}\left(\frac{v-w}{h}\right)^{\ell}K\left(\frac{v-w}{h}\right)dw\right|^{2+\eta}f_U(v-v_2)dvf_X(v_2)dv_2 \\
= \; & C'h^{-2-\eta}\sum_{\ell=0}^{p}\int\left|-he^{-itv}\int e^{ithv_1}v_1^{\ell}K(v_1)dv_1\right|^{2+\eta}f_U(v-v_2)dvf_X(v_2)dv_2 \\
= \; & C'h^{-2-\eta}\sum_{\ell=0}^{p}\int\left|-he^{-itv}i^{-\ell}\phi_K^{(\ell)}(th)\right|^{2+\eta}f_U(v-v_2)dvf_X(v_2)dv_2 \\
= \; & C'\sum_{\ell=0}^{p}\left|\phi_K^{(\ell)}(th)\right|^{2+\eta}.
\end{aligned}$$

Therefore, $E|\phi_{P_{W,j}f_W}(t)|^{2+\eta}$ is bounded by a non-random quantity of the same order in h as $\sum_{\ell=0}^{p}\left|\phi_K^{(\ell)}(th)\right|^{2+\eta}$.

Second, by the definition of $Q_{W,j}(w)$, we have

$$Q_{W,j}(w)f_W(w) = -\sum_{\ell=0}^{p}\sum_{\ell'=1}^{p}h^{\ell'-\ell}\tilde{R}_W^{0,\ell}(w)\frac{m^{*(\ell')}(w)}{\ell'!}(W_j-w)^{\ell+\ell'}K_h(W_j-w).$$

It follows that

$$\begin{aligned}
& E|\phi_{Q_{W,j}f_W}(t)|^{2+\eta} \\
\asymp\ & \sum_{\ell=0}^{p}\sum_{\ell'=1}^{p} E\left|\int e^{itw} h^{\ell'-\ell}\tilde{R}_W^{0,\ell}(w)\frac{m^{*(\ell')}(w)}{\ell'!}(W-w)^{\ell+\ell'}K_h(W-w)\right|^{2+\eta} \\
\leq\ & C\sum_{\ell=0}^{p}\sum_{\ell'=1}^{p} h^{(\ell'-\ell)(2+\eta)} E\left|\int e^{itw}(W-w)^{\ell+\ell'}K_h(W-w)dw\right|^{2+\eta} \\
=\ & C\sum_{\ell=0}^{p}\sum_{\ell'=1}^{p} h^{2\ell'(2+\eta)}\left|\int e^{ithv}(-1)^{1+\ell+\ell'}v^{\ell+\ell'}K(v)dv\right|^{2+\eta} \\
=\ & C\sum_{\ell=0}^{p}\sum_{\ell'=1}^{p} h^{2\ell'(2+\eta)}\left|i^{-(\ell+\ell')}\phi_K^{(\ell+\ell')}(th)\right|^{2+\eta} \\
=\ & C\sum_{\ell=0}^{p}\sum_{\ell'=1}^{p} h^{2\ell'(2+\eta)}\left|\phi_K^{(\ell+\ell')}(th)\right|^{2+\eta}. \qquad (C.15)
\end{aligned}$$

Therefore, $E|\phi_{Q_{W,j}f_W}(t)|^{2+\eta}$ is bounded by a non-random quantity of the same order in h as $\sum_{\ell=0}^{p}\sum_{\ell'=1}^{p} h^{2\ell'(2+\eta)}\left|\phi_K^{(\ell+\ell')}(th)\right|^{2+\eta}$.

C.6.2. *Normality with ordinary smooth U*

Now we are ready to tackle the orders of $E|\tilde{P}_{n,j}(x)|^{2+\eta}$ and $E|\tilde{Q}_{n,j}(x)|^{2+\eta}$. For ordinary smooth measurement error,

$$\begin{aligned}
& E|\tilde{P}_{n,j}(x)|^{2+\eta} \\
=\ & E\left|\frac{1}{2\pi}\int e^{-itx}\frac{\phi_{P_{W,j}f_W}(t)}{\phi_U(t)}dt\right|^{2+\eta} \\
\leq\ & E\left\{\left(\frac{1}{2\pi}\right)^{2+\eta}\int\frac{|\phi_{P_{W,j}f_W}(t)|^{2+\eta}}{|\phi_U(t)|^{2+\eta}}dt\right\} \\
=\ & \left(\frac{1}{2\pi}\right)^{2+\eta}\int\frac{E|\phi_{P_{W,j}f_W}(t)|^{2+\eta}}{|\phi_U(t)|^{2+\eta}}dt \\
=\ & \left(\frac{1}{2\pi}\right)^{2+\eta}\left\{\int_{|t|\leq M}\frac{E|\phi_{P_{W,j}f_W}(t)|^{2+\eta}}{|\phi_U(t)|^{2+\eta}}dt+\int_{|t|>M}\frac{E|\phi_{P_{W,j}f_W}(t)|^{2+\eta}}{|\phi_U(t)|^{2+\eta}}dt\right\} \\
\leq\ & \left(\frac{1}{2\pi}\right)^{2+\eta}\left[\left\{\inf_{|t|\leq M}|\phi_U(t)|^{2+\eta}\right\}^{-1}\int_{|t|\leq M}E|\phi_{P_{W,j}f_W}(t)|^{2+\eta}dt+\int_{|t|>M}\frac{E|\phi_{P_{W,j}f_W}(t)|^{2+\eta}}{\left(\frac{c}{2}|t|^{-b}\right)^{2+\eta}}dt\right]. \qquad (C.16)
\end{aligned}$$

Using the result in Section C.6.1, we have that (C.16) is bounded from above by

$$C\int_{|t|\leq M}\sum_{\ell=0}^{p}\left|\phi_K^{(\ell)}(th)\right|^{2+\eta}dt+C'\int_{|t|>M}|t|^{b(2+\eta)}\sum_{\ell=0}^{p}\left|\phi_K^{(\ell)}(th)\right|^{2+\eta}dt,$$

where the first term in the sum is

$$Ch^{-1}\sum_{\ell=0}^{p}\int_{|s|\le Mh}\left|\int e^{isu}u^{\ell}K(u)du\right|ds \le Ch^{-1}\sum_{\ell=0}^{p}\int_{|s|\le Mh}\int\left|u^{\ell}K(u)\right|dudt,$$

where the integral is of order $O(h)$, and thus the first term is bounded by a finite constant. And the second term is equal to

$$C'h^{-b(2+\eta)-1}\int_{|s|>Mh}|s|^{b(2+\eta)}\sum_{\ell=0}^{p}\left|\phi_K^{(\ell)}(s)\right|^{2+\eta}ds,$$

which is of order $h^{-b(2+\eta)-1}$ under the assumption that $\int\{|s|^b|\phi_K^{(\ell)}(s)|\}^{2+\eta}ds<\infty$, for $\ell=0,1,\dots,p$. Hence, $E|\tilde{P}_{n,j}(x)|^{2+\eta}$ is bounded by a quantity of order $h^{-b(2+\eta)-1}$.

As for $E|\tilde{Q}_{n,j}(x)|^{2+\eta}$, we have

$$\begin{aligned}
&E|\tilde{Q}_{n,j}(x)|^{2+\eta}\\
\le\ & C\left\{\int_{|t|\le M}\frac{E|\phi_{Q_{W,j}f_W}(t)|^{2+\eta}}{|\phi_U(t)|^{2+\eta}}dt+\int_{|t|>M}\frac{E|\phi_{Q_{W,j}f_W}(t)|^{2+\eta}}{|\phi_U(t)|^{2+\eta}}dt\right\}\\
\le\ & C_1\int_{|t|\le M}E|\phi_{Q_{W,j}f_W}(t)|^{2+\eta}dt+C_2\int_{|t|>M}|t|^{b(2+\eta)}E|\phi_{Q_Wf_W}(t)|^{2+\eta}dt\\
\asymp\ & h^{2(2+\eta)+1}+C_3\sum_{\ell=0}^{p}\sum_{\ell'=1}^{p}h^{2\ell'(2+\eta)}\int_{|s|\le Mh}h^{-b(2+\eta)-1}|s|^{b(2+\eta)}|\phi_K(s)^{(\ell+\ell')}|^{2+\eta}dw,
\end{aligned}$$

which is of the order $h^{(2-b)(2+\eta)-1}$ assuming that $\int|t|^{b(2+\eta)}|\phi_K^{(k)}(t)|^{2+\eta}dt<\infty$ for $k=1,\dots,2p$.

Finally, the order of $E(\tilde{U}_{n,j}^2)$ is the same as that of the variance of $\hat{m}(x)_{\text{HZ}}$, which is h^{-2b-1}. Combing the above three parts of the derivations, we conclude that $E|\tilde{U}_{n,j}|^{2+\eta}=O(h^{-b(2+\eta)-1})$ and $E(\tilde{U}_{n,j}^2)=Ch^{-2b-1}\{1+o(1)\}$. Therefore, if $\eta\ge 2$ and $(nh)^{-\eta/2}\to 0$ as $n\to\infty$, (C.2) holds. This completes the proof of the asymptotic normality of $\hat{m}(x)_{\text{HZ}}$ when the density of U is ordinary smooth.

C.6.3. *Normality with super smooth U*

When U is super smooth, we assume $\phi_K(t)$ supported on $[-1,1]$. The main change from the derivations in Section C.6.2 is how to partition the range of integrations.

For the order of $E|\tilde{P}_{n,j}(x)|^{2+\eta}$, we have

$$\begin{aligned}
&E|\tilde{P}_{n,j}(x)|^{2+\eta}\\
\le\ & \left(\frac{1}{2\pi}\right)^{2+\eta}\left\{\int_{|t|\le M}\frac{E|\phi_{P_{W,j}f_W}(t)|^{2+\eta}}{|\phi_U(t)|^{2+\eta}}dt+\int_{M<|t|\le 1/h}\frac{E|\phi_{P_{W,j}f_W}(t)|^{2+\eta}}{|\phi_U(t)|^{2+\eta}}dt\right\}\\
\le\ & \left(\frac{1}{2\pi}\right)^{2+\eta}\left[\left\{\inf_{|t|\le M}|\phi_U(t)|^{2+\eta}\right\}^{-1}\int_{|t|\le M}E|\phi_{P_{W,j}f_W}(t)|^{2+\eta}dt\right.\\
&\left.+\int_{M<|t|\le 1/h}\frac{E|\phi_{P_{W,j}f_W}(t)|^{2+\eta}}{|d_0|t|^{b_0}\exp(-|t|^b/d_2)/2|^{2+\eta}}dt\right],
\end{aligned}$$

which is bounded from above by, using the result in Section C.6.1,

$$\begin{aligned}
& C\int_{|t|\le M}\sum_{\ell=0}^{p}|\phi_K^{(\ell)}(th)|^{2+\eta}dt \\
& +C\int_{M<|t|\le 1/h}|t|^{-b_0(2+\eta)}\exp\{(2+\eta)|t|^b/d_2\}\sum_{\ell=0}^{p}|\phi_K^{(\ell)}(th)|^{2+\eta}dt \\
= \; & Ch^{-1}\int_{|s|\le Mh}\sum_{\ell=0}^{p}|\phi_K^{(\ell)}(s)|^{2+\eta}dt \\
& +Ch^{b_0(2+\eta)-1}\int_{Mh<|s|\le 1}|s|^{-b_0(2+\eta)}\exp\{(2+\eta)|s|^b/(h^bd_2)\}\sum_{\ell=0}^{p}|\phi_K^{(\ell)}(s)|^{2+\eta}dt,
\end{aligned}$$

of which the first term is $O(1)$ under the assumption that $\|\phi_K^{\ell}(t)\|_\infty < \infty$, for $\ell = 0, 1, \ldots, p$, and the second term is bounded from above by, under the same assumption,

$$\begin{aligned}
& Ch^{b_0(2+\eta)-1}\exp\{(2+\eta)h^{-b}/d_2\}\int_{Mh<|s|\le 1}|s|^{-b_0(2+\eta)}ds. \\
= & \begin{cases} Ch^{b_0(2+\eta)-1}\exp\{(2+\eta)h^{-b}/d_2\} & \text{if } b_0 < 1/(2+\eta), \\ Ch^{-1}\exp\{(2+\eta)h^{-b}/d_2\} & \text{if } b_0 = 1/(2+\eta), \\ C\exp\{(2+\eta)h^{-b}/d_2\} & \text{if } b_0 > 1/(2+\eta). \end{cases}
\end{aligned}$$

Similarly, one can show that

$$E|\tilde{Q}_{n,j}(x)|^{2+\eta} \le \begin{cases} Ch^{(2+b_0)(2+\eta)-1}\exp\{(2+\eta)h^{-b}/d_2\} & \text{if } b_0 < 1/(2+\eta), \\ Ch^{2(2+\eta)-1}\exp\{(2+\eta)h^{-b}/d_2\} & \text{if } b_0 = 1/(2+\eta), \\ Ch^{2(2+\eta)}\exp\{(2+\eta)h^{-b}/d_2\} & \text{if } b_0 > 1/(2+\eta). \end{cases}$$

Hence,

$$\begin{aligned}
E|\tilde{U}_{n,j}(x)|^{2+\eta} \; & \le \begin{cases} Ch^{b_0(2+\eta)-1}\exp\{(2+\eta)h^{-b}/d_2\} & \text{if } b_0 < 1/(2+\eta), \\ Ch^{-1}\exp\{(2+\eta)h^{-b}/d_2\} & \text{if } b_0 = 1/(2+\eta), \\ C\exp\{(2+\eta)h^{-b}/d_2\} & \text{if } b_0 > 1/(2+\eta), \end{cases} \\
& \le Ch^{(2+\eta)b_3-1}\exp\{(2+\eta)h^{-b}/d_2\},
\end{aligned}$$

where $b_3 = b_0 I(b_0 < 0.5)$.

Using the variance result for $\hat{m}_{\mathrm{HZ}}(x)$ for the super smooth U, we have $E|\tilde{U}_{n,j}^2(x)| \le Ch^{2b_3-2}\exp(2h^{-b}/d_2)$. Putting these together, we have

$$\frac{E|\tilde{U}_{n,j}(x)|^{2+\eta}}{n^{\eta/2}E|\tilde{U}_{n,j}^2(x)|^{(2+\eta)/2}} \asymp h^{1+\eta/2}/n^{\eta/2} \to 0,$$

as $n \to \infty$, for any $\eta > 0$. Hence, the normality of $\hat{m}_{\mathrm{HZ}}(x)$ for the case with super smooth U is proved.

Appendix D: Additional simulation studies without assuming measurement error distribution known

To address the practical scenario where the measurement error distribution is unknown, we consider two strategies in Section 7 in the main article. The first strategy, which we recommend, is to assume Laplace measurement error with characteristic function given by $\phi_U(t) = 1/\{1 + (\sigma_u^2/2)t^2\}$, in which σ_u^2 is estimated using repeated measures following equation (4.3) in Carroll, Ruppert, Stefanski, and Crainiceanu (2006). The second strategy, which is inferior to the first strategy according to Figure 9 in the main article, is to estimate $\phi_U(t)$ following the approach in Delaigle, Hall, and Meister (2008). Suppose there are two repeated measures, $W_{j,1}$ and $W_{j,2}$, for each true covariate value X_j, for $j = 1, \ldots, n$, then, assuming a symmetric measurement error distribution, this approach yields an estimated characteristic function of the measurement error associated with $W_{j,k}$ $(k = 1, 2)$ given by $\hat{\phi}_{U_1}(t) = \sqrt{\sum_{j=1}^n \cos\{it(W_{j,1} - W_{j,2})\}/n}$. Then we define $W_j = (W_{j,1}+W_{j,2})/2$ as the error-contaminated surrogate of X_j, for $j = 1, \ldots, n$, and the estimated characteristic function associated with the measurement error in $W_j = X_j+U_j$ is given by $\hat{\phi}_U(t) = \{\hat{\phi}_{U_1}(t/2)\}^2$.

Adopting the first strategy, Figures D.1–D.3 provide the results for our estimator and the DFC estimator under cases (C1), (C3), and (C4) considered in Section 6.3 in the main article, respectively. These are parallel to Figures 1, 3, and 4 in the main article, where one assumes a known measurement error distribution. Contrasting these two sets of figures, one can see that estimating σ_u^2 has very little impact on the estimates.

For illustration purpose, we demonstrate in Figure D.4 our estimate resulting from the first strategy and our estimate employing the second strategy to account for an unknown measurement error distribution under (C1). This comparison shows that using $\hat{\phi}_U(t)$ in the estimate usually leads to more biased estimates with higher variability than using an assumed Laplace characteristic function with σ_u^2 estimated.

References

Billingsley, P. (1979), *Probability and measure*, Wiley.

Carroll, R. J., Ruppert, D., Stefanski, L. A., and Crainiceanu, C. M. (2006), *Measurement Error in Non-linear Models: A Modern Perspective*, (2nd ed.), Chapman and Hall/CRC, Boca Raton.

Delaigle, A., Fan, J., and Carroll, R. (2009), 'A design-adaptive local polynomial estimator for the error-in-variables problem' *Journal of the American Statistical Association*, 104, 348–359.

Delaigle, A., Hall, P., and Meister, A. (2008), 'On deconvolution with repeated measurements', *Annals of Statistics*, 36, 665–685.

Fan, J. and Gijbels, I. (1996), *Local polynomial modelling and its applications*, Chapman and Hall/CRC, Boca Raton.

Fan, J. and Truong, Y. K. (1993), 'Nonparametric regression with errors in variables', *The Annals of Statistics*, 21, 1900–1925.

Ruppert, D. and Wand, M. P. (1994), 'Multivariate locally weighted least squares regression', *The Annals of Statistics*, 22, 1346–1370.

Stefanski, L. A. and Carroll, R. J. (1990), 'Deconvoluting kernel density estimators', *Statistics*, 21, 169–184.

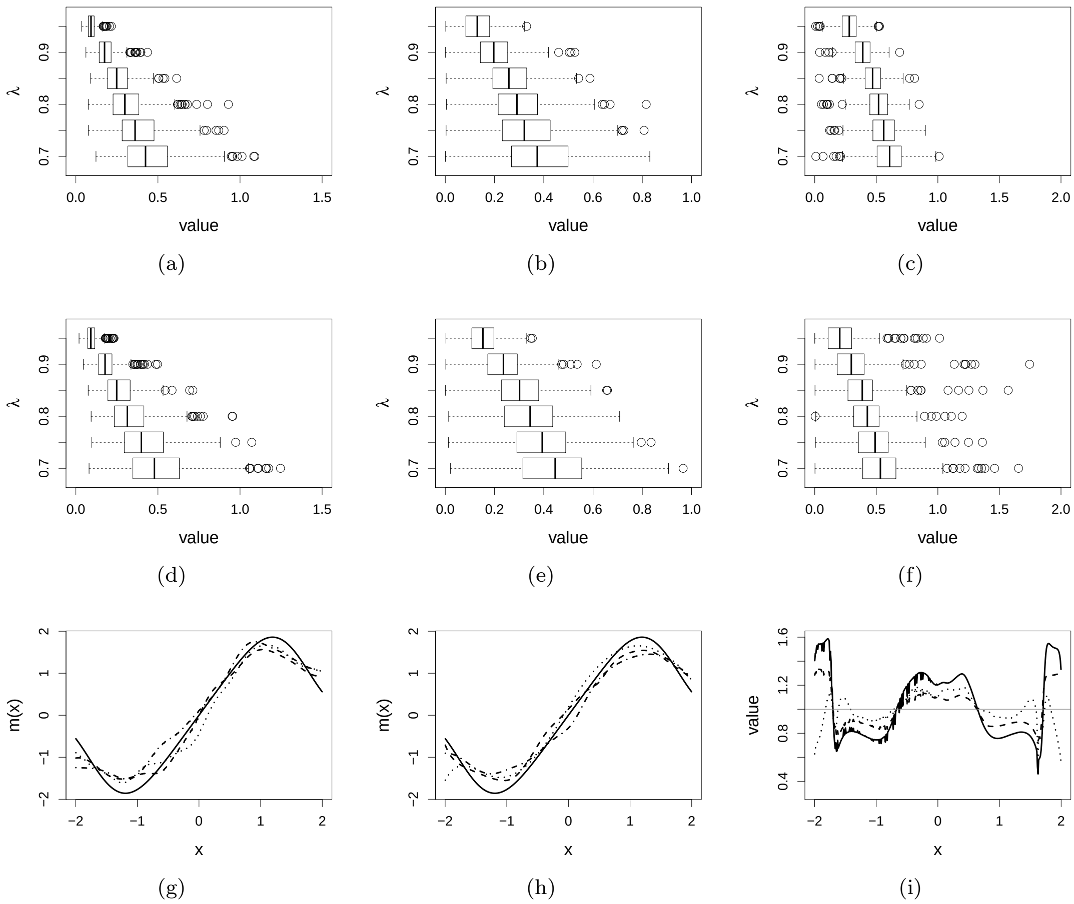

Figure D.1. Simulation results under (C1) using the theoretical optimal h, assuming Laplace U with σ_u^2 estimated using repeated measures. Panels (a) & (d): boxplots of ISEs versus λ for $\hat{m}_{\mathrm{HZ}}(x)$ and $\hat{m}_{\mathrm{DFC}}(x)$, respectively. Panels (b) & (e): boxplots of PAE(1) versus λ for $\hat{m}_{\mathrm{HZ}}(1)$ and $\hat{m}_{\mathrm{DFC}}(1)$, respectively. Panels (c) & (f): boxplots of PAE(2) versus λ for $\hat{m}_{\mathrm{HZ}}(2)$ and $\hat{m}_{\mathrm{DFC}}(2)$, respectively. Panels (g) & (h): quantile curves when $\lambda = 0.85$ for $\hat{m}_{\mathrm{HZ}}(x)$ and $\hat{m}_{\mathrm{DFC}}(x)$, respectively, based on ISEs (dashed lines for the first quartile, dotted lines for the second quartile, and dot-dashed lines for the third quartile, solid lines for the truth). Panel (i): PmAER (dashed line), PsdAER (dotted line), and PMSER (solid line) versus x when $\lambda = 0.85$; the horizontal reference line highlights the value 1.

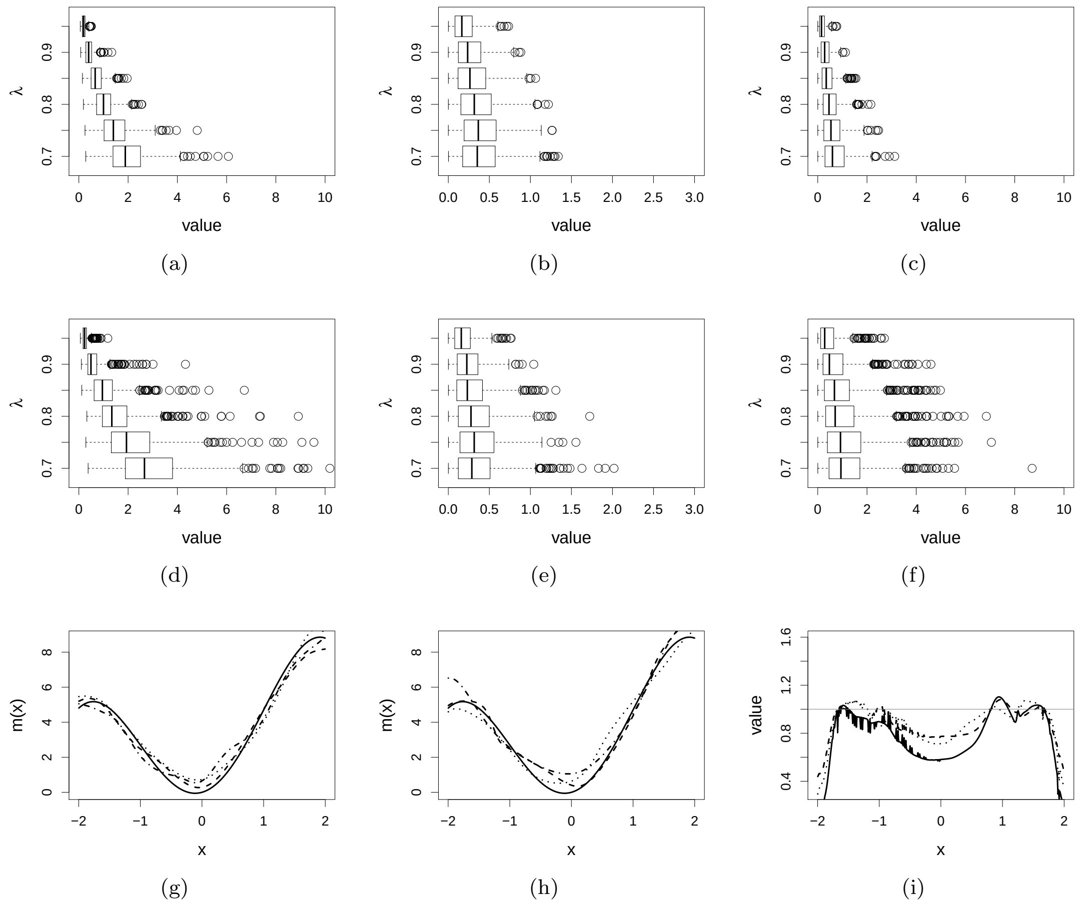

Figure D.2. Simulation results under (C3) using the theoretical optimal h, assuming Laplace U with σ_u^2 estimated using repeated measures. Panels (a) & (d): boxplots of ISEs versus λ for $\hat{m}_{\mathrm{HZ}}(x)$ and $\hat{m}_{\mathrm{DFC}}(x)$, respectively. Panels (b) & (e): boxplots of PAE(1) versus λ for $\hat{m}_{\mathrm{HZ}}(1)$ and $\hat{m}_{\mathrm{DFC}}(1)$, respectively. Panels (c) & (f): boxplots of PAE(2) versus λ for $\hat{m}_{\mathrm{HZ}}(2)$ and $\hat{m}_{\mathrm{DFC}}(2)$, respectively. Panels (g) & (h): quantile curves when $\lambda = 0.8$ for $\hat{m}_{\mathrm{HZ}}(x)$ and $\hat{m}_{\mathrm{DFC}}(x)$, respectively, based on ISEs (dashed lines for the first quartile, dotted lines for the second quartile, dot-dashed lines for the third quartile, and solid lines for the truth). Panel (i): PmAER (dashed line), PsdAER (dotted line), and PMSER (solid line) versus x when $\lambda = 0.8$; the horizontal reference line highlights the value 1.

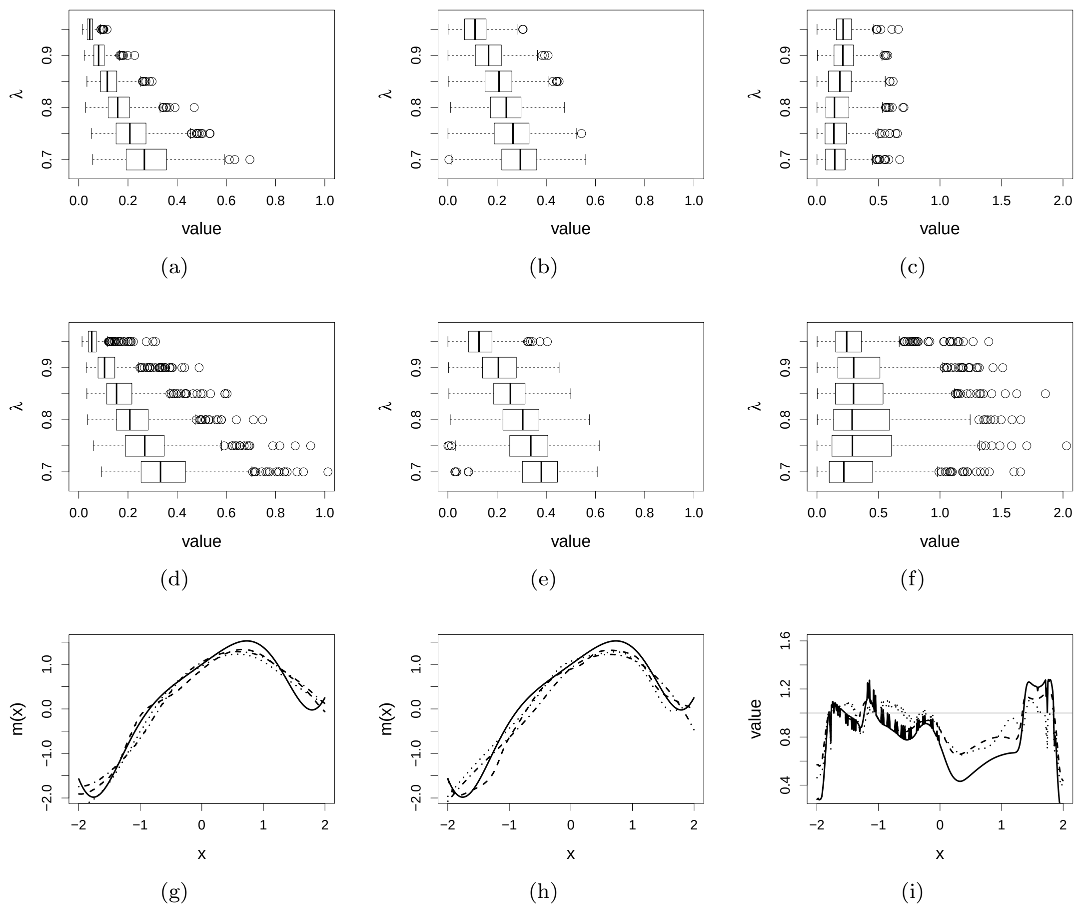

Figure D.3. Simulation results under (C4) using the theoretical optimal h, assuming Laplace U with σ_u^2 estimated using repeated measures. Panels (a) & (d): boxplots of ISEs versus λ for $\hat{m}_{\mathrm{HZ}}(x)$ and $\hat{m}_{\mathrm{DFC}}(x)$, respectively. Panels (b) & (e): boxplots of PAE(1) versus λ for $\hat{m}_{\mathrm{HZ}}(1)$ and $\hat{m}_{\mathrm{DFC}}(1)$, respectively. Panels (c) & (f): boxplots of PAE(2) versus λ for $\hat{m}_{\mathrm{HZ}}(2)$ and $\hat{m}_{\mathrm{DFC}}(2)$, respectively. Panels (g) & (h): quantile curves when $\lambda = 0.8$ for $\hat{m}_{\mathrm{HZ}}(x)$ and $\hat{m}_{\mathrm{DFC}}(x)$, respectively, based on ISEs (dashed lines for the first quartile, dotted lines for the second quartile, dot-dashed lines for the third quartile, and solid lines for the truth). Panel (i): PmAER(dashed line), PsdAER (dotted line), and PMSER (solid line) versus x when $\lambda = 0.8$; the horizontal reference line highlights the value 1.

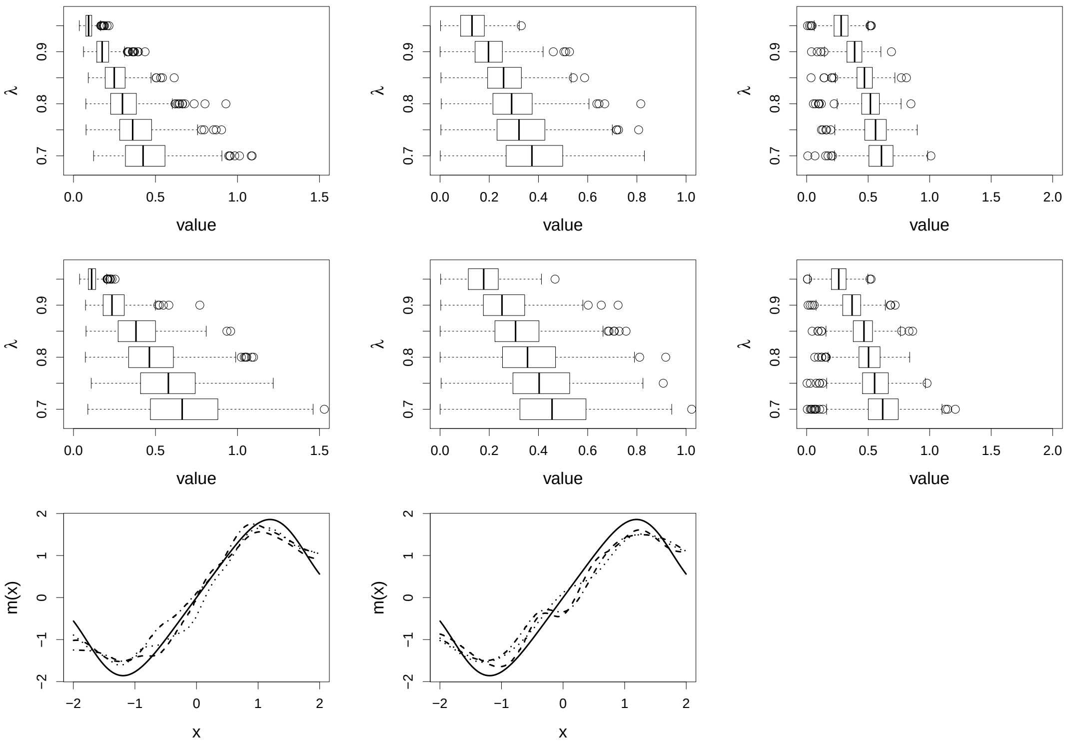

Figure D.4. Simulation results under (C1) using the theoretical optimal h without assuming measurement error distribution known. The first row is identical to the first row in Figure D.1, presenting (from left to right) the boxplots of ISE, PAE(1), and PAE(2), respectively, associated with $\hat{m}_{\mathrm{HZ}}(x)$ when one assumes Laplace U with σ_u^2 estimated. The second row presents the counterpart boxplots when one estimates $\phi_U(t)$ as described in the second strategy. The third row contains the quantile curves when $\lambda = 0.85$ for $\hat{m}_{\mathrm{HZ}}(x)$ resulting from the first strategy (same as panel (g) in Figure D.1) on the left, and the counterpart quantile curves resulting from the second strategy of accounting for unknown $\phi_U(t)$ on the right.