Supplementary Material for "Efficient Diagnostics for Parametric Regression Models with Distortion Measurement Errors Incorporating Dimension-reduction"

Zhihua Sun[1], Feifei Chen[2] and Hua Liang[3]

[1] *University of Chinese Academy of Sciences,* [2] *Beijing Normal University, and* [3] *George Washington University*

In this document, we present the preliminary lemmas and the detailed proofs of Theorems 1 and 2. Additional simulation studies and real data analyses of Boston house price data are also included.

S1 Preliminary lemmas

Lemma 1. *Suppose that Conditions (C1)–(C2) and (C7)–(C8) hold. For the estimators $\hat{\psi}_n(u)$ and $\hat{\gamma}_n(u)$, we have the following asymptotic representations:*

$$\sqrt{nh_n}\left\{\hat{\psi}_n(u) - \psi(u)\mathrm{E}(Y) - \frac{1}{2}\psi^{(2)}(u)\mathrm{E}(Y)\kappa_{21}h_n^2\right\}$$
$$= \frac{1}{\sqrt{nh_n}f_u(u)}\sum_{i=1}^{n} K\left(\frac{U_i - u}{h_n}\right)\left\{\widetilde{Y}_i - \psi(U_i)\mathrm{E}(Y)\right\} + o_p(1) \quad \text{(S1.1)}$$

and

$$\sqrt{nh_n}\left\{\hat{\gamma}_n(u)-\gamma(u)\mathrm{E}(\mathbf{X})-\frac{1}{2}\gamma^{(2)}(u)\mathrm{E}(\mathbf{X})\kappa_{21}h_n^2\right\}$$
$$=\frac{1}{\sqrt{nh_n}f_u(u)}\sum_{i=1}^{n}K\left(\frac{U_i-u}{h_n}\right)\{\widetilde{\mathbf{X}}_i-\gamma(U_i)\mathrm{E}(\mathbf{X})\}+o_p(1), \quad \text{(S1.2)}$$

where $\kappa_{21}=\int u^2K(u)du$.

Proof: We only prove (S1.1). The proof of (S1.2) is similar. Let $\lambda(u)=\mathrm{E}(\widetilde{Y}|U=u)$. By the condition that U is independent of $(Y,\mathbf{X}^\top,\mathbf{Z}^\top)^\top$, we have $\lambda(u)=\psi(u)\mathrm{E}(Y)$ and $\lambda^{(2)}(u)=\psi^{(2)}(u)\mathrm{E}(Y)$. Note that $\hat{\psi}_n(u)$ is the local linear estimation of $\lambda(u)$. By Theorem 2.7 in Li and Racine (2007), (S1.1) holds. □

Lemma 2. *Suppose that Conditions (C1)–(C2) and (C7)–(C8) hold. Let* $\Lambda(\mathbf{X},\mathbf{Z})$ *be a continuous function satisfying* $\mathrm{E}|\Lambda(\mathbf{X},\mathbf{Z})|<\infty$. *Then, we have the following asymptotic representations:*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Lambda(\mathbf{X}_i,\mathbf{Z}_i)(\hat{Y}_i-Y_i)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathrm{E}\{Y\Lambda(\mathbf{X},\mathbf{Z})\}}{\mathrm{E}(Y)}\{\widetilde{Y}_i-Y_i\}+o_p(1), \quad \text{(S1.3)}$$

and for $r=1,\ldots,p$,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Lambda(\mathbf{X}_i,\mathbf{Z}_i)(\hat{X}_{ri}-X_{ri})=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathrm{E}\{X_r\Lambda(\mathbf{X},\mathbf{Z})\}}{\mathrm{E}(X_r)}\{\widetilde{X}_{ri}-X_{ri}\}+o_p(1). \quad \text{(S1.4)}$$

Proof: We only prove (S1.3). Note that $\widetilde{Y}_{m,n}=n^{-1}\sum_{i=1}^{n}\widetilde{Y}_i$, thus

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Lambda(\mathbf{X}_i,\mathbf{Z}_i)(\hat{Y}_i-Y_i) \quad = \quad \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Lambda(\mathbf{X}_i,\mathbf{Z}_i)\left\{\frac{\widetilde{Y}_i\widetilde{Y}_{m,n}}{\hat{\psi}_n(U_i)}-Y_i\right\}$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Lambda(\mathbf{X}_i,\mathbf{Z}_i)\left\{\frac{\widetilde{Y}_i\{\mathrm{E}(Y)+\widetilde{Y}_{m,n}-\mathrm{E}(Y)\}}{\hat{\psi}_n(U_i)}-Y_i\right\}\\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Lambda(\mathbf{X}_i,\mathbf{Z}_i)\left\{\frac{\widetilde{Y}_i\mathrm{E}(Y)}{\hat{\psi}_n(U_i)}-Y_i\right\}\\
&\quad +\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Lambda(\mathbf{X}_i,\mathbf{Z}_i)\frac{\widetilde{Y}_i}{\hat{\psi}_n(U_i)}\left\{\widetilde{Y}_{m,n}-\mathrm{E}(Y)\right\}\\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Lambda(\mathbf{X}_i,\mathbf{Z}_i)\left\{\frac{\widetilde{Y}_i\mathrm{E}(Y)}{\hat{\psi}_n(U_i)}-Y_i\right\}\\
&\quad +\frac{1}{n\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{n}\Lambda(\mathbf{X}_i,\mathbf{Z}_i)\frac{\widetilde{Y}_i}{\hat{\psi}_n(U_i)}\left\{\widetilde{Y}_j-\mathrm{E}(Y)\right\}\\
&=: I_1+I_2. \qquad\qquad \text{(S1.5)}
\end{aligned}
$$

Firstly, we investigate the property of I_1, it follows from Lemma 1 that

$$
I_1 = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Lambda(\mathbf{X}_i,\mathbf{Z}_i)Y_i\left\{\frac{\psi(U_i)\mathrm{E}(Y)-\hat{\psi}_n(U_i)}{\psi(U_i)\mathrm{E}(Y)}\right\}+o_p(1).
$$

By the law of large numbers and Conditions (C7)–(C8), we can validate that $S_0(u,h_n) = f_u(u)+o_p(1), S_1(u,h_n) = h_n^2\int u^2K(u)f_u'(u)du+o_p(h_n^2)$, $S_2(u,h_n) = h_n^2\int u^2K(u)du+o_p(h_n^2)$. Thus, we have

$$
\begin{aligned}
I_1 &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\Lambda(\mathbf{X}_i,\mathbf{Z}_i)Y_i}{\psi(U_i)\mathrm{E}(Y)}\frac{1}{nh_nf_u(U_i)}\sum_{j=1}^{n}K\left(\frac{U_j-U_i}{h_n}\right)\{\psi(U_i)\mathrm{E}(Y)-\widetilde{Y}_j\}+o_p(1)\\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\Lambda(\mathbf{X}_i,\mathbf{Z}_i)Y_i}{\psi(U_i)\mathrm{E}(Y)}\frac{\mathrm{E}(Y)}{nh_nf_u(U_i)}\sum_{j=1}^{n}K\left(\frac{U_j-U_i}{h_n}\right)\{\psi(U_i)-\psi(U_j)\}\\
&\quad +\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\Lambda(\mathbf{X}_i,\mathbf{Z}_i)Y_i}{\psi(U_i)\mathrm{E}(Y)}\frac{1}{nh_nf_u(U_i)}\sum_{j=1}^{n}K\left(\frac{U_j-U_i}{h_n}\right)\{\psi(U_j)\mathrm{E}(Y)-\widetilde{Y}_j\}+o_p(1)\\
&=: \Omega_1+\Omega_2. \qquad\qquad \text{(S1.6)}
\end{aligned}
$$

By the fact that

$$\frac{1}{nh_n f_u(U_i)}\sum_{j=1}^{n} K\left(\frac{U_j-U_i}{h_n}\right)\{\psi(U_i)-\psi(U_j)\} = \frac{1}{2}\psi^{(2)}(U_i)\kappa_{21}h_n^2 + o_p(h_n^2)$$

with κ_{21} shown in Lemma 1, we can validate that

$$\Omega_1 = O_p(\sqrt{n}h_n^2) = o_p(1). \tag{S1.7}$$

Notice further that $\mathrm{E}\{Y\Lambda(\mathbf{X},\mathbf{Z})|U=U_j\} = \mathrm{E}\{Y\Lambda(\mathbf{X},\mathbf{Z})\}$ by the independence condition. Then, we can get

$$\begin{aligned}
\Omega_2 &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\{\psi(U_j)\mathrm{E}(Y)-\widetilde{Y}_j\}\frac{1}{nh_n}\sum_{i=1}^{n}\frac{\Lambda(\mathbf{X}_i,\mathbf{Z}_i)Y_i}{f_u(U_i)\psi(U_i)\mathrm{E}(Y)}K\left(\frac{U_j-U_i}{h_n}\right)\\
&= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\{\psi(U_j)\mathrm{E}(Y)-\psi(U_j)Y_j\}\frac{\mathrm{E}\{Y\Lambda(\mathbf{X},\mathbf{Z})|U=U_j\}}{\psi(U_j)\mathrm{E}(Y)} + o_p(1)\\
&= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\{\mathrm{E}(Y)-Y_j\}\frac{\mathrm{E}\{Y\Lambda(\mathbf{X},\mathbf{Z})\}}{\mathrm{E}(Y)} + o_p(1).
\end{aligned} \tag{S1.8}$$

It follows from (S1.6), (S1.7) and (S1.8) that

$$I_1 = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\{\mathrm{E}(Y)-Y_j\}\frac{\mathrm{E}\{Y\Lambda(\mathbf{X},\mathbf{Z})\}}{\mathrm{E}(Y)} + o_p(1). \tag{S1.9}$$

Now we investigate the property of I_2. Note that

$$I_2 = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left\{\widetilde{Y}_j-\mathrm{E}(Y)\right\}\frac{1}{n}\sum_{i=1}^{n}\Lambda(\mathbf{X}_i,\mathbf{Z}_i)\frac{\widetilde{Y}_i}{\hat{\psi}_n(U_i)}.$$

Similar to the proof of (S1.9), we have

$$I_2 = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\{\widetilde{Y}_j-\mathrm{E}(Y)\}\frac{\mathrm{E}\{Y\Lambda(\mathbf{X},\mathbf{Z})\}}{\mathrm{E}(Y)} + o_p(1). \tag{S1.10}$$

Combining (S1.5), (S1.9) and (S1.10), we have proved (S1.3). □

Remark 1. Lemma 2 is slightly different from Lemma B.2 in Zhang et al. (2012) in that we relaxed the conditions that $Var\{\psi(U)\} = 1$ and $Var\{\gamma_r(u)\} = 1$ for $r = 1, \ldots, p$, which are needed in Zhang et al. (2012).

Lemma 3. *Under Conditions (C1)–(C8), when* $\mathcal{H}_{1n} : Y = g(\mathbf{X}, \mathbf{Z}, \beta) + n^{-1/2}S(\mathbf{X}, \mathbf{Z}) + \varepsilon$ *hold, we have*

$$\begin{aligned} \sqrt{n}(\hat{\beta}_n - \beta) &= \frac{\Sigma^{-1}}{\sqrt{n}} \sum_{i=1}^{n} \frac{\mathrm{E}\{Y \dot{g}_\beta(\mathbf{X}, \mathbf{Z}, \beta)\}}{\mathrm{E}(Y)} \{\widetilde{Y}_i - Y_i\} + \frac{\Sigma^{-1}}{\sqrt{n}} \sum_{i=1}^{n} \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta)\varepsilon_i \\ &\quad + \Sigma^{-1}\mathrm{E}\{\dot{g}_\beta(\mathbf{X}, \mathbf{Z}, \beta) S(\mathbf{X}, \mathbf{Z})\} + \frac{\Sigma^{-1}}{\sqrt{n}} \sum_{i=1}^{n} \Sigma_x(\widetilde{\mathbf{X}}_i - \mathbf{X}_i) + o_p(1) \end{aligned}$$

with Σ_x *and* Σ *presented in Appendix A.*

Proof: From the objective function (2.1), the estimator $\hat{\beta}_n$ satisfies the following estimating equation:

$$\sum_{i=1}^{n} \dot{g}_\beta(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \hat{\beta}_n)\{\hat{Y}_i - g(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \hat{\beta}_n)\} = 0.$$

Let $B_n = n^{-1/2} \sum_{i=1}^{n} \dot{g}_\beta(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \hat{\beta}_n)\{\hat{Y}_i - g(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \hat{\beta}_n)\}$. Then, it can be decomposed into two parts:

$$\begin{aligned} B_n &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta)\{\hat{Y}_i - g(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \hat{\beta}_n)\} \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{\dot{g}_\beta(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \hat{\beta}_n) - \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta)\}\{\hat{Y}_i - g(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \hat{\beta}_n)\} \\ &=: B_{n1} + B_{n2}. \end{aligned} \tag{S1.11}$$

The first term B_{n1} can further be decomposed into four terms as follows:

$$B_{n1} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta)(\hat{Y}_i - Y_i) + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta)\{Y_i - g(\mathbf{X}_i, \mathbf{Z}_i, \beta)\}$$

$$
\begin{aligned}
& +\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta)\{g(\mathbf{X}_i, \mathbf{Z}_i, \beta) - g(\mathbf{X}_i, \mathbf{Z}_i, \hat{\beta}_n)\} \\
& +\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta)\{g(\mathbf{X}_i, \mathbf{Z}_i, \hat{\beta}_n) - g(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \hat{\beta}_n)\} \\
=: \;& B_{n1,1} + B_{n1,2} + B_{n1,3} + B_{n1,4}. \qquad \text{(S1.12)}
\end{aligned}
$$

For $B_{n1,1}$, it follows from (S1.3) that

$$
B_{n1,1} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\mathrm{E}\{Y \dot{g}_\beta(\mathbf{X}, \mathbf{Z}, \beta)\}}{\mathrm{E}(Y)} \{\widetilde{Y}_i - Y_i\} + o_p(1). \qquad \text{(S1.13)}
$$

For $B_{n1,2}$, under the alternative hypothetical models $\mathcal{H}_{1n} : Y = g(\mathbf{X}, \mathbf{Z}, \beta) + n^{-1/2} S(\mathbf{X}, \mathbf{Z}) + \varepsilon$, by the law of large numbers, we have

$$
\begin{aligned}
B_{n1,2} &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta)\varepsilon_i + \frac{1}{n} \sum_{i=1}^{n} \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta) S(\mathbf{X}_i, \mathbf{Z}_i) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta)\varepsilon_i + \mathrm{E}\{\dot{g}_\beta(\mathbf{X}, \mathbf{Z}, \beta) S(\mathbf{X}, \mathbf{Z})\} + o_p(1). \qquad \text{(S1.14)}
\end{aligned}
$$

By the law of large numbers, we can validate that

$$
B_{n1,3} = -\mathrm{E}\{\dot{g}_\beta(\mathbf{X}, \mathbf{Z}, \beta)\dot{g}_\beta(\mathbf{X}, \mathbf{Z}, \beta)^\top\} \sqrt{n}(\hat{\beta}_n - \beta) + o_p(1). \qquad \text{(S1.15)}
$$

For the fourth term $B_{n1,4}$, it can be decomposed into

$$
\begin{aligned}
B_{n1,4} &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta)\dot{g}_x(\mathbf{X}_i, \mathbf{Z}_i, \hat{\beta}_n)^\top(\mathbf{X}_i - \hat{\mathbf{X}}_i) + o_p(1) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta)\dot{g}_x(\mathbf{X}_i, \mathbf{Z}_i, \beta)^\top(\mathbf{X}_i - \hat{\mathbf{X}}_i) \\
&\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta)\{\dot{g}_x(\mathbf{X}_i, \mathbf{Z}_i, \hat{\beta}_n) - \dot{g}_x(\mathbf{X}_i, \mathbf{Z}_i, \beta)\}^\top(\mathbf{X}_i - \hat{\mathbf{X}}_i) \\
&=: B_{n1,4}^{[1]} + B_{n1,4}^{[2]}.
\end{aligned}
$$

By (S1.4) and the definition of Σ_x in Appendix A, we can obtain that

$$B_{n1,4}^{[1]} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Sigma_x(\widetilde{\mathbf{X}}_i - \mathbf{X}_i) + o_p(1).$$

For $B_{n1,4}^{[2]}$, we have

$$B_{n1,4}^{[2]} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta)(\mathbf{X}_i - \hat{\mathbf{X}}_i)^\top \ddot{g}_{x,\beta}(\mathbf{X}_i, \mathbf{Z}_i, \beta)(\hat{\beta}_n - \beta) + o_p(1),$$

where $\ddot{g}_{x,\beta}(\mathbf{X}_i, \mathbf{Z}_i, \beta)$ is the second partial derivative of $g(\mathbf{X}_i, \mathbf{Z}_i, \beta)$ related to x and β. We can validate that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta)(\mathbf{X}_i - \hat{\mathbf{X}}_i)^\top \ddot{g}_{x,\beta}(\mathbf{X}_i, \mathbf{Z}_i, \beta) = O_p(1)$$

by (S1.4). Then, we can obtain that $B_{n1,4}^{[2]} = O_p(n^{-1/2}) = o_p(1)$. Therefore, it yields

$$B_{n1,4} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Sigma_x(\widetilde{\mathbf{X}}_i - \mathbf{X}_i) + o_p(1). \tag{S1.16}$$

Thus, from (S1.12)–(S1.16), we obtain the following result for B_{n1},

$$\begin{aligned} B_{n1} = & \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\mathrm{E}\{Y \dot{g}_\beta(\mathbf{X}, \mathbf{Z}, \beta)\}}{\mathrm{E}(Y)} \{\widetilde{Y}_i - Y_i\} + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta)\varepsilon_i \\ & + \mathrm{E}\{\dot{g}_\beta(\mathbf{X}, \mathbf{Z}, \beta) S(\mathbf{X}, \mathbf{Z})\} + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Sigma_x(\widetilde{\mathbf{X}}_i - \mathbf{X}_i) \\ & - \mathrm{E}\{\dot{g}_\beta(\mathbf{X}, \mathbf{Z}, \beta)\dot{g}_\beta(\mathbf{X}, \mathbf{Z}, \beta)^\top\}\sqrt{n}(\hat{\beta}_n - \beta) + o_p(1). \end{aligned} \tag{S1.17}$$

In the following, we evaluate B_{n2}. It can be decomposed into two parts:

$$B_{n2} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{\dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \hat{\beta}_n) - \dot{g}_\beta(\mathbf{X}_i, \mathbf{Z}_i, \beta)\}\{\hat{Y}_i - g(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \hat{\beta}_n)\}$$

$$
\begin{aligned}
&+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{\dot{g}_{\beta}(\hat{\mathbf{X}}_i,\mathbf{Z}_i,\hat{\beta}_n)-\dot{g}_{\beta}(\mathbf{X}_i,\mathbf{Z}_i,\hat{\beta}_n)\}\{\hat{Y}_i-g(\hat{\mathbf{X}}_i,\mathbf{Z}_i,\hat{\beta}_n)\}\\
=: \;& B_{n2,1}+B_{n2,2}. \qquad \text{(S1.18)}
\end{aligned}
$$

Note that

$$
B_{n2,1}=\frac{1}{n}\sum_{i=1}^{n}\ddot{g}_{\beta,\beta}(\mathbf{X}_i,\mathbf{Z}_i,\beta)\{\hat{Y}_i-g(\hat{\mathbf{X}}_i,\mathbf{Z}_i,\hat{\beta}_n)\}\sqrt{n}(\hat{\beta}_n-\beta)+o_p(1).
$$

By the similar method to obtain (S1.17), we can prove that

$$
\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\ddot{g}_{\beta,\beta}(\mathbf{X}_i,\mathbf{Z}_i,\beta)\{\hat{Y}_i-g(\hat{\mathbf{X}}_i,\mathbf{Z}_i,\hat{\beta}_n)\}=O_p(1).
$$

Therefore, we obtain that

$$
B_{n2,1}=O_p\left(n^{-1/2}\right)=o_p(1). \qquad \text{(S1.19)}
$$

Now we investigate $B_{n2,2}$.

$$
\begin{aligned}
B_{n2,2} &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\ddot{g}_{\beta,x}(\mathbf{X}_i,\mathbf{Z}_i,\hat{\beta}_n)(\hat{\mathbf{X}}_i-\mathbf{X}_i)\{\hat{Y}_i-g(\hat{\mathbf{X}}_i,\mathbf{Z}_i,\hat{\beta}_n)\}+o_p(1)\\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\ddot{g}_{\beta,x}(\mathbf{X}_i,\mathbf{Z}_i,\beta)(\hat{\mathbf{X}}_i-\mathbf{X}_i)\{\hat{Y}_i-g(\hat{\mathbf{X}}_i,\mathbf{Z}_i,\hat{\beta}_n)\}\\
&\quad +\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{\ddot{g}_{\beta,x}(\mathbf{X}_i,\mathbf{Z}_i,\hat{\beta}_n)-\ddot{g}_{\beta,x}(\mathbf{X}_i,\mathbf{Z}_i,\beta)\}(\hat{\mathbf{X}}_i-\mathbf{X}_i)\{\hat{Y}_i-g(\hat{\mathbf{X}}_i,\mathbf{Z}_i,\hat{\beta}_n)\}\\
&\quad +o_p(1)\\
&=: B_{n2,2}^{[1]}+B_{n2,2}^{[2]}+o_p(1). \qquad \text{(S1.20)}
\end{aligned}
$$

For the first term $B_{n2,2}^{[1]}$, we have

$$
B_{n2,2}^{[1]} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\ddot{g}_{\beta,x}(\mathbf{X}_i,\mathbf{Z}_i,\beta)(\hat{\mathbf{X}}_i-\mathbf{X}_i)(\hat{Y}_i-Y_i)
$$

$$
\begin{aligned}
&+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\ddot{g}_{\beta,x}(\mathbf{X}_i,\mathbf{Z}_i,\beta)(\hat{\mathbf{X}}_i-\mathbf{X}_i)\{Y_i-g(\mathbf{X}_i,\mathbf{Z}_i,\beta)\}\\
&+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\ddot{g}_{\beta,x}(\mathbf{X}_i,\mathbf{Z}_i,\beta)(\hat{\mathbf{X}}_i-\mathbf{X}_i)\{g(\mathbf{X}_i,\mathbf{Z}_i,\beta)-g(\mathbf{X}_i,\mathbf{Z}_i,\hat{\beta}_n)\\
&+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\ddot{g}_{\beta,x}(\mathbf{X}_i,\mathbf{Z}_i,\beta)(\hat{\mathbf{X}}_i-\mathbf{X}_i)\{g(\mathbf{X}_i,\mathbf{Z}_i,\hat{\beta}_n)-g(\hat{\mathbf{X}}_i,\mathbf{Z}_i,\hat{\beta}_n)\}\\
=:\;& B_{n2,2}^{[1,1]}+B_{n2,2}^{[1,2]}+B_{n2,2}^{[1,3]}+B_{n2,2}^{[1,4]}. \qquad\qquad \text{(S1.21)}
\end{aligned}
$$

For $B_{n2,2}^{[1,1]}$, by Slutsky theorem and the definition of Δ_{ni} for $i=1,\ldots,n$, we have

$$
\begin{aligned}
B_{n2,2}^{[1,1]} &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\ddot{g}_{\beta,x}(\mathbf{X}_i,\mathbf{Z}_i,\beta)\Delta_{ni}\left\{\frac{\widetilde{Y}_i\mathrm{E}(Y)}{\hat{\psi}_n(U_i)}-Y_i\right\}+o_p(1)\\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\ddot{g}_{\beta,x}(\mathbf{X}_i,\mathbf{Z}_i,\beta)Y_i\Delta_{ni}\frac{\{\psi(U_i)\mathrm{E}(Y)-\hat{\psi}_n(U_i)\}}{\mathrm{E}(\widetilde{Y}|U=U_i)}+o_p(1).
\end{aligned}
$$

By Theorem 6 in Masry (1996), we have $\sup_u|\hat{\gamma}_{nr}(u)-\gamma_r(u)\mathrm{E}(X_r)| = O_P(\{\ln n/(nh_n)\}^{1/2}+h_n^2)$ and $\sup_u|\hat{\psi}_n(u)-\psi(u)\mathrm{E}(Y)| = O_P(\{\ln n/(nh_n)\}^{1/2}+h_n^2)$. By Condition (C8), we can further prove that

$$
B_{n2,2}^{[1,1]} = o_p(1). \qquad\qquad \text{(S1.22)}
$$

For $B_{n2,2}^{[1,2]}$, by the result of (S1.2), we have

$$
\begin{aligned}
B_{n2,2}^{[1,2]} &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\ddot{g}_{\beta,x}(\mathbf{X}_i,\mathbf{Z}_i,\beta)\varepsilon_i\Delta_{ni}+\frac{1}{n}\sum_{i=1}^{n}\ddot{g}_{\beta,x}(\mathbf{X}_i,\mathbf{Z}_i,\beta)S(\mathbf{X}_i,\mathbf{Z}_i)(\hat{\mathbf{X}}_i-\mathbf{X}_i)\\
&\quad +o_p(1)\\
&=: B_{n2,2}^{[1,2,1]}+B_{n2,2}^{[1,2,2]}+o_p(1).
\end{aligned}
$$

Recalling the definition of Δ_{ni} and $\tilde{\Delta}_{ij}$ for $i,j=1,\ldots,n$, we can obtain that

$$B_{n2,2}^{[1,2,1]} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\ddot{g}_{\beta,x}(\mathbf{X}_i, \mathbf{Z}_i, \beta)\varepsilon_i}{nh_n f_u(U_i)} \tilde{\Delta}_{ij} K\left(\frac{U_j - U_i}{h_n}\right) + o_p(1). \quad \text{(S1.23)}$$

The term $B_{n2,2}^{[1,2,1]}$ is a $1 \times P$ vector. We can prove that the second moment of each component converges to zero. Thus, we can obtain that

$$B_{n2,2}^{[1,2,1]} \quad = \quad o_p(1).$$

From Lemma 2, we can validate that $B_{n2,2}^{[1,2,2]} = o_p(n^{-1/2}) = o_p(1)$. Therefore, we have

$$B_{n2,2}^{[1,2]} \quad = \quad o_p(1). \quad \text{(S1.24)}$$

According to Lemma 2, it can be validated that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \ddot{g}_{\beta,x}(\mathbf{X}_i, \mathbf{Z}_i, \beta)(\hat{\mathbf{X}}_i - \mathbf{X}_i)\dot{g}_{\beta}(\mathbf{X}_i, \mathbf{Z}_i, \beta)^\top = O_p(1).$$

Thus, we can prove that

$$\begin{aligned} B_{n2,2}^{[1,3]} &= -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \ddot{g}_{\beta,x}(\mathbf{X}_i, \mathbf{Z}_i, \beta)(\hat{\mathbf{X}}_i - \mathbf{X}_i)\dot{g}_{\beta}(\mathbf{X}_i, \mathbf{Z}_i, \beta)^\top(\hat{\beta}_n - \beta) + o_p(1) \\ &= o_p(1). \quad \text{(S1.25)} \end{aligned}$$

For $B_{n2,2}^{[1,4]}$, it can be decomposed into two parts:

$$\begin{aligned} B_{n2,2}^{[1,4]} &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \ddot{g}_{\beta,x}(\mathbf{X}_i, \mathbf{Z}_i, \beta)(\hat{\mathbf{X}}_i - \mathbf{X}_i)(\hat{\mathbf{X}}_i - \mathbf{X}_i)^\top \dot{g}_x(\mathbf{X}_i, \mathbf{Z}_i, \beta) \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \ddot{g}_{\beta,x}(\mathbf{X}_i, \mathbf{Z}_i, \beta)(\hat{\mathbf{X}}_i - \mathbf{X}_i)(\hat{\mathbf{X}}_i - \mathbf{X}_i)^\top \end{aligned}$$

$$\{\dot{g}_x(\mathbf{X}_i, \mathbf{Z}_i, \hat{\beta}_n) - \dot{g}_x(\mathbf{X}_i, \mathbf{Z}_i, \beta)\} + o_p(1).$$

Similarly to the proof of $B_{n2,2}^{[1,1]} = o_p(1)$, we can prove the first term

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \ddot{g}_{\beta,x}(\mathbf{X}_i, \mathbf{Z}_i, \beta)(\hat{\mathbf{X}}_i - \mathbf{X}_i)(\hat{\mathbf{X}}_i - \mathbf{X}_i)^\top \dot{g}_x(\mathbf{X}_i, \mathbf{Z}_i, \beta) = o_p(1).$$

Furthermore, we can prove that

$$\begin{aligned}&\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \ddot{g}_{\beta,x}(\mathbf{X}_i, \mathbf{Z}_i, \beta)(\hat{\mathbf{X}}_i - \mathbf{X}_i)(\hat{\mathbf{X}}_i - \mathbf{X}_i)^\top \{\dot{g}_x(\mathbf{X}_i, \mathbf{Z}_i, \hat{\beta}_n) - \dot{g}_x(\mathbf{X}_i, \mathbf{Z}_i, \beta)\}\\ &= o_p\left(n^{-1/2}\right) = o_p(1).\end{aligned}$$

Therefore, we have the result that

$$B_{n2,2}^{[1,4]} = o_p(1). \tag{S1.26}$$

From (S1.21)–(S1.26), we can validate that

$$B_{n2,2}^{[1]} = o_p(1). \tag{S1.27}$$

For $B_{n2,2}^{[2]}$, we have

$$B_{n2,2}^{[2]} = (\hat{\beta}_n - \beta)^\top \frac{1}{\sqrt{n}} \sum_{i=1}^{n} g_{\beta,x,\beta}^{(3)}(\mathbf{X}_i, \mathbf{Z}_i, \beta)(\hat{\mathbf{X}}_i - \mathbf{X}_i)\{\hat{Y}_i - g(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \hat{\beta}_n)\},$$

where $g_{\beta,x,\beta}^{(3)}(\mathbf{X}_i, \mathbf{Z}_i, \beta)$ is three order partial derivative of the function g. Similar to the proof of (S1.27), we can prove that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} g_{\beta,x,\beta}^{(3)}(\mathbf{X}_i, \mathbf{Z}_i, \beta)(\hat{\mathbf{X}}_i - \mathbf{X}_i)\{\hat{Y}_i - g(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \hat{\beta}_n)\} = o_p(1).$$

Then, we have

$$B_{n2,2}^{[2]} = (\hat{\beta}_n - \beta)^\top o_p\left(n^{-1/2}\right) = o_p(1). \tag{S1.28}$$

From (S1.20) and (S1.27)–(S1.28), it yields $B_{n2,2} = o_p(1)$. Furthermore, by (S1.18)–(S1.19), it can be validated that $B_{n2} = o_p(1)$. The above result, combining with (S1.11) and (S1.17), yields

$$\begin{aligned} B_n &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\mathrm{E}\{Y\dot{g}_\beta(\mathbf{X},\mathbf{Z},\beta)\}}{\mathrm{E}(Y)}\{\widetilde{Y}_i - Y_i\} + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\dot{g}_\beta(\mathbf{X}_i,\mathbf{Z}_i,\beta)\varepsilon_i \\ &\quad + \mathrm{E}\{\dot{g}_\beta(\mathbf{X},\mathbf{Z},\beta)S(\mathbf{X},\mathbf{Z})\} + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Sigma_x(\widetilde{\mathbf{X}}_i - \mathbf{X}_i) \\ &\quad - \mathrm{E}\{\dot{g}_\beta(\mathbf{X},\mathbf{Z},\beta)\dot{g}_\beta(\mathbf{X},\mathbf{Z},\beta)^\top\}\sqrt{n}(\hat{\beta}_n - \beta) + o_p(1) \end{aligned}$$

Letting $B_n = 0$ and solving the equation for $(\hat{\beta}_n - \beta)$, we complete the proof of the lemma. □

S2 The proofs of the theorems

Proof of Theorem 1: By setting the deviation function $S(\mathbf{X},\mathbf{Z})$ to be zero, we obtain the result of Theorem 1 from Theorem 2. We omit the details of the proof. □

Proof of Theorem 2: (I) The proof of Part (1): We first prove the result for $\mathcal{M}_{n,pro}(t)$. By the definition of $\mathcal{M}_{n,pro}(t)$, we have the following decomposition:

$$\begin{aligned} \mathcal{M}_{n,pro}(t) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\hat{Y}_i - g(\hat{\mathbf{X}}_i,\mathbf{Z}_i,\hat{\beta}_n))\mathbf{1}(\nu_i^\top\theta \le t) \\ &\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\hat{Y}_i - g(\hat{\mathbf{X}}_i,\mathbf{Z}_i,\hat{\beta}_n))\{\mathbf{1}(\mathbf{V}_i^\top\theta \le t) - \mathbf{1}(\nu_i^\top\theta \le t) \end{aligned}$$

$$=: \mathcal{D}_n(t) + \mathcal{G}_n(t).$$

We first consider the term $\mathcal{D}_n(t)$:

$$\begin{aligned}\mathcal{D}_n(t) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)\mathbf{1}(\nu_i^\top\theta \le t) + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{Y_i - g(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \hat{\beta}_n)\}\mathbf{1}(\nu_i^\top\theta \le t) \\ &=: \mathcal{D}_{n1}(t) + \mathcal{D}_{n2}(t). \end{aligned} \quad \text{(S2.1)}$$

For the term $\mathcal{D}_{n1}(t)$, by Slutsky theorem, we have the following result:

$$\begin{aligned}\mathcal{D}_{n1}(t) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{\frac{\widetilde{Y}_i\widetilde{Y}_{m,n}}{\hat{\psi}_n(U_i)} - Y_i\right\}\mathbf{1}(\nu_i^\top\theta \le t) + o_p(1) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{\frac{\widetilde{Y}_i\mathrm{E}(Y) - Y_i\hat{\psi}_n(U_i)}{\hat{\psi}_n(U_i)}\right\}\mathbf{1}(\nu_i^\top\theta \le t) + o_p(1) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\psi(U_i)\mathrm{E}(Y) - \hat{\psi}_n(U_i)}{\mathrm{E}(\widetilde{Y}|U_i)}Y_i\mathbf{1}(\nu_i^\top\theta \le t) + o_p(1). \end{aligned} \quad \text{(S2.2)}$$

We can further validate that

$$\begin{aligned}\mathcal{D}_{n1}(t) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{Y_i\mathbf{1}(\nu_i^\top\theta \le t)}{\mathrm{E}(\widetilde{Y}|U_i)}\frac{1}{nh_nf_u(U_i)}\sum_{j=1}^{n}K\left(\frac{U_i - U_j}{h_n}\right)\{\psi(U_i)\mathrm{E}(Y) - \widetilde{Y}_j\} + o_p(1) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}Y_i\mathbf{1}(\nu_i^\top\theta \le t)\frac{1}{nh_nf_u(U_i)}\sum_{j=1}^{n}\mathrm{E}(Y)K\left(\frac{U_i - U_j}{h_n}\right)\frac{\psi(U_i) - \psi(U_j)}{\mathrm{E}(\widetilde{Y}|U_i)} \\ &\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}Y_i\mathbf{1}(\nu_i^\top\theta \le t)\frac{1}{nh_nf_u(U_i)}\sum_{j=1}^{n}K\left(\frac{U_i - U_j}{h_n}\right)\frac{\psi(U_j)\mathrm{E}(Y) - \widetilde{Y}_j}{\mathrm{E}(\widetilde{Y}|U_i)} + o_p(1) \\ &=: \mathcal{D}_{n1,1}(t) + \mathcal{D}_{n1,2}(t) + o_p(1). \end{aligned}$$

By taking Taylor expansion of $\psi(U_i) - \psi(U_j)$, we employ Condition (C8) and then prove that $\mathcal{D}_{n1,1}(t) = O_p(\sqrt{n}h_n^2) = o_p(1)$. Therefore, $\mathcal{D}_{n1,1}(t)$ can be expressed as follows.

$$\mathcal{D}_{n1}(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{Y_i\mathbf{1}(\nu_i^\top\theta \le t)}{\mathrm{E}(\widetilde{Y}|U_i)}\frac{1}{nh_nf_u(U_i)}\sum_{j=1}^{n}K\left(\frac{U_i - U_j}{h_n}\right)\{\psi(U_j)\mathrm{E}(Y) - \widetilde{Y}_j\} + o_p(1)$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\frac{\psi(U_j)\mathrm{E}(Y)-\psi(U_j)Y_j}{nh_nf_u(U_i)}\sum_{i=1}^{n}\frac{Y_i\mathbf{1}(\nu_i^{\top}\theta\le t)K\left(\frac{U_i-U_j}{h_n}\right)}{\mathrm{E}(\widetilde{Y}|U_i)}+o_p(1)\\
&= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\frac{\psi(U_j)\{\mathrm{E}(Y)-Y_j\}}{\mathrm{E}(\widetilde{Y}|U_j)}\mathrm{E}\{Y\mathbf{1}(\nu^{\top}\theta\le t)|U_j\}+o_p(1)\\
&= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\frac{\mathrm{E}(Y)-Y_j}{\mathrm{E}(Y)}\mathrm{E}\{Y\mathbf{1}(\nu^{\top}\theta\le t)|U_j\}+o_p(1). \qquad \text{(S2.3)}
\end{aligned}
$$

In the following, we investigate $\mathcal{D}_{n2}(t)$. Under the alternative hypothetical models $\mathcal{H}_{1n}: Y=g(\mathbf{X},\mathbf{Z},\beta)+n^{-1/2}S(\mathbf{X},\mathbf{Z})+\varepsilon$, it can be split into

$$
\begin{aligned}
\mathcal{D}_{n2}(t) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\varepsilon_i\mathbf{1}(\nu_i^{\top}\theta\le t)+\frac{1}{n}\sum_{i=1}^{n}S(\mathbf{X}_i,\mathbf{Z}_i)\mathbf{1}(\nu_i^{\top}\theta\le t)\\
&\quad+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{g(\mathbf{X}_i,\mathbf{Z}_i,\beta)-g(\mathbf{X}_i,\mathbf{Z}_i,\hat{\beta}_n)\}\mathbf{1}(\nu_i^{\top}\theta\le t)\\
&\quad+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{g(\mathbf{X}_i,\mathbf{Z}_i,\hat{\beta}_n)-g(\hat{\mathbf{X}}_i,\mathbf{Z}_i,\hat{\beta}_n)\}\mathbf{1}(\nu_i^{\top}\theta\le t)\\
&=: \mathcal{D}_{n2,1}(t)+\mathcal{D}_{n2,2}(t)+\mathcal{D}_{n2,3}(t)+\mathcal{D}_{n2,4}(t). \qquad \text{(S2.4)}
\end{aligned}
$$

By the law of large numbers, it is easy to observe that

$$
\mathcal{D}_{n2,2}(t) = \mathrm{E}\{S(\mathbf{X},\mathbf{Z})\mathbf{1}(\nu^{\top}\theta\le t)\}+o_p(1). \qquad \text{(S2.5)}
$$

It follows that

$$
\begin{aligned}
\mathcal{D}_{n2,3}(t) &= -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\dot{g}_{\beta}(\mathbf{X}_i,\mathbf{Z}_i,\beta)^{\top}(\hat{\beta}_n-\beta)\mathbf{1}(\nu_i^{\top}\theta\le t)+o_p(1)\\
&= -\mathrm{E}\{\dot{g}_{\beta}(\mathbf{X},\mathbf{Z},\beta)^{\top}\mathbf{1}(\nu^{\top}\theta\le t)\}\sqrt{n}(\hat{\beta}_n-\beta)+o_p(1). \qquad \text{(S2.6)}
\end{aligned}
$$

For $\mathcal{D}_{n2,4}(t)$, we have

$$
\mathcal{D}_{n2,4}(t) = -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\dot{g}_x(\mathbf{X}_i,\mathbf{Z}_i,\beta)^{\top}(\hat{\mathbf{X}}_i-\mathbf{X}_i)\mathbf{1}(\nu_i^{\top}\theta\le t)
$$

$$
\begin{aligned}
& -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{\dot{g}_x(\mathbf{X}_i,\mathbf{Z}_i,\hat{\beta}_n)-\dot{g}_x(\mathbf{X}_i,\mathbf{Z}_i,\beta)\}^\top(\hat{\mathbf{X}}_i-\mathbf{X}_i)\mathbf{1}(\nu_i^\top\theta\le t)+o_p(1)\\
= \;& \mathcal{D}^{[1]}_{n2,4}(t)+\mathcal{D}^{[2]}_{n2,4}(t)+o_p(1).
\end{aligned}
$$

For the first term $\mathcal{M}^{[1]}_{n2,4}(t)$, it follows from Lemma 2 that

$$
\mathcal{D}^{[1]}_{n2,4}(t)=\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\mathrm{E}\{(\mathbf{X}\otimes\dot{g}_x(\mathbf{X},\mathbf{Z},\beta)\oslash\mathrm{E}(\mathbf{X}))^\top\mathbf{1}(\nu^\top\theta\le t)|U_j\}(\widetilde{\mathbf{X}}_j-\mathbf{X}_j)+o_p(1).
$$

Note that

$$
\mathcal{D}^{[2]}_{n2,4}(t)=-(\hat{\beta}_n-\beta)^\top\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{\ddot{g}_{x,\beta}(\mathbf{X}_i,\mathbf{Z}_i,\hat{\beta}_n)(\hat{\mathbf{X}}_i-\mathbf{X}_i)\mathbf{1}(\nu_i^\top\theta\le t)+o_p(1).
$$

Similarly to the proof of (S2.7), we prove that

$$
\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{\ddot{g}_{x,\beta}(\mathbf{X}_i,\mathbf{Z}_i,\hat{\beta}_n)(\hat{\mathbf{X}}_i-\mathbf{X}_i)\mathbf{1}(\nu_i^\top\theta\le t)=O_p(1).
$$

Further by the results of Lemma 3, we can prove that $\mathcal{D}^{[2]}_{n2,4}(t)=o_p(1)$. Therefore, we conclude that

$$
\begin{aligned}
\mathcal{D}_{n2,4}(t) \;=\;& \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\mathrm{E}\left\{(\mathbf{X}\otimes\dot{g}_x(\mathbf{X},\mathbf{Z},\beta)\oslash\mathrm{E}(\mathbf{X}))^\top\mathbf{1}(V^\top\theta\le t)|U_j\right\}(\widetilde{\mathbf{X}}_j-\mathbf{X}_j)\\
&+o_p(1). \qquad\qquad \text{(S2.7)}
\end{aligned}
$$

It follows from (S2.4)–(S2.7) that

$$
\begin{aligned}
\mathcal{D}_{n2}(t) \;=\;& \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{1}(\nu_i^\top\theta\le t)\varepsilon_i-\Gamma_1(t)\sqrt{n}(\hat{\beta}_n-\beta)+\mathrm{E}\{S(\mathbf{X},\mathbf{Z})\mathbf{1}(\nu^\top\theta\le t)\}\\
&+\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\mathrm{E}\{(\mathbf{X}\otimes\dot{g}_x(\mathbf{X},\mathbf{Z},\beta)\oslash\mathrm{E}(\mathbf{X}))^\top\mathbf{1}(V^\top\theta\le t)|U_j\}(\widetilde{\mathbf{X}}_j-\mathbf{X}_j)\\
&+o_p(1) \qquad\qquad \text{(S2.8)}
\end{aligned}
$$

with $\Gamma_1(t)$ presented in Appendix A. By (S2.1), (S2.3) and (S2.8), we validate that

$$\begin{aligned}
\mathcal{D}_n(t) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{1}(\nu_i^\top\theta\le t)\varepsilon_i+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\widetilde{Y}_i-Y_i}{\mathrm{E}(Y)}\mathrm{E}\{Y\mathbf{1}(\nu^\top\theta\le t)|U_i\}\\
&\quad-\Gamma_1(t)\sqrt{n}(\hat\beta_n-\beta)+\ \mathrm{E}\{S(\mathbf{X},\mathbf{Z})\mathbf{1}(\nu^\top\theta\le t)\}+o_p(1)\\
&\quad+\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\mathrm{E}\left\{(\mathbf{X}\otimes\dot g_x(\mathbf{X},\mathbf{Z},\beta)\oslash\mathrm{E}(\mathbf{X}))^\top\mathbf{1}(\nu^\top\theta\le t)|U_j\right\}(\widetilde{\mathbf{X}}_j-\mathbf{X}_j)\\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathcal{IF}_{(v,\theta)}(Y_i,\mathbf{X}_i,\mathbf{Z}_i,\nu_i,U_i)+\mathcal{DR}_t+o_p(1)
\end{aligned}$$

with $\mathcal{IF}_{(v,\theta)}(Y,\mathbf{X},\mathbf{Z},\nu,U)$ and $\mathcal{DR}_t$ defined in Appendix A.

By $\mathbf{1}(\mathbf{V}_i^\top\theta\le t)=\mathbf{1}(\nu_i^\top\theta\le(\nu_i-\mathbf{V}_i)^\top\theta+t)$, for $\mathcal{G}_{n,pro}(t)$, we have

$$\begin{aligned}
\mathcal{G}_{n,pro}(t) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\hat Y_i-g(\hat{\mathbf{X}}_i,\mathbf{Z}_i,\beta))\{\mathbf{1}(\nu_i^\top\theta\le(\nu_i-\mathbf{V}_i)^\top\theta+t)-\mathbf{1}(\nu_i^\top\theta\le t)\}\\
&\quad+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(g(\hat{\mathbf{X}}_i,\mathbf{Z}_i,\beta)-g(\hat{\mathbf{X}}_i,\mathbf{Z}_i,\hat\beta_n))\{\mathbf{1}(\nu_i^\top\theta\le(\nu_i-\mathbf{V}_i)^\top\theta+t)-\mathbf{1}(\nu_i^\top\theta\le t)\}\\
&=: \mathcal{G}_{n1}(t)+\mathcal{G}_{n2}(t).
\end{aligned}$$

Denote the distribution of $\nu^\top\theta$ by $F_{\nu^\top\theta}(\cdot)$. We first split $\mathcal{G}_{n1}(t)$ into two parts:

$$\begin{aligned}
\mathcal{G}_{n1}(t) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\hat Y_i-g(\hat{\mathbf{X}}_i,\mathbf{Z}_i,\beta))[\mathbf{1}(t\le\nu_i^\top\theta\le(\nu_i-\mathbf{V}_i)^\top\theta+t)-\{F_{\nu^\top\theta}((\nu_i-\mathbf{V}_i)^\top\theta+t))\\
&\quad-F_{\nu^\top\theta}(t)\}]+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\hat Y_i-g(\hat{\mathbf{X}}_i,\mathbf{Z}_i,\beta))\{F_{\nu^\top\theta}((\nu_i-\mathbf{V}_i)^\top\theta+t))-F_{\nu^\top\theta}(t)\}\\
&=: \mathcal{G}_{n11}(t)+\mathcal{G}_{n12}(t).
\end{aligned}$$

For $\mathcal{G}_{n11}(t)$, we have

$$\mathcal{G}_{n11}(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\hat Y_i-Y_i)[\mathbf{1}(t\le\nu_i^\top\theta\le(\nu_i-\mathbf{V}_i)^\top\theta+t)-F_{\nu^\top\theta}((\nu_i-\mathbf{V}_i)^\top\theta+t)]$$

$$
\begin{aligned}
&+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(Y_i - g(\mathbf{X}_i, \mathbf{Z}_i, \beta))[\mathbf{1}(t \le \nu_i^\top\theta \le (\nu_i - \mathbf{V}_i)^\top\theta + t) - F_{\nu^\top\theta}((\nu_i - \mathbf{V}_i)^\top\theta + t)] \\
&+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(g(\mathbf{X}_i, \mathbf{Z}_i, \beta) - g(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \beta))[\mathbf{1}(t \le \nu_i^\top\theta \le (\nu_i - \mathbf{V}_i)^\top\theta + t) \\
&-F_{\nu^\top\theta}((\nu_i - \mathbf{V}_i)^\top\theta + t)] =: \mathcal{G}_{n11,1}(t) + \mathcal{G}_{n11,2}(t) + \mathcal{G}_{n11,3}(t).
\end{aligned}
$$

Similarly to the decomposition of (S2.2), we have

$$
\begin{aligned}
\mathcal{G}_{n11,1}(t) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\psi(U_i)\mathrm{E}(Y) - \hat{\psi}_n(U_i)}{\mathrm{E}(\widetilde{Y}|U_i)}Y_i[\mathbf{1}(t \le \nu_i^\top\theta \le (\nu_i - \mathbf{V}_i)^\top\theta + t) \\
&\quad -F_{\nu^\top\theta}((\nu_i - \mathbf{V}_i)^\top\theta + t)] + o_p(1).
\end{aligned}
$$

Recalling Lemma 2, we can validated that $\mathcal{G}_{n11,1}(t) = O_p(\sqrt{n}h_n^2) = o_p(1)$. Note that $\mathrm{E}[\mathcal{G}_{n11,2}(t)] = 0$. Furthermore, we can prove that $\mathrm{E}[\mathcal{G}_{n11,2}(t)]^2$ converges to zero as $n \to \infty$. So we have $\mathcal{G}_{n11,2}(t) = o_p(1)$. Next we consider $\mathcal{G}_{n11,3}(t)$:

$$
\begin{aligned}
\mathcal{G}_{n11,3}(t) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(g(\mathbf{X}_i, \mathbf{Z}_i, \beta) - g(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \beta))[\mathbf{1}(t \le \nu_i^\top\theta \le (\nu_i - \mathbf{V}_i)^\top\theta + t) \\
&\quad -F_{\nu^\top\theta}((\nu_i - \mathbf{V}_i)^\top\theta + t)] \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\dot{g}_x(\mathbf{X}_i, \mathbf{Z}_i, \beta)(\mathbf{X}_i - \hat{\mathbf{X}}_i)[\mathbf{1}(t \le \nu_i^\top\theta \le (\nu_i - \mathbf{V}_i)^\top\theta + t) \\
&\quad -F_{\nu^\top\theta}((\nu_i - \mathbf{V}_i)^\top\theta + t)].
\end{aligned}
$$

Similar to the proof of Lemma 2, we can prove that $\mathcal{G}_{n11,3}(t) = o_p(1)$. So We have $\mathcal{G}_{n11}(t) = o_p(1)$. Note that $\mathcal{G}_{n12}(t) = n^{-1/2}\sum_{i=1}^{n}(\hat{Y}_i - g(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \beta))f_{\nu^\top\theta}(t)$ $(\nu_i - \mathbf{V}_i)^\top\theta + o_p(1)$ where $f_{\nu^\top\theta}(t)$ is the density of $\nu^\top\theta$. By the similar method to prove $B_{n2,2}^{[2]} = o_p(1)$, we can prove $\mathcal{G}_{n12}(t) = o_p(1)$. Therefore, we have validated that $\mathcal{G}_{n1}(t) = o_p(1)$. For $\mathcal{G}_{n2}(t)$, we have

$$\begin{aligned}\mathcal{G}_{n2}(t) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\dot{g}_\beta(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \beta)^\top\{\mathbf{1}(\nu_i^\top\theta \le (\nu_i - \mathbf{V}_i)^\top\theta + t) - \mathbf{1}(\nu_i^\top\theta \le t)\}(\beta - \hat{\beta}_n) \\ &\quad + o_p(1).\end{aligned}$$

We can validate that $n^{-1/2}\sum_{i=1}^{n}\dot{g}_\beta(\hat{\mathbf{X}}_i, \mathbf{Z}_i, \beta)^\top\{\mathbf{1}(\nu_i^\top\theta \le (\nu_i - \mathbf{V}_i)^\top\theta + t) - \mathbf{1}(\nu_i^\top\theta \le t)\} = O_p(1)$ and $\beta - \hat{\beta}_n = O_p(n^{-1/2}) = O_p(1)$. Thus we obtain that $\mathcal{G}_{n2}(t) = o_p(1)$ and then $\mathcal{G}_n(t) = o_p(1)$. Therefore the following result is true: $\mathcal{M}_{n,pro}(t) = n^{-1/2}\sum_{i=1}^{n}\mathcal{IF}_{(v,\theta)}(Y_i, \mathbf{X}_i, \mathbf{Z}_i, \nu_i, U_i) + \mathcal{DR}_t + o_p(1)$.

By using the similar method in Sun et al. (2018), we can validate that the functional class $\Delta_t = \{\mathcal{IF}_{(t,\theta)}(Y, \mathbf{X}, \mathbf{Z}, \nu, U) + \mathcal{DR}_t : t \in \mathcal{R}\}$ is a V-C class of functions. Thus we can further prove the weak convergence of $\mathcal{M}_{n,pro}(t)$ as shown in Theorem 2. By the principle of continuous mapping, we can prove convergence properties of $\mathcal{T}_{n,CvM}$ and $\mathcal{T}_{n,KS}$.

(II) The proof of Part (2): By rewriting (4.1) as $Y = g(\mathbf{X}, \mathbf{Z}, \beta) + n^{-1/2}S^*(\mathbf{X}, \mathbf{Z}) + \varepsilon$ with $S^*(\mathbf{X}, \mathbf{Z}) = C_n n^{1/2} S(\mathbf{X}, \mathbf{Z})$, and applying the results of Theorem 2, then can be validated that $\mathcal{T}_{n,CvM} \to \infty$ and $\mathcal{T}_{n,KS} \to \infty$, since the new drift function $\mathcal{DR}_t^*(\mathbf{X}, \mathbf{Z}, \nu) = C_n n^{1/2}\mathcal{DR}_t \to \infty$ as $n \to \infty$.

$\square$

S3 Additional simulation studies

In this section, we report additional simulation results to evaluate the finite sample performance of the proposed method. The following setting is considered:

Setting 1. A five-dimensional nonlinear candidate model with part of covariates are observed distortedly is taken into account:

$$Y = \exp(\beta_1^\top \mathbf{X} + \beta_2^\top \mathbf{Z}) + C\beta_1^\top \mathbf{X} + \varepsilon, \tag{S3.1}$$

where $\mathbf{X} \sim \mathcal{U}_3[1,2]$, $\mathbf{Z} \sim \mathcal{U}_2[1,2]$, $\beta_1 = (1,1,-1)^\top$, and $\beta_2 = (-1,-1)^\top$. The distorting functions are specified as $\gamma_1(U) = 1 + 0.3\cos(2\pi U)$, $\gamma_2(U) = 1 + 0.2(U^2 - 1/3)$, and $\gamma_3(U) = U^2 + 2/3$. The constant C is selected to be $0.0, 0.1, 0.2, 0.3, 0.4$. All other configurations are the same as those in the main text of the article.

We calculate the empirical sizes and powers for model (S3.1) and present the results in Table 1.

S4 Analyses of Boston house price data

In the following, we further employ the proposed and existing methods to analyze the Boston house price data set (Harrison Jr and Rubinfeld, 1978; Şentürk and Müller, 2005; Xie and Zhu, 2019)(http://lib.stat.cmu.edu/datasets/).

Table 1: Results for Setting 1. Empirical sizes and powers of $\mathcal{T}_{n,CvM}$, $\mathcal{T}^U_{n,CvM}$, $\mathcal{T}^N_{n,CvM}$, $\mathcal{T}^U_{n,KS}$, $\mathcal{T}^N_{n,KS}$, $\mathcal{T}^{ZLF}_n$, and $\mathcal{T}^{ZX}_n$ at the 5% significance level for the five-dimensional model (S3.1).

Model	n	C	$\mathcal{T}_{n,CvM}$	$\mathcal{T}^U_{n,CvM}$	$\mathcal{T}^N_{n,CvM}$	$\mathcal{T}^U_{n,KS}$	$\mathcal{T}^N_{n,KS}$	$\mathcal{T}^{ZLF}_n$	$\mathcal{T}^{ZX}_n$
(S3.1)	100	0.0	0.058	0.050	0.056	0.058	0.054	0.036	0
		0.1	0.194	0.192	0.190	0.102	0.092	0.076	0
		0.2	0.364	0.360	0.354	0.234	0.228	0.172	0
		0.3	0.602	0.578	0.604	0.518	0.526	0.262	0
		0.4	0.704	0.670	0.682	0.666	0.670	0.304	0
	200	0.0	0.058	0.058	0.054	0.056	0.056	0.036	0
		0.1	0.318	0.304	0.286	0.124	0.124	0.082	0
		0.2	0.658	0.624	0.644	0.376	0.378	0.262	0
		0.3	0.848	0.846	0.854	0.710	0.710	0.398	0
		0.4	0.930	0.922	0.920	0.874	0.874	0.550	0
	300	0.0	0.056	0.054	0.058	0.048	0.052	0.040	0
		0.1	0.398	0.384	0.402	0.114	0.108	0.122	0
		0.2	0.830	0.810	0.798	0.404	0.404	0.360	0
		0.3	0.964	0.956	0.956	0.808	0.814	0.580	0
		0.4	0.980	0.976	0.982	0.944	0.940	0.698	0.002

The data set contains 506 observations and 14 variables. We aim for checking the adequacy of the candidate linear model of the response variable Y: the median value of owner-occupied homes in \$1000's (MEDV) and the other 12 variables: the per capita crime rate (CRIM)(X), proportion of residential land zoned for lots over 25,000 square feet(Z_1), the proportion of non-retail business acres per town(Z_2), Charles River dummy variable ($= 1$ if tract bounds river, 0 otherwise)(Z_3), nitric oxides concentration (parts per 10 million)(Z_4), average number of rooms per dwelling(Z_5), the proportion of owner-occupied units built prior to 1940(Z_6), weighted distances to five Boston employment centers(Z_7), index of accessibility to radial highways(Z_8), full-value property-tax rate per \$10,000(Z_9), the pupil-teacher ratio by town(Z_{10}), and $1000(B_k - 0.63)^2$ where B_k is the proportion of blacks by town(Z_{11}). Our interest is in testing the following hypothesis:

$$\mathcal{H}_{03} : \mathrm{E}(Y|X, \mathbf{Z}) = \beta_0 + \beta_1 X + \beta_2^\top \mathbf{Z} \tag{S4.1}$$

with $\mathbf{Z} = (Z_1, \ldots, Z_{11})^\top$.

We make the same assumptions as those of Xie and Zhu (2019) that both MEDV and CRIM are distorted by the confounding variable: the proportion of the population of lower status (LSTAT). By the same settings of the kernel function, bandwidth, the number of bootstrap repetitions and the value of m in the random approximation procedures as those in simulation studies. We obtain the p-values of the seven test methods, which are displayed in Table 2. The results in Table 2 demonstrate that the proposed tests suggest rejecting the null hypothetical linear model in (S4.1), while the tests of Zhang et al. (2015) and Zhao and Xie (2018) cannot reject the null hypothesis (S4.1).

We draw the scatter plots of the calibrated MEDV and the estimated residuals versus the estimated regression function in Figure 1. The evidence that the estimated residual curve deviates significantly from a horizontal line in Figure 1 (b) indicates that the linear model (S4.1) is inadequate for this data set. Compared with the tests $\mathcal{T}_n^{ZLF}$ and $\mathcal{T}_n^{ZX}$, the proposed tests are more powerful and may provide more accurate results.

Table 2: The p-values of the tests for the analyses of Boston house price data set. The test $\mathcal{T}_n^{ZLF}$ yields three p-values corresponding to three weighting functions: $\sin(X), \exp(X)$ and $\cos(X)$.

Study/Model	$\mathcal{T}_{n,CvM}$	$\mathcal{T}_{n,CvM}^U$	$\mathcal{T}_{n,CvM}^N$	$\mathcal{T}_{n,KS}^U$	$\mathcal{T}_{n,KS}^N$	$\mathcal{T}_n^{ZLF}$	$\mathcal{T}_n^{ZX}$
Boston/(S4.1)	0	0	0	0.003	0.004	(0.225 0.501 0.812)	0.158

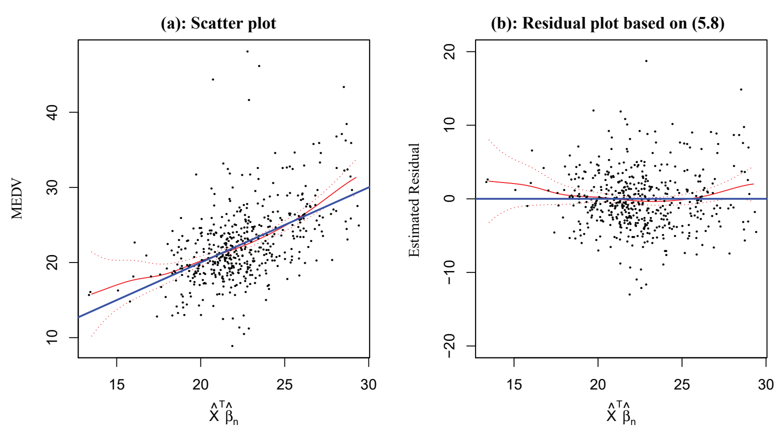

Figure 1: Scatter plots of calibrated variable of MEDV (a) and the estimated residuals (b) versus the estimated regression function along with linear fittings (thick lines) and nonparametric estimated curves (solid lines) with 95% confidence bands (dotted lines).

References

Şentürk, D. and Müller, H.-G. (2005). Covariate adjusted correlation analysis via varying coefficient models. *Scand. J. Statist.* **32**, 365–383.

Harrison Jr, D. and Rubinfeld, D. L. (1978). Hedonic housing prices and the demand for clean air. *Journal of Environmental Economics and Management* **5**, 81–102.

Li, Q. and Racine, J. S. (2007). *Nonparametric Econometrics.* Princeton University Press, Princeton, NJ. Theory and Practice.

Masry, E. (1996). Multivariate local polynomial regression for time series: uniform strong consistency and rates. *J. Time Ser. Anal.* **17**, 571–599.

Sun, Z., Ye, X. and Sun, L. (2018). Consistent test for parametric models with right-censored data using projections. *Comput. Statist. Data Anal.* **118**, 112–125.

Xie, C. and Zhu, L. (2019). A goodness-of-fit test for variable-adjusted models. *Comput. Statist. Data Anal.* **138**, 27–48.

Zhang, J., Li, G.-R. and Feng, Z.-H. (2015). Checking the adequacy for a distortion errors-in-variables parametric regression model. *Comput. Statist. Data Anal.* **83**, 52–64.

Zhang, J., Zhu, L. and Liang, H. (2012). Nonlinear models with measurement errors subject to single-indexed distortion. *J. Multivariate Anal.* **112**, 1–23.

Zhao, J. and Xie, C. (2018). A nonparametric test for covariate-adjusted models. *Statist. Probab. Lett.* **133**, 65–70.