Supplementary Appendix

Supplement to : Zongwu Cai, Ying Fang, Ming Lin, and Jia Su, "Inferences for a Partially Varying Coefficient Model with Endogenous Regressors".

Lemma 1. *Let* (X_1, Y_1),..., (X_n, Y_n) *be i.i.d random vectors, where the* Y_i*'s are scalar random variables. Assume that* $\sup_x \int |y|^s f(x,y)dy < \infty$ *for some* $s > 2$*, where* $f(x,y)$ *denotes the joint density of* (X_i, Y_i)*. Let* $K(x)$ *be a bounded positive function with bounded support, satisfying a Lipschitz condition, and let* $K_h(x) = K(x/h)/h$*. Given that* $n^{2\epsilon-1}h \to \infty$ *for some* $\varepsilon < 1 - s^{-1}$*, then*

$$\sup_x \Big| \frac{1}{n} \sum_{k=1}^{n} K_h(X_i - x)Y_i - E[K_h(X_i - x)Y_i] \Big| = O_p \left[\frac{\log(1/h)}{nh} \right]^{1/2}.$$

Proof of Lemma 1: The proof is given by Proposition 4 in Mack and Silverman (1982). $\square$

Lemma 2. *Suppose that Assumptions A1-A7 hold, we have*

$$\left[\mathbf{D}_Z^T(u)\mathbf{H}(u)\mathbf{D}_Z(u)\right]^{-1} \mathbf{D}_Z^T(u)\mathbf{H}(u)\mathbf{D}_X(u) = \begin{pmatrix} \Pi_X(u) & 0 \\ 0 & \Pi_X(u) \end{pmatrix} + O_p(c_n)$$

holds uniformly for $u \in \Omega$*, where* $c_n = \left[\frac{\log(1/h)}{nh}\right]^{1/2} + h$*.*

Proof of Lemma 2: Note that

$$\begin{aligned} & \left[\mathbf{D}_Z^T(u)\mathbf{H}(u)\mathbf{D}_Z(u)\right]^{-1} \mathbf{D}_Z^T(u)\mathbf{H}(u)\mathbf{D}_X(u) - \begin{pmatrix} \Pi_X(u) & 0 \\ 0 & \Pi_X(u) \end{pmatrix} \qquad \text{(A.1)} \\ = & \left[\mathbf{D}_Z^T(u)\mathbf{H}(u)\mathbf{D}_Z(u)\right]^{-1} \mathbf{D}_Z^T(u)\mathbf{H}(u) \left[\mathbf{D}_X(u) - \mathbf{D}_Z(u) \begin{pmatrix} \Pi_X(u) & 0 \\ 0 & \Pi_X(u) \end{pmatrix}\right] \\ = & \left[\mathbf{D}_Z^T(u)\mathbf{H}(u)\mathbf{D}_Z(u)\right]^{-1} \\ & \mathbf{D}_Z^T(u)\mathbf{H}(u) \begin{pmatrix} X_1^T - Z_1^T\Pi_X(u) & \frac{u_1-u}{h}\left[X_1^T - Z_1^T\Pi_X(u)\right] \\ \vdots & \vdots \\ X_n^T - Z_n^T\Pi_X(u) & \frac{u_n-u}{h}\left[X_n^T - Z_n^T\Pi_X(u)\right] \end{pmatrix} \end{aligned}$$

By Lemma 1, we have

$$
\begin{aligned}
&\frac{1}{n}\mathbf{D}_Z^T(u)\mathbf{H}(u)\mathbf{D}_Z(u) \qquad\qquad (A.2)\\
= \;& \frac{1}{n}\begin{pmatrix} \sum_{k=1}^n Z_k Z_k^T K_h(u_k-u) & \sum_{k=1}^n Z_k Z_k^T \frac{u_k-u}{h} K_h(u_k-u) \\ \sum_{k=1}^n Z_k Z_k^T \frac{u_k-u}{h} K_h(u_k-u) & \sum_{k=1}^n Z_k Z_k^T \frac{(u_k-u)^2}{h^2} K_h(u_k-u) \end{pmatrix} \\
= \;& \begin{pmatrix} E\left[Z_k Z_k^T K_h(u_k-u)\right] & E\left[Z_k Z_k^T \frac{u_k-u}{h} K_h(u_k-u)\right] \\ E\left[Z_k Z_k^T \frac{u_k-u}{h} K_h(u_k-u)\right] & E\left[Z_k Z_k^T \frac{(u_k-u)^2}{h^2} K_h(u_k-u)\right] \end{pmatrix} + O_p\left[\frac{\log(1/h)}{nh}\right]^{1/2} \\
= \;& \begin{pmatrix} E(Z_k Z_k^T \mid u_k=u)f(u) + O_p(h^2) & O_p(h) \\ O_p(h) & \mu_2 E(Z_k Z_k^T \mid u_k=u)f(u) + O_p(h^2) \end{pmatrix} + O_p\left[\frac{\log(1/h)}{nh}\right]^{1/2} \\
= \;& \begin{pmatrix} E(Z_k Z_k^T \mid u_k=u)f(u) & 0 \\ 0 & \mu_2 E(Z_k Z_k^T \mid u_k=u)f(u) \end{pmatrix} + O_p(c_n),
\end{aligned}
$$

where $\mu_2 = \int t^2 K(t)dt$. Similarly, we can show that

$$
\begin{aligned}
&\frac{1}{n}\mathbf{D}_Z^T(u)\mathbf{H}(u)\begin{pmatrix} X_1^T - Z_1^T\Pi_X(u) & \frac{u_1-u}{h}\left[X_1^T - Z_1^T\Pi_X(u)\right] \\ \vdots & \vdots \\ X_n^T - Z_n^T\Pi_X(u) & \frac{u_n-u}{h}\left[X_n^T - Z_n^T\Pi_X(u)\right] \end{pmatrix} \qquad (A.3)\\
= \;& \frac{1}{n}\mathbf{D}_Z^T(u)\mathbf{H}(u)\begin{pmatrix} v_{X,1}^T & \frac{u_1-u}{h} v_{X,1}^T \\ \vdots & \vdots \\ v_{X,n}^T & \frac{u_n-u}{h} v_{X,n}^T \end{pmatrix} = O_P(c_n).
\end{aligned}
$$

Combine (A.1), (A.2) and (A.3), we obtain that

$$
\left[\mathbf{D}_Z^T(u)\mathbf{H}(u)\mathbf{D}_Z(u)\right]^{-1}\mathbf{D}_Z^T(u)\mathbf{H}(u)\mathbf{D}_X(u) - \begin{pmatrix} \Pi_X(u) & 0 \\ 0 & \Pi_X(u) \end{pmatrix} = O_p(c_n)
$$

uniformly for $u \in \Omega$. This completes the proof. □

Lemma 3 *Suppose that Assumptions A1-A7 hold, then*

$$
\frac{1}{n}\widehat{\mathbf{D}}_X(u)^T\mathbf{H}(u)\widehat{\mathbf{D}}_X(u) = \begin{pmatrix} \mathbf{\Phi}(u) & 0 \\ 0 & \mu_2\mathbf{\Phi}(u) \end{pmatrix} + O_p(c_n)
$$

uniformly for $u \in \Omega$, where $\mu_2 = \int t^2 K(t)dt$. Furthermore,

$$
\left[\frac{1}{n}\widehat{\mathbf{D}}_X(u)^T\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\right]^{-1} = \begin{pmatrix} \mathbf{\Phi}^{-1}(u) & 0 \\ 0 & \mu_2^{-1}\mathbf{\Phi}^{-1}(u) \end{pmatrix} + O_p(c_n).
$$

Proof of Lemma 3: Note that

$$
\begin{aligned}
& \frac{1}{n}\widehat{\mathbf{D}}_X(u)^T\mathbf{H}(u)\widehat{\mathbf{D}}_X(u) \\
= \; & \mathbf{D}_X^T(u)\mathbf{H}(u)\mathbf{D}_Z(u)\left[\mathbf{D}_Z^T(u)\mathbf{H}(u)\mathbf{D}_Z(u)\right]^{-1}\left[\frac{1}{n}\mathbf{D}_Z^T(u)\mathbf{H}(u)\mathbf{D}_X(u)\right] \\
= \; & \mathbf{D}_X^T(u)\mathbf{H}(u)\mathbf{D}_Z(u)\left[\mathbf{D}_Z^T(u)\mathbf{H}(u)\mathbf{D}_Z(u)\right]^{-1} \\
& \qquad \frac{1}{n}\begin{pmatrix} \sum_{k=1}^n Z_kX_k^TK_h(u_k-u) & \sum_{k=1}^n Z_kX_k^T\frac{u_k-u}{h}K_h(u_k-u) \\ \sum_{k=1}^n Z_kX_k^T\frac{u_k-u}{h}K_h(u_k-u) & \sum_{k=1}^n Z_kX_k^T\frac{(u_k-u)^2}{h^2}K_h(u_k-u) \end{pmatrix} \\
= \; & \mathbf{D}_X^T(u)\mathbf{H}(u)\mathbf{D}_Z(u)\left[\mathbf{D}_Z^T(u)\mathbf{H}(u)\mathbf{D}_Z(u)\right]^{-1} \\
& \qquad \left[\begin{pmatrix} E(Z_kX_k^T \mid u_k=u)f(u) & 0 \\ 0 & \mu_2E(Z_kX_k^T \mid u_k=u)f(u) \end{pmatrix} + O_p(c_n)\right].
\end{aligned}
$$

By Lemma 2, we have

$$
\begin{aligned}
& \frac{1}{n}\widehat{\mathbf{D}}_X(u)^T\mathbf{H}(u)\widehat{\mathbf{D}}_X(u) \\
= \; & \begin{pmatrix} \Pi_X^T(u) & 0 \\ 0 & \Pi_X^T(u) \end{pmatrix}\begin{pmatrix} E(Z_kX_k^T \mid u_k=u)f(u) & 0 \\ 0 & \mu_2E(Z_kX_k^T \mid u_k=u)f(u) \end{pmatrix} + O_p(c_n) \\
= \; & \begin{pmatrix} \Pi_X^T(u) & 0 \\ 0 & \Pi_X^T(u) \end{pmatrix}\begin{pmatrix} E(Z_kZ_k^T \mid u_k=u)\Pi_X(u)f(u) & 0 \\ 0 & \mu_2E(Z_kZ_k^T \mid u_k=u)\Pi_X(u)f(u) \end{pmatrix} + O_p(c_n) \\
= \; & \begin{pmatrix} \mathbf{\Phi}(u) & 0 \\ 0 & \mu_2\mathbf{\Phi}(u) \end{pmatrix} + O_p(c_n).
\end{aligned}
$$

This completes the proof. □

Lemma 4 *Suppose that Assumptions A1-A7 hold, then we have*

$$
\frac{1}{n}\widehat{\mathbf{W}}^T(\mathbf{I}-\widehat{\mathbf{S}})^T(\mathbf{I}-\widehat{\mathbf{S}})\widehat{\mathbf{W}} \xrightarrow{p} \Sigma_1 \quad \text{and} \quad \frac{1}{n}\widehat{\mathbf{W}}^T(\mathbf{I}-\widehat{\mathbf{S}})^T(\mathbf{I}-\widetilde{\mathbf{S}})\mathbf{W} \xrightarrow{p} \Sigma_1,
$$

where $\Sigma_1 = E\big(\Upsilon_k\Upsilon_k^T\big)$ *with* $\Upsilon_k = \Pi_W^T(u_k)Z_k - \mathbf{\Psi}^T(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_k$.

Proof of Lemma 4: Note that

$$
\begin{aligned}
\frac{1}{n}\widehat{\mathbf{W}}^T(\mathbf{I}-\widehat{\mathbf{S}})^T(\mathbf{I}-\widehat{\mathbf{S}})\widehat{\mathbf{W}} &= \frac{1}{n}(\widehat{\mathbf{W}}-\widehat{\mathbf{S}}\widehat{\mathbf{W}})^T(\widehat{\mathbf{W}}-\widehat{\mathbf{S}}\widehat{\mathbf{W}}), \\
\frac{1}{n}\widehat{\mathbf{W}}^T(\mathbf{I}-\widehat{\mathbf{S}})^T(\mathbf{I}-\widetilde{\mathbf{S}})\mathbf{W} &= \frac{1}{n}(\widehat{\mathbf{W}}-\widehat{\mathbf{S}}\widehat{\mathbf{W}})^T(\mathbf{W}-\widetilde{\mathbf{S}}\mathbf{W}),
\end{aligned}
$$

We first consider the terms $\widehat{\mathbf{S}}\widehat{\mathbf{W}}$ and $\widetilde{\mathbf{S}}\mathbf{W}$. By the similar argument in Lemma 2, we obtain that $\widehat{X}_k = Z_k^T \Pi_X(u_k) + O_p(c_n)$ and $\widehat{W}_k = Z_k^T \Pi_W(u_k) + O_p(c_n)$. Then by Lemma 3, we have

$$\begin{aligned}
\widehat{\mathbf{S}}\widehat{\mathbf{W}} &= \begin{pmatrix} \begin{pmatrix}\widehat{X}_1^T & 0\end{pmatrix} \left[\frac{1}{n}\widehat{\mathbf{D}}_X(u_1)^T\mathbf{H}(u_1)\widehat{\mathbf{D}}_X(u_1)\right]^{-1} \frac{1}{n}\widehat{\mathbf{D}}_X(u_1)^T\mathbf{H}(u_1)\widehat{\mathbf{W}} \\ \vdots \\ \begin{pmatrix}\widehat{X}_n^T & 0\end{pmatrix} \left[\frac{1}{n}\widehat{\mathbf{D}}_X(u_n)^T\mathbf{H}(u_n)\widehat{\mathbf{D}}_X(u_n)\right]^{-1} \frac{1}{n}\widehat{\mathbf{D}}_X(u_n)^T\mathbf{H}(u_n)\widehat{\mathbf{W}} \end{pmatrix} \\
&= \begin{pmatrix} Z_1^T\Pi_X(u_1)\mathbf{\Phi}^{-1}(u_1)\mathbf{\Psi}(u_1) \\ \vdots \\ Z_n^T\Pi_X(u_1)\mathbf{\Phi}^{-1}(u_n)\mathbf{\Psi}(u_n) \end{pmatrix} + O_p(c_n),
\end{aligned}$$

and

$$\begin{aligned}
\widetilde{\mathbf{S}}\mathbf{W} &= \begin{pmatrix} \begin{pmatrix}X_1^T & 0\end{pmatrix} \left[\frac{1}{n}\widehat{\mathbf{D}}_X(u_1)^T\mathbf{H}(u_1)\widehat{\mathbf{D}}_X(u_1)\right]^{-1} \frac{1}{n}\widehat{\mathbf{D}}_X(u_1)^T\mathbf{H}(u_1)\mathbf{W} \\ \vdots \\ \begin{pmatrix}X_n^T & 0\end{pmatrix} \left[\frac{1}{n}\widehat{\mathbf{D}}_X(u_n)^T\mathbf{H}(u_n)\widehat{\mathbf{D}}_X(u_n)\right]^{-1} \frac{1}{n}\widehat{\mathbf{D}}_X(u_n)^T\mathbf{H}(u_n)\mathbf{W} \end{pmatrix} \\
&= \begin{pmatrix} \left[Z_1^T\Pi_X(u_1) + v_{X,1}^T\right]\mathbf{\Phi}^{-1}(u_1)\mathbf{\Psi}(u_1) \\ \vdots \\ \left[Z_n^T\Pi_X(u_1) + v_{X,n}^T\right]\mathbf{\Phi}^{-1}(u_n)\mathbf{\Psi}(u_n) \end{pmatrix} + O_p(c_n).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\frac{1}{n}\widehat{\mathbf{W}}^T(\mathbf{I}-\widehat{\mathbf{S}})^T(\mathbf{I}-\widehat{\mathbf{S}})\widehat{\mathbf{W}} \\
&= \frac{1}{n}(\widehat{\mathbf{W}}-\widehat{\mathbf{S}}\widehat{\mathbf{W}})^T(\widehat{\mathbf{W}}-\widehat{\mathbf{S}}\widehat{\mathbf{W}}) \\
&= \frac{1}{n}\sum_{k=1}^n \left[\Pi_W^T(u_k)Z_k - \mathbf{\Psi}^T(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_k\right] \\
&\qquad \times \left[Z_k^T\Pi_W(u_k) - Z_k^T\Pi_X(u_k)\mathbf{\Phi}^{-1}(u_k)\mathbf{\Psi}(u_k)\right] + O_p(c_n) \\
&\xrightarrow{p} \Sigma_1.
\end{aligned}$$

Since $E(Z_k v_k^T \mid u_k) = 0$, where $v_k = (v_{X,k}^T, v_{W,k}^T)^T$, we also have

$$\begin{aligned}
&\frac{1}{n}\widehat{\mathbf{W}}^T(\mathbf{I}-\widehat{\mathbf{S}})^T(\mathbf{I}-\widetilde{\mathbf{S}})\mathbf{W} \\
&= \frac{1}{n}(\widehat{\mathbf{W}}-\widehat{\mathbf{S}}\widehat{\mathbf{W}})^T(\mathbf{W}-\widetilde{\mathbf{S}}\mathbf{W}) \\
&= \frac{1}{n}\sum_{k=1}^n \left[\Pi_W^T(u_k)Z_k - \mathbf{\Psi}^T(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_k\right] \\
&\qquad \times \left[Z_k^T\Pi_W(u_k) + v_{W,k}^T - Z_k^T\Pi_X(u_k)\mathbf{\Phi}^{-1}(u_k)\mathbf{\Psi}(u_k) - v_{X,k}^T\mathbf{\Phi}^{-1}(u_k)\mathbf{\Psi}(u_k)\right] + O_p(c_n) \\
&= \frac{1}{n}\sum_{k=1}^n \left[\Pi_W^T(u_k)Z_k - \mathbf{\Psi}^T(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_k\right] \\
&\qquad \times \left[Z_k^T\Pi_W(u_k) - Z_k^T\Pi_X(u_k)\mathbf{\Phi}^{-1}(u_k)\mathbf{\Psi}(u_k)\right] + O_p(c_n) \\
&\xrightarrow{p} \Sigma_1.
\end{aligned}$$

We complete the proof. $\square$

Theorem 1. *Suppose that Assumptions A1-A7 hold. When $nh^4 \to 0$, we have*

$$\sqrt{n}\left(\widetilde{\beta} - \beta\right) \xrightarrow{d} N\left(0, \Sigma_1^{-1}\Sigma_1^{*}\Sigma_1^{-1}\right),$$

where $\Sigma_1 = E\big(\Upsilon_k \Upsilon_k^T\big)$ and $\Sigma_1^* = E\big(\Upsilon_k \varepsilon_k^2 \Upsilon_k^T\big)$ with

$$\Upsilon_k = \Pi_W^T(u_k)Z_k - \mathbf{\Psi}^T(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_k.$$

Furthermore, if ε_k is conditionally homoskedastic; that is, $E(\varepsilon_k^2 \mid u_k, Z_k) = \sigma_\varepsilon^2$, then, we have

$$\sqrt{n}\left(\widetilde{\beta} - \beta\right) \xrightarrow{d} N\left(0, \sigma_\varepsilon^2 \Sigma_1^{-1}\right)$$

since $\Sigma_1^* = \Sigma_1 \sigma_\varepsilon^2$.

Proof of Theorem 1: The estimator $\widetilde{\beta}$ is given by

$$\begin{aligned}
\widetilde{\beta} &= \{\widehat{\mathbf{W}}^T(\mathbf{I} - \widehat{\mathbf{S}})^T(\mathbf{I} - \widetilde{\mathbf{S}})\mathbf{W}\}^{-1}\widehat{\mathbf{W}}^T(\mathbf{I} - \widehat{\mathbf{S}})^T(\mathbf{I} - \widetilde{\mathbf{S}})\mathbf{Y} \\
&= \{\widehat{\mathbf{W}}^T(\mathbf{I} - \widehat{\mathbf{S}})^T(\mathbf{I} - \widetilde{\mathbf{S}})\mathbf{W}\}^{-1}\widehat{\mathbf{W}}^T(\mathbf{I} - \widehat{\mathbf{S}})^T(\mathbf{I} - \widetilde{\mathbf{S}})(\mathbf{W}\beta + \mathbf{M} + \boldsymbol{\varepsilon}),
\end{aligned}$$

then

$$\begin{aligned}
\sqrt{n}\left(\widetilde{\beta} - \beta\right) &= \left\{\frac{1}{n}\widehat{\mathbf{W}}^T(\mathbf{I} - \widehat{\mathbf{S}})^T(\mathbf{I} - \widetilde{\mathbf{S}})\mathbf{W}\right\}^{-1} \qquad\qquad \text{(A.4)} \\
&\qquad \left[\frac{1}{\sqrt{n}}\widehat{\mathbf{W}}^T(\mathbf{I} - \widehat{\mathbf{S}})^T(\mathbf{I} - \widetilde{\mathbf{S}})\mathbf{M} + \frac{1}{\sqrt{n}}\widehat{\mathbf{W}}^T(\mathbf{I} - \widehat{\mathbf{S}})^T(\mathbf{I} - \widetilde{\mathbf{S}})\boldsymbol{\varepsilon}\right].
\end{aligned}$$

The first term converges to Σ_1^{-1} by Lemma 4, we now consider the second term $\frac{1}{\sqrt{n}}\widehat{\mathbf{W}}^T(\mathbf{I} - \widehat{\mathbf{S}})^T(\mathbf{I} - \widetilde{\mathbf{S}})\mathbf{M}$ and the last term $\frac{1}{\sqrt{n}}\widehat{\mathbf{W}}^T(\mathbf{I} - \widehat{\mathbf{S}})^T(\mathbf{I} - \widetilde{\mathbf{S}})\boldsymbol{\varepsilon}$.

First, note that

$$\frac{1}{\sqrt{n}}\widehat{\mathbf{W}}^T(\mathbf{I} - \widehat{\mathbf{S}})^T(\mathbf{I} - \widetilde{\mathbf{S}})\mathbf{M} = \frac{1}{\sqrt{n}}(\widehat{\mathbf{W}} - \widehat{\mathbf{S}}\widehat{\mathbf{W}})^T(\mathbf{M} - \widetilde{\mathbf{S}}\mathbf{M}).$$

Define $r(u_k) = \overline{A}(u_k) - A(u_k)$, where $\overline{A}(u_k)$ is the vector consisting of the first p rows of $\left[\widehat{\mathbf{D}}_X(u_k)^T\mathbf{H}(u_k)\widehat{\mathbf{D}}_X(u_k)\right]^{-1}\widehat{\mathbf{D}}_X(u_k)^T\mathbf{H}(u_k)\mathbf{M}$. Then

$$\begin{aligned}
\widetilde{\mathbf{S}}\mathbf{M} &= \begin{pmatrix} \begin{pmatrix} X_1^T & \mathbf{0_{1\times p}} \end{pmatrix} \left[\widehat{\mathbf{D}}_X(u_1)^T\mathbf{H}(u_1)\widehat{\mathbf{D}}_X(u_1)\right]^{-1}\widehat{\mathbf{D}}_X(u_1)^T\mathbf{H}(u_1)\mathbf{M} \\ \vdots \\ \begin{pmatrix} X_n^T & \mathbf{0_{1\times p}} \end{pmatrix} \left[\widehat{\mathbf{D}}_X(u_n)^T\mathbf{H}(u_n)\widehat{\mathbf{D}}_X(u_n)\right]^{-1}\widehat{\mathbf{D}}_X(u_n)^T\mathbf{H}(u_n)\mathbf{M} \end{pmatrix} \\
&= \begin{pmatrix} X_1^T[A(u_1) + r(u_1)] \\ \vdots \\ X_n^T[A(u_n) + r(u_n)] \end{pmatrix}.
\end{aligned}$$

By the similar argument in Lemma 4, it can be shown that $r(u_k) = O_p(c_n)$ uniformly for $u_k \in \Omega$. Note that $\sqrt{n}\, O_p(c_n^2) = o_p(1)$, we have

$$
\begin{aligned}
& \frac{1}{\sqrt{n}} \widehat{\mathbf{W}}^T (\mathbf{I} - \widehat{\mathbf{S}})^T (\mathbf{I} - \widetilde{\mathbf{S}}) \mathbf{M} \\
= \; & \frac{1}{\sqrt{n}} (\widehat{\mathbf{W}} - \widehat{\mathbf{S}} \widehat{\mathbf{W}})^T (\mathbf{M} - \widetilde{\mathbf{S}} \mathbf{M}) \\
= \; & \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \left[\Pi_W^T(u_k) Z_k - \boldsymbol{\Psi}(u_k)^T \boldsymbol{\Phi}^{-1}(u_k) \Pi_X^T(u_k) Z_k + O_p(c_n) \right] \left\{ X_k^T A(u_k) - X_k^T [A(u_k) + r(u_k)] \right\} \\
= \; & \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \left[\Pi_W^T(u_k) Z_k - \boldsymbol{\Psi}(u_k)^T \boldsymbol{\Phi}^{-1}(u_k) \Pi_X^T(u_k) Z_k + O_p(c_n) \right] X_k^T r(u_k) \\
= \; & \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \left[\Pi_W^T(u_k) Z_k - \boldsymbol{\Psi}(u_k)^T \boldsymbol{\Phi}^{-1}(u_k) \Pi_X^T(u_k) Z_k \right] X_k^T r(u_k) + o_p(1) \\
= \; & \frac{1}{\sqrt{n}} \sum_{k=1}^{n} L_k(u_k) r(u_k) + o_p(1),
\end{aligned}
$$

where

$$
L_k(u_k) = \left[\Pi_W^T(u_k) Z_k - \boldsymbol{\Psi}(u_k)^T \boldsymbol{\Phi}^{-1}(u_k) \Pi_X^T(u_k) Z_k \right] X_k^T.
$$

Let $r_{-k}(u_k)$ be the leave-one-out version of $r(u_k)$ without using the k-th observation, that is, replace all $\mathbf{H}(u_k)$'s in $r(u_k)$ by $\mathbf{H}_{-k}(u_k) = \operatorname{diag}\big(K_h(u_1 - u_k), \cdots, K_h(u_{k-1} - u_k), 0, K_h(u_{k+1} - u_k), \cdots, K_h(u_n - u_k)\big)$. It can be shown that $r(u_k) - r_{-k}(u_k) = O_p\left(\frac{1}{nh}\right)$. Since $nh^2 \to \infty$, we have

$$
\frac{1}{\sqrt{n}} \widehat{\mathbf{W}}^T (\mathbf{I} - \widehat{\mathbf{S}})^T (\mathbf{I} - \widetilde{\mathbf{S}}) \mathbf{M} = \frac{1}{\sqrt{n}} \sum_{k=1} L_k(u_k) r_{-k}(u_k) + o_p(1). \tag{A.5}
$$

Conditional on u_k, $L_k(u_k)$ and $r_{-k}(u_k)$ are independent, and $E\left[L_k(u_k) \mid u_k\right] = 0$, so

$$
E\left[\frac{1}{\sqrt{n}} \sum_{k=1}^{n} L_k(u_k) r_{-k}(u_k) \right] = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} E\left\{ E\left[L_k(u_k) \mid u_k\right] E\left[r_{-k}(u_k) \mid u_k\right] \right\} = 0. \tag{A.6}
$$

and

$$
\begin{aligned}
& \operatorname{Var}\left[\frac{1}{\sqrt{n}} \sum_{k=1}^{n} L_k(u_k) r_{-k}(u_k) \right] \\
= \; & \frac{1}{n} \sum_{k=1}^{n} \operatorname{Var}\left[L_k(u_k) r_{-k}(u_k)\right] + \frac{1}{n} \sum_{j \neq k} \operatorname{Cov}\left[L_j(u_j) r_{-j}(u_j), L_k(u_k) r_{-k}(u_k)\right] \\
= \; & \frac{1}{n} \sum_{j \neq k} E\left[L_j(u_j) r_{-j}(u_j) L_k(u_k) r_{-k}(u_k)\right] + o(1).
\end{aligned}
$$

We further let $r_{-jk}(u_k)$ be the jackknife estimator of $r(u_k)$ without using the j,k-th observations, then for $j \neq k$,

$$\begin{aligned}
& E\left[L_j(u_j)r_{-j}(u_j)L_k(u_k)r_{-k}(u_k) \mid u_j, u_k\right] \\
= \;& E\left[L_j(u_j)r_{-jk}(u_j)L_k(u_k)r_{-jk}(u_k) \mid u_j, u_k\right] \\
& \qquad\qquad +E\left\{L_j(u_j)\left[r_{-j}(u_j)-r_{-jk}(u_j)\right]L_k(u_k)r_{-jk}(u_k) \mid u_j, u_k\right\} \\
& \qquad\qquad +E\left\{L_j(u_j)r_{-jk}(u_j)L_k(u_k)\left[r_{-k}(u_k)-r_{-jk}(u_k)\right] \mid u_j, u_k\right\} \\
& \qquad\qquad +E\left\{L_j(u_j)\left[r_{-j}(u_j)-r_{-jk}(u_j)\right]L_k(u_k)\left[r_{-k}(u_k)-r_{-jk}(u_k)\right] \mid u_j, u_k\right\} \\
= \;& 0+0+0+O\left(\frac{1}{n^2h^2}\right).
\end{aligned}$$

Therefore, we have

$$\operatorname{Var}\left[\frac{1}{\sqrt{n}}\sum_{k=1}^{n} L_k(u_k)r_{-k}(u_k)\right] = \frac{1}{n}\cdot n^2 \cdot O\left(\frac{1}{n^2h^2}\right) + o(1) \to 0. \tag{A.7}$$

From (A.5), (A.6) and (A.7), we can obtain that

$$\frac{1}{\sqrt{n}}\widehat{\mathbf{W}}^T(\mathbf{I}-\widehat{\mathbf{S}})^T(\mathbf{I}-\widetilde{\mathbf{S}})\mathbf{M} = o_p(1). \tag{A.8}$$

Next, note that

$$\begin{aligned}
\widetilde{\mathbf{S}}\boldsymbol{\varepsilon} &= \begin{pmatrix} \left(X_1^T \;\; 0\right)\left[\widehat{\mathbf{D}}_X(u_1)^T\mathbf{H}(u_1)\widehat{\mathbf{D}}_X(u_1)\right]^{-1}\widehat{\mathbf{D}}_X(u_1)^T\mathbf{H}(u_1)\boldsymbol{\varepsilon} \\ \vdots \\ \left(X_n^T \;\; 0\right)\left[\widehat{\mathbf{D}}_X(u_n)^T\mathbf{H}(u_n)\widehat{\mathbf{D}}_X(u_n)\right]^{-1}\widehat{\mathbf{D}}_X(u_n)^T\mathbf{H}(u_n)\boldsymbol{\varepsilon} \end{pmatrix} \\
&= \begin{pmatrix} X_1^T O_p(c_n) \\ \vdots \\ X_n^T O_p(c_n) \end{pmatrix},
\end{aligned}$$

we have

$$\begin{aligned}
& \frac{1}{\sqrt{n}}\widehat{\mathbf{W}}^T(\mathbf{I}-\widehat{\mathbf{S}})^T(\mathbf{I}-\widetilde{\mathbf{S}})\boldsymbol{\varepsilon} \\
= \;& \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\left[\Pi_W^T(u_k)Z_k - \boldsymbol{\Psi}(u_k)^T\boldsymbol{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_k + O_p(c_n)\right]\left[\varepsilon_k - X_k^T O_p(c_n)\right] \\
= \;& \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\left[\Pi_W^T(u_k)Z_k - \boldsymbol{\Psi}^T(u_k)\boldsymbol{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_k\right]\varepsilon_k + \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\varepsilon_k O_p(c_n) \\
& \quad -\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\left[\Pi_W^T(u_k)Z_k - \boldsymbol{\Psi}^T(u_k)\boldsymbol{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_k\right]X_k O_p(c_n) + o_p(1).
\end{aligned}$$

Since $E(\varepsilon_k \mid u_k) = 0$ and $E\left\{\left[\Pi_W^T(u_k)Z_k - \mathbf{\Psi}^T(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_k\right] X_k \mid u_k\right\} = 0$, by the similar argument of obtaining (A.8), we have

$$
\begin{aligned}
& \frac{1}{\sqrt{n}}\widehat{\mathbf{W}}^T(\mathbf{I}-\widehat{\mathbf{S}})^T(\mathbf{I}-\widetilde{\mathbf{S}})\boldsymbol{\varepsilon} \qquad\qquad \text{(A.9)}\\
= & \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\left[\Pi_W^T(u_k)Z_k - \mathbf{\Psi}^T(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_k\right]\varepsilon_k + o_p(1) + o_p(1) + o_p(1)\\
\xrightarrow{d} & N\left(0, \Sigma_1^*\right).
\end{aligned}
$$

By Lemma 4 and (A.4), (A.8), (A.9), we have

$$
\sqrt{n}\left(\widetilde{\beta}-\beta\right) \xrightarrow{d} N\left(0, \Sigma_1^{-1}\Sigma_1^*\Sigma_1^{-1}\right).
$$

This completes the proof. □

Theorem 2. *Suppose that $\overline{\beta}$ is a $\sqrt{n}$-consistent estimate of β and under Assumptions A1-A7, we have*

$$
\sqrt{nh}\left[\widetilde{A}(u) - A(u) - \frac{1}{2}h^2\mu_2 A''(u)\big(1+o_p(1)\big)\right] \xrightarrow{d} N\left(0, \Sigma_2\right),
$$

where $\mu_2 = \int t^2K(t)dt$ and $\Sigma_2 = \nu_0\mathbf{\Phi}^{-1}(u)\mathbf{\Lambda}(u)\mathbf{\Phi}^{-1}(u)$ with $\nu_0 = \int K^2(t)dt$. If ε_k is conditionally homoskedastic, then, the asymptotic variance reduces to $\Sigma_2 = \nu_0\sigma_\varepsilon^2\mathbf{\Phi}^{-1}(u)$ since $\mathbf{\Lambda}(u) = \sigma_\varepsilon^2\mathbf{\Phi}(u)$. Furthermore, if $nh^5 \to 0$, we have

$$
\sqrt{nh}\left[\widetilde{A}(u) - A(u)\right] \xrightarrow{d} N\left(0, \Sigma_2\right). \qquad \text{(A.10)}
$$

Proof of Theorem 2: Note that $\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u) = \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\mathbf{D}_X(u)$, we have

$$
\begin{aligned}
& \sqrt{nh}\left[\begin{pmatrix}\widetilde{A}(u)\\ h\widetilde{A}'(u)\end{pmatrix} - \begin{pmatrix}A(u)\\ hA'(u)\end{pmatrix}\right]\\
= & \sqrt{nh}\left\{\left[\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\right]^{-1}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\left(\mathbf{Y}-\mathbf{W}\overline{\beta}\right) - \begin{pmatrix}A(u)\\ hA'(u)\end{pmatrix}\right\}\\
= & \sqrt{nh}\left[\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\right]^{-1}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\left[\mathbf{W}(\beta-\overline{\beta}) + \mathbf{M} + \boldsymbol{\varepsilon} - \widehat{\mathbf{D}}_X(u)\begin{pmatrix}A(u)\\ hA'(u)\end{pmatrix}\right]\\
= & \sqrt{nh}\left[\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\right]^{-1}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\left[\mathbf{W}(\beta-\overline{\beta}) + \mathbf{M} + \boldsymbol{\varepsilon} - \mathbf{D}_X(u)\begin{pmatrix}A(u)\\ hA'(u)\end{pmatrix}\right]\\
= & \left[\frac{1}{n}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\right]^{-1}\\
& \quad \sqrt{\frac{h}{n}}\left\{\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\mathbf{W}(\beta-\overline{\beta}) + \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\left[\mathbf{M} - \mathbf{D}_X(u)\begin{pmatrix}A(u)\\ hA'(u)\end{pmatrix}\right] + \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\boldsymbol{\varepsilon}\right\}\\
\triangleq & \left[\frac{1}{n}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\right]^{-1}(J_1 + J_2 + J_3).
\end{aligned}
$$

We first consider the term J_1. By Lemma 3 and the square root consistency of $\overline{\beta}$, we have

$$\begin{aligned} J_1 &= \sqrt{\frac{h}{n}}\,\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\mathbf{W}(\beta-\overline{\beta}) \\ &= \sqrt{h}\left[\frac{1}{n}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\mathbf{W}\right]\left[\sqrt{n}(\beta-\overline{\beta})\right] = o_p(1). \end{aligned} \tag{A.11}$$

We now move to the term J_2. By Lemma 2,

$$\begin{aligned} J_2 &= \sqrt{\frac{h}{n}}\,\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\left[\begin{pmatrix} X_1^T A(u_1) \\ \vdots \\ X_n^T A(u_n)\end{pmatrix} - \mathbf{D}_X(u)\begin{pmatrix} A(u) \\ hA'(u)\end{pmatrix}\right] \\ &= \sqrt{\frac{h}{n}}\left[\begin{pmatrix} \Pi_X^T(u) & 0 \\ 0 & \Pi_X^T(u)\end{pmatrix} + o_p(1)\right]\mathbf{D}_Z^T(u)\mathbf{H}(u) \\ &\qquad \left[\begin{pmatrix} X_1^T A(u_1) \\ \vdots \\ X_n^T A(u_n)\end{pmatrix} - \begin{pmatrix} X_1^T A(u) + X_1^T A'(u)(u_1-u) \\ \vdots \\ X_1^T A(u) + X_1^T A'(u)(u_n-u)\end{pmatrix}\right] \\ &= \sqrt{nh}\left[\begin{pmatrix} \Pi_X^T(u) & 0 \\ 0 & \Pi_X^T(u)\end{pmatrix} + o_p(1)\right]\frac{1}{n}\mathbf{D}_Z^T(u)\mathbf{H}(u)\left[\begin{pmatrix} X_1^T A''(\xi_1)(u_1-u)^2/2 \\ \vdots \\ X_1^T A''(\xi_n)(u_n-u)^2/2\end{pmatrix}\right] \\ &= \sqrt{nh}\begin{pmatrix} \frac{1}{2}h^2\mathbf{\Phi}(u)\left[\mu_2 A''(u) + o_p(1)\right] \\ o_p(h^2)\end{pmatrix}. \end{aligned} \tag{A.12}$$

Here ξ_k is a point between u_k and u.

Finally, we consider the term J_3. Note that

$$\begin{aligned} J_3 &= \sqrt{\frac{h}{n}}\,\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\boldsymbol{\varepsilon} \\ &= \left[\begin{pmatrix} \Pi_X^T(u) & 0 \\ 0 & \Pi_X^T(u)\end{pmatrix} + o_p(1)\right]\sqrt{\frac{h}{n}}\,\mathbf{D}_Z^T(u)\mathbf{H}(u)\boldsymbol{\varepsilon} \\ &= \begin{pmatrix} \Pi_X^T(u) & 0 \\ 0 & \Pi_X^T(u)\end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{n}}\sum_{k=1}^n \sqrt{h}\,Z_k\varepsilon_k K_h(u_k-u) \\ \frac{1}{\sqrt{n}}\sum_{k=1}^n \sqrt{h}\,Z_k\varepsilon_k \frac{u_k-u}{h}K_h(u_k-u)\end{pmatrix} + o_p(1) \end{aligned}$$

Here, we only focus on the variance of $\widehat{A}(u)$, then we have

$$\begin{aligned} \mathrm{Var}\left[\sqrt{h}Z_k\varepsilon_k K_h(u_k-u)\right] &= E\left[h\,Z_k\varepsilon_k^2 Z_k^T K_h^2(u_k-u)\right] \\ &= \int K^2(t)\,dt\,E\left[Z_k\varepsilon_k^2 Z_k^T \mid u_k=u\right]f(u) + o_p(1) \\ &= \nu_0 E\left[Z_k\varepsilon_k^2 Z_k^T \mid u_k=u\right]f(u) + o_p(1). \end{aligned}$$

Hence, we have

$$J_3 \xrightarrow{d} N\big(0, \nu_0\mathbf{\Lambda}(u)\big). \tag{A.13}$$

By Lemma 3, we have

$$\left[\frac{1}{n}\widehat{\mathbf{D}}_X(u)^T\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\right]^{-1} \xrightarrow{p} \begin{pmatrix} \mathbf{\Phi}^{-1}(u) & 0 \\ 0 & \mu_2^{-1}\mathbf{\Phi}^{-1}(u) \end{pmatrix},$$

together with (A.11), (A.12) and (A.13), we have

$$\sqrt{nh}\left[\widetilde{A}(u) - A(u) - \frac{1}{2}h^2\mu_2 A''(u)\big(1+o_p(1)\big)\right] \xrightarrow{d} N\left(0, \nu_0\mathbf{\Phi}^{-1}(u)\mathbf{\Lambda}(u)\mathbf{\Phi}^{-1}(u)\right).$$

This completes the proof. □

Lemma 5 *Suppose Assumptions A1-A7 hold, then under the null hypothesis,*

$$\frac{1}{n}\mathrm{RSS}_1 \xrightarrow{p} E\left\{\left[\varepsilon_k + v_{X,k}^T A_0(u_k;\theta)\right]^2 f(u_k)\right\}.$$

Proof of Lemma 5: We have

$$\begin{aligned}
\frac{1}{n}\mathrm{RSS}_1 &= \frac{1}{n}\sum_{k=1}^n \frac{1}{n}\left[\mathbf{Y} - \widehat{\mathbf{D}}_X(u_k)\begin{pmatrix} \widetilde{A}(u_k) \\ h\widetilde{A}'(u_k)\end{pmatrix} - \mathbf{W}\overline{\beta}\right]^T \\
&\qquad \mathbf{H}(u_k)\left[\mathbf{Y} - \widehat{\mathbf{D}}_X(u_k)\begin{pmatrix} \widetilde{A}(u_k) \\ h\widetilde{A}'(u_k)\end{pmatrix} - \mathbf{W}\overline{\beta}\right] \\
&= \frac{1}{n}\sum_{k=1}^n \frac{1}{n}\left[\mathbf{Y} - \widehat{\mathbf{D}}_X(u_k)\begin{pmatrix} A_0(u_k;\theta) \\ hA_0'(u_k;\theta)\end{pmatrix} - \mathbf{W}\beta\right]^T \\
&\qquad \mathbf{H}(u_k)\left[\mathbf{Y} - \widehat{\mathbf{D}}_X(u_k)\begin{pmatrix} A_0(u_k;\theta) \\ hA_0'(u_k;\theta)\end{pmatrix} - \mathbf{W}\beta\right] + o_p(1).
\end{aligned}$$

According to Lemma 2, we have

$$\begin{aligned}
\frac{1}{n}\mathrm{RSS}_1 &= \frac{1}{n}\sum_{k=1}^n \frac{1}{n}\left[\mathbf{Y} - \mathbf{D}_Z(u_k)\begin{pmatrix} \Pi(u_k)\,A_0(u_k;\theta) \\ \Pi(u_k)\,hA_0'(u_k;\theta)\end{pmatrix} - \mathbf{W}\beta\right]^T \\
&\qquad \mathbf{H}(u_k)\left[\mathbf{Y} - \mathbf{D}_Z(u_k)\begin{pmatrix} \Pi(u_k)\,A_0(u_k;\theta) \\ \Pi(u_k)\,hA_0'(u_k;\theta)\end{pmatrix} - \mathbf{W}\beta\right] + o_p(1) \\
&= \frac{1}{n}\sum_{k=1}^n \frac{1}{n}\begin{pmatrix} \varepsilon_1 + v_{X,1}^T A_0(u_1,\theta) \\ \vdots \\ \varepsilon_n + v_{X,n}^T A_0(u_n,\theta)\end{pmatrix}^T \mathbf{H}(u_k)\begin{pmatrix} \varepsilon_1 + v_{X,1}^T A_0(u_1,\theta) \\ \vdots \\ \varepsilon_n + v_{X,n}^T A_0(u_n,\theta)\end{pmatrix} + o_p(1) \\
&= \frac{1}{n}\sum_{k=1}^n E\left\{\left[\varepsilon_k + v_{X,k}^T A_0(u_k;\theta)\right]^2 \mid u_k\right\} f(u_k) + o_p(1) \\
&= E\left\{\left[\varepsilon_k + v_{X,k}^T A_0(u_k;\theta)\right]^2 f(u_k)\right\} + o_p(1).
\end{aligned}$$

The proof is complete. □

Lemma 6 *Let $\overline{\theta}$ and $\overline{\beta}$ be $\sqrt{n}$-consistent estimates of θ and β under the null hypothesis, respectively. Define*

$$S_0(u) = \left[\mathbf{Y} - \widehat{\mathbf{D}}_X(u)\begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix} - \mathbf{W}\overline{\beta}\right]^T \mathbf{H}(u)\left[\mathbf{Y} - \widehat{\mathbf{D}}_X(u)\begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix} - \mathbf{W}\overline{\beta}\right],$$

$$S_1(u) = \left[\mathbf{Y} - \widehat{\mathbf{D}}_X(u)\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} - \mathbf{W}\overline{\beta}\right]^T \mathbf{H}(u)\left[\mathbf{Y} - \widehat{\mathbf{D}}_X(u)\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} - \mathbf{W}\overline{\beta}\right],$$

where

$$\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} = \left[\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\right]^{-1}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\left(\mathbf{Y} - \mathbf{W}\overline{\beta}\right).$$

Suppose that Assumptions A1-A7 hold and $nh^4 \to 0$, then under the null hypothesis,

$$S_0(u) - S_1(u) = P(u) + T_1(u) + T_2(u) + o_p(1/\sqrt{h}),$$

where

$$\begin{aligned} P(u) &= \frac{1}{n}\boldsymbol{\varepsilon}^T\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\left[\frac{1}{n}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\right]^{-1}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\boldsymbol{\varepsilon}, \\ T_1(u) &= \frac{1}{n}(\beta - \overline{\beta})^T\mathbf{W}^T\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\left[\frac{1}{n}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\right]^{-1}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\boldsymbol{\varepsilon}, \\ T_2(u) &= \frac{1}{n}(\theta - \overline{\theta})^T\frac{\partial A_0^T(u;\theta)}{\partial\theta}\mathbf{X}^T\mathbf{H}(u)\widehat{\mathbf{D}}_X(u) \\ &\qquad \left[\frac{1}{n}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\right]^{-1}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\boldsymbol{\varepsilon}. \end{aligned}$$

Proof of Lemma 6: Note that we have

$$\begin{aligned} & S_0(u) \\ = & \left\{\mathbf{Y} - \widehat{\mathbf{D}}_X(u)\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} - \mathbf{W}\overline{\beta} + \widehat{\mathbf{D}}_X(u)\left[\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} - \begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix}\right]\right\}^T \\ & \mathbf{H}(u)\left\{\mathbf{Y} - \widehat{\mathbf{D}}_X(u)\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} - \mathbf{W}\overline{\beta} + \widehat{\mathbf{D}}_X(u)\left[\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} - \begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix}\right]\right\} \\ = & S_1(u) + 2\left[\mathbf{Y} - \widehat{\mathbf{D}}_X(u)\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} - \mathbf{W}\overline{\beta}\right]^T\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\left[\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} - \begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix}\right] \\ & + \left[\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} - \begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix}\right]^T\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\left[\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} - \begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix}\right]. \end{aligned}$$

Because

$$\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} = \left[\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\right]^{-1}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\left(\mathbf{Y} - \mathbf{W}\overline{\beta}\right),$$

so the cross term

$$\begin{aligned}
& \left[\mathbf{Y} - \widehat{\mathbf{D}}_X(u)\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} - \mathbf{W}\overline{\beta}\right]^T \mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\left[\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} - \begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix}\right] \\
= \;& \left(\mathbf{Y} - \mathbf{W}\overline{\beta}\right)^T \left\{\mathbf{I}_n - \mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\left[\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\right]^{-1}\widehat{\mathbf{D}}_X^T(u)\right\} \\
& \qquad \mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\left[\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} - \begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix}\right] \\
= \;& \left(\mathbf{Y} - \mathbf{W}\overline{\beta}\right)^T \left[\mathbf{H}(u)\widehat{\mathbf{D}}_X(u) - \mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\right]\left[\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} - \begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix}\right] \\
= \;& 0.
\end{aligned}$$

Furthermore, because $\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u) = \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\mathbf{D}_X(u)$, we have

$$\begin{aligned}
& S_0(u) - S_1(u) \\
= \;& \left[\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} - \begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix}\right]^T \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\left[\begin{pmatrix} \widetilde{A}(u) \\ h\widetilde{A}'(u) \end{pmatrix} - \begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix}\right] \\
= \;& \left[\mathbf{Y} - \mathbf{W}\overline{\beta} - \widehat{\mathbf{D}}_X(u)\begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix}\right]^T \mathbf{H}(u)\widehat{\mathbf{D}}_X(u) \\
& \qquad \left[\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\right]^{-1}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\left[\mathbf{Y} - \mathbf{W}\overline{\beta} - \widehat{\mathbf{D}}_X(u)\begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix}\right] \\
= \;& \frac{1}{\sqrt{n}}\left[\mathbf{Y} - \mathbf{W}\overline{\beta} - \mathbf{D}_X(u)\begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix}\right]^T \mathbf{H}(u)\widehat{\mathbf{D}}_X(u) \\
& \qquad \left[\frac{1}{n}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\widehat{\mathbf{D}}_X(u)\right]^{-1}\frac{1}{\sqrt{n}}\widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\left[\mathbf{Y} - \mathbf{W}\overline{\beta} - \mathbf{D}_X(u)\begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix}\right].
\end{aligned}$$

Note that $\overline{\beta}$ and $\overline{\theta}$ are $\sqrt{n}$-consistent estimates. When $nh^4 \to 0$, we have

$$\begin{aligned}
& \frac{1}{\sqrt{n}} \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u) \left[\mathbf{Y} - \mathbf{W}\overline{\beta} - \mathbf{D}_X(u) \begin{pmatrix} A_0(u;\overline{\theta}) \\ hA_0'(u;\overline{\theta}) \end{pmatrix} \right] \\
= & \frac{1}{\sqrt{n}} \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u) \left[\mathbf{Y} - \mathbf{W}\beta - \begin{pmatrix} X_1^T A_0(u_1;\theta) \\ \vdots \\ X_n^T A_0(u_n;\theta) \end{pmatrix} \right] + \frac{1}{\sqrt{n}} \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\mathbf{W}(\beta - \overline{\beta}) \\
& + \frac{1}{\sqrt{n}} \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u) \begin{pmatrix} X_1^T \left[A_0(u_1;\theta) - A_0(u;\overline{\theta}) - A_0'(u;\overline{\theta})(u_1 - u) \right] \\ \vdots \\ X_n^T \left[A_0(u_n;\theta) - A_0(u;\overline{\theta}) - A_0'(u;\overline{\theta})(u_n - u) \right] \end{pmatrix} \\
= & \frac{1}{\sqrt{n}} \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\boldsymbol{\varepsilon} + \frac{1}{\sqrt{n}} \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\mathbf{W}(\beta - \overline{\beta}) \\
& + \frac{1}{\sqrt{n}} \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\mathbf{X}\frac{\partial A_0(u;\theta)}{\partial\theta}(\theta - \overline{\theta}) + O_p\left[\sqrt{n} \left(\frac{1}{n} + \frac{1}{\sqrt{n}} h + h^2 \right) \right] \\
= & \frac{1}{\sqrt{n}} \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\boldsymbol{\varepsilon} + \frac{1}{\sqrt{n}} \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\mathbf{W}(\beta - \overline{\beta}) \\
& + \frac{1}{\sqrt{n}} \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\mathbf{X}\frac{\partial A_0(u;\theta)}{\partial\theta}(\theta - \overline{\theta}) + o_p(1).
\end{aligned}$$

Since $\frac{1}{\sqrt{n}} \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\mathbf{W}(\beta - \overline{\beta}) = O_p(1)$ and $\frac{1}{\sqrt{n}} \widehat{\mathbf{D}}_X^T(u)\mathbf{H}(u)\mathbf{X}\frac{\partial A_0(u;\theta)}{\partial\theta}(\theta - \overline{\theta}) = O_p(1)$, then

$$\begin{aligned}
S_0(u) - S_1(u) &= P(u) + T_1(u) + T_2(u) + O_p(1) \\
&= P(u) + T_1(u) + T_2(u) + o_p(1/\sqrt{h}).
\end{aligned}$$

This completes the proof.

Theorem 3. *Suppose that Assumptions A1-A7 hold, when $nh^4 \to 0$, under the null hypothesis, we have*

$$\sigma_n^{-1} \{\lambda_n - \mu_n\} \xrightarrow{d} N(0,1),$$

where

$$\mu_n = \frac{(\nu_0 + \mu_2^{-1}\nu_2)\, E\Big\{ \mathrm{Trace}\left\{ \boldsymbol{\Phi}^{-1}(u_k)\boldsymbol{\Lambda}(u_k) \right\} \Big\}}{h\, E\left\{ \left[\varepsilon_k + v_{X,k}^T A_0(u_k;\theta) \right]^2 f(u_k) \right\}}$$

$$\text{and} \quad \sigma_n^2 = \frac{2\int g^2(t)\,dt\, E\Big\{ \mathrm{Trace}\left\{ \boldsymbol{\Phi}^{-1}(u_k)\boldsymbol{\Lambda}(u_k)\boldsymbol{\Phi}^{-1}(u_k)\boldsymbol{\Lambda}(u_k) f(u_k) \right\} \Big\}}{h\, E^2\left\{ \left[\varepsilon_k + v_{X,k}^T A_0(u_k;\theta) \right]^2 f(u_k) \right\}}$$

with $\nu_0 = \int K^2(t)\,dt$, $\nu_2 = \int t^2 K^2(t)\,dt$, $\mu_2 = \int t^2 K(t)\,dt$ and $g(t) = \int K(s)K(t+s)\,ds + \mu_2^{-1}\int s(t+s)K(s)K(t+s)\,ds$. Furthermore, if ε_k is conditionally homoskedastic, then,

$\mathbf{\Phi}^{-1}(u_k)\mathbf{\Lambda}(u_k) = \sigma_\varepsilon^2 \mathbf{I}_p$ *so that*

$$\mu_n = \frac{(\nu_0 + \mu_2^{-1}\nu_2)\, p\, \sigma_\varepsilon^2}{h\, E\left\{\left[\varepsilon_k + v_{X,k}^T A_0(u_k;\theta)\right]^2 f(u_k)\right\}}$$

$$\textit{and} \quad \sigma_n^2 = \frac{2\int g^2(t)\, dt\, p\, \sigma_\varepsilon^4\, E\left[f(u_k)\right]}{h\, E^2\left\{\left[\varepsilon_k + v_{X,k}^T A_0(u_k;\theta)\right]^2 f(u_k)\right\}}.$$

Proof of Theorem 3: Note that

$$\begin{aligned} \mathrm{RSS}_0 - \mathrm{RSS}_1 &= \frac{1}{n}\sum_{k=1}^{n}\left[S_0(u_k) - S_1(u_k)\right] \\ &= \mathbf{P} + \mathbf{T}_1 + \mathbf{T}_2 + o_p(1/\sqrt{h}), \end{aligned}$$

where $\mathbf{P} = \frac{1}{n}\sum_{k=1}^n P(u_k)$, $\mathbf{T}_1 = \frac{1}{n}\sum_{k=1}^n T_1(u_k)$ and $\mathbf{T}_2 = \frac{1}{n}\sum_{k=1}^n T_2(u_k)$.

By Lemma 3, we have

$$\begin{aligned}
\mathbf{T}_1 &= \frac{1}{\sqrt{h}}\sqrt{n}(\beta_0 - \overline{\beta})^T \frac{1}{n}\sum_{k=1}^{n}\left[\frac{1}{n}\mathbf{W}^T\mathbf{H}(u_k)\widehat{\mathbf{D}}_X(u_k)\right] \\
&\qquad \left[\frac{1}{n}\widehat{\mathbf{D}}_X^T(u_k)\mathbf{H}(u_k)\widehat{\mathbf{D}}_X(u_k)\right]^{-1}\sqrt{\frac{h}{n}}\,\widehat{\mathbf{D}}_X^T(u_k)\mathbf{H}(u_k)\boldsymbol{\varepsilon} \\
&= \frac{1}{\sqrt{h}}\sqrt{n}(\beta_0 - \overline{\beta})^T \frac{1}{n}\sum_{k=1}^{n}\left(\mathbf{\Psi}^T(u_k),\, 0\right) \\
&\qquad \begin{pmatrix} \mathbf{\Phi}^{-1}(u_k) & 0 \\ 0 & \mu_2^{-1}\mathbf{\Phi}^{-1}(u_k)\end{pmatrix}\sqrt{\frac{h}{n}}\begin{pmatrix} \Pi_X^T(u_k) & 0 \\ 0 & \Pi_X^T(u)\end{pmatrix}\mathbf{D}_Z^T(u_k)\mathbf{H}(u_k)\boldsymbol{\varepsilon} + o_p(1/\sqrt{h}) \\
&= \frac{1}{n}(\beta_0 - \overline{\beta})^T\sum_{k=1}^{n}\left(\mathbf{\Psi}^T(u_k),\, 0\right) \\
&\qquad \begin{pmatrix} \mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k) & 0 \\ 0 & \mu_2^{-1}\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u)\end{pmatrix}\mathbf{D}_Z^T(u_k)\mathbf{H}(u_k)\boldsymbol{\varepsilon} + o_p(1/\sqrt{h}) \\
&= (\beta_0 - \overline{\beta})^T\frac{1}{n}\sum_{k=1}^{n}\left[\mathbf{\Psi}^T(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k),\, 0\right]\,\mathbf{D}_Z^T(u_k)\mathbf{H}(u_k)\boldsymbol{\varepsilon} + o_p(1/\sqrt{h}) \\
&= (\beta_0 - \overline{\beta})^T\frac{1}{n}\sum_{k=1}^{n}\mathbf{\Psi}^T(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k)\sum_{i=1}^{n}Z_i\varepsilon_i K_h(u_i - u_k) + o_p(1/\sqrt{h}) \\
&= \sqrt{n}\,(\beta_0 - \overline{\beta})^T \\
&\qquad \frac{1}{\sqrt{n}}\sum_{i=1}^{n}Z_i\varepsilon_i\left[\frac{1}{n}\sum_{k=1}^{n}\mathbf{\Psi}^T(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k)K_h(u_k - u_i)\right] + o_p(1/\sqrt{h}) \\
&= O_p(1) + o_p(1/\sqrt{h}) = o_p(1/\sqrt{h}).
\end{aligned}$$

Similarly, we can obtain that

$$\mathbf{T}_2 = o_p(1/\sqrt{h}).$$

Next, we focus on $\mathbf{P}$. We have

$$
\begin{aligned}
\mathbf{P} &= \frac{1}{nh}\sum_{k=1}^{n}\sqrt{\frac{h}{n}}\,\boldsymbol{\varepsilon}^T\mathbf{H}(u_k)\widehat{\mathbf{D}}_X(u_k)\left[\frac{1}{n}\,\widehat{\mathbf{D}}_X^T(u_k)\mathbf{H}(u_k)\widehat{\mathbf{D}}_X(u_k)\right]^{-1}\sqrt{\frac{h}{n}}\,\widehat{\mathbf{D}}_X^T(u_k)\mathbf{H}(u_k)\boldsymbol{\varepsilon} \\
&= \frac{1}{nh}\sum_{k=1}^{n}\sqrt{\frac{h}{n}}\,\boldsymbol{\varepsilon}^T\mathbf{H}(u_k)\mathbf{D}_Z(u_k) \\
&\quad \begin{pmatrix} \Pi_X(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k) & 0 \\ 0 & \mu_2^{-1}\Pi_X(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k) \end{pmatrix}\sqrt{\frac{h}{n}}\,\mathbf{D}_Z^T(u)\mathbf{H}(u)\boldsymbol{\varepsilon} + O_p(c_n/h) \\
&= \frac{1}{n^2}\sum_{k=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\Big[K_h(u_i-u_k)K_h(u_j-u_k)\varepsilon_i\varepsilon_j Z_i^T\Pi_X(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_j \\
&\quad +\mu_2^{-1}K_h(u_i-u_k)\,\frac{u_i-u_k}{h}\,K_h(u_j-u_k)\,\frac{u_j-u_k}{h}\,\varepsilon_i\varepsilon_j Z_i^T\Pi_X(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_j\Big] + o_p(1/\sqrt{h}) \\
&= \frac{1}{n^2}\sum_{k=1}^{n}\sum_{i=1}^{n}\Big[K_h^2(u_i-u_k)\varepsilon_i^2 Z_i^T\Pi_X(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_i \\
&\qquad +\mu_2^{-1}K_h^2(u_i-u_k)\left(\frac{u_i-u_k}{h}\right)^2\varepsilon_i^2 Z_i^T\Pi_X(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_i\Big] \\
&+ \frac{2}{n^2}\sum_{k=1}^{n}\sum_{i<j}\Big[K_h(u_i-u_k)K_h(u_j-u_k)\varepsilon_i\varepsilon_j Z_i^T\Pi_X(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_j \\
&\quad +\mu_2^{-1}K_h(u_i-u_k)\,\frac{u_i-u_k}{h}\,K_h(u_j-u_k)\,\frac{u_j-u_k}{h}\,\varepsilon_i\varepsilon_j Z_i^T\Pi_X(u_k)\mathbf{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_j\Big] + o_p(1/\sqrt{h}) \\
&= \mathbf{Q}_1 + \mathbf{Q}_2 + o_p(1/\sqrt{h}).
\end{aligned}
$$

Define $\nu_0 = \int K^2(t)\,dt$ and $\nu_2 = \int t^2 K^2(t)\,dt$, then

$$\begin{aligned}
\mathbf{Q}_1 &= \frac{1}{n^2}\sum_{k=1}^{n}\sum_{i=1}^{n}\Big[K_h^2(u_i-u_k)\varepsilon_i^2 Z_i^T \Pi_X(u_k)\boldsymbol{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_i \\
&\qquad +\mu_2^{-1}K_h^2(u_i-u_k)\left(\frac{u_i-u_k}{h}\right)^2 \varepsilon_i^2 Z_i^T \Pi_X(u_k)\boldsymbol{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_i\Big] \\
&= \frac{1}{nh}\sum_{k=1}^{n}\frac{1}{n}\sum_{i\neq k}\Big\{\frac{1}{h}K^2[(u_i-u_k)/h]\varepsilon_i^2 Z_i^T \Pi_X(u_k)\boldsymbol{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_i \\
&\qquad +\mu_2^{-1}\frac{1}{h}K^2[(u_i-u_k)/h]\left(\frac{u_i-u_k}{h}\right)^2 \varepsilon_i^2 Z_i^T \Pi_X(u_k)\boldsymbol{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_i\Big\} + O_p\left(\frac{1}{nh^2}\right) \\
&= \frac{1}{nh}\sum_{k=1}^{n}\Big\{\nu_0 E\left[\varepsilon_i^2 Z_i^T \Pi_X(u_k)\boldsymbol{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_i \mid u_i=u_k\right] f(u_k) \\
&\qquad +\mu_2^{-1}\nu_2 E\left[\varepsilon_i^2 Z_i^T \Pi_X(u_k)\boldsymbol{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_i \mid u_i=u_k\right] f(u_k)\Big\} + o_p(1/\sqrt{h}) \\
&= \frac{\nu_0+\mu_2^{-1}\nu_2}{nh}\sum_{k=1}^{n}\mathrm{Trace}\left\{\boldsymbol{\Phi}^{-1}(u_k)\Pi_X^T(u_k)E\left[Z_i\varepsilon_i^2 Z_i^T \mid u_i=u_k\right]\Pi_X(u_k)f(u_k)\right\} + o_p(1/\sqrt{h}) \\
&= \frac{\nu_0+\mu_2^{-1}\nu_2}{h}E\Big\{\mathrm{Trace}\left\{\boldsymbol{\Phi}^{-1}(u_k)\boldsymbol{\Lambda}(u_k)\right\}\Big\} + o_p(1/\sqrt{h}).
\end{aligned} \tag{A.14}$$

Now we consider $\mathbf{Q}_2$. We have $E(\mathbf{Q}_2)=0$ and

$$\begin{aligned}
\mathbf{Q}_2 &= \frac{2}{n^2}\sum_{i<j}\sum_{k\neq i,j}\Big[K_h(u_k-u_i)K_h(u_k-u_j)\varepsilon_i\varepsilon_j Z_i^T \Pi_X(u_k)\boldsymbol{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_j \\
&\quad +\mu_2^{-1}K_h(u_k-u_i)\frac{u_k-u_i}{h}K_h(u_k-u_j)\frac{u_k-u_j}{h}\varepsilon_i\varepsilon_j Z_i^T \Pi_X(u_k)\boldsymbol{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_j\Big] + O_p\left(\frac{1}{nh^2}\right) \\
&= \frac{2}{nh}\sum_{i<j}\frac{h}{n}\sum_{k\neq i,j}\Big[K_h(u_k-u_i)K_h(u_k-u_j)\varepsilon_i\varepsilon_j Z_i^T \Pi_X(u_k)\boldsymbol{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_j \\
&\quad +\mu_2^{-1}K_h(u_k-u_i)\frac{u_k-u_i}{h}K_h(u_k-u_j)\frac{u_k-u_j}{h}\varepsilon_i\varepsilon_j Z_i^T \Pi_X(u_k)\boldsymbol{\Phi}^{-1}(u_k)\Pi_X^T(u_k)Z_j\Big] + o_p\left(1/\sqrt{h}\right) \\
&= \frac{2}{nh}\sum_{i<j}\Big\{\int K(t)K\left(\frac{u_i-u_j}{h}+t\right)dt\, f(u_i)\,\varepsilon_i\varepsilon_j Z_i^T \Pi_X(u_i)\boldsymbol{\Phi}^{-1}(u_i)\Pi_X^T(u_i)Z_j[1+O_p(c_n)] \\
&+\mu_2^{-1}\int t\left(\frac{u_i-u_j}{h}+t\right)K(t)K\left(\frac{u_i-u_j}{h}+t\right)dt\, f(u_i)\,\varepsilon_i\varepsilon_j Z_i^T \Pi_X(u_i)\boldsymbol{\Phi}^{-1}(u_i)\Pi_X^T(u_i)Z_j[1+O_p(c_n)]\Big\} \\
&\qquad + o_p(1/\sqrt{h}) \\
&= \frac{2}{nh}\sum_{i<j} g\left(\frac{u_i-u_j}{h}\right) f(u_i)\,\varepsilon_i\varepsilon_j Z_i^T \Pi_X(u_i)\boldsymbol{\Phi}^{-1}(u_i)\Pi_X^T(u_i)Z_j[1+O_p(c_n)] + o_p(1/\sqrt{h}),
\end{aligned}$$

where

$$g(s) = \int K(t)K(s+t)\,dt + \mu_2^{-1}\int t(s+t)K(t)K(s+t)\,dt.$$

Consider the variance of $\mathbf{Q}_2$, then

$$
\begin{aligned}
& \operatorname{Var}\left[\frac{2}{nh}\sum_{i<j} g\left(\frac{u_i-u_j}{h}\right) f(u_i)\,\varepsilon_i\varepsilon_j Z_i^T \Pi_X(u_i)\mathbf{\Phi}^{-1}(u_i)\Pi_X^T(u_i)Z_j\right] \\
= \;& \frac{2(n-1)}{nh^2} E\left[g^2\left(\frac{u_i-u_j}{h}\right) f^2(u_i)\,\varepsilon_i^2\varepsilon_j^2\left[Z_i^T \Pi_X(u_i)\mathbf{\Phi}^{-1}(u_i)\Pi_X^T(u_i)Z_j\right]^2\right] \\
= \;& \frac{2(n-1)}{nh^2} E\left[g^2\left(\frac{u_i-u_j}{h}\right) f^2(u_i)\,\varepsilon_i^2\varepsilon_j^2 Z_j^T \Pi_X(u_i)\mathbf{\Phi}^{-1}(u_i)\Pi_X^T(u_i)Z_i Z_i^T \Pi_X(u_i)\mathbf{\Phi}^{-1}(u_i)\Pi_X^T(u_i)Z_j\right] \\
= \;& \frac{2(n-1)}{nh^2} E\Big[g^2\left(\frac{u_i-u_j}{h}\right) \varepsilon_j^2 Z_j^T \Pi_X(u_i)\mathbf{\Phi}^{-1}(u_i)\Pi_X^T(u_i) \\
& \qquad\qquad E\left(Z_i\varepsilon_i^2 Z_i^T \mid u_i,u_j,\varepsilon_j,Z_j\right)\Pi_X(u_i)f^2(u_i)\,\mathbf{\Phi}^{-1}(u_i)\Pi_X^T(u_i)Z_j\Big] \\
= \;& \frac{2(n-1)}{nh^2} E\Big[g^2\left(\frac{u_i-u_j}{h}\right) \varepsilon_j^2 Z_j^T \Pi_X(u_i)\mathbf{\Phi}^{-1}(u_i)\mathbf{\Lambda}(u_i)\mathbf{\Phi}^{-1}(u_i)\Pi_X^T(u_i)Z_j f(u_i)\Big] \\
= \;& \frac{2(n-1)}{nh^2} h\left\{\int g^2(t)dt\, E\Big[\varepsilon_j^2 Z_j^T \Pi_X(u_j)\mathbf{\Phi}^{-1}(u_j)\mathbf{\Lambda}(u_j)\mathbf{\Phi}^{-1}(u_j)\Pi_X^T(u_j)Z_j f^2(u_j)\Big] + O_p(h)\right\} \\
= \;& \frac{2\int g^2(t)dt}{h} E\Big\{\text{Trace}\left\{\mathbf{\Phi}^{-1}(u_j)\mathbf{\Lambda}(u_j)\mathbf{\Phi}^{-1}(u_j)\Pi_X^T(u_j)Z_j\varepsilon_j^2 Z_j^T\Pi_X(u_j)f^2(u_j)\right\}\Big\} + O_p(1) \\
= \;& \frac{2\int g^2(t)dt}{h} E\Big\{\text{Trace}\left\{\mathbf{\Phi}^{-1}(u_j)\mathbf{\Lambda}(u_j)\mathbf{\Phi}^{-1}(u_j)\mathbf{\Lambda}(u_j)f(u_j)\right\}\Big\} + O_p(1).
\end{aligned}
$$

By the argument similar to Theorem 5 in Fan et al. (2001), we have

$$
\frac{\mathbf{Q}_2}{\sqrt{\frac{2\int g^2(t)dt}{h} E\Big\{\text{Trace}\left\{\mathbf{\Phi}^{-1}(u_j)\mathbf{\Lambda}(u_j)\mathbf{\Phi}^{-1}(u_j)\mathbf{\Lambda}(u_j)f(u_j)\right\}\Big\}}} \xrightarrow{d} N(0,1). \tag{A.15}
$$

Finally, we have

$$
\begin{aligned}
\lambda_n &= \frac{\text{RSS}_0 - \text{RSS}_1}{\text{RSS}_1/n} \\
&= \frac{\mathbf{Q}_1+\mathbf{Q}_2+\mathbf{T}_1+\mathbf{T}_2}{\text{RSS}_1/n} + o_p(1/\sqrt{h}) \\
&= \frac{\mathbf{Q}_1+\mathbf{Q}_2}{\text{RSS}_1/n} + o_p(1/\sqrt{h}).
\end{aligned} \tag{A.16}
$$

By Lemma 5 and (A.14), (A.15), (A.16), we complete the proof. □

References

Fan, J., Zhang, C., and Zhang, J. (2001), "Generalized likelihood ratio statistics and Wilks phenomenon," *Annals of Statistics*, 29, 153–193.

Mack, Y. P. and Silverman, B. W. (1982), "Weak and strong uniform consistency of kernel regression estimates," *Probability Theory and Related Fields*, 61, 405–415.