Online Supplement for 'Inference and Specification Testing in Threshold Regression with Endogeneity'

by Ping Yu[42], Qin Liao[43] and Peter C. B. Phillips[44]

We first collect notation for future reference. The $n \times 1$ vectors Y and $\boldsymbol{\varepsilon}$ stack the variables y_i and ε_i, the $n \times \overline{d}$ matrices X, $X_{\leq\gamma}$ and $X_{>\gamma}$ stack the vectors $\mathbf{x}_i'$, $\mathbf{x}_i' 1(q_i \leq \gamma)$ and $\mathbf{x}_i' 1(q_i > \gamma)$, and the $n \times d_z$ matrices Z, $Z_{\leq\gamma}$ and $Z_{>\gamma}$ are similarly defined. The symbol $\approx$ means asymptotic equivalence in the sense that higher order terms are neglected, $=_d$ signifies equality in distribution, C means a positive constant that may change at each occurrence, and $\rightsquigarrow$ signifies weak convergence of the respective probability measures over an associated compact metric space.

To aid intuition in the development of Methods II and III, we let $\Delta(x, q) = \mathbf{x}'\delta$ throughout the proofs for these two methods.

Supplement A: Proofs

Proof of Theorem 1. We first show the consistency of $\widehat{\theta}$. If $\widehat{\gamma}$ is consistent, then a standard argument can be applied to show that

$$\left(\widehat{\beta}, \widehat{\delta}\right) = \left(\widehat{\beta}(\widehat{\gamma}), \widehat{\delta}(\widehat{\gamma})\right)$$

with

$$\begin{pmatrix} \widehat{\beta}(\gamma) \\ \widehat{\delta}(\gamma) \end{pmatrix} = \left[\left(\mathbf{X}_{\leq\gamma}' Z\right) \widehat{W} \left(Z'\mathbf{X}_{\leq\gamma}\right)\right]^{-1} \left[\left(\mathbf{X}_{\leq\gamma}' Z\right) \widehat{W} \left(Z'Y\right)\right]$$

is consistent, where $\mathbf{X}_{\leq\gamma} = (X, X_{\leq\gamma})$. So we now concentrate on the consistency of $\widehat{\gamma}$. First note that the concentrated objective function of (12) after plugging in $\left(\widehat{\beta}(\gamma), \widehat{\delta}(\gamma)\right)$ is

$$\widehat{Q}_n(\gamma) = \widehat{g}_n(\gamma)' \widehat{W} \widehat{g}_n(\gamma).$$

Here,

$$\begin{aligned}
\widehat{g}_n(\gamma) &= \tfrac{1}{n} Z' \left(Y - \mathbf{X}_{\leq\gamma}\left[\left(\tfrac{1}{n}\mathbf{X}_{\leq\gamma}' Z\right) \widehat{W} \left(\tfrac{1}{n} Z'\mathbf{X}_{\leq\gamma}\right)\right]^{-1} \left[\left(\tfrac{1}{n}\mathbf{X}_{\leq\gamma}' Z\right) \widehat{W} \left(\tfrac{1}{n} Z'Y\right)\right]\right) \\
&= \left\{I_l - \left(\tfrac{1}{n} Z'\mathbf{X}_{\leq\gamma}\right)\left[\left(\tfrac{1}{n}\mathbf{X}_{\leq\gamma}' Z\right) \widehat{W} \left(\tfrac{1}{n} Z'\mathbf{X}_{\leq\gamma}\right)\right]^{-1} \left(\tfrac{1}{n}\mathbf{X}_{\leq\gamma}' Z\right) \widehat{W}\right\} \left(\tfrac{1}{n} Z'\mathbf{X}_{\leq\gamma_0}\theta_0 + \tfrac{1}{n} Z'\boldsymbol{\varepsilon}\right) \\
&= \left\{I_l - \left(\tfrac{1}{n} Z'\mathbf{X}_{\leq\gamma}\right)\left[\left(\tfrac{1}{n}\mathbf{X}_{\leq\gamma}' Z\right) \widehat{W} \left(\tfrac{1}{n} Z'\mathbf{X}_{\leq\gamma}\right)\right]^{-1} \left(\tfrac{1}{n}\mathbf{X}_{\leq\gamma}' Z\right) \widehat{W}\right\} \left(\tfrac{1}{n} Z' X_{\leq\gamma_0}\delta_n + \tfrac{1}{n} Z'\boldsymbol{\varepsilon}\right),
\end{aligned}$$

and the second equality holds because the first $\overline{d}$ columns of $Z'\mathbf{X}_{\leq\gamma_0}$ are the same as those of $Z'\mathbf{X}_{\leq\gamma}$.

We apply Theorem 2.1 of Newey and McFadden (1994) to prove the result. First, we can show

$$\sup_{\gamma\in\Gamma} \left\| \tfrac{1}{\|\delta_n\|} \widehat{g}_n(\gamma) - g_0(\gamma) \right\| \xrightarrow{p} 0,$$

[42] Faculty of Business and Economics, The University of Hong Kong, Pokfulam Road, Hong Kong; corresponding author email: pingyu@hku.hk.

[43] Faculty of Business and Economics, The University of Hong Kong, Pokfulam Road, Hong Kong; email: liaoq@connect.hku.hk.

[44] Cowles Foundation for Research in Economics, Yale University, POBox 208281, New Haven, CT, USA; email: peter.phillips@yale.edu

where

$$g_0(\gamma) = \left(I - \underline{G}_\gamma \left(\underline{G}'_\gamma W \underline{G}_\gamma\right)^{-1} \underline{G}'_\gamma W\right) \underline{G}_{2,\gamma_0} c.$$

To see why, note that by a Glivenko-Cantelli argument,

$$I_l - \left(\tfrac{1}{n} Z'\mathbf{X}_{\le\gamma}\right) \left[\left(\tfrac{1}{n}\mathbf{X}'_{\le\gamma} Z\right) \widehat{W} \left(\tfrac{1}{n} Z'\mathbf{X}_{\le\gamma}\right)\right]^{-1} \left(\tfrac{1}{n}\mathbf{X}'_{\le\gamma} Z\right) \widehat{W} \xrightarrow{p} I_l - \underline{G}_\gamma \left(\underline{G}'_\gamma W \underline{G}_\gamma\right)^{-1} \underline{G}'_\gamma W$$

uniformly in γ. Also, $\frac{1}{n} Z' X_{\le\gamma_0} \frac{\delta_n}{\|\delta_n\|} \xrightarrow{p} \underline{G}_{2,\gamma_0} c$ and $\frac{1}{\|\delta_n\|} \frac{1}{n} Z'\boldsymbol{\varepsilon} = O_p\left(\frac{1}{\sqrt{n}\|\delta_n\|}\right) = o_p(1)$, the result follows. Second, by the CMT,

$$\begin{aligned}
\widehat{Q}_n(\gamma) / \|\delta_n\|^2 &\xrightarrow{p} c'\underline{G}'_{2,\gamma_0} \left(I_l - W\underline{G}_\gamma \left(\underline{G}'_\gamma W \underline{G}_\gamma\right)^{-1} \underline{G}'_\gamma\right) W \left(I_l - \underline{G}_\gamma \left(\underline{G}'_\gamma W \underline{G}_\gamma\right)^{-1} \underline{G}'_\gamma W\right) \underline{G}_{2,\gamma_0} c \\
&= c'\underline{G}'_{2,\gamma_0} \left(W - W\underline{G}_\gamma \left(\underline{G}'_\gamma W \underline{G}_\gamma\right)^{-1} \underline{G}'_\gamma W\right) \underline{G}_{2,\gamma_0} c \\
&= c' R'_{2,\gamma_0} \left(I - R_\gamma \left(R'_\gamma R_\gamma\right)^{-1} R'_\gamma\right) R_{2,\gamma_0} c \\
&= \left\|R_{2,\gamma_0} c\right\|^2 - \left\|P_{R_\gamma}\left(R_{2,\gamma_0} c\right)\right\|^2 =: Q_0(\gamma)
\end{aligned}$$

uniformly in γ, where $R_\gamma = W^{1/2}\underline{G}_\gamma$ and $R_{2,\gamma} = W^{1/2}\underline{G}_{2,\gamma}$. Obviously, $Q_0(\gamma_0) = 0$. Also, P_{R_γ} is a projection on a $2\bar{d}$-dimensional space, while $R_{2,\gamma_0} c$ is a $l(> 2\bar{d})$-dimensional vector, so as long as $R_{\gamma_0} c$ does not fall in *span* (R_γ) when $\gamma \neq \gamma_0$, $Q_0(\gamma) > 0$. This requirement is satisfied by virtue of Assumption IV.

We can now adjust Theorem 7.2 of Newey and McFadden (1994) to derive the asymptotic distribution of $\widehat{\theta}$. We only point out the difference in the proof. Replace G by $G_n r_n$ and $\theta - \theta_0$ by $r_n^{-1}(\theta - \theta_0)$, where $G_n = -\left(\ \mathbb{E}[\mathbf{z}\mathbf{x}'], \mathbb{E}\left[\mathbf{z}\mathbf{x}'_{\le\gamma_0}\right], \mathbb{E}[\mathbf{z}\mathbf{x}'|q = \gamma_0]\,\delta_n f_q(\gamma_0)\ \right)$ and $r_n = \begin{pmatrix} I_{2\bar{d}} & \mathbf{0} \\ \mathbf{0} & 1/\|\delta_n\| \end{pmatrix}$. Then $H = -r_n G'_n W G_n r_n \to -G'WG$ and $\widehat{D} = -r_n G'_n \widehat{W} \sqrt{n}\widehat{g}_n(\theta_0) \xrightarrow{d} N(\mathbf{0}, G'W\Omega WG)$. What remains is to show that for any $h_n \to 0$,

$$\sup\nolimits_{\|(\theta-\theta_0)\|\le h_n} \sqrt{n}\left\|\widehat{g}_n(\theta) - \widehat{g}_n(\theta_0) - g_0(\theta_0)\right\| \xrightarrow{p} 0.$$

This stochastic equicontinuity result is obvious because $\widehat{g}_n(\theta)$ is generated by a VC subgraph class of functions. We mention that this part of proof is similar to the 'convergence rate and asymptotic normality' part in the proof of Theorem 1 of Seo and Shin (2016). Their consistency proof is marred by a typo, which has a material effect. Specifically, at the end of page 181, the probability limit should be $I - A(\gamma)\cdots$ rather than $I + A(\gamma)\cdots$. This is why they did not specify an identification assumption such as that in Assumption IV.

■

Proof of Theorem 2. First assume $\widetilde{W} = \text{diag}\left\{\widetilde{W}_1, \widetilde{W}_2\right\}$ with $\widetilde{W}_1 \xrightarrow{p} W_1 > 0$ and $\widetilde{W}_2 \xrightarrow{p} W_2 > 0$. Then $\widetilde{Q}_n(\theta)$ can be expressed as a sum of two quadratic forms:

$$\widetilde{Q}_n(\theta) = \widetilde{Q}_{1n}(\theta_1) + \widetilde{Q}_{2n}(\theta_2),$$

where

$$\begin{aligned}
&\widetilde{Q}_{1n}(\theta_1) = \widetilde{m}_{1n}(\theta_1)' \widetilde{W}_1 \widetilde{m}_{1n}(\theta_1) \text{ and } \widetilde{Q}_{1n}(\theta_1) = \widetilde{m}_{2n}(\theta_2)' \widetilde{W}_2 \widetilde{m}_{2n}(\theta_2), \\
&\widetilde{m}_{1n}(\theta_1) = \tfrac{1}{n}\textstyle\sum_{i=1}^n \mathbf{z}_i\left(y_i - \mathbf{x}'_i\beta_1\right) 1(q_i \le \gamma), \\
&\widetilde{m}_{2n}(\theta_2) = \tfrac{1}{n}\textstyle\sum_{i=1}^n \mathbf{z}_i\left(y_i - \mathbf{x}'_i\beta_2\right) 1(q_i > \gamma),
\end{aligned}$$

with $\theta_1 = (\beta_1', \gamma)'$ and $\theta_2 = (\beta_2', \gamma)'$. It is not hard to see that

$$\begin{pmatrix} \widetilde{\beta}_1(\gamma) \\ \widetilde{\beta}_2(\gamma) \end{pmatrix} = \begin{pmatrix} \left[X_{\le\gamma}' Z \widetilde{W}_1 (Z' X_{\le\gamma})\right]^{-1} \left[(X_{\le\gamma}' Z)\widetilde{W}_1 (Z' Y_{\le\gamma})\right] \\ \left[X_{>\gamma}' Z \widetilde{W}_2 (Z' X_{>\gamma})\right]^{-1} \left[(X_{>\gamma}' Z)\widetilde{W}_2 (Z' Y_{>\gamma})\right] \end{pmatrix},$$

so if $\widetilde{\gamma}$ is consistent, then both $\widetilde{\beta}_1$ and $\widetilde{\beta}_2$ are consistent. So we concentrate on the consistency of $\widetilde{\gamma}$.

By similar analysis to the proof of Theorem 1,

$$\widetilde{Q}_n(\gamma) \xrightarrow{p} \|\widetilde{r}_{1,\gamma,\theta_0}\|^2 - \left\|P_{\widetilde{R}_{1,\gamma}}(\widetilde{r}_{1,\gamma,\theta_0})\right\|^2 + \|\widetilde{r}_{2,\gamma,\theta_0}\|^2 - \left\|P_{\widetilde{R}_{2,\gamma}}(\widetilde{r}_{2,\gamma,\theta_0})\right\|^2,$$

where $\widetilde{r}_{1,\gamma,\theta_0} = W_1^{1/2}\mathbb{E}[\mathbf{z}y_{\le\gamma}]$, $\widetilde{r}_{2,\gamma,\theta_0} = W_2^{1/2}\mathbb{E}[\mathbf{z}y_{>\gamma}]$, $\widetilde{R}_{1,\gamma} = W_1^{1/2}\mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma}'\right]$ and $\widetilde{R}_{2,\gamma} = W_2^{1/2}\mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma}'\right]$. Note that

$$\begin{aligned}
\begin{pmatrix} \mathbb{E}[\mathbf{z}y_{\le\gamma}] \\ \mathbb{E}[\mathbf{z}y_{>\gamma}] \end{pmatrix} &= \begin{pmatrix} \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma\wedge\gamma_0}'\right] & \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\gamma_0<\le\gamma}'\right] \\ \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\gamma<\le\gamma_0}'\right] & \mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma\vee\gamma_0}'\right] \end{pmatrix} \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \mathbb{E}[\mathbf{z}\varepsilon_{\le\gamma}] \\ \mathbb{E}[\mathbf{z}\varepsilon_{>\gamma}] \end{pmatrix} \\
&= \begin{pmatrix} \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma}'\right] & \mathbf{0} \\ \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\gamma<\le\gamma_0}'\right] & \mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma_0}'\right] \end{pmatrix} \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \mathbb{E}[\mathbf{z}\varepsilon_{\le\gamma}] \\ \mathbb{E}[\mathbf{z}\varepsilon_{>\gamma}] \end{pmatrix} \text{ if } \gamma \le \gamma_0 \\
&= \begin{pmatrix} \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma_0}'\right] & \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\gamma_0<\le\gamma}'\right] \\ \mathbf{0} & \mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma}'\right] \end{pmatrix} \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \mathbb{E}[\mathbf{z}\varepsilon_{\le\gamma}] \\ \mathbb{E}[\mathbf{z}\varepsilon_{>\gamma}] \end{pmatrix} \text{ if } \gamma > \gamma_0.
\end{aligned}$$

We consider the following two cases.

(i) q is exogenous (i.e., q is included in $\mathbf{z}$, and $\mathbb{E}[\varepsilon|\mathbf{z}] = 0$). In this case, $\mathbb{E}[\mathbf{z}\varepsilon_{\le\gamma}] = \mathbb{E}[\mathbf{z}\varepsilon_{>\gamma}] = \mathbf{0}$. First suppose $\gamma \le \gamma_0$. Then $\widetilde{r}_{1,\gamma,\theta_0} = W_1^{1/2}\mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma}'\right]\beta_{10}$ and $\widetilde{r}_{2,\gamma,\theta_0} = W_2^{1/2}\left(\mathbb{E}\left[\mathbf{z}\mathbf{x}_{\gamma<\le\gamma_0}'\right]\beta_{10} + \mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma_0}'\right]\beta_{20}\right)$. The question is whether we can find a_1 and a_2 such that $\mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma}'\right]a_1 = \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma}'\right]\beta_{10}$, and $\mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma}'\right]a_2 = \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\gamma<\le\gamma_0}'\right]\beta_{10} + \mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma_0}'\right]\beta_{20} = \mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma}'\right]\beta_{10} + \mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma_0}'\right](\beta_{20} - \beta_{10})$ or $\mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma}'\right](\beta_{10} - a_2) = \mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma_0}'\right]\delta_0$. We can let $a_1 = \beta_{10}$, but if $l > (d+1)$, such an a_2 is impossible by Assumption IV$'$(i). Next suppose $\gamma > \gamma_0$. Then we try to select a_1 and a_2 such that $\mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma}'\right]a_1 = \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma_0}'\right]\beta_{10} + \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\gamma_0<\le\gamma}'\right]\beta_{20} = \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma}'\right]\beta_{20} + \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma_0}'\right](\beta_{10} - \beta_{20})$ or $\mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma}'\right](a_1 - \beta_{20}) = \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma_0}'\right]\delta_0$ and $\mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma}'\right]a_2 = \mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma}'\right]\beta_{20}$. We can let $a_2 = \beta_{20}$, but such an a_2 is impossible by Assumption IV$'$(i).

(ii) q is endogenous and satisfies only $\mathbb{E}[\mathbf{z}\varepsilon 1(q \le \gamma_0)] = \mathbf{0}$ and $\mathbb{E}[\mathbf{z}\varepsilon 1(q > \gamma_0)] = \mathbf{0}$. Again, first suppose $\gamma \le \gamma_0$. Then we try to select a_1 and a_2 such that $\mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma}'\right]a_1 = \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma}'\right]\beta_{10} + \mathbb{E}[\mathbf{z}\varepsilon_{\le\gamma}]$ and $\mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma}'\right]a_2 = \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\gamma<\le\gamma_0}'\right]\beta_{10} + \mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma_0}'\right]\beta_{20} + \mathbb{E}[\mathbf{z}\varepsilon_{>\gamma}]$ or $\mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma}'\right](a_1 - \beta_{10}) = \mathbb{E}[\mathbf{z}\varepsilon_{\le\gamma}]$ and $\mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma}'\right](\beta_{10} - a_2) + \mathbb{E}[\mathbf{z}\varepsilon_{>\gamma}] = \mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma_0}'\right]\delta_0$. Such an a_1 and a_2 are impossible by Assumption IV$'$(ii). Next suppose $\gamma > \gamma_0$. Then we try to select a_1 and a_2 such that $\mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma}'\right]a_1 = \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma_0}'\right]\beta_{10} + \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\gamma_0<\le\gamma}'\right]\beta_{20} + \mathbb{E}[\mathbf{z}\varepsilon_{\le\gamma}] = \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma}'\right]\beta_{20} + \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma_0}'\right](\beta_{10} - \beta_{20}) + \mathbb{E}[\mathbf{z}\varepsilon_{\le\gamma}]$ and $\mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma}'\right]a_2 = \mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma}'\right]\beta_{20} + \mathbb{E}[\mathbf{z}\varepsilon_{>\gamma}]$ or $\mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma}'\right](a_1 - \beta_{20}) = \mathbb{E}\left[\mathbf{z}\mathbf{x}_{\le\gamma_0}'\right]\delta_0 + \mathbb{E}[\mathbf{z}\varepsilon_{\le\gamma}]$ and $\mathbb{E}\left[\mathbf{z}\mathbf{x}_{>\gamma}'\right](a_2 - \beta_{20}) = \mathbb{E}[\mathbf{z}\varepsilon_{>\gamma}]$. Such an a_1 and a_2 are impossible by Assumption IV$'$(ii).

Now consider the case where q is exogenous and is independent of $(\mathbf{z}', \mathbf{x}')'$. Suppose $\gamma \le \gamma_0$. We can set $a_2 = \frac{(1-F_q(\gamma_0))(\beta_{20}-\beta_{10})}{1-F_q(\gamma)} + \beta_{10} = \frac{(1-F_q(\gamma_0))\beta_{20}+(F_q(\gamma_0)-F_q(\gamma))\beta_{10}}{1-F_q(\gamma)}$, which is $\theta_*^{(2)}(\lambda)$ in Proposition 1(ii) of HHB where $q \sim U[0,1]$. Now suppose $\gamma > \gamma_0$. We can set $a_1 = \frac{F_q(\gamma)\beta_{10}+(F_q(\gamma)-F_q(\gamma_0))\beta_{20}}{F_q(\gamma)}$, which is $\theta_*^{(1)}(\lambda)$ in Proposition 1(ii) of HHB where $q \sim U[0,1]$.

If $\widetilde{W}$ is a general positive definite matrix, then

$$\widetilde{Q}_n(\gamma) \xrightarrow{p} \|\widetilde{r}_{\gamma,\theta_0}\|^2 - \left\|P_{\widetilde{R}_\gamma}(\widetilde{r}_{\gamma,\theta_0})\right\|^2,$$

where $\widetilde{R}_\gamma = W^{1/2}\underline{\widetilde{G}}_\gamma$ with $\underline{\widetilde{G}}_\gamma = \begin{pmatrix} \mathbb{E}[\mathbf{z}\mathbf{x}'_{\le\gamma}] & 0 \\ 0 & \mathbb{E}[\mathbf{z}\mathbf{x}'_{>\gamma}] \end{pmatrix}$, and $\widetilde{r}_{\gamma,\theta_0} = W^{1/2}\begin{pmatrix} \mathbb{E}[\mathbf{z}y_{\le\gamma}] \\ \mathbb{E}[\mathbf{z}y_{>\gamma}] \end{pmatrix}$. In case (i), if there does not exist a_1 and a_2 such that $\begin{pmatrix} \mathbb{E}[\mathbf{z}\mathbf{x}'_{\le\gamma}] \\ 0 \end{pmatrix} a_1 + \begin{pmatrix} 0 \\ \mathbb{E}[\mathbf{z}\mathbf{x}'_{>\gamma}] \end{pmatrix} a_2 = \begin{pmatrix} \mathbb{E}[\mathbf{z}\mathbf{x}'_{\le\gamma}] & \mathbf{0} \\ \mathbb{E}\left[\mathbf{z}\mathbf{x}'_{\gamma<\le\gamma_0}\right] & \mathbb{E}\left[\mathbf{z}\mathbf{x}'_{>\gamma_0}\right] \end{pmatrix} \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix}$ or $\mathbb{E}[\mathbf{z}\mathbf{x}'_{\le\gamma}]\, a_1 = \mathbb{E}[\mathbf{z}\mathbf{x}'_{\le\gamma}]\,\beta_{10}$ and $\mathbb{E}[\mathbf{z}\mathbf{x}'_{>\gamma}]\, a_2 = \mathbb{E}\left[\mathbf{z}\mathbf{x}'_{\gamma<\le\gamma_0}\right]\beta_{10} + \mathbb{E}\left[\mathbf{z}\mathbf{x}'_{>\gamma_0}\right]\beta_{20}$ for any $\gamma < \gamma_0$, and $\begin{pmatrix} \mathbb{E}[\mathbf{z}\mathbf{x}'_{\le\gamma}] \\ 0 \end{pmatrix} a_1 + \begin{pmatrix} 0 \\ \mathbb{E}[\mathbf{z}\mathbf{x}'_{>\gamma}] \end{pmatrix} a_2 = \begin{pmatrix} \mathbb{E}\left[\mathbf{z}\mathbf{x}'_{\le\gamma_0}\right] & \mathbb{E}\left[\mathbf{z}\mathbf{x}'_{\gamma_0<\le\gamma}\right] \\ \mathbf{0} & \mathbb{E}[\mathbf{z}\mathbf{x}'_{>\gamma}] \end{pmatrix} \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix}$ or $\mathbb{E}[\mathbf{z}\mathbf{x}'_{\le\gamma}]\, a_1 = \mathbb{E}\left[\mathbf{z}\mathbf{x}'_{\le\gamma_0}\right]\beta_{10} + \mathbb{E}\left[\mathbf{z}\mathbf{x}'_{\gamma_0<\le\gamma}\right]\beta_{20}$ and $\mathbb{E}[\mathbf{z}\mathbf{x}'_{>\gamma}]\, a_2 = \mathbb{E}[\mathbf{z}\mathbf{x}'_{>\gamma}]\,\beta_{20}$ for any $\gamma > \gamma_0$, then γ_0 is identified. These conditions are exactly the same as in the diagonal $\widetilde{W}$ case. Similarly, in case (ii), the same conditions as in the diagonal $\widetilde{W}$ case are required to identify γ_0. ■

Proof of Corollary 2. By a Glivenko-Cantelli theorem, we have the probability limit of $J_n(\theta)$ to be

$$J(\theta) := E[g_i(\theta)]' W E[g_i(\theta)].$$

We want to show that for any $\gamma \in \Gamma$, we can find β_{22}, δ such that $E[g_i(\theta)] = 0$, i.e., $J(\gamma) := \min_{\beta_{22},\delta} J(\theta) = 0$, so γ is not identifiable.

Note that

$$E[g_i(\theta)] = E\begin{bmatrix} z_{it_0}\Delta y_{it_0} \\ \vdots \\ z_{iT}\Delta y_{iT} \end{bmatrix} - E\begin{bmatrix} z_{it_0}\Delta x'_{it_0} \\ \vdots \\ z_{iT}\Delta x'_{iT} \end{bmatrix}\beta_{22} - E\begin{bmatrix} z_{it_0}1_{it_0}(\gamma)'X_{it_0} \\ \vdots \\ z_{iT}1_{iT}(\gamma)'X_{iT} \end{bmatrix}\delta,$$

where

$$\begin{aligned} E\begin{bmatrix} z_{it_0}\Delta y_{it_0} \\ \vdots \\ z_{iT}\Delta y_{iT} \end{bmatrix} &= E\begin{bmatrix} z_{it_0}\left(\beta'_{220}\Delta x_{it_0} + \delta'_0 X_{it_0}1_{it_0}(\gamma_0) + \Delta v_{it_0}\right) \\ \vdots \\ z_{iT}\left(\beta'_{220}\Delta x_{iT} + \delta'_0 X_{iT}1_{iT}(\gamma_0) + \Delta v_{iT}\right) \end{bmatrix} \\ &= E\begin{bmatrix} z_{it_0}\Delta x'_{it_0} \\ \vdots \\ z_{iT}\Delta x'_{iT} \end{bmatrix}\beta_{220} + E\begin{bmatrix} z_{it_0}1_{it_0}(\gamma_0)'X_{it_0} \\ \vdots \\ z_{iT}1_{iT}(\gamma_0)'X_{iT} \end{bmatrix}\delta_0, \end{aligned}$$

and

$$\begin{aligned} E\begin{bmatrix} z_{it_0}1_{it_0}(\gamma)'X_{it_0} \\ \vdots \\ z_{iT}1_{iT}(\gamma)'X_{iT} \end{bmatrix} &= E\begin{bmatrix} z_{it_0}\left[\left(1, x'_{it_0}\right)F_q(\gamma) - \left(1, x'_{i,t_0-1}\right)F_q(\gamma)\right] \\ \vdots \\ z_{iT}\left[(1, x'_{iT})F_q(\gamma) - (1, x'_{iT})F_q(\gamma)\right] \end{bmatrix} \\ &= E\begin{bmatrix} z_{it_0}\Delta x_{it_0} \\ \vdots \\ z_{iT}\Delta x'_{iT} \end{bmatrix}F_q(\gamma). \end{aligned}$$

As a result,

$$E\left[g_i\left(\theta\right)\right]=E\left[\begin{array}{c} z_{it_0}\Delta x_{it_0}' \\ \vdots \\ z_{iT}\Delta x_{iT}' \end{array}\right]\left(\beta_{220}+F_q\left(\gamma_0\right)\delta_{20}-\beta_{22}-F_q\left(\gamma\right)\delta_2\right),$$

where $\delta=\left(\delta_1,\delta_2'\right)'$. There are infinite many possible β_{22} and δ's such that $\beta_{220}+F_q\left(\gamma_0\right)\delta_{20}-\beta_{22}-F_q\left(\gamma\right)\delta_2=\mathbf{0}$, e.g., $\beta_{22}=\beta_{220}$, and $\delta=\left(0,\frac{F_q(\gamma_0)}{F_q(\gamma)}\delta_{20}\right)$. ■

Proof of Theorem 3. Proposition 1 proves the consistency of $\widehat{\gamma}$, and Proposition 2 proves $\widehat{\gamma}-\gamma_0=O_p\left((n/h)^{-1/2}\right)$, so we can apply the argmax continuous mapping theorem (see, e.g., Theorem 3.2.2 of Van der Vaart and Wellner (1996)) to establish the asymptotic distribution of $\sqrt{n/h}\left(\widehat{\gamma}-\gamma_0\right)$. From Proposition 3, the finite-dimensional limit distributions of $nh\left(\widehat{Q}_n(\gamma_0^v)-\widehat{Q}_n(\gamma_0)\right)$ are the same as those of $-v^2\mathbb{E}[\Delta_f(x_i)\Delta_i f(x_i)|q_i=\gamma_0]f_q(\gamma_0)k_+'(0)+2vU$, where $U\sim\mathcal{N}\left(0,\mathbb{E}\left[\Delta_f^2(x_i)f^2(x_i)(\sigma_+^2(x_i)+\sigma_-^2(x_i))|q_i=\gamma_0\right]f_q(\gamma_0)\xi_{(1)}\right)$. Combining this with the stochastic equicontinuity result in Lemma 6, we have

$$nh\left(\widehat{Q}_n(\gamma_0^v)-\widehat{Q}_n(\gamma_0)\right)\rightsquigarrow -v^2\mathbb{E}[\Delta_f(x_i)\Delta_i f(x_i)|q_i=\gamma_0]f_q(\gamma_0)k_+'(0)+2Uv.$$

Thus,

$$\begin{aligned}\sqrt{n/h}\left(\widehat{\gamma}-\gamma_0\right)\xrightarrow{d}v^* &= \arg\max_v\left\{-v^2\mathbb{E}[\Delta_f(x_i)\Delta_i f(x_i)|q_i=\gamma_0]f_q(\gamma_0)k_+'(0)+2Uv\right\}\\ &= \frac{U}{\mathbb{E}[\Delta_f(x_i)\Delta_i f(x_i)|q_i=\gamma_0]f_q(\gamma_0)k_+'(0)}\sim\mathcal{N}(0,\Sigma).\end{aligned}$$

■

Proof of Theorem 4. We prove the theorem in two steps. First, we show that $\widehat{\Sigma}\left(\gamma_0\right)\xrightarrow{d}\Sigma$, where $\widehat{\Sigma}\left(\gamma_0\right)$ is replacing $\widehat{\gamma}$ in $\widehat{\Sigma}$ by γ_0. Second, we show that $\widehat{\Sigma}-\widehat{\Sigma}\left(\gamma_0\right)\xrightarrow{p}0$. The first result is shown in Proposition 4 and the second is shown in Proposition 5. ■

Proof of Corollary 3. From the proof of Theorem 3 and the CMT,

$$\begin{aligned}&nh\frac{\mathbb{E}[\Delta_f(x_i)\Delta_i f(x_i)|q_i=\gamma_0]k_+'(0)}{\mathbb{E}\left[\Delta_f^2(x_i)f^2(x_i)(\sigma_+^2(x_i)+\sigma_-^2(x_i))|q_i=\gamma_0\right]\xi_{(1)}}\sup_v\left\{\widehat{Q}_n(\gamma_0^v)-\widehat{Q}_n(\gamma_0)\right\}\\ \xrightarrow{d}\ &\frac{\mathbb{E}[\Delta_f(x_i)\Delta_i f(x_i)|q_i=\gamma_0]k_+'(0)}{\mathbb{E}\left[\Delta_f^2(x_i)f^2(x_i)(\sigma_+^2(x_i)+\sigma_-^2(x_i))|q_i=\gamma_0\right]\xi_{(1)}}\sup_v\left\{-v^2\mathbb{E}[\Delta_f(x_i)\Delta_i f(x_i)|q_i=\gamma_0]f_q(\gamma_0)k_+'(0)+2Uv\right\}\\ =\ &\frac{\mathbb{E}[\Delta_f(x_i)\Delta_i f(x_i)|q_i=\gamma_0]k_+'(0)}{\mathbb{E}\left[\Delta_f^2(x_i)f^2(x_i)(\sigma_+^2(x_i)+\sigma_-^2(x_i))|q_i=\gamma_0\right]\xi_{(1)}}\frac{U^2}{f_q(\gamma_0)\mathbb{E}[\Delta_f(x_i)\Delta_i f(x_i)|q_i=\gamma_0]|k_+'(0)}\\ =\ &\frac{U^2}{\mathbb{E}\left[\Delta_f^2(x_i)f^2(x_i)(\sigma_+^2(x_i)+\sigma_-^2(x_i))|q_i=\gamma_0\right]f_q(\gamma_0)\xi_{(1)}}\sim\chi_1^2.\end{aligned}$$

By the proof of Theorem 4 and Slutsky's theorem, $\widehat{LR}_n\left(\gamma_0\right)\xrightarrow{d}\chi_1^2$. ■

Proof of Theorem 5. Assume the densities of $(x',q)'$ and e are known. Since the minimax risk for a larger class of probability models is no smaller than that for a smaller class of probability models, the lower bound for a particular distributional assumption also holds for a wider class of distributions. To simplify the calculation, assume e_i is iid $N(0,1)$ and $(x_i',q_i)'$ is iid uniform on $\mathcal{X}\times[0,1]$. Such a specification also appears in Fan (1993) where it is called the assumption of richness of joint densities. We will use the technique in Sun (2005) to develop our results. This technique is also implicitly used in Stone (1980) and the essential part of the technique can be cast in the language of Neyman-Pearson testing.

Let P,Q be probability measures defined on the same measurable space $(\Omega,\mathcal{A})$ with the affinity between the two measures defined as usual to be

$$\pi(P,Q) = \inf\left(\mathbb{E}_P[\phi] + \mathbb{E}_Q[1-\phi]\right),$$

where the infimum is taken over the measurable function ϕ such that $0 \le \phi \le 1$. In other words, $\pi(P,Q)$ is the smallest sum of type I and type II errors of any test between P and Q. It is a natural measure of the difficulty of distinguishing P and Q. Suppose μ is a measure dominating both P and Q with corresponding densities p and q. It follows from the Neyman-Pearson lemma that the infimum is achieved by setting $\phi = 1(p \le q)$ and then

$$\pi(P,Q) = \int 1(p \le q) p d\mu + \int 1(p > q) q d\mu = 1 - \tfrac{1}{2}\int |p-q|\, d\mu \equiv 1 - \tfrac{1}{2}\|P-Q\|_1,$$

where $\|\cdot\|_1$ is the L_1 distance between two probability measures. Now consider a pair of probability models $P, Q \in \mathcal{P}(s,B)$ such that $|\gamma(P) - \gamma(Q)| \ge \epsilon$.

For any estimator $\widehat{\gamma}$, we have

$$1\left(\|\widehat{\gamma} - \gamma(P)\| > \epsilon/2\right) + 1\left(\|\widehat{\gamma} - \gamma(Q)\| > \epsilon/2\right) \ge 1.$$

Let

$$\phi = \frac{1(|\widehat{\gamma}-\gamma(P)|>\epsilon/2)}{1(|\widehat{\gamma}-\gamma(P)|>\epsilon/2)+1(|\widehat{\gamma}-\gamma(Q)|>\epsilon/2)}.$$

Then $0 \le \phi \le 1$ and

$$\sup_{P\in\mathcal{P}(s,B)} P\left(|\widehat{\gamma} - \gamma(P)| > \epsilon/2\right) \ge \tfrac{1}{2}\left\{P\left(|\widehat{\gamma} - \gamma(P)| > \epsilon/2\right) + Q\left(|\widehat{\gamma} - \gamma(Q)| > \epsilon/2\right)\right\} \ge \tfrac{1}{2}\mathbb{E}_P[\phi] + \tfrac{1}{2}\mathbb{E}_Q[1-\phi].$$

Therefore

$$\inf_{\widehat{\gamma}} \sup_{P\in\mathcal{P}(s,B)} P\left(|\widehat{\gamma} - \gamma(P)| > \epsilon/2\right) \ge \tfrac{1}{2}\pi(P,Q)$$

for any P and Q such that $|\gamma(P) - \gamma(Q)| \ge \epsilon$. So we need only search for the pair (P,Q) which minimize $\pi(P,Q)$ subject to the constraint $|\gamma(P) - \gamma(Q)| \ge \epsilon$. To obtain a lower bound with a sequence of independent observations, let $(\Omega, \mathcal{A})$ be the product space and $\mathcal{P}(s,B)$ be the family of product probabilities on such a space. Then for any pair of finite-product measures $P = \prod_{i=1}^n P_i$ and $Q = \prod_{i=1}^n Q_i$, the minimax risk satisfies

$$\inf_{\widehat{\gamma}} \sup_{P\in\mathcal{P}(s,B)} P\left(|\widehat{\gamma} - \gamma(P)| > \epsilon/2\right) \ge \tfrac{1}{2}\left(1 - \tfrac{1}{2}\left\|\prod_{i=1}^n P_i - \prod_{i=1}^n Q_i\right\|_1\right)$$

provided that $|\gamma(P) - \gamma(Q)| \ge \epsilon$. From Pollard (1993), if $dQ_i/dP_i = 1 + \Delta_i(\cdot)$, then

$$\left\|\prod_{i=1}^n P_i - \prod_{i=1}^n Q_i\right\|_1 \le \exp\left(\sum_{i=1}^n \nu_i^2\right) - 1,$$

where $\nu_i^2 = \mathbb{E}_{P_i}[\Delta_i^2(\cdot)]$ is finite. So

$$\inf_{\widehat{\gamma}} \sup_{P\in\mathcal{P}(s,B)} \mathbb{P}\left(|\widehat{\gamma} - \gamma(P)| > \epsilon/2\right) \ge \tfrac{1}{2}\left(\tfrac{3}{2} - \exp\left(\sum_{i=1}^n \nu_i^2\right)\right) \tag{29}$$

provided that $|\gamma(P) - \gamma(Q)| \ge \epsilon$.

It remains to find probabilities P and Q that are difficult to distinguish by the data set $\{(x_i', q_i, y_i)\}_{i=1}^n$.

Under P, the data is generated according to

$$y_i = g_P(x_i, q_i) + \Delta m_P(x_i, q_i)\, 1(q_i \le \gamma_P) + e_i,$$

where $\Delta m_P(x_i, q_i) = \delta_{\alpha P} + x_i'\delta_{xP} + q_i\delta_{qP}$, and under Q, g_P, Δm_P and γ_P are changed to g_Q, Δm_Q and γ_Q, respectively. The point here is that only Δm_P instead of δ_P matters for our purpose. We now specify g, Δm and γ for each model. First suppose $n^{\frac{s}{2s+1}}\rho_n \to \infty$. For P, let $g_P = 0$, $\Delta m_P = 0$, and $\gamma_P = 0$ without loss of generality; for Q, let

$$g_Q(x,q) = 0,\ \Delta m_Q(x,q) = \xi\rho_n,\ \gamma_Q = \left(\xi n \rho_n^2\right)^{-1},$$

where ξ is a positive constant. Obviously, $g_Q(x,q) \in \mathcal{C}_s(B, \mathcal{X} \times [0,1])$ for some $B > 0$, so it remains to compute the L_1 distance between the two measures. Let the density of Q_i with respect to P_i be $1 + \Delta_i(\cdot)$, then

$$\Delta_i(x_i, q_i, y_i) = \begin{cases} \phi(y_i - \Delta m_Q(x_i, q_i))/\phi(y_i) - 1, & \text{if } q_i \in [0, \gamma_Q], \\ 0, & \text{otherwise} \end{cases}$$

where $\phi(\cdot)$ is the standard normal pdf. Therefore,

$$\begin{aligned}
&\mathbb{E}_{P_i}\left[\Delta_i^2\right] \\
&= \int_0^{\gamma_Q}\int_0^1 \cdots \int_0^1 \int_{-\infty}^{\infty} \left[\phi(y - \Delta m_Q(x,q))/\phi(y) - 1\right]^2 \phi(y) f(x,q)\, dy dx dq \\
&= \int_0^{\gamma_Q}\int_0^1 \cdots \int_0^1 \int_{-\infty}^{\infty} \phi(y - \Delta m_Q(x,q))^2/\phi(y)\, dy dx dq - 2\int_0^{\gamma_Q}\int_0^1 \cdots \int_0^1 \int_{-\infty}^{\infty} \phi(y - \Delta m_Q(x,q))\, dy dx dq + \gamma_Q \\
&= \int_0^{\gamma_Q}\int_0^1 \cdots \int_0^1 \int_{-\infty}^{\infty} \phi(y - \Delta m_Q(x,q))^2/\phi(y)\, dy dx dq - \gamma_Q.
\end{aligned}$$

Plugging in the standard normal pdf yields

$$\begin{aligned}
\mathbb{E}_{P_i}[\Delta_i^2] &= \int_0^{\gamma_Q}\int_0^1 \cdots \int_0^1 \int_{-\infty}^{\infty} \tfrac{1}{\sqrt{2\pi}} \exp\left\{-\tfrac{2(y - \Delta m_Q(x,q))^2}{2} + \tfrac{y^2}{2}\right\} dy dx dq - \gamma_Q \\
&= \int_0^{\gamma_Q}\int_0^1 \cdots \int_0^1 \exp\left\{\Delta m_Q(x,q)^2\right\} dx dq - \gamma_Q \\
&= \int_0^{\gamma_Q} \exp\left(\xi^2\rho_n^2\right) dq - \gamma_Q \\
&= \gamma_Q\left[\exp\left(\xi^2\rho_n^2\right) - 1\right] = \gamma_Q \xi^2 \rho_n^2 (1 + o(1)) \le \tfrac{\xi}{n},
\end{aligned}$$

when n is large enough.

When ξ is small enough, say $\xi \le \log(5/4)$, we have

$$\exp\left(\sum_{i=1}^n \nu_i^2\right) \le \exp(\xi) < \tfrac{5}{4}.$$

It follows from (29) that

$$\inf_{\widehat{\gamma}} \sup_{P \in \mathcal{P}(s,B)} \mathbb{P}\left(|\widehat{\gamma} - \gamma(P)| > \tfrac{\epsilon}{2}\left(n\rho_n^2\right)^{-1}\right) \ge \tfrac{1}{2}\left(\tfrac{3}{2} - \tfrac{5}{4}\right) = \tfrac{1}{8} \ge C,$$

on choosing $C \le 1/8$, where $\frac{\epsilon}{2}\left(n\rho_n^2\right)^{-1}$ appears because $|\gamma(P) - \gamma(Q)| = \left(\xi n\rho_n^2\right)^{-1} \ge \epsilon\left(n\rho_n^2\right)^{-1}$ for a small ϵ.

We next suppose $n^{\frac{s}{2s+1}}\rho_n = O(1)$. Let P and Q be the same as above except that in Q,

$$g_Q(x,q) = -\xi\eta^s \varphi_q\left(\tfrac{q - \gamma_Q}{\eta}\right),\ \Delta m_Q(x,q) = \xi\eta^s,\ \gamma_Q = \xi,$$

where $\eta = n^{-1/(2s+1)}$, φ_q is an infinitely differentiable function in q satisfying (i) $\varphi_q(v) = 0$ for $v \ge 0$, (ii)

$\varphi_q(v) = 1$, for $v \leq -1$, and (iii) $\varphi_q(v) \in (0,1)$ for $v \in (-1,0)$. It is not hard to check that $g_Q(x,q) \in \mathcal{C}_s(B, \mathcal{X} \times [0,1])$ for some $B > 0$. By similar steps above, we can show

$$\mathbb{E}_{P_i}[\Delta_i^2] \leq \frac{\xi^2}{2n}$$

when n is large enough. By choosing ξ appropriately, we have

$$\inf_{\widehat{\gamma}} \sup_{P \in \mathcal{P}(s,B)} \mathbb{P}\left(|\widehat{\gamma} - \gamma(P)| > \tfrac{\epsilon}{2}\right) \geq C,$$

where we choose $\epsilon \leq \xi$. ■

Proof of Theorem 6. We apply Theorem 2.7 of Kim and Polland (1990) to derive the asymptotic distribution of $n\rho_n^2(\widehat{\gamma} - \gamma_0)$. Note that $n\rho_n^2(\widehat{\gamma} - \gamma_0) = \arg\max_v nh\left(\widehat{Q}_n(\gamma_0^v) - \widehat{Q}_n(\gamma_0)\right) =: \arg\max_v \{C_n(v)\}$, where $\gamma_0^v = \gamma_0 + \frac{v}{n\rho_n^2}$.

(i) $C_n(v) \rightsquigarrow C(v) \in \mathbf{C}_{\max}(\mathbb{R})$, where

$$C(v) = \Sigma^{1/2} W(v) - 2k_+(0) f_q(\gamma_0) D\,|v|\,,$$

$W(v) := W_1(-v)1(v \leq 0) + W_2(v)1(v > 0)$ is a two-sided Brownian motion, $D = \lim_{n\to\infty} D_n$, and $\Sigma(v) = \lim_{n\to\infty} \Sigma_n$ with Σ_n defined in Proposition 8. $\mathbf{C}_{\max}(\mathbb{R})$ is defined as the subset of continuous functions $x(\cdot) \in \mathbf{B}_{\mathrm{loc}}(\mathbb{R})$ for which (i) $x(t) \to -\infty$ as $|t| \to \infty$ and (ii) $x(t)$ achieves its maximum at a unique point in $\mathbb{R}$, and $\mathbf{B}_{\mathrm{loc}}(\mathbb{R})$ is the space of all locally bounded real functions on $\mathbb{R}$, endowed with the uniform metric on compacta. The weak convergence can be proved by combining Proposition 8 and Lemma 12. We now check $C(v) \in \mathbf{C}_{\max}(\mathbb{R})$. It is not hard to check $C(v)$ is continuous, has a unique minimum (see Lemma 2.6 of Kim and Pollard (1990)), and $\lim_{|v|\to\infty} C(v) = -\infty$ almost surely (which is true since $\lim_{|v|\to\infty} W_\ell(v)/|v| = 0$ almost surely).

(ii) $n\rho_n^2(\widehat{\gamma} - \gamma_0) = O_p(1)$. This is proved in Proposition 7.

So

$$n\rho_n^2(\widehat{\gamma} - \gamma_0) \xrightarrow{d} \arg\max_v \{C(v)\}\,.$$

Making the change-of-variables $v = \frac{V_1}{f_q(\gamma_0)D^2} r$, and noting the distributional equality $W_\ell(a^2 r) =_d aW_\ell(r)$, we can rewrite the asymptotic distribution as

$$\begin{aligned}
&\arg\max_v \{C(v)\} \\
&= \frac{V_1}{f_q(\gamma_0)D^2} \arg\max_r \left\{C\left(\frac{V_1}{f_q(\gamma_0)D^2} r\right)\right\} \\
&= \frac{V_1}{f_q(\gamma_0)D^2} \arg\max_r \begin{cases} 4k_+(0)\sqrt{f_q(\gamma_0)V_1}\, W_1\left(-\frac{V_1}{f_q(\gamma_0)D^2} r\right) - 2k_+(0) f_q(\gamma_0) D \left|\frac{V_1}{f_q(\gamma_0)D^2} r\right|, & \text{if } r \leq 0, \\ 4k_+(0)\sqrt{f_q(\gamma_0)V_2}\, W_2\left(\frac{V_1}{f_q(\gamma_0)D^2} r\right) - 2k_+(0) f_q(\gamma_0) D \left|\frac{V_1}{f_q(\gamma_0)D^2} r\right|, & \text{if } r > 0, \end{cases} \\
&= \frac{V_1}{f_q(\gamma_0)D^2} \arg\max_r \begin{cases} \frac{V_1}{D} W_1(-r) - \frac{1}{2}\frac{V_1}{D}|r|\,, & \text{if } r \leq 0, \\ \frac{\sqrt{V_1 V_2}}{D} W_2(r) - \frac{1}{2}\frac{V_1}{D}|r|\,, & \text{if } r > 0, \end{cases} \\
&= \frac{V_1}{f_q(\gamma_0)D^2} \arg\max_r \begin{cases} W_1(-r) - \frac{1}{2}|r|\,, & \text{if } r \leq 0, \\ \sqrt{\frac{V_2}{V_1}} W_2(r) - \frac{1}{2}|r|\,, & \text{if } r > 0, \end{cases} =: \frac{V_1}{f_q(\gamma_0)D^2} \Lambda(\lambda)\,.
\end{aligned}$$

■

Proof of Corollary 4. We mimic the proof of Theorem 6. Note that $nh^{d-1}\Delta_o^2(\widetilde{\gamma} - \gamma_0) = \arg\max_v$

$nh^d\left(\widetilde{Q}_n(\gamma_0^v)-\widetilde{Q}_n(\gamma_0)\right) =: \arg\max_v \{C_{on}(v)\}$, where $\gamma_0^v=\gamma_0+\frac{v}{nh^{d-1}\Delta_o^2}$.

(i) $C_{on}(v) \rightsquigarrow C_o(v)\in\mathbf{C}_{\max}(\mathbb{R})$, where

$$C_o(v)=\Sigma_o(v)^{1/2}W(v)-2k_+(0)f(x_o,\gamma_0)^2|v|.$$

The weak convergence can be proved by combining Proposition 11 and Lemma 17.

(ii) $nh^{d-1}\Delta_o^2(\widetilde{\gamma}-\gamma_0)=O_p(1)$. This is proved in Proposition 10.

So

$$nh^{d-1}\Delta_o^2(\widetilde{\gamma}-\gamma_0)\xrightarrow{d}\arg\max_v\{C_o(v)\}.$$

Making the change-of-variables $v=\frac{\sigma_-^2(x_o)}{f(x_o,\gamma_0)}r$, and noting the distributional equality $W_\ell(a^2r)=aW_\ell(r)$, we can rewrite the asymptotic distribution as

$$\begin{aligned}
&\arg\max_v\{C_o(v)\}\\
=&\frac{\sigma_-^2(x_o)}{f(x_o,\gamma_0)}\arg\max_r\left\{C_o\left(\frac{\sigma_-^2(x_o)}{f(x_o,\gamma_0)}r\right)\right\}\\
=&\frac{\sigma_-^2(x_o)}{f(x_o,\gamma_0)}\arg\max_r\begin{cases}4k_+(0)\kappa\sqrt{f(x_o,\gamma_0)^3\sigma_-^2(x_o)}W_1\left(\frac{\sigma_-^2(x_o)}{f(x_o,\gamma_0)}r\right)-2k_+(0)f(x_o,\gamma_0)^2\left|\frac{\sigma_-^2(x_o)}{f(x_o,\gamma_0)}r\right|, & \text{if } r\le 0,\\ 4k_+(0)\kappa\sqrt{f(x_o,\gamma_0)^3\sigma_+^2(x_o)}W_2\left(\frac{\sigma_-^2(x_o)}{f(x_o,\gamma_0)}r\right)-2k_+(0)f(x_o,\gamma_0)^2\left|\frac{\sigma_-^2(x_o)}{f(x_o,\gamma_0)}r\right|, & \text{if } r>0,\end{cases}\\
=&\frac{\sigma_-^2(x_o)}{f(x_o,\gamma_0)}\arg\max_r\begin{cases}f(x_o,\gamma_0)\sigma_-^2(x_o)W_1(-r)-\frac{1}{2\kappa}f(x_o,\gamma_0)\sigma_-^2(x_o)|r|, & \text{if } r\le 0,\\ f(x_o,\gamma_0)\sigma_+(x_o)\sigma_-(x_o)W_2(r)-\frac{1}{2\kappa}f(x_o,\gamma_0)\sigma_-^2(x_o)|r|, & \text{if } r>0,\end{cases}\\
=&\frac{\sigma_-^2(x_o)}{f(x_o,\gamma_0)}\arg\max_r\begin{cases}W_1(-r)-\frac{1}{2\kappa}|r|, & \text{if } r\le 0,\\ \frac{\sigma_+(x_o)}{\sigma_-(x_o)}W_2(r)-\frac{1}{2\kappa}|r|, & \text{if } r>0,\end{cases} =: \frac{\sigma_-^2(x_o)}{f(x_o,\gamma_0)}\Lambda(\lambda_o,\kappa).
\end{aligned}$$

■

Proof of Corollary 5. From the proof of Theorem 6 and the CMT, we have

$$nh\left(\widehat{Q}_n(\gamma_0^v)-\widehat{Q}_n(\gamma_0)\right)\xrightarrow{d}\sup_v\{C(v)\}\ ,$$

where

$$\begin{aligned}
\sup_v\{C(v)\} &= \sup_r\begin{cases}4k_+(0)\sqrt{f_q(\gamma_0)V_1}W_1\left(-\frac{V_1}{f_q(\gamma_0)D^2}r\right)-2k_+(0)f_q(\gamma_0)D\left|\frac{V_1}{f_q(\gamma_0)D^2}r\right|, & \text{if } r\le 0,\\ 4k_+(0)\sqrt{f_q(\gamma_0)V_2}W_2\left(\frac{V_1}{f_q(\gamma_0)D^2}r\right)-2k_+(0)f_q(\gamma_0)D\left|\frac{V_1}{f_q(\gamma_0)D^2}r\right|, & \text{if } r>0,\end{cases}\\
&=4k_+(0)\sup_r\begin{cases}\frac{V_1}{D}W_1(-r)-\frac{1}{2}\frac{V_1}{D}|r|, & \text{if } r\le 0,\\ \frac{\sqrt{V_1V_2}}{D}W_2(r)-\frac{1}{2}\frac{V_1}{D}|r|, & \text{if } r>0,\end{cases}\\
&=4k_+(0)\frac{V_1}{D}\sup_r\begin{cases}W_1(-r)-\frac{1}{2}|r|, & \text{if } r\le 0,\\ \sqrt{\lambda}W_2(r)-\frac{1}{2}|r|, & \text{if } r>0.\end{cases}
\end{aligned}$$

So $\sup_v\{C(v)\}=4k_+(0)\frac{V_1}{D}\max\{M_1,M_2\}=:4k_+(0)\frac{V_1}{D}M$, where $M_1=\sup_{r\le 0}\left\{W_1(-r)-\frac{1}{2}|r|\right\}$, $M_2=\sup_{r\ge 0}\left\{\sqrt{\lambda}W_2(r)-\frac{1}{2}|r|\right\}$, and M_1 and M_2 are independent. From Bhattacharya and Brockwell (1976), M_1 follows the standard exponential function, and M_2 follows an exponential distribution with mean λ. It follows that

$$P(M\le x)=P(M_1\le x,M_2\le x)=P(M_1\le x)P(M_2\le x)=(1-e^{-x})(1-e^{-x/\lambda}).$$

By Slutsky's theorem, the required result follows. ■

Proof of Theorem 7. Because

$$\begin{aligned}\widehat{e}_i &= y_i - \mathbf{x}_i'\widehat{\beta} - \mathbf{x}_i'\widehat{\delta}1\left(q_i \le \widehat{\gamma}\right) \\ &= u_i + \left[m_i - \mathbf{x}_i'\widehat{\beta} - \mathbf{x}_i'\widehat{\delta}1\left(q_i \le \widehat{\gamma}\right)\right] \\ &\equiv u_i + D_i,\end{aligned}$$

we decompose $I_n^{(1)}$ as

$$\begin{aligned}I_n^{(1)} &= \tfrac{nh^{d/2}}{n(n-1)}\sum_i\sum_{j\neq i}\left[D_iD_j + u_iu_j + 2u_iD_j\right]K_{h,ij} \\ &\equiv I_{1n}^{(1)} + I_{2n}^{(1)} + I_{3n}^{(1)}.\end{aligned}$$

We complete the proof by examining $I_{1n}^{(1)}, I_{2n}^{(1)}, I_{3n}^{(1)}$, and showing that $v_n^{(1)2} = \Sigma^{(1)} + o_p\left(1\right)$ under $H_0^{(1)}$ and the local alternative and $v_n^{(1)2} = O_p(1)$ under $H_1^{(1)}$. Throughout this proof, $z_i = \left(x_i', q_i, u_i\right)'$ and $\mathbb{E}_i\left[\cdot\right] = \mathbb{E}\left[\cdot|x_i, q_i\right]$.

It is shown in Proposition 12 that $I_{1n}^{(1)} = O_p(h^{d/2})$ under $H_0^{(1)}$ and converges to Δ under the local alternative. It can also be shown that $I_{3n}^{(1)} = O_p(h^{d/2})$ under $H_0^{(1)}$ and is dominated by I_{1n} under the alternative, see, e.g., Zheng (1996). Proposition 14 shows that $I_{2n}^{(1)} \xrightarrow{d} N\left(0, \Sigma^{(1)}\right)$, and Proposition 15 shows the results related to $v_n^{(1)2}$. The proof is then complete. ■

Proof of Theorem 8. First, decompose $I_n^{(2)}$ by using (10):

$$\begin{aligned}&I_n^{(2)} \\ &= \tfrac{nh^{d/2}}{n(n-1)}\sum_i\sum_{j\neq i}1_i^\Gamma 1_j^\Gamma\left\{\left(m_i - \widehat{m}_i\right)\left(m_j - \widehat{m}_j\right) + u_iu_j + \widehat{u}_i\widehat{u}_j + 2u_i\left(m_j - \widehat{m}_j\right) - 2\widehat{u}_i\left(m_j - \widehat{m}_j\right) - 2u_i\widehat{u}_j\right\}K_{h,ij} \\ &\equiv I_{1n}^{(2)} + I_{2n}^{(2)} + I_{3n}^{(2)} + 2I_{4n}^{(2)} - 2I_{5n}^{(2)} - 2I_{6n}^{(2)}.\end{aligned}$$

We complete the proof by examining $I_{1n}^{(2)}, \cdots, I_{6n}^{(2)}$, and showing that $v_n^{(2)2} = \Sigma^{(2)} + o_p\left(1\right)$ under both $H_0^{(2)}$ and $H_1^{(2)}$. Throughout this proof, $z_i = \left(x_i', q_i, u_i\right)'$ and $\mathbb{E}_i\left[\cdot\right] = \mathbb{E}\left[\cdot|x_i, q_i\right]$. We show that $I_{2n}^{(2)}$ contributes to the asymptotic distribution under the null, and $I_{1n}^{(2)}$ contributes to the power under the local alternative. All other terms will not contribute to the asymptotic distribution under either the null or the alternative; that proof just extends Propositions 3, 4, 5 and 6 in Appendix B of Porter and Yu (2011), so it is omitted here. The remaining part of the proof concentrates on $I_{1n}^{(2)}$ and $I_{2n}^{(2)}$, and we only briefly mention the results for the other terms since these are obtained in a similar fashion.

First, $I_{2n}^{(2)}$, $I_{3n}^{(2)}$ and $I_{6n}^{(2)}$ are invariant under $H_0^{(2)}$ and $H_1^{(2)}$. It can be shown that $I_{3n}^{(2)}$ and $I_{6n}^{(2)}$ are both $o_p(1)$. Proposition 14 shows that $I_{2n}^{(2)} \xrightarrow{d} N\left(0, \Sigma^{(2)}\right)$.

Under $H_0^{(2)}$, Proposition 13 shows that $I_{1n}^{(2)} = o_{P_m}\left(1\right)$, and it can also be shown that I_{4n} and I_{5n} are both $o_{P_m}\left(1\right)$ uniformly in $m(\cdot) \in H_0$.

Under $H_1^{(2)}$, it can be shown that $I_{4n}^{(2)}$ and $I_{5n}^{(2)}$ are dominated by $I_{1n}^{(2)}$, and Proposition 13 shows that $I_{1n}^{(2)} = O_p\left(nh^{d/2}b\right)$ under $H_1^{(2)}$. The local power can be easily obtained from the proof of Proposition 13.

Finally, Proposition 16 shows that $v_n^{(2)2} = \Sigma^{(2)} + o_p(1)$ under both $H_0^{(2)}$ and $H_1^{(2)}$. So the proof is complete. ■

Proof of Theorem 9. This proof is similar but more tedious than the proofs of Theorem 7 and 8. Note that $\Phi\left(z\right)$ is a continuous function. By Pólya's theorem, it suffices to show that for any fixed value of $z \in \mathbb{R}$, $\left|P\left(T_n^{(\ell)*} \le z|\mathcal{F}_n\right) - \Phi(z)\right| = o_p\left(1\right)$.

For the first test, let

$$D_i^* = \mathbf{x}_i'\widehat{\beta} + \mathbf{x}_i'\widehat{\delta}1\left(q_i \le \widehat{\gamma}\right) - \mathbf{x}_i'\widehat{\beta}^* - \mathbf{x}_i'\widehat{\delta}^*1\left(q_i \le \widehat{\gamma}^*\right),$$

where $\left(\widehat{\beta}^*, \widehat{\delta}^*, \widehat{\gamma}^*\right)$ is the least squares estimator using the data $\{y_i^*, x_i, q_i\}_{i=1}^n$. Then

$$\begin{aligned} I_n^{(1)*} &= \tfrac{nh^{d/2}}{n(n-1)} \sum_i \sum_{j\neq i} \left[D_i^* D_j^* + u_i^* u_j^* + 2u_i^* D_j^*\right] K_{h,ij} \\ &\equiv I_{1n}^{(1)*} + I_{2n}^{(1)*} + I_{3n}^{(1)*}. \end{aligned}$$

The theorem is proved if we can show that $I_{in}^{(1)*}|\mathcal{F}_n = o_p(1)$ for $i = 1$ and 3 and $I_{2n}^{(1)*}/v_n^{(1)*}|\mathcal{F}_n \to \mathcal{N}(0,1)$ in probability. The first part can be proved as in the proof of Theorem 7, and, for the second part, see the discussion below.

For the second test, denote $m_i^* = \widehat{y}_i$ and define $\widehat{m}_i^*$ and $\widehat{u}_i^*$ by

$$\widehat{m}_i^* = \tfrac{1}{n-1} \sum_{j\neq i} m_j^* L_{b,ij} \Big/ \widehat{f}_i,$$

and

$$\widehat{u}_i^* = \tfrac{1}{n-1} \sum_{j\neq i} u_j^* L_{b,ij} \Big/ \widehat{f}_i.$$

Then using $\widehat{e}_i^* = y_i^* - \widehat{y}_i^* = m_i^* + u_i^* - (\widehat{m}_i^* + \widehat{u}_i^*)$, we get

$$\begin{aligned} &I_n^{(2)*} \\ &= \tfrac{nh^{d/2}}{n(n-1)} \sum_i \sum_{j\neq i} 1_i^\Gamma 1_j^\Gamma \left\{(m_i^* - \widehat{m}_i^*)\left(m_j^* - \widehat{m}_j^*\right) + u_i^* u_j^* + \widehat{u}_i^* \widehat{u}_j^* + 2u_i^*\left(m_j^* - \widehat{m}_j^*\right) - 2\widehat{u}_i^*\left(m_j^* - \widehat{m}_j^*\right) - 2u_i^* \widehat{u}_j^*\right\} K_{h,ij} \\ &\equiv I_{1n}^{(2)*} + I_{2n}^{(2)*} + I_{3n}^{(2)*} + 2I_{4n}^{(2)*} - 2I_{5n}^{(2)*} - 2I_{6n}^{(2)*}. \end{aligned}$$

The theorem is proved if we can show that $I_{in}^{(2)*}|\mathcal{F}_n = o_p(1)$ for $i = 1,3,4,5,6$ and $I_{2n}^{(2)*}/v_n^{(2)*}|\mathcal{F}_n \to N(0,1)$ in probability. The first part is similar to that of Theorem 8 under $H_0^{(2)}$. However, note that $m^*(\cdot)|\mathcal{F}_n$ as defined above satisfies $H_0^{(2)}$ even if $m(\cdot)$ is from $H_1^{(2)}$; see Gu et al. (2007) for a similar analysis in testing omitted variables. But there is some differences in showing the second part.

First, because $u_i^*|\mathcal{F}_n$ are mean zero and mutually independent and have variance $\widehat{e}_i^2$,

$$\tfrac{nh^{d/2}}{n(n-1)} \sum_i \sum_{j\neq i} 1_i^\Gamma 1_j^\Gamma u_i^* u_j^* K_{h,ij} = \tfrac{2nh^{1/2}}{n(n-1)} \sum_i \sum_{j>i} 1_i^\Gamma 1_j^\Gamma u_i^* u_j^* K_{h,ij} \equiv \sum_i \sum_{j>i} U_{n,ij}^*$$

is a second order degenerate U-statistic with conditional variance

$$\tfrac{2h^d}{n(n-1)} \sum_i \sum_{j\neq i} 1_i^\Gamma 1_j^\Gamma \widehat{e}_i^2 \widehat{e}_j^2 K_{h,ij}^2 = v_n^2.$$

Because $U_{n,ij}^*$ depends on i and j, we use the central limit theorem of de Jong (1987) for generalized quadratic forms rather than Hall (1984) to find the asymptotic distribution of $I_{2n}^{(2)*}$. From his Proposition 3.2, we know $I_{2n}^{(2)*}/v_n^{(2)}|\mathcal{F}_n \to \mathcal{N}(0,1)$ in probability as long as

$$\begin{aligned} &G_I^* = \sum_i \sum_{j>i} \mathbb{E}^*\left[U_{n,ij}^{*4}\right] = o_p\left(v_n^{(2)4}\right), \\ &G_{II}^* = \sum_i \sum_{j>i} \sum_{l>j>i} \mathbb{E}^*\left[U_{n,ij}^{*2} U_{n,il}^{*2} + U_{n,ji}^{*2} U_{n,jl}^{*2} + U_{n,li}^{*2} U_{n,lj}^{*2}\right] = o_p\left(v_n^{(2)4}\right), \\ &G_{IV}^* = \sum_i \sum_{j>i} \sum_{k>j>i} \sum_{l>k>j>i} \mathbb{E}^*\left[U_{n,ij}^* U_{n,ik}^* U_{n,lj}^* U_{n,lk}^* + U_{n,ij}^* U_{n,il}^* U_{n,kj}^* U_{n,kl}^* + U_{n,ik}^* U_{n,il}^* U_{n,jk}^* U_{n,jl}^*\right] = o_p\left(v_n^{(2)4}\right). \end{aligned}$$

It is straightforward to show that

$$G_I^* = O_p\left(\left(n^2 h^d\right)^{-1}\right), G_{II}^* = O_p(n^{-1}), G_{IV}^* = O_p(h^d),$$

see, e.g., the proof of Theorem 2 of Hsiao et al. (2007), so the result follows by $v_n^{(2)4} = O_p(1)$. Next, it is easy to check that $\mathbb{E}^*\left[v_n^{(2)*2}\right] = v_n^{(2)2} + o_p(1)$, and $Var^*\left(v_n^{(2)*2}\right) = o_p(1)$. Thus $I_{2n}^{(2)*}/v_n^{(2)*}|\mathcal{F}_n \to \mathcal{N}(0,1)$ in probability. The analysis for $I_{2n}^{(1)*}$ is similar. ■

Supplement B: Propositions

Proposition 1 $\widehat{\gamma} - \gamma_0 = O_p(h)$.

Proof. We apply Lemma 4 of Porter and Yu (2015) to prove the result. By Lemma B.1 of Newey (1994), we have

$$\sup_{\gamma\in\Gamma}\left|\widehat{Q}_n(\gamma) - Q_n(\gamma)\right| = O_p\left(\sqrt{\ln n/nh^d}\right) \xrightarrow{p} 0 \ ,$$

where

$$Q_n(\gamma) = \int \begin{bmatrix} \int_{-1}^{0}\int K^x(u_x,x)\, k_-(u_q)\, m(x+u_xh, \gamma+u_qh)\, f(x+u_xh, \gamma+u_qh)\, du_x du_q \\ -\int_0^1\int K^x(u_x,x)\, k_+(u_q)\, m(x+u_xh, \gamma+u_qh)\, f(x+u_xh, \gamma+u_qh)\, du_x du_q \end{bmatrix}^2 f(x)\, dx \ .$$

Let $\mathcal{N}_n = [\gamma_0 - h, \gamma_0 + h]$ and $\gamma_n = \arg\max_{\gamma\in\Gamma} Q_n(\gamma)$; then it is easy to show that $\sup_{\gamma\in\Gamma\backslash\mathcal{N}_n} Q_n(\gamma) = O\left(h^2\right)$. But for $\gamma \in \mathcal{N}_n$, the result is different. Specifically, let $\gamma = \gamma_0 + ah$, $a \in (0,1)$. Then

$$Q_n(\gamma) = \int \begin{bmatrix} \int_{-1}^{0}\int K^x(u_x,x)\, k_-(u_q)\, g(x+u_xh, \gamma+u_qh)\, f(x+u_xh, \gamma+u_qh)\, du_x du_q \\ \int_{-1}^{-a}\int K^x(u_x,x)\, k_-(u_q)\, (1, x'+hu_x', \gamma+u_qh)\, \delta_0 f(x+u_xh, \gamma+u_qh)\, du_x du_q \\ -\int_0^1\int K^x(u_x,x)\, k_+(u_q)\, g(x+u_xh, \gamma+u_qh)\, f(x+u_xh, \gamma+u_qh)\, du_x du_q \end{bmatrix}^2 f(x)\, dx \ .$$

The differences of the first and the third terms in brackets are still $O\left(h^2\right)$, so the second term will dominate. With Assumption I, we have $\int_x\left[\int K^x(u_x,x)\,(1,x,\gamma_0)'\,\delta_0 f(x,\gamma_0)\,du_x\right]^2 f(x)\,dx > C$. Under Assumption K, if $a \in (0,1)$, then $\int_{-1}^{-a} k_-(u_q)\,du_q \leq 1$ and $\int_{-1}^{-a} k_-(u_q)\,du_q$ is a nonincreasing function of a for $a \in (0,1)$. As a result, $Q_n(\gamma)$ is a nonincreasing function of a for $a \in (0,1)$ up to $O\left(h^2\right)$. Similarly, $Q_n(\gamma)$ is a nondecreasing function of a for $a \in (-1,0)$ up to $O\left(h^2\right)$. So $Q_n(\gamma)$ is maximized at some $\gamma_n \in \mathcal{N}_n$ such that $Q_n(\gamma_n) > \sup_{\gamma\in\Gamma\backslash\mathcal{N}_n}|Q_n(\gamma)| + C/2$ for n large enough. The result of interest then follows. ■

Proposition 2 $\widehat{\gamma} - \gamma_0 = O_p\left((n/h)^{-1/2}\right)$.

Proof. We apply the standard shelling method to obtain the result. Specifically, for each n, the parameter space is partitioned into the 'shells' $S_{l,n} = \left\{\gamma : 2^{l-1} \leq (n/h)^{1/2}|\gamma - \gamma_0| \leq 2^l\right\}$ with l ranging over the integers. If $(n/h)^{1/2}|\widehat{\gamma} - \gamma_0|$ is larger than 2^L for a given integer L, then $\widehat{\gamma}$ is in one of the shells $S_{l,n}$ with $l \geq L$. In that case the supremum of the map $\gamma \to \widehat{Q}_n(\gamma) - \widehat{Q}_n(\gamma_0)$ over this shell is nonnegative by the property of $\widehat{\gamma}$. Note that

$$
\begin{aligned}
& P\left((n/h)^{1/2}\left|\widehat{\gamma}-\gamma_0\right|>2^L\right) \\
\leq & P\left(\sup_{2^L<(n/h)^{1/2}|\widehat{\gamma}-\gamma_0|<(n/h)^{1/2}h}\left(\frac{1}{n}\sum_{i=1}^n \widehat{\Delta}_i^2(\gamma)-\frac{1}{n}\sum_{i=1}^n \widehat{\Delta}_i^2(\gamma_0)\right)\geq 0\right)+P\left(\left|\widehat{\gamma}-\gamma_0\right|\geq h\right) \\
\leq & \sum_{l=L}^{\frac{1}{2}\log_2(nh)} P\left(\sup_{S_{l,n}}\frac{1}{n}\sum_{i=1}^n \widehat{\Delta}_i^2(\gamma)\geq \frac{1}{n}\sum_{i=1}^n \widehat{\Delta}_i^2(\gamma_0)\right)+P\left(\left|\widehat{\gamma}-\gamma_0\right|\geq h\right) \\
\leq & \sum_{l=L}^{\frac{1}{2}\log_2(nh)} P\left(\sup_{S_{l,n}}\frac{1}{n}\sum_{i=1}^n \widehat{\Delta}_i^2(\gamma)1(\Delta_i>0)\geq \frac{1}{n}\sum_{i=1}^n \widehat{\Delta}_i^2(\gamma_0)1(\Delta_i>0)\right) \\
& +\sum_{l=L}^{\frac{1}{2}\log_2(nh)} P\left(\sup_{S_{l,n}}\frac{1}{n}\sum_{i=1}^n \widehat{\Delta}_i^2(\gamma)1(\Delta_i<0)\geq \frac{1}{n}\sum_{i=1}^n \widehat{\Delta}_i^2(\gamma_0)1(\Delta_i<0)\right) \\
& +P\left(\left|\widehat{\gamma}-\gamma_0\right|\geq h\right) \\
=: & T1+T2+T3.
\end{aligned}
$$

As $T3$ converges to zero by Proposition 1 and $T2$ is similar to $T1$, we only use $T1$ to illustrate the derivations in the following discussion. Since

$$
\begin{aligned}
T1 \leq & \sum_{l=L}^{\frac{1}{2}\log_2(nh)} P\left(\sup_{S_{l,n}}\frac{1}{n}\sum_{i=1}^n\left(\widehat{\Delta}_i(\gamma)-\widehat{\Delta}_i(\gamma_0)\right)1(\Delta_i>0)>0\right) \\
& +\sum_{l=L}^{\frac{1}{2}\log_2(nh)} P\left(\sup_{S_{l,n}}\frac{1}{n}\sum_{i=1}^n\left(\widehat{\Delta}_i(\gamma)+\widehat{\Delta}_i(\gamma_0)\right)1(\Delta_i>0)<0\right),
\end{aligned}
$$

we focus on the first term because the second term is easier to analyze given that $\Delta_i>0$. To simplify notations, $1(\Delta_i>0)$ is neglected in the remaining proof. Notice that

$$
\begin{aligned}
& \frac{1}{n}\sum_{i=1}^n\left(\widehat{\Delta}_i(\gamma)-\widehat{\Delta}_i(\gamma_0)\right) \\
= & \frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j=1,j\neq i}^n\left(y_jK_{h,ij}^{\gamma-}-y_jK_{h,ij}^{\gamma+}\right)-\frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j=1,j\neq i}^n\left(y_jK_{h,ij}^{\gamma_0-}-y_jK_{h,ij}^{\gamma_0+}\right) \\
= & \frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j=1,j\neq i}^n\left(m_jK_{h,ij}^{\gamma-}-m_jK_{h,ij}^{\gamma+}\right)-\frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j=1,j\neq i}^n\left(m_jK_{h,ij}^{\gamma_0-}-m_jK_{h,ij}^{\gamma_0+}\right) \\
& +\frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j=1,j\neq i}^n\left(u_jK_{h,ij}^{\gamma-}-u_jK_{h,ij}^{\gamma+}\right)-\frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j=1,j\neq i}^n\left(u_jK_{h,ij}^{\gamma_0-}-u_jK_{h,ij}^{\gamma_0+}\right) \\
=: & D1+D2,
\end{aligned}
$$

where $m_j=g_j+\left(1,x_j',q_j\right)\delta_0 1\left(q_j\leq\gamma_0\right)$ with $g_j=g\left(x_j,q_j\right)$. Suppose $\gamma_0<\gamma<\gamma_0+h$, then for some $C>0$, we have

$$
\begin{aligned}
D1 = & \frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j=1,j\neq i}^n g_j\left(K_{h,ij}^{\gamma-}-K_{h,ij}^{\gamma_0-}\right)-\frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j=1,j\neq i}^n g_j\left(K_{h,ij}^{\gamma+}-K_{h,ij}^{\gamma_0+}\right) \\
& +\frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j=1,j\neq i}^n\left(1,x_j^{'},q_j\right)\delta_0\left(K_{h,ij}^{\gamma-}-K_{h,ij}^{\gamma_0-}\right)1\left(q_j\leq\gamma_0\right) \\
\leq & -C\left(\frac{\gamma-\gamma_0}{h}\right)^2
\end{aligned}
$$

with probability approaching 1 by calculating the mean and variance of $D1$ in its U-projection, where the difference of the first two terms contribute only $O_p\left(\left|\gamma-\gamma_0\right|^2\right)$, and the third term contributes to $-C\left(\frac{\gamma-\gamma_0}{h}\right)^2$ because for each i, $K_{h,ij}^{\gamma-}$ covers $O_p\left(n\left(\gamma-\gamma_0\right)\right)$ terms less than $K_{h,ij}^{\gamma_0-}$ given that $\gamma>\gamma_0$ and

$k_{\pm}(0)=0$. In consequence, for $\eta \le h$,

$$P\left(\sup_{|\gamma-\gamma_0|<\eta} \tfrac{1}{n}\sum_{i=1}^{n}\left(\widehat{\Delta}_i(\gamma)-\widehat{\Delta}_i(\gamma_0)\right)>0\right) \le P\left(\sup_{|\gamma-\gamma_0|<\eta} D2 > C\left(\tfrac{\gamma-\gamma_0}{h}\right)^2\right).$$

Next,

$$\begin{aligned}
& D2 \\
&= \tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} u_j\left(K_{h,ij}^{\gamma-}-K_{h,ij}^{\gamma_0-}\right)1(q_j\le\gamma_0)+\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} u_j\left(K_{h,ij}^{\gamma_0+}-K_{h,ij}^{\gamma+}\right)1(q_j>\gamma) \\
&\quad + \tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} u_j\left(K_{h,ij}^{\gamma-}+K_{h,ij}^{\gamma_0+}\right)1(\gamma_0<q_j\le\gamma) \\
&=: P1+P2+P3,
\end{aligned} \tag{30}$$

and we can apply Lemma 8.4 of Newey and Mcfadden (1994) to bound $D2$. Since the first two terms are similar, we just check the first term and the last term. For the first term, set

$$m_n(z_i,z_j)=u_j\left(K_{h,ij}^{\gamma-}-K_{h,ij}^{\gamma_0-}\right)1(q_j\le\gamma_0)\ ,$$

where $z_i=(u_i,x_i',q_i)'$, and $m_n(z_i,z_j)=0$ for any $i=j$. So we have

$$n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n} m_n(z_i,z_j)=P1\ ,$$

$$\mathbb{E}[m_n(z_i,z_j)]=0,\ \mathbb{E}_i[m_n(z_i,z_j)]=0,\ \mathbb{E}[|m_n(z_1,z_1)|]/n=0\ ,$$

and

$$\begin{aligned}
& \mathbb{E}\left[\|m_n(z_1,z_2)\|^2\right] \\
&= \mathbb{E}\left[E\left[\left(\tfrac{u_j}{h^d}K^x\left(\tfrac{x_j-x_i}{h},x_i\right)\left[k^-\left(\tfrac{q_j-\gamma}{h}\right)-k^-\left(\tfrac{q_j-\gamma_0}{h}\right)\right]1(q_j\le\gamma_0)\right)^2|x_i\right]\right] \\
&= \mathbb{E}\left[E\left[\tfrac{u_j^2}{h^{2d}}K^x\left(\tfrac{x_j-x_i}{h},x_i\right)^2\left[k^-\left(\tfrac{q_j-\gamma}{h}\right)-k^-\left(\tfrac{q_j-\gamma_0}{h}\right)\right]^2 1(q_j\le\gamma_0)|x_i\right]\right] \\
&= \mathbb{E}\left[\tfrac{1}{h^{2d}}\int_{-\infty}^{\gamma_0}\int_x\int_u u_j^2 K^x\left(\tfrac{x_j-x_i}{h},x_i\right)^2\left[k^-\left(\tfrac{q_j-\gamma}{h}\right)-k^-\left(\tfrac{q_j-\gamma_0}{h}\right)\right]^2 f(u_j|x_j,q_j)f(x_j,q_j)\,du_j dx_j dq_j\right] \\
&= \mathbb{E}\left[\tfrac{1}{h^d}\int_{-1}^{0}\int_{u_x}\sigma^2(x_i+u_xh,\gamma_0+u_qh)K^x(u_x,x_i)^2\left[k^-\left(u_q+\tfrac{\gamma_0-\gamma}{h}\right)-k^-(u_q)\right]^2 f(x_i+u_xh,\gamma_0+u_qh)\,du_x du_q\right] \\
&\le \tfrac{C}{h^d}\left(\tfrac{\gamma_0-\gamma}{h}\right)^2
\end{aligned}$$

under Assumption I and the fact that $\left|k'_-(\cdot)\right|<\infty$ and $K^x(\cdot)<\infty$ over their supports, where $\mathbb{E}_i[\cdot]=\mathbb{E}[\cdot|z_i]$, $\sigma^2(x_j,q_j)=\mathbb{E}\left[u_j^2|x_j,q_j\right]$. Hence, by Proposition 1,

$$\left(\mathbb{E}\left[\|m_n(z_1,z_2)\|^2\right]\right)^{1/2}\Big/ n \le C\left|\tfrac{\gamma_0-\gamma}{h}\right|\tfrac{1}{nh^{d/2}}=O_p(1)\,O_p\left(\tfrac{1}{nh^{d/2}}\right)=o_p(1)$$

under Assumption H. As a result,

$$P1=\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} u_j\left(K_{h,ij}^{\gamma-}-K_{h,ij}^{\gamma_0-}\right)1(q_j\le\gamma_0)=n^{-1}\sum_{j=1}^{n}\mathbb{E}[m_n(z_i,z_j)|z_j]+o_p(1)\ ,$$

where

$$\begin{aligned}
&\mathbb{E}_j\left[m_n\left(z_i,z_j\right)\right] \\
&= \tfrac{u_j}{h^d} 1\left(q_j \le \gamma_0\right)\left[k^-\left(\tfrac{q_j-\gamma}{h}\right)-k^-\left(\tfrac{q_j-\gamma_0}{h}\right)\right]\textstyle\int_x K^x\left(\tfrac{x_j-x_i}{h},x_i\right)f\left(x_i\right)dx_i \\
&= \tfrac{u_j}{h} 1\left(q_j \le \gamma_0\right)\left[k^-\left(\tfrac{q_j-\gamma}{h}\right)-k^-\left(\tfrac{q_j-\gamma_0}{h}\right)\right]\textstyle\int_x K^x\left(u_x,x_i\right)f\left(x_i-u_xh\right)du_x \\
&\le C\tfrac{u_j}{h} 1\left(q_j \le \gamma_0\right)\left[k^-\left(\tfrac{q_j-\gamma}{h}\right)-k^-\left(\tfrac{q_j-\gamma_0}{h}\right)\right]f\left(x_i\right).
\end{aligned}$$

Hence,

$$\begin{aligned}
Var\left(P1\right) &= Var\left(n^{-1}\textstyle\sum_{j=1}^n \mathbb{E}_j\left[m_n\left(z_i,z_j\right)\right]+o_p\left(1\right)\right) \\
&\le n^{-1}\textstyle\int_{-\infty}^{\gamma_0}\int_x\int_u C\tfrac{u_j^2}{h^2}\left[k^-\left(\tfrac{q_j-\gamma}{h}\right)-k^-\left(\tfrac{q_j-\gamma_0}{h}\right)\right]^2 f\left(x_j\right)^2 f\left(u_j|x_j,q_j\right)f\left(x_j,q_j\right)du_jdx_jdq_j \\
&\le C\tfrac{1}{nh}\textstyle\int_{-1}^0\int_x \sigma^2\left(x_j,\gamma_0+u_qh\right)\left[k^-\left(u_q+\tfrac{\gamma_0-\gamma}{h}\right)-k^-\left(u_q\right)\right]^2 f\left(x_j,\gamma_0+u_qh\right)f\left(x_j\right)^2dx_jdu_q \\
&\le C\tfrac{1}{nh}\left(\tfrac{\gamma_0-\gamma}{h}\right)^2.
\end{aligned}$$

Similarly for $P3$, we can set

$$P3=\tfrac{1}{n(n-1)}\textstyle\sum_{i=1}^n\sum_{j=1,j\ne i}^n u_j\left(K_{h,ij}^{\gamma-}+K_{h,ij}^{\gamma_0+}\right)1\left(\gamma_0<q_j\le\gamma\right)=n^{-2}\sum_{i=1}^n\sum_{j=1}^n m_n\left(z_i,z_j\right)$$

with

$$\mathbb{E}\left[m_n\left(z_i,z_j\right)\right]=0,\ \mathbb{E}_i\left[m_n\left(z_i,z_j\right)\right]=0,\ \mathbb{E}\left[\left\|m_n\left(z_1,z_1\right)\right\|\right]/n=0\ \ ,$$

and

$$\begin{aligned}
&\mathbb{E}\left[\left\|m_n\left(z_1,z_2\right)\right\|^2\right] \\
&=\mathbb{E}\left[E\left[\left(\tfrac{u_j}{h^d}K^x\left(\tfrac{x_j-x_i}{h},x_i\right)\left[k^-\left(\tfrac{q_j-\gamma}{h}\right)-k^+\left(\tfrac{q_j-\gamma_0}{h}\right)\right]1\left(\gamma_0<q_j\le\gamma\right)\right)^2|x_i\right]\right] \\
&=\mathbb{E}\left[\tfrac{1}{h^d}\textstyle\int_0^{\frac{\gamma-\gamma_0}{h}}\int_{u_x}\sigma^2\left(x_i+u_xh,\gamma_0+u_qh\right)K^x\left(u_x,x_i\right)^2\left[k^-\left(u_q+\tfrac{\gamma_0-\gamma}{h}\right)-k^+\left(u_q\right)\right]^2 f\left(x_i+u_xh,\gamma_0+u_qh\right)du_xdu_q\right] \\
&\le \tfrac{C}{h^d}\left|\tfrac{\gamma_0-\gamma}{h}\right|^3.
\end{aligned}$$

By Proposition 1, we have

$$\left(\mathbb{E}\left[\left\|m_n\left(z_1,z_2\right)\right\|^2\right]\right)^{1/2}\Big/ n\le C\left|\tfrac{\gamma_0-\gamma}{h}\right|^{3/2}\tfrac{1}{nh^{d/2}}=O_p\left(1\right)O_p\left(\tfrac{1}{nh^{d/2}}\right)=o_p\left(1\right)$$

under Assumption H. As a result,

$$P3=\tfrac{1}{n(n-1)}\textstyle\sum_{i=1}^n\sum_{j=1,j\ne i}^n u_j\left(K_{h,ij}^{\gamma-}-K_{h,ij}^{\gamma_0-}\right)1\left(q_j\le\gamma_0\right)=n^{-1}\sum_{j=1}^n\mathbb{E}_j\left[m_n\left(z_i,z_j\right)\right]+o_p\left(1\right)$$

with

$$\begin{aligned}
&\mathbb{E}_j\left[m_n\left(z_i,z_j\right)\right] \\
&= \tfrac{u_j}{h^d}\left[k^-\left(\tfrac{q_j-\gamma}{h}\right)-k^+\left(\tfrac{q_j-\gamma_0}{h}\right)\right]1\left(\gamma_0<q_j\le\gamma\right)\textstyle\int_x K^x\left(\tfrac{x_j-x_i}{h},x_i\right)f\left(x_i\right)dx_i \\
&= \tfrac{u_j}{h^d}\left[k^-\left(\tfrac{q_j-\gamma}{h}\right)-k^+\left(\tfrac{q_j-\gamma_0}{h}\right)\right]1\left(\gamma_0<q_j\le\gamma\right)\textstyle\int_x K^x\left(u_x,x_i\right)f\left(x_i-u_xh\right)du_x \\
&\le C\tfrac{u_j}{h^d}\left[k^-\left(\tfrac{q_j-\gamma}{h}\right)-k^+\left(\tfrac{q_j-\gamma_0}{h}\right)\right]1\left(\gamma_0<q_j\le\gamma\right)f\left(x_i\right).
\end{aligned}$$

Hence,

$$\begin{aligned}
Var\left(P3\right) &\approx Var\left(n^{-1}\sum_{j=1}^{n}\mathbb{E}_j\left[m_n\left(z_i,z_j\right)\right]\right)\\
&\leq C\frac{1}{nh}\int_0^{\frac{\gamma-\gamma_0}{h}}\int_x \sigma^2\left(x_j,\gamma_0+u_qh\right)\left[k^-\left(u_q+\frac{\gamma_0-\gamma}{h}\right)-k^+\left(u_q\right)\right]^2 f\left(x_j,\gamma_0+u_qh\right)f\left(x_j\right)^2dx_jdu_q\\
&\leq C\frac{1}{nh}\left|\frac{\gamma_0-\gamma}{h}\right|^3.
\end{aligned}$$

Since, conditional on x_i, the three summations on the right hand side of (30) are independent, we obtain

$$\begin{aligned}
Var\left(D2\right) &= Var\left(P1\right)+Var\left(P2\right)+Var\left(P3\right)\\
&\leq C\frac{1}{nh}\left[\left(\frac{\gamma_0-\gamma}{h}\right)^2+\left|\frac{\gamma_0-\gamma}{h}\right|^3\right]\leq C\frac{(\gamma-\gamma_0)^2}{nh^3}
\end{aligned}$$

uniformly for $|\gamma-\gamma_0|<\eta$. In consequence,

$$\begin{aligned}
&P\left(\sup_{|\gamma-\gamma_0|<\eta}\frac{1}{n}\sum_{i=1}^{n}\left(\widehat{\Delta}_i\left(\gamma\right)-\widehat{\Delta}_i\left(\gamma_0\right)\right)>0\right)\\
&\leq C\mathbb{E}\left[\left(\sup_{|\gamma-\gamma_0|<\eta}D2\right)^2\right]\Big/\left[C\left(\frac{\gamma-\gamma_0}{h}\right)^2\right]^2\leq C\frac{(\gamma-\gamma_0)^2}{nh^3}\Big/\frac{(\gamma-\gamma_0)^4}{h^4}=\frac{Ch}{n(\gamma-\gamma_0)^2}
\end{aligned}$$

by Markov's inequality. So

$$\begin{aligned}
&\sum_{l=L}^{\frac{1}{2}\log_2(nh)}P\left(\sup_{S_{l,n}}\frac{1}{n}\sum_{i=1}^{n}\left(\widehat{\Delta}_i\left(\gamma\right)-\widehat{\Delta}_i\left(\gamma_0\right)\right)1\left(\Delta_i>0\right)>0\right)\\
&\leq\sum_{l\geq L}\frac{Ch}{n\left(2^l/\sqrt{n/h}\right)^2}=C\sum_{l\geq L}\frac{1}{4^l}\to 0
\end{aligned}$$

as $L\to\infty$. The proof is completed. ■

Proposition 3 *For v on any compact set,*

$$\mathbb{E}\left[nh\left(\widehat{Q}_n\left(\gamma_0^v\right)-\widehat{Q}_n(\gamma_0)\right)\right]=-v^2\mathbb{E}\left[\Delta_f\left(x_i\right)\Delta_i f(x_i)|q_i=\gamma_0\right]f_q(\gamma_0)k_+'(0)+o(1),$$

$$\begin{aligned}
&Cov\left(nh\left(\widehat{Q}_n\left(\gamma_0^{v_1}\right)-\widehat{Q}_n\left(\gamma_0\right)\right),nh\left(\widehat{Q}_n\left(\gamma_0^{v_2}\right)-\widehat{Q}_n\left(\gamma_0\right)\right)\right)\\
&=4\mathbb{E}\left[\Delta_f^2(x_i)f^2\left(x_i\right)\left(\sigma_+^2(x_i)+\sigma_-^2(x_i)\right)|q_i=\gamma_0\right]f_q(\gamma_0)v_1v_2\xi_{(1)}+o(v_1v_2),
\end{aligned}$$

and

$$\frac{1}{\sqrt{Var\left(nh\left(\widehat{Q}_n\left(\gamma_0^v\right)-\widehat{Q}_n(\gamma_0)\right)\right)}}\left(nh\left(\widehat{Q}_n\left(\gamma_0^v\right)-\widehat{Q}_n(\gamma_0)\right)-\mathbb{E}\left[nh\left(\widehat{Q}_n\left(\gamma_0^v\right)-\widehat{Q}_n(\gamma_0)\right)\right]\right)\xrightarrow{d}N\left(0,1\right),$$

where $\gamma_0^v=\gamma_0+\frac{v}{\sqrt{n/h}}$.

Proof. Note that

$$\widehat{Q}_n\left(\gamma_0^v\right)-\widehat{Q}_n(\gamma_0)=\frac{1}{n}\sum_{i=1}^{n}\left(\widehat{\Delta}_i\left(\gamma_0^v\right)^2-\widehat{\Delta}_i\left(\gamma_0\right)^2\right)=\frac{1}{n}\sum_{i=1}^{n}\left(\widehat{\Delta}_i\left(\gamma_0^v\right)+\widehat{\Delta}_i\left(\gamma_0\right)\right)\left(\widehat{\Delta}_i\left(\gamma_0^v\right)-\widehat{\Delta}_i\left(\gamma_0\right)\right).$$

By Lemma B.1 of Newey(1994), we can show that $\left|\widehat{\Delta}_i\left(\gamma_0^v\right)-\Delta_f(x_i)\right|\xrightarrow{p}0$ uniformly in i and v, where $\Delta_f(x_i):=(1,x_i',\gamma_0)\,\delta_0 f(x_i,\gamma_0)=O_p(1)$. So $\left|\widehat{\Delta}_i\left(\gamma_0^v\right)+\widehat{\Delta}_i\left(\gamma_0\right)-2\Delta_f(x_i)\right|\xrightarrow{p}0$ uniformly in i and v. Next we focus on the other term. For simplicity, let $v > 0$. Now,

$$
\begin{aligned}
&\widehat{\Delta}_i\left(\gamma_0^v\right)-\widehat{\Delta}_i\left(\gamma_0\right)\\
&=\left(\tfrac{1}{n-1}\sum_{j=1,j\neq i}^{n} y_j K_{h,ij}^{\gamma_0^v-}-\tfrac{1}{n-1}\sum_{j=1,j\neq i}^{n} y_j K_{h,ij}^{\gamma_0^v+}\right)-\left(\tfrac{1}{n-1}\sum_{j=1,j\neq i}^{n} y_j K_{h,ij}^{\gamma_0-}-\tfrac{1}{n-1}\sum_{j=1,j\neq i}^{n} y_j K_{h,ij}^{\gamma_0+}\right)\\
&=\left(\tfrac{1}{n-1}\sum_{j=1,j\neq i}^{n} y_j K_{h,ij}^{\gamma_0^v-}-\tfrac{1}{n-1}\sum_{j=1,j\neq i}^{n} y_j K_{h,ij}^{\gamma_0-}\right)-\left(\tfrac{1}{n-1}\sum_{j=1,j\neq i}^{n} y_j K_{h,ij}^{\gamma_0^v+}-\tfrac{1}{n-1}\sum_{j=1,j\neq i}^{n} y_j K_{h,ij}^{\gamma_0+}\right)\\
&=:\tfrac{1}{n-1}\sum_{j=1,j\neq i}^{n}\left(T_{1ij}+T_{2ij}+T_{3ij}+T_{4ij}+T_{5ij}+T_{6ij}\right),
\end{aligned}
$$

where

$$
\begin{aligned}
T_{1ij}&=g\left(x_j,q_j\right)\left(K_{h,ij}^{\gamma_0^v-}-K_{h,ij}^{\gamma_0-}\right),\\
T_{2ij}&=-g\left(x_j,q_j\right)\left(K_{h,ij}^{\gamma_0^v+}-K_{h,ij}^{\gamma_0+}\right),\\
T_{3ij}&=u_j\left(K_{h,ij}^{\gamma_0^v-}-K_{h,ij}^{\gamma_0-}\right),\\
T_{4ij}&=u_j\left(K_{h,ij}^{\gamma_0^v+}-K_{h,ij}^{\gamma_0+}\right),\\
T_{5ij}&=-\left(1,x_j',q_j\right)\delta_0 K_{h,ij}^{\gamma_0-}1\left(\gamma_0-h\le q_j\le\gamma_0\right),\\
T_{6ij}&=-\left(1,x_j',q_j\right)\delta_0 K_{h,ij}^{\gamma_0^v-}1\left(\gamma_0^v-h\le q_j\le\gamma_0\right).
\end{aligned}
$$

By Lemma 1, we have

$$
\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}2\Delta_f\left(x_i\right)\left(T_{1ij}+T_{2ij}\right)\approx 0,
$$

by Lemma 2, we have

$$
\begin{aligned}
&\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}2\Delta_f\left(x_i\right)\left(T_{3ij}+T_{4ij}\right)\\
&\approx\tfrac{2}{nh}\sum_{i=1}^{n}\Delta_f\left(x_i\right)f\left(x_i\right)u_i\left[\left(k^-\left(\tfrac{q_i-\gamma_0^v}{h}\right)-k^-\left(\tfrac{q_i-\gamma_0}{h}\right)\right)-\left(k^+\left(\tfrac{q_i-\gamma_0^v}{h}\right)-k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right)\right],
\end{aligned}
$$

and by Lemma 3, we have

$$
\begin{aligned}
&\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}2\Delta_f\left(x_i\right)\left(T_{5ij}+T_{6ij}\right)\\
&\approx\tfrac{2}{nh}\sum_{i=1}^{n}\Delta_f\left(x_i\right)f\left(x_i\right)\left(1,x_i',q_i\right)\delta_0\left[k^-\left(\tfrac{q_i-\gamma_0^v}{h}\right)1\left(\gamma_0^v-h\le q_j\le\gamma_0\right)-k^-\left(\tfrac{q_i-\gamma_0}{h}\right)1\left(\gamma_0-h\le q_i\le\gamma_0\right)\right].
\end{aligned}
$$

Combining results we have

$$
\begin{aligned}
&nh\left(\widehat{Q}_n\left(\gamma_0^v\right)-\widehat{Q}_n(\gamma_0)\right)\\
&=2\sum_{i=1}^{n}\Delta_f\left(x_i\right)f\left(x_i\right)u_i\left[\left(k^-\left(\tfrac{q_i-\gamma_0^v}{h}\right)-k^-\left(\tfrac{q_i-\gamma_0}{h}\right)\right)-\left(k^+\left(\tfrac{q_i-\gamma_0^v}{h}\right)-k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right)\right]\\
&\quad+2\sum_{i=1}^{n}\Delta_f\left(x_i\right)f\left(x_i\right)\left(1,x_i',q_i\right)\delta_0\left[k^-\left(\tfrac{q_i-\gamma_0^v}{h}\right)1\left(\gamma_0^v-h\le q_j\le\gamma_0\right)-k^-\left(\tfrac{q_i-\gamma_0}{h}\right)1\left(\gamma_0-h\le q_i\le\gamma_0\right)\right]\\
&\quad+o_p\left(1\right)\\
&=:S_1+S_2+o_p\left(1\right).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&\mathbb{E}\left[nh\left(\widehat{Q}_n\left(\gamma_0^v\right)-\widehat{Q}_n(\gamma_0)\right)\right]\approx\mathbb{E}\left[S_2\left(v\right)\right]\\
&=2n\int_{x_i}\int_{\gamma_0^v-h}^{\gamma_0}\Delta_f\left(x_i\right)f\left(x_i\right)\left(1,x_j',q_j\right)\delta_0 k^-\left(\tfrac{q_i-\gamma_0^v}{h}\right)f\left(x_i,q_i\right)dq_i dx_i\\
&\quad-2n\int_{x_i}\int_{\gamma_0-h}^{\gamma_0}\Delta_f\left(x_i\right)f\left(x_i\right)\left(1,x_j',q_j\right)\delta_0 k^-\left(\tfrac{q_i-\gamma_0}{h}\right)f\left(x_i,q_i\right)dq_i dx_i\\
&=2n\int_{x_i}\int_{\frac{v}{\sqrt{nh}}-1}^{0}\Delta_f\left(x_i\right)f\left(x_i\right)\left(1,x_i',\gamma_0+u_q h\right)\delta_0 k^-\left(u_q-\tfrac{v}{\sqrt{nh}}\right)f\left(x_i,\gamma_0+u_q h\right)du_q dx_i\\
&\quad-2n\int_{x_i}\int_{-1}^{0}\Delta_f\left(x_i\right)f\left(x_i\right)\left(1,x_i',\gamma_0+u_q h\right)\delta_0 k^-\left(u_q\right)f\left(x_i,\gamma_0+u_q h\right)du_q dx_i\\
&\approx -v^2 f_q\left(\gamma_0\right)\mathbb{E}\left[\Delta_f\left(x_i\right)\Delta_i f\left(x_i\right)|q_i=\gamma_0\right]k_+'\left(0\right),
\end{aligned}
$$

where the first equality comes from the zero conditional mean property of the error term, and the last one applies $|k'_-(0)| = k'_+(0)$. By Lemma 4 and Lemma 5, as well as the exogeneity property of error term u_i, we have

$$\begin{aligned}
&Cov\left(nh\left(\widehat{Q}_n(\gamma_0^{v_1}) - \widehat{Q}_n(\gamma_0)\right), nh\left(\widehat{Q}_n(\gamma_0^{v_2}) - \widehat{Q}_n(\gamma_0)\right)\right) \\
&= Cov\left(S_1(v_1), S_1(v_2)\right) + Cov\left(S_2(v_1), S_2(v_2)\right) + Cov\left(S_1(v_1), S_2(v_2)\right) + Cov\left(S_1(v_2), S_2(v_1)\right) \\
&\approx 4\mathbb{E}\left[\Delta_f^2(x_i) f^2(x_i)\left(\sigma_+^2(x_i) + \sigma_-^2(x_i)\right)|q_i = \gamma_0\right] f_q(\gamma_0) v_1 v_2 \xi_{(1)}.
\end{aligned}$$

Roughly speaking, $S_2(v)$ contributes to the mean process of $nh\left(\widehat{Q}_n(\gamma_0^v) - \widehat{Q}_n(\gamma_0)\right)$, and $S_1(v)$ contributes to the variance process.

To show the weak convergence, we apply the Lyapunov CLT by checking the Lyapunov condition. Specifically, we show that

$$\begin{aligned}
&\textstyle\sum_{i=1}^n \mathbb{E}\left[\Delta_f^4(x_i) f(x_i)^4 u_i^4 \left[\left(k^-\left(\frac{q_i-\gamma_0^v}{h}\right) - k^-\left(\frac{q_i-\gamma_0}{h}\right)\right) - \left(k^+\left(\frac{q_i-\gamma_0^v}{h}\right) - k^+\left(\frac{q_i-\gamma_0}{h}\right)\right)\right]^4\right] \\
&= O_p\left(\frac{nh}{(nh)^2}\right) = o_p(1)
\end{aligned}$$

as $n \to \infty$. ■

Proposition 4 $\widehat{\Sigma}(\gamma_0) \xrightarrow{d} \Sigma$.

Proof. By standard arguments, we have

$$\frac{\frac{1}{n}\sum_{i=1}^n k_h(q_i-\gamma_0)\Delta_f^2(x_i)f^2(x_i)u_i^2}{\frac{1}{n}\sum_{i=1}^n k_h(q_i-\gamma_0)} - \mathbb{E}\left[\Delta_f^2(x_i)f^2(x_i)(\sigma_+^2(x_i)+\sigma_-^2(x_i))|q_i=\gamma_0\right] \xrightarrow{p} 0$$

and

$$\frac{\frac{1}{n}\sum_{i=1}^n k_h(q_i-\gamma_0)\Delta_f(x_i)\Delta_i f(x_i)}{\frac{1}{n}\sum_{i=1}^n k_h(q_i-\gamma_0)} - \mathbb{E}\left[\Delta_f(x_i)\Delta_i f(x_i)|q_i=\gamma_0\right] \xrightarrow{p} 0.$$

So all we need to show is

$$\tfrac{1}{n}\textstyle\sum_{i=1}^n k_h(q_i - \widehat{\gamma})\widehat{\Delta}_i^2(\gamma_0)\widehat{f}^2(x_i)\widehat{u}_i(\gamma_0)^2 - \frac{1}{n}\sum_{i=1}^n k_h(q_i-\gamma_0)\Delta_f^2(x_i)f^2(x_i)u_i^2 \xrightarrow{p} 0$$

and

$$\tfrac{1}{n}\textstyle\sum_{i=1}^n k_h(q_i-\gamma_0)\widehat{\Delta}_i^2(\gamma_0)\widehat{f}^{-1}(x_i,\gamma_0)\widehat{f}(x_i) - \frac{1}{n}\sum_{i=1}^n k_h(q_i-\gamma_0)\Delta_f(x_i)\Delta_i f(x_i) \xrightarrow{p} 0,$$

which are implied by

$$\widehat{f}(x_i,\gamma_0) - f(x_i,\gamma_0) \xrightarrow{p} 0, \widehat{\Delta}_i(\gamma_0) - \Delta_f(x_i) \xrightarrow{p} 0, \widehat{f}(x_i) - f(x_i) \xrightarrow{p} 0 \text{ and } \widehat{u}_i(\gamma_0) - u_i \xrightarrow{p} 0$$

uniformly in $x_i \in \mathcal{X}$.

In the following, we take $\widehat{u}_i(\gamma_0) - u_i \xrightarrow{p} 0$ for illustration since others are easier to show. By Lemma B.3 of Newey (1994),

$$\sup_{x_i}\left|\tfrac{1}{n-1}\textstyle\sum_{j=1,j\neq i}^n K_h(x_j-x_i)k_h^{\pm}(q_j-\gamma_0)y_j - \mathbb{E}\left[y_j|x_i,\gamma_0^{\pm}\right]f(x_i,\gamma_0)\right| = O_p\left(\sqrt{\ln n/nh^d} + h\right) = o_p(1),$$

which implies

$$\sup_{x_i}\left|\widehat{m}_{\pm}(x_i,\gamma_0) - m_{\pm}(x_i,\gamma_0)\right| \xrightarrow{p} 0.$$

As a result,

$$
\begin{aligned}
&\sup_{x_i, |q_i-\gamma_0|\le h} |\widehat{u}_i(\gamma_0) - u_i| \\
&= \sup_{x_i, |q_i-\gamma_0|\le h} |y_i - \widehat{m}_-(x_i,\gamma_0) 1(q_i \le \gamma_0) - \widehat{m}_+(x_i,\gamma_0) 1(q_i > \gamma_0) - u_i| \\
&= \sup_{x_i, |q_i-\gamma_0|\le h} |[m_-(x_i,q_i) - \widehat{m}_-(x_i,\gamma_0)] 1(q_i \le \gamma_0) + [m_+(x_i,q_i) - \widehat{m}_+(x_i,\gamma_0)] 1(q_i > \gamma_0)| \\
&\le \sup_{x_i} |m_\pm(x_i,\gamma_0) - \widehat{m}_\pm(x_i,\gamma_0)| + \sup_{x_i, |q_i-\gamma_0|\le h} |m_\pm(x_i,q_i) - m_\pm(x_i,\gamma_0)| \\
&\xrightarrow{p} 0.
\end{aligned}
$$

■

Proposition 5 $\widehat{\Sigma} - \widehat{\Sigma}(\gamma_0) \xrightarrow{p} 0.$

Proof. To derive the result, we need to show

$$\tfrac{1}{n}\sum_{i=1}^n k_h(q_i-\widehat{\gamma})\widehat{\Delta}_i^2(\widehat{\gamma})\widehat{f}^2(x_i)\widehat{u}_i^2 - \tfrac{1}{n}\sum_{i=1}^n k_h(q_i-\gamma_0)\widehat{\Delta}_i^2(\gamma_0)\widehat{f}^2(x_i)\widehat{u}_i(\gamma_0)^2 \xrightarrow{p} 0$$

and

$$\tfrac{1}{n}\sum_{i=1}^n k_h(q_i-\widehat{\gamma})\widehat{\Delta}_i^2(\widehat{\gamma})\widehat{f}^{-1}(x_i,\widehat{\gamma})\widehat{f}(x_i) - \tfrac{1}{n}\sum_{i=1}^n k_h(q_i-\gamma_0)\widehat{\Delta}_i^2(\gamma_0)\widehat{f}^{-1}(x_i,\gamma_0)\widehat{f}(x_i) \xrightarrow{p} 0.$$

Take the first result for illustration since the second is simpler. First, we show that

$$\tfrac{1}{n}\sum_{i=1}^n k_h(q_i-\gamma_0)\widehat{\Delta}_i^2(\widehat{\gamma})\widehat{f}^2(x_i)\widehat{u}_i^2 - \tfrac{1}{n}\sum_{i=1}^n k_h(q_i-\gamma_0)\widehat{\Delta}_i^2(\gamma_0)\widehat{f}^2(x_i)\widehat{u}_i(\gamma_0)^2 \xrightarrow{p} 0. \tag{31}$$

Since

$$
\begin{aligned}
&\sup_{x_i} \left|\widehat{f}(x_i,\widehat{\gamma}) - \widehat{f}(x_i,\gamma_0)\right| \\
&= \left|\tfrac{1}{n-1}\sum_{j=1,j\ne i}^n K_h(x_j-x_i) k_h(q_j-\widehat{\gamma}) - \tfrac{1}{n-1}\sum_{j=1,j\ne i}^n K_h(x_j-x_i) k_h(q_j-\gamma_0)\right| \\
&= \left|\tfrac{1}{(n-1)h}\sum_{j=1,j\ne i}^n K_h(x_j-x_i)\left(k\left(\tfrac{q_j-\widehat{\gamma}}{h}\right) - k\left(\tfrac{q_j-\gamma_0}{h}\right)\right)\right| \\
&= O\left(\tfrac{v}{\sqrt{nh}}\right) = o_p(1),
\end{aligned}
\tag{32}
$$

and, by a similar argument as that in (32),

$$\sup_{x_i}\left|\widehat{m}_\pm(x_i,\widehat{\gamma})\widehat{f}_\pm(x_i,\widehat{\gamma}) - \widehat{m}_\pm(x_i,\gamma_0)\widehat{f}_\pm(x_i,\gamma_0)\right| = o_p(1),$$

we have

$$
\begin{aligned}
&\sup_{x_i}\left|\widehat{\Delta}_i(\widehat{\gamma}) - \widehat{\Delta}_i(\gamma_0)\right| \\
&\le \sup_{x_i}\left|\widehat{m}_-(x_i,\widehat{\gamma})\widehat{f}_-(x_i,\widehat{\gamma}) - \widehat{m}_-(x_i,\gamma_0)\widehat{f}_-(x_i,\gamma_0)\right| + \sup_{x_i}\left|\widehat{m}_+(x_i,\widehat{\gamma})\widehat{f}_+(x_i,\widehat{\gamma}) - \widehat{m}_+(x_i,\gamma_0)\widehat{f}_+(x_i,\gamma_0)\right| \\
&\xrightarrow{p} 0.
\end{aligned}
\tag{33}
$$

With the results in (32), (33) and $1(\gamma_0 < q_i \le \widehat{\gamma}) = O_p\left(\frac{1}{\sqrt{n/h}}\right) = o_p(1)$, we have

$$
\begin{aligned}
&\sup_{x_i}|\widehat{u}_i - \widehat{u}_i(\gamma_0)| \\
&= \sup_{x_i}|\widehat{m}_-(x_i,\widehat{\gamma}) 1(q_i \le \widehat{\gamma}) - \widehat{m}_-(x_i,\gamma_0) 1(q_i \le \gamma_0) + \widehat{m}_+(x_i,\widehat{\gamma}) 1(q_i > \widehat{\gamma}) - \widehat{m}_+(x_i,\gamma_0) 1(q_i > \gamma_0)| \\
&\le \sup_{x_i}|\widehat{m}_-(x_i,\widehat{\gamma}) - \widehat{m}_-(x_i,\gamma_0)| 1(q_i \le \widehat{\gamma}) + \sup_{x_i}|\widehat{m}_-(x_i,\gamma_0)| 1(\gamma_0 < q_i \le \widehat{\gamma}) \\
&\quad + \sup_{x_i}|\widehat{m}_+(x_i,\widehat{\gamma}) - \widehat{m}_+(x_i,\gamma_0)| 1(q_i > \widehat{\gamma}) + \sup_{x_i}|\widehat{m}_+(x_i,\gamma_0)| 1(\widehat{\gamma} < q_i \le \gamma_0) \xrightarrow{p} 0.
\end{aligned}
\tag{34}
$$

Combining (32)-(34), (31) is obtained.

Secondly, we show

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{i=1}^{n} k_h(q_i-\widehat{\gamma})\widehat{\Delta}_i^2(\widehat{\gamma})\widehat{f}^2(x_i)\widehat{u}_i^2-\frac{1}{n}\sum_{i=1}^{n} k_h(q_i-\gamma_0)\widehat{\Delta}_i^2(\widehat{\gamma})\widehat{f}^2(x_i)\widehat{u}_i^2\right| \\
&=\left|\frac{1}{n}\sum_{i=1}^{n}\left(k_h(q_i-\widehat{\gamma})-k_h(q_i-\gamma_0)\right)\widehat{\Delta}_i^2(\widehat{\gamma})\widehat{f}^2(x_i)\widehat{u}_i^2\right|+o_p(1) \\
&=O_p\left(\frac{1}{\sqrt{nh}}\frac{1}{nh}\sum_{i=1}^{n} 1\left(|q_i-\gamma_0|\le h\right)\widehat{\Delta}_i^2(\widehat{\gamma})\widehat{f}^2(x_i)\widehat{u}_i^2\right)+o_p(1).
\end{aligned}$$

Hence the required result is derived. ■

Proposition 6 $\widehat{\gamma}-\gamma_0=O_p(h)$.

Proof. The proof mimics that of Proposition 1. By replacing Assumption K, G and H with K′, G′ and H′, we now have

$$\sup_{\gamma\in\Gamma}\left|\rho_n^{-2}\widehat{Q}_n(\gamma)-Q_n(\gamma)\right|=O_p\left(\sqrt{\ln n/nh^d}\right)\xrightarrow{p}0,$$

where $Q_n(\gamma)$ contains only the middle term in Proposition 1, and the first and third terms disappear because their difference is $O\left(h^{2s}/\rho_n^2\right)=o(1)$. Now, for $\gamma\in\Gamma\backslash\mathcal{N}_n$, $\sup_{\gamma\in\Gamma\backslash\mathcal{N}_n}Q_n(\gamma)=o(1)$. For $\gamma\in\mathcal{N}_n$, $Q_n(\gamma)$ is a nondecreasing (nonincreasing) function of a for $a\in(-1,0)$ ($a\in(0,1)$) up to $o(1)$ and $\sup_{\gamma\in\mathcal{N}_n}|Q_n(\gamma)|=O(1)$. So $Q_n(\gamma)$ is maximized at some $\gamma_n\in\mathcal{N}_n$ such that $Q_n(\gamma_n)>\sup_{\gamma\in\Gamma\backslash\mathcal{N}_n}|Q_n(\gamma)|+C/2$ for n large enough. The result of interest is then derived. ■

Proposition 7 $\widehat{\gamma}-\gamma_0=O_p\left((n\rho_n^2)^{-1}\right)$.

Proof. This proof mimics that of Proposition 2 with the term $\sqrt{n/h}$ replaced by $n\rho_n^2$; Suppose $\gamma_0<\gamma<\gamma_0+h$, now we have

$$\begin{aligned}
D1\ &=\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} g_j\left(K_{h,ij}^{\gamma-}-K_{h,ij}^{\gamma_0-}\right)-\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} g_j\left(K_{h,ij}^{\gamma+}-K_{h,ij}^{\gamma_0+}\right) \\
&\quad+\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\left(1,x_j',q_j\right)\delta_n\left(K_{h,ij}^{\gamma-}-K_{h,ij}^{\gamma_0-}\right)\mathbf{1}\left(q_j\le\gamma_0\right) \\
&\le -C\rho_n\left|\frac{\gamma-\gamma_0}{h}\right|
\end{aligned}$$

for some $C>0$ with probability approaching 1 by calculating the mean and variance of D1 in its U-projection, where the difference of the first two terms contribute only $O_p\left(|\gamma-\gamma_0|\,h^s\right)$, the third term contributes to $-C\rho_n\left|\frac{\gamma-\gamma_0}{h}\right|$. Since $\rho_n/h^s\to\infty$, for $\eta\le h$,

$$P\left(\sup_{|\gamma-\gamma_0|<\eta}\frac{1}{n}\sum_{i=1}^{n}\left(\widehat{\Delta}_i(\gamma)-\widehat{\Delta}_i(\gamma_0)\right)>0\right)\le P\left(\sup_{|\gamma-\gamma_0|<\eta}D2>C\rho_n\left|\frac{\gamma-\gamma_0}{h}\right|\right).$$

With a different kernel function in Assumption K′ and the same formula of $D2$, we now have

$$Var(P1)\le C\frac{1}{nh}\left(\frac{\gamma_0-\gamma}{h}\right)^2, Var(P2)\le C\frac{1}{nh}\left(\frac{\gamma_0-\gamma}{h}\right)^2 \text{ and } Var(P3)\le C\frac{1}{nh}\left|\frac{\gamma_0-\gamma}{h}\right|.$$

As a result,

$$\begin{aligned}
Var(D2)\ &=Var(P1)+Var(P2)+Var(P3) \\
&\le C\frac{1}{nh}\left[\left(\frac{\gamma_0-\gamma}{h}\right)^2+\left|\frac{\gamma_0-\gamma}{h}\right|\right]\le\frac{C}{nh}\left|\frac{\gamma_0-\gamma}{h}\right|
\end{aligned}$$

uniformly for $|\gamma-\gamma_0|<\eta$. In consequence,

$$\begin{aligned} P\left(\sup_{|\gamma-\gamma_0|<\eta}\frac{1}{n}\sum_{i=1}^n\left(\widehat{\Delta}_i\left(\gamma\right)-\widehat{\Delta}_i\left(\gamma_0\right)\right)>0\right) &\leq C\mathbb{E}\left[\left(\sup_{|\gamma-\gamma_0|<\eta}D2\right)^2\right]\Big/\left[\rho_n\left|\frac{\gamma_0-\gamma}{h}\right|\right]^2 \\ &\leq \frac{C}{nh}\left|\frac{\gamma_0-\gamma}{h}\right|\Big/\rho_n^2\left(\frac{\gamma_0-\gamma}{h}\right)^2=\frac{C}{n\rho_n^2|\gamma-\gamma_0|} \end{aligned}$$

by Markov's inequality. So

$$\begin{aligned} &\textstyle\sum_{l=L}^{\log_2\left(nh\rho_n^2\right)}P\left(\sup_{S_{l,n}}\frac{1}{n}\sum_{i=1}^n\left(\widehat{\Delta}_i\left(\gamma\right)-\widehat{\Delta}_i\left(\gamma_0\right)\right)1\left(\Delta_i>0\right)>0\right) \\ &\leq \textstyle\sum_{l\geq L}\frac{C}{n\rho_n^2\cdot 2^l/n\rho_n^2}=C\sum_{l\geq L}\frac{1}{2^l}\to 0 \end{aligned}$$

as $L\to\infty$, and the proof is completed. ■

Proposition 8 *On any compact set of v,*

$$\mathbb{E}\left[nh\left(\widehat{Q}_n\left(\gamma_0^v\right)-\widehat{Q}_n(\gamma_0)\right)\right]=-2k_+(0)f_q(\gamma_0)D_n\left|v\right|+o(v),$$

$$\begin{aligned} &Cov\left(nh\left(\widehat{Q}_n\left(\gamma_0^{v_1}\right)-\widehat{Q}_n\left(\gamma_0\right)\right),nh\left(\widehat{Q}_n\left(\gamma_0^{v_2}\right)-\widehat{Q}_n\left(\gamma_0\right)\right)\right) \\ &=\begin{cases}\Sigma_n v_2+o\left(v_2\right), & \text{if } v_1\geq v_2\geq 0 \text{ or } v_1\leq v_2\leq 0\ , \\ 0, & \text{otherwise,}\end{cases} \end{aligned}$$

and

$$\Sigma_n=\begin{cases}16k_+^2\left(0\right)f_q\left(\gamma_0\right)V_{n1}, & \text{if } v\leq 0, \\ 16k_+^2\left(0\right)f_q\left(\gamma_0\right)V_{n2}, & \text{if } v>0,\end{cases}$$

and the finite-dimensional limit distributions of $nh\left(\widehat{Q}_n\left(\gamma_0^v\right)-\widehat{Q}_n(\gamma_0)\right)$ are the same as those of $C\left(v\right)$, where $\gamma_0^v=\gamma_0+\frac{v}{n\rho_n^2}$, and $C\left(v\right)$ is defined in the proof of Theorem 6.

Proof. We mimic the proof of Proposition 3. Now, $\left|\widehat{\Delta}_i\left(\gamma_0^v\right)-\Delta_f(x_i)\right|\xrightarrow{p}0$ uniformly in i and v, where $\Delta_f(x_i):=(1,x_i',\gamma_0)\delta_n f(x_i,\gamma_0)=O_p(\rho_n)$. Decompose $\widehat{\Delta}_i\left(\gamma_0^v\right)-\widehat{\Delta}_i\left(\gamma_0\right)$ into the same six terms as in the proof of Proposition 3 only with δ_0 replaced by δ_n. By Lemma 7, we have

$$\frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j=1,j\neq i}^n 2\Delta_f\left(x_i\right)\left(T_{1ij}+T_{2ij}\right)\approx 0,$$

by Lemma 8, we have

$$\begin{aligned} &\frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j=1,j\neq i}^n 2\Delta_f\left(x_i\right)\left(T_{3ij}+T_{4ij}\right) \\ &\approx\frac{2}{nh}\sum_{i=1}^n\Delta_f\left(x_i\right)f\left(x_i\right)u_i\left[\left(k^-\left(\frac{q_i-\gamma_0^v}{h}\right)-k^-\left(\frac{q_i-\gamma_0}{h}\right)\right)-\left(k^+\left(\frac{q_i-\gamma_0^v}{h}\right)-k^+\left(\frac{q_i-\gamma_0}{h}\right)\right)\right] \\ &=: S_2(v)/nh \end{aligned}$$

and by Lemma 9, we have

$$\begin{aligned} &\frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j=1,j\neq i}^n 2\Delta_f\left(x_i\right)\left(T_{5ij}+T_{6ij}\right) \\ &\approx\frac{2}{nh}\sum_{i=1}^n\Delta_f\left(x_i\right)f\left(x_i\right)\left(1,x_j',q_j\right)\delta_n\left[k^-\left(\frac{q_i-\gamma_0^v}{h}\right)1\left(\gamma_0^v-h\leq q_j\leq\gamma_0\right)-k^-\left(\frac{q_i-\gamma_0}{h}\right)1\left(\gamma_0-h\leq q_j\leq\gamma_0\right)\right] \\ &=: S_1(v)/nh. \end{aligned}$$

As a result,

$$\mathbb{E}\left[nh\left(\widehat{Q}_n(\gamma_0^v)-\widehat{Q}_n(\gamma_0)\right)\right]\approx\mathbb{E}\left[S_2\left(v\right)\right]\approx 2k_+\left(0\right)f_q\left(\gamma_0\right)\mathbb{E}\left[\Delta_i^2 f\left(x_i,q_i\right)f\left(x_i\right)|q_i=\gamma_0\right]v/\rho_n^2.$$

Combining these results and the fact that $k_-(0)=k_+(0)$, we obtain the first equation in the proposition.

By Lemma 10 and Lemma 11, we have

$$\begin{aligned}&Cov\left(nh\left(\widehat{Q}_n(\gamma_0^{v_1})-\widehat{Q}_n(\gamma_0)\right),nh\left(\widehat{Q}_n(\gamma_0^{v_2})-\widehat{Q}_n(\gamma_0)\right)\right)\\&=Cov\left(S_1\left(v_1\right),S_1\left(v_2\right)\right)+Cov\left(S_2\left(v_1\right),S_2\left(v_2\right)\right)+Cov\left(S_1\left(v_1\right),S_2\left(v_2\right)\right)+Cov\left(S_1\left(v_2\right),S_2\left(v_1\right)\right)\\&\approx\Sigma_n v_2.\end{aligned}$$

Just as in Method II, $S_2(v)$ contributes to the mean process of $nh\left(\widehat{Q}_n(\gamma_0^{v_1})-\widehat{Q}_n(\gamma_0)\right)$, and $S_1(v)$ contributes to the variance process.

To show fidi convergence, we apply the Cramér-Wold device, combined with the Lyapunov CLT. Specifically, we check the Lyapunov condition that

$$\begin{aligned}&\sum_{i=1}^n\mathbb{E}\left[\Delta_f^4\left(x_i\right)f^4\left(x_i\right)u_i^4\left[\left(k^-\left(\tfrac{q_i-\gamma_0^v}{h}\right)-k^-\left(\tfrac{q_i-\gamma_0}{h}\right)\right)-\left(k^+\left(\tfrac{q_i-\gamma_0^v}{h}\right)-k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right)\right]^4\right]\\&=O_p\left(\tfrac{nh\rho_n^4}{nh\rho_n^2}\right)=o_p\left(1\right)\end{aligned}$$

as $n\to\infty$. Then the proposition is proved. ■

Proposition 9 $\widetilde{\gamma}-\gamma_0=O_p\left(h\right).$

Proof. The proof is similar to that of Proposition 6. But now

$$Q_n\left(\gamma\right)=\left[\begin{array}{c}\int_{-1}^0\int K^x\left(u_x,x_o\right)k_-\left(u_q\right)m\left(x_o+u_xh,\gamma+u_qh\right)f\left(x_o+u_xh,\gamma+u_qh\right)du_xdu_q\\-\int_0^1\int K^x\left(u_x,x_o\right)k_+\left(u_q\right)m\left(x_o+u_xh,\gamma+u_qh\right)f\left(x_o+u_xh,\gamma+u_qh\right)du_xdu_q\end{array}\right]^2,$$

and we require $\Delta_o/h^s\to\infty$ to make the proof go through. ■

Proposition 10 $\widetilde{\gamma}-\gamma_0=O_p\left((nh^{d-1}\Delta_o^2)^{-1}\right).$

Proof. The proof is the same as that of Proposition 7. We only pay attention to the role that $\Delta_o\to 0$ plays to make the proof go through. ■

Proposition 11 *On any compact set of v,*

$$\mathbb{E}\left[nh^d\left(\widetilde{Q}_n\left(\gamma_0^v\right)-\widetilde{Q}_n(\gamma_0)\right)\right]=-2k_+(0)f\left(x_o,\gamma_0\right)^2|v|+o(v),$$

$$\begin{aligned}&Cov\left(nh^d\left(\widetilde{Q}_n\left(\gamma_0^{v_1}\right)-\widetilde{Q}_n\left(\gamma_0\right)\right),nh^d\left(\widetilde{Q}_n\left(\gamma_0^{v_2}\right)-\widetilde{Q}_n\left(\gamma_0\right)\right)\right)\\&=\begin{cases}\Sigma_o\left(v_2\right)v_2+o\left(v_2\right), & \text{if } v_1\geq v_2\geq 0 \text{ or } v_1\leq v_2\leq 0,\\ 0, & \text{otherwise,}\end{cases}\end{aligned}$$

and the finite-dimensional limit distributions of $nh^d\left(\widetilde{Q}_n\left(\gamma_0^v\right)-\widetilde{Q}_n(\gamma_0)\right)$ are the same as those of $C_o\left(v\right)$, *where* $\gamma_0^v = \gamma_0 + \frac{v}{nh^{d-1}\Delta_o^2}$,

$$\Sigma_o(v) = \begin{cases} 16k_+^2(0)\,\sigma_-^2(x_o)\,f(x_o,\gamma_0)^3\,\kappa^2, & \text{if } v \le 0, \\ 16k_+^2(0)\,\sigma_+^2(x_o)\,f(x_o,\gamma_0)^3\,\kappa^2, & \text{if } v > 0, \end{cases}$$

and $C_o(v)$ *is defined in the proof of Corollary 4.*

Proof. Mimic the proof of Proposition 8. Now,

$$\widetilde{Q}_n(\gamma_0^v) - \widetilde{Q}_n(\gamma_0) = \widehat{\Delta}_o(\gamma_0^v)^2 - \widehat{\Delta}_o(\gamma_0)^2 = \left(\widehat{\Delta}_o(\gamma_0^v) + \widehat{\Delta}_o(\gamma_0)\right)\left(\widehat{\Delta}_o(\gamma_0^v) - \widehat{\Delta}_o(\gamma_0)\right).$$

By Lemma B.1 of Newey (1994), we can show that $\left|\widehat{\Delta}_o(\gamma_0^v) - \Delta_f(x_o)\right| \xrightarrow{p} 0$ uniformly in i and v, where $\Delta_f(x_o) := (1, x_o', \gamma_0)\delta_0 f(x_o,\gamma_0) = O_p(\Delta_o)$, so $\left|\widehat{\Delta}_o(\gamma_0^v) + \widehat{\Delta}_o(\gamma_0) - 2\Delta_f(x_o)\right| \xrightarrow{p} 0$ uniformly in i and v. We then only need to focus on the other term. For simplicity, let $v > 0$. Now,

$$\widehat{\Delta}_o(\gamma_0^v) - \widehat{\Delta}_o(\gamma_0) =: \tfrac{1}{n}\textstyle\sum_{j=1}^n (T_{1j} + T_{2j} + T_{3j} + T_{4j} + T_{5j} + T_{6j}),$$

where $\{T_{kj}\}_{k=1}^6$ are defined similarly as in Proposition 8, only with $K_{h,ij}^{\gamma-}$ and $K_{h,ij}^{\gamma+}$ replaced by $K_{h,j}^{\gamma-}$ and $K_{h,j}^{\gamma+}$, respectively. By Lemma 13, we have

$$\tfrac{1}{n}\textstyle\sum_{j=1}^n 2\Delta_f(x_o)(T_{1j} + T_{2j}) \approx 0,$$

and by Lemma 14, we have

$$\tfrac{1}{n}\textstyle\sum_{j=1}^n 2\Delta_f(x_o)(T_{5j} + T_{6j}) \approx -2k_+(0)\frac{v}{nh^d} f(x_o,\gamma_0)^2.$$

So

$$\begin{aligned} &nh^d\left(\widetilde{Q}_n(\gamma_0^v) - \widetilde{Q}_n(\gamma_0)\right) \\ &= h^d \textstyle\sum_{j=1}^n 2\Delta_f(x_o)(T_{1j} + T_{2j}) + h^d \sum_{j=1}^n 2\Delta_f(x_o)(T_{3j} + T_{4j}) + h^d \sum_{j=1}^n 2\Delta_f(x_o)(T_{5j} + T_{6j}) \\ &= o_p(1) + S_1(v) + S_2(v). \end{aligned}$$

Hence, with the zero mean assumption of the error term, we have

$$\mathbb{E}\left[nh^d\left(\widetilde{Q}_n(\gamma_0^v) - \widetilde{Q}_n(\gamma_0)\right)\right] = \mathbb{E}[S_2(v)] = -2k_+(0) f(x_o,\gamma_0)^2 v + o_p(v).$$

By Lemmas 13, 15 and 16, and the exogeneity property of the error terms, we have

$$\begin{aligned} &Cov\left(nh^d\left(\widetilde{Q}_n(\gamma_0^{v_1}) - \widetilde{Q}_n(\gamma_0)\right), nh^d\left(\widetilde{Q}_n(\gamma_0^{v_2}) - \widetilde{Q}_n(\gamma_0)\right)\right) \\ &= Cov(S_1(v_1), S_1(v_2)) + Cov(S_2(v_1), S_2(v_2)) + Cov(S_1(v_1), S_2(v_2)) + Cov(S_1(v_2), S_2(v_1)) \\ &\approx \Sigma_{on} v_2. \end{aligned}$$

To show the weak convergence, we apply the Lyapunov CLT by checking the Lyapunov condition. Specifically, we show that

$$\textstyle\sum_{j=1}^n \mathbb{E}\left[h^{4d}\Delta_f^4(x_o)\,u_i^4\left[\left(K_{h,j}^{\gamma_0^v-} - K_{h,j}^{\gamma_0-}\right) - \left(K_{h,j}^{\gamma_0^v+} - K_{h,j}^{\gamma_0+}\right)\right]^4\right] = O_p\left(\frac{nh^{3d}\Delta_o^4}{nh^d\Delta_o^2}\right) = o_p(1).$$

■

Proposition 12 *$I_{1n}^{(1)}$ is $o_p(1)$ under $H_0^{(1)}$, and is $O_p\left(nh^{d/2}\right)$ under $H_1^{(1)}$.*

Proof. Note that

$$\begin{aligned} I_{1n} &= \tfrac{nh^{d/2}}{n(n-1)} \sum_i \sum_{j\neq i} D_i D_j K_{h,ij} \\ &= \tfrac{nh^{d/2}}{n(n-1)} \sum_i \sum_{j\neq i} \left[m_i - \mathbf{x}_i'\widehat{\beta} - \mathbf{x}_i'\widehat{\delta}1\left(q_i \leq \widehat{\gamma}\right)\right]\left[m_j - \mathbf{x}_j'\widehat{\beta} - \mathbf{x}_j'\widehat{\delta}1\left(q_j \leq \widehat{\gamma}\right)\right] K_{h,ij}. \end{aligned}$$

Under $H_0^{(1)}$, $m_i = \mathbf{x}_i'\beta_0 + \mathbf{x}_i'\delta_0 1\left(q_i \leq \gamma_0\right)$, so that

$$\begin{aligned} & m_i - \mathbf{x}_i'\widehat{\beta} - \mathbf{x}_i'\widehat{\delta}1\left(q_i \leq \widehat{\gamma}\right) \\ &= \mathbf{x}_i'\left(\beta_0 - \widehat{\beta}\right) + \mathbf{x}_i'\left(\delta_0 - \widehat{\delta}\right)1\left(q_i \leq \widehat{\gamma}\wedge\gamma_0\right) \\ &\quad + \mathbf{x}_i'\delta_0 1\left(\widehat{\gamma} < q_i \leq \gamma_0\right) - \mathbf{x}_i'\widehat{\delta}1\left(\gamma_0 < q_i \leq \widehat{\gamma}\right). \end{aligned}$$

As a result, I_{1n} has ten terms with a typical term of the form

$$T1 = \left(\widehat{\beta} - \beta_0\right)'\left[\tfrac{nh^{d/2}}{n(n-1)} \sum_i \sum_{j\neq i} K_{h,ij}\mathbf{x}_i\mathbf{x}_j'\right]\left(\widehat{\beta} - \beta_0\right)$$

or

$$T2 = \delta_0'\left[\tfrac{nh^{d/2}}{n(n-1)} \sum_i \sum_{j\neq i} K_{h,ij}\mathbf{x}_i\mathbf{x}_j' 1\left(\widehat{\gamma} < q_i \leq \gamma_0\right)1\left(\widehat{\gamma} < q_j \leq \gamma_0\right)\right]\delta_0.$$

Given that $\widehat{\beta} - \beta_0 = O_p(n^{-1/2})$, $\widehat{\delta} - \delta_0 = O_p(n^{-1/2})$, and $\widehat{\gamma} - \gamma_0 = O_p(n^{-1})$, it is easy to show that $T1 = O_p(h^{d/2})$ and $T2 = O_p(h^{d/2})$ since $\frac{1}{n(n-1)}\sum\limits_i\sum\limits_{j\neq i} K_{h,ij}\mathbf{x}_i\mathbf{x}_j' = O_p(1)$ and $\frac{1}{n(n-1)}\sum\limits_i\sum\limits_{j\neq i} K_{h,ij}\mathbf{x}_i\mathbf{x}_j' 1\left(\widehat{\gamma} < q_i \leq \gamma_0\right)1\left(\widehat{\gamma} < q_j \leq \gamma_0\right) = O_p\left(n^{-1}\right)$.

We now analyze I_{1n} under $H_1^{(1)}$. There are three cases. Let

$$\left(\beta_o', \delta_o', \gamma_o\right)' = \arg\inf_{\beta,\delta,\gamma} \mathbb{E}\left[\left(y - \mathbf{x}'\beta - \mathbf{x}'\delta 1\left(q \leq \gamma\right)\right)^2\right].$$

If $\delta_o = 0$, then $\overline{m}(x,q) = \mathbf{x}'\beta_o$ and the model degenerates to the case analyzed in Zheng (1996). If $\delta_{xo} = \mathbf{0}$ and $\delta_{\alpha o} + \gamma_o\delta_{qo} = 0$, then $\overline{m}(x,q)$ takes the CTR form of Chan and Tsay (1998). It follows that $\widehat{\beta} - \beta_o = O_p(n^{-1/2})$, $\widehat{\delta} - \delta_o = O_p(n^{-1/2})$, and $\widehat{\gamma} - \gamma_o = O_p(n^{-1/2})$. If $\delta_{xo} \neq \mathbf{0}$ or $\delta_{\alpha o} + \gamma_o\delta_{qo} \neq 0$, then $\widehat{\beta} - \beta_o = O_p(n^{-1/2})$, $\widehat{\delta} - \delta_o = O_p(n^{-1/2})$, and $\widehat{\gamma} - \gamma_o = O_p(n^{-1})$. See Yu (2017) for these results. We concentrate on the last case. Now,

$$I_{1n} = \tfrac{nh^{d/2}}{n(n-1)} \sum_i \sum_{j\neq i} \left(m_i - \overline{m}_i\right)\left(m_j - \overline{m}_j\right) K_{h,ij}(1 + o_p(1)),$$

where $\overline{m}_i = \mathbf{x}_i'\beta_o + \mathbf{x}_i'\delta_o 1\left(q_i \leq \gamma_o\right)$, so we need only calculate $\mathbb{E}\left[\left(m_i - \overline{m}_i\right)\left(m_j - \overline{m}_j\right)K_{h,ij}\right]$, which is equal to

$$\begin{aligned} & \int \left(m_i - \overline{m}_i\right)\left(m_j - \overline{m}_j\right) K_{h,ij} f_i f_j dx_i dq_i dx_j dq_j \\ \approx & \int \left(m_i - \overline{m}_i\right)^2 K^x(u_x, x_i) k(u_q) f_i^2 dx_i dq_i du_x du_q \\ = & \int \left(m_i - \overline{m}_i\right)^2 f_i^2 dx_i dq_i, \end{aligned}$$

The result follows. ■

Proposition 13 *$I_{1n}^{(2)}$ is $o_{P_m}(1)$ uniformly in m under $H_0^{(2)}$, and is $O_p\left(nh^{d/2}b\right)$ under $H_1^{(2)}$.*

Proof. Given that $\widehat{f}_i^{-1} = f_i^{-1} + o_p(1)$ and f_i is bounded uniformly over $(x_i, q_i) \in \mathcal{X} \times \Gamma$,

$$\begin{array}{l} I_{1n} \\ = \frac{nh^{d/2}}{n(n-1)} \sum_i \sum_{j \neq i} 1_i^\Gamma 1_j^\Gamma (m_i - \widehat{m}_i) \widehat{f}_i (m_j - \widehat{m}_j) \widehat{f}_j K_{h,ij} \left(\widehat{f}_i^{-1} \widehat{f}_j^{-1} \right) \\ \approx \frac{nh^{d/2}}{n(n-1)^3} \sum_i \sum_{j \neq i} \sum_{l \neq i} \sum_{k \neq j} 1_i^\Gamma (m_i - m_l) L_{b,il} 1_j^\Gamma (m_j - m_k) L_{b,jk} K_{h,ij} f_i^{-1} f_j^{-1} \\ = O_p \left(\frac{nh^{d/2}}{n(n-1)^3} \sum_i \sum_{j \neq i} \sum_{l \neq i} \sum_{k \neq j} 1_i^\Gamma (m_i - m_l) L_{b,il} 1_j^\Gamma (m_j - m_k) L_{b,jk} K_{h,ij} \right), \end{array} \tag{35}$$

Mimicking the proof of Proposition A.1 of Fan and Li (1996), we can show that under $H_0^{(2)}$, $I_{1n} = O_p\left(nh^{d/2} b^{2\eta}\right) = o_p(1)$. The only new result we need to employ is that $\left|\mathbb{E}_1\left[(m_2 - m_1) L_{b,21}\right]\right| = O_p\left(b^\eta\right)$, which is accomplished in Lemma 18.

We now analyze I_{1n} under $H_1^{(2)}$. It can be shown that the case where i, j, l, k are all different from each other dominates in the formula of the second equality of (35), so

$$I_{1n} \approx O_p\left(nh^{d/2} \mathbb{E}\left[1_1^\Gamma (m_1 - m_2) L_{b,12} 1_3^\Gamma (m_3 - m_4) L_{b,34} K_{h,13} f_1^{-1} f_3^{-1}\right]\right).$$

Because $h/b \to 0$, we can treat $(x_1, q_1) = (x_3, q_3)$. Specifically,

$$\begin{array}{l} \quad \mathbb{E}\left[1_1^\Gamma (m_1 - m_2) L_{b,12} 1_3^\Gamma (m_3 - m_4) L_{b,34} K_{h,13} f_1^{-1} f_3^{-1}\right] \\ = \mathbb{E}\left[\begin{array}{c} 1_1^\Gamma (m_1 - m_2) L_{b,12} f_1^{-1} \int 1(q_1 + u_q h \in \Gamma)(m((x_1, q_1) + uh) - m_4) \\ \frac{1}{b^d} L^x\left(\frac{x_4 - x_1 - u_x h}{b}, x_1 + u_x h\right) l\left(\frac{q_4 - q_1 - u_q h}{b}\right) K^x(u_x, x_1) k(u_q) du \end{array} \right] \\ \approx \mathbb{E}\left[1_1^\Gamma (m_1 - m_2) L_{b,12} (m_1 - m_4) L_{b,14} f_1^{-1}\right] \\ = \mathbb{E}\left\{1_1^\Gamma f_1^{-1} \left\{\mathbb{E}_1\left[(m_1 - m_2) L_{b,12}\right]\right\}^2\right\} \\ = \int_{\underline{\gamma}}^{\overline{\gamma}} \int \left[\int (m(x_1, q_1) - m(x_2\, q_2)) \frac{1}{b^d} L^x\left(\frac{x_2 - x_1}{b}, x_1\right) l\left(\frac{q_2 - q_1}{b}\right) f(x_2, q_2) dx_2 dq_2\right]^2 dx_1 dq_1 \\ \approx O\left(b^{2\eta}\right) + \int_{\gamma_0 - b}^{\gamma_0 + b} \int \left[\int (m(x_1, q_1) - m(x_1 + u_x b, q_1 + u_q b)) L^x(u_x, x_1) l(u_q) f(x_1 + u_x b, q_1 + u_q b) du\right]^2 dx_1 dq_1 \\ \approx O\left(b^{2\eta}\right) + \int_{\gamma_0 - b}^{\gamma_0 + b} \int \left[\begin{array}{c} \int_{\frac{\gamma_0 - q_1}{b}}^1 (m(x_1, q_1) - m(x_1, q_1 + u_q b)) l(u_q) du \\ + \int_{-1}^{\frac{\gamma_0 - q_1}{b}} (m(x_1, q_1) - m(x_1, q_1 + u_q b)) l(u_q) du_q \end{array} \right]^2 f(x_1, \gamma_0)^2 dx_1 dq_1 \end{array}$$

where $u = (u_x, u_q)$. Under $H_1^{(2)}$,

$$\begin{array}{l} \quad \int_{\gamma_0 - b}^{\gamma_0 + b} \int \left[\begin{array}{c} \int_{\frac{\gamma_0 - q_1}{b}}^1 (m(x_1, q_1) - m(x_1, q_1 + u_q b)) l(u_q) du \\ + \int_{-1}^{\frac{\gamma_0 - q_1}{b}} (m(x_1, q_1) - m(x_1, q_1 + u_q b)) l(u_q) du_q \end{array} \right]^2 f(x_1, \gamma_0)^2 dx_1 dq_1 \\ \approx \int_{\gamma_0}^{\gamma_0 + b} \int \left[-\int_{\frac{\gamma_0 - q_1}{b}}^1 m'_+(x_1) u_q b l(u_q) du_q + \int_{-1}^{\frac{\gamma_0 - q_1}{b}} (m_+(x_1) - m_-(x_1) + C u_q b) l(u_q) du_q \right]^2 f(x_1, \gamma_0)^2 dx_1 dq_1 \\ \quad + \int_{\gamma_0 - b}^{\gamma_0} \int \left[\int_{\frac{\gamma_0 - q_1}{b}}^1 (m_-(x_1) - m_+(x_1) + C u_q b) l(u_q) du_q - \int_{-1}^{\frac{\gamma_0 - q_1}{b}} m'_-(x_1) u_q b l(u_q) du_q \right]^2 f(x_1, \gamma_0)^2 dx_1 dq_1 \\ \approx b \int_0^1 \int \left[\int_{-1}^{-v} (m_+(x_1) - m_-(x_1)) l(u_q) du_q \right]^2 dx_1 dv \\ \quad + b \int_0^1 \int \left[\int_v^1 (m_+(x_1) - m_-(x_1)) l(u_q) du_q \right]^2 f(x_1, \gamma_0)^2 dx_1 dv \\ = 2b \int (m_+(x_1) - m_-(x_1))^2 f(x_1, \gamma_0)^2 dx \int_0^1 \left(\int_v^1 l(u_q) du_q \right)^2 dv, \end{array}$$

where $m'_\pm(x) = \lim_{\gamma \to \gamma_0 \pm} \partial m(x, \gamma)/\partial \gamma$, and $m_\pm(x) = \lim_{\gamma \to \gamma_0 \pm} m(x, \gamma)$. The result follows. ■

Proposition 14 *$I_{2n}^{(1)} \xrightarrow{d} N\left(0, \Sigma^{(1)}\right)$ and $I_{2n}^{(2)} \xrightarrow{d} N\left(0, \Sigma^{(2)}\right)$.*

Proof. We only prove the second result since the first is proved in a similar way.

$$\begin{aligned} I_{2n}^{(2)} &= \tfrac{nh^{d/2}}{n(n-1)} \sum_i \sum_{j\neq i} 1_i^{\Gamma} 1_j^{\Gamma} u_i u_j K_{h,ij} \\ &\equiv \tfrac{nh^{d/2}}{n(n-1)} \sum_i \sum_{j\neq i} H_n(z_i, z_j) \equiv nh^{d/2} U_n, \end{aligned}$$

where U_n is a second order degenerate U-statistic with kernel function H_n. We can apply theorem 1 of Hall (1984) to find its asymptotic distribution. Two conditions need to be checked: (i) $\mathbb{E}[H_n^2(z_1, z_2)] < \infty$; (ii)

$$\frac{\mathbb{E}\left[G_n^2(z_1,z_2)\right]+n^{-1}\mathbb{E}[H_n^4(z_1,z_2)]}{\mathbb{E}^2[H_n^2(z_1,z_2)]} \to 0 \text{ as } n \to \infty,$$

where $G_n(z_1, z_2) = \mathbb{E}[H_n(z_3, z_1)H_n(z_3, z_2)|z_1, z_2]$. Because these checks follow in a similar way to lemma 3.3a of Zheng (1996) they are omitted here to save space. In conclusion

$$nU_n \Big/ \sqrt{2\mathbb{E}[H_n^2(z_1, z_2)]} \xrightarrow{d} N(0, 1).$$

It is easy to check that

$$\begin{aligned} \mathbb{E}\left[H_n^2(z_1,z_2)\right] &= \mathbb{E}\left[1_1^{\Gamma} 1_2^{\Gamma} K_{h,12}^2 \mathbb{E}[u_1^2|x_1,q_1]\mathbb{E}[u_2^2|x_2,q_2]\right] \\ &= \int_{\underline{\gamma}}^{\overline{\gamma}} \int \int_{\underline{\gamma}}^{\overline{\gamma}} \int \tfrac{1}{h^{2d}} K^x\left(\tfrac{x_2-x_1}{h}, x_1\right)^2 k^2\left(\tfrac{q_2-q_1}{h}\right) \sigma^2(x_1,q_1)\sigma^2(x_2,q_2) f(x_1,q_1) f(x_2,q_2) dx_2 dq_2 dx_1 dq_1 \\ &= \int_{\underline{\gamma}}^{\overline{\gamma}} \int \int_{\frac{\underline{\gamma}-q}{h}}^{\frac{\overline{\gamma}-q}{h}} \int \tfrac{1}{h^d} K^x\left(u_x, x\right)^2 k^2(u_q) \sigma^2(x,q)\sigma^2(x+u_x h, q+u_q h) f(x,q) f(x+u_x h, q+u_q h) du dx dq \\ &\approx \tfrac{1}{h^d} \int_{\underline{\gamma}}^{\overline{\gamma}} \int \left[\int K^x\left(u_x, x\right)^2 k^2(u_q) du\right] \sigma^4(x,q) f^2(x,q) dx dq + o\left(\tfrac{1}{h^d}\right) \\ &\approx \tfrac{1}{h^d} \left[\int k^{2d}(u) du\right] \int_{\underline{\gamma}}^{\overline{\gamma}} \int \sigma^4(x,q) f^2(x,q) dx dq = \tfrac{1}{h^d} \tfrac{\Sigma^{(2)}}{2}, \end{aligned}$$

so the result follows. ■

Proposition 15 *$v_n^{(1)2} = \Sigma^{(1)} + o_p(1)$ under $H_0^{(1)}$ and the local alternative and $v_n^{(1)2} = O_p(1)$ under $H_1^{(1)}$.*

Proof. It can be shown that

$$\begin{aligned} &\tfrac{h^d}{n(n-1)} \sum_i \sum_{j\neq i} K_{h,ij}^2 \widehat{e}_i^2 \widehat{e}_j^2 \\ &= \tfrac{h^d}{n(n-1)} \sum_i \sum_{j\neq i} K_{h,ij}^2 \left(u_i + m_i - \overline{m}_i\right)^2 \left(u_j + m_j - \overline{m}_j\right)^2 + o_p(1) \\ &= h^d \mathbb{E}\left[K_{h,ij}^2 \left(u_i + m_i - \overline{m}_i\right)^2 \left(u_j + m_j - \overline{m}_j\right)^2\right] + o_p(1) \\ &= h^d \mathbb{E}\left[K_{h,ij}^2 \mathbb{E}_i\left[\left(u_i + m_i - \overline{m}_i\right)^2\right] \mathbb{E}_j\left[\left(u_j + m_j - \overline{m}_j\right)^2\right]\right] + o_p(1) \\ &= \int\int \left[\int K^x\left(u_x, x_i\right)^2 k^2(u_q) du\right] \left(\sigma_i^2 + \left(m_i - \overline{m}_i\right)^2\right)^2 f_i^2 dx_i dq_i + o_p(1) \\ &= \int k^{2d}(u) du \mathbb{E}\left[f\left(x,q\right)\left(\sigma^2(x,q) + \left(m - \overline{m}\right)^2\right)^2\right] + o_p(1), \end{aligned}$$

where $\sigma_i^2 = \sigma^2(x_i, q_i)$. Under $H_0^{(1)}$, $m - \overline{m} = 0$. Under the local alternative, $\mathbb{E}\left[f\left(x,q\right)\left(m-\overline{m}\right)^4\right] = O\left(n^{-2}h^{-d}\right) = o(1)$, and under $H_1^{(1)}$, $\mathbb{E}\left[f\left(x,q\right)\left(m-\overline{m}\right)^4\right] = O(1)$. ■

Proposition 16 *$v_n^{(2)2} = \Sigma^{(2)} + o_p(1)$ under both $H_0^{(2)}$ and $H_1^{(2)}$.*

Proof. By steps similar to the last proposition, we have

$$\tfrac{h^d}{n(n-1)} \sum_i \sum_{j\neq i} 1_i^{\Gamma} 1_j^{\Gamma} K_{h,ij}^2 \widehat{e}_i^2 \widehat{e}_j^2 = \int k^{2d}(u) du \mathbb{E}\left[1_i^{\Gamma} f\left(x_i, q_i\right) \left(\sigma^2(x_i,q_i) + \left(m_i - \overline{m}_i\right)^2\right)^2\right] + o_p(1),$$

where $\overline{m}_i$ is redefined as $\mathbb{E}_i\left[\widehat{m}_i\right]$. Note that $\mathbb{E}\left[1_i^{\Gamma} f\left(x_i, q_i\right)\left(m_i-\overline{m}_i\right)^4\right]$ is at most $O(b)$ since $m_i-\overline{m}_i$ contributes only for $q \in[\gamma-b, \gamma+b]$. ■

Supplement C: Lemmas

Lemma 1 $\frac{1}{n(n-1)} \sum_i^n \sum_{j=1, j \neq i}^n 2 \Delta_f\left(x_i\right)\left(T_{1 i j}+T_{2 i j}\right) \approx 0$.

Proof. For $2 \Delta_f\left(x_i\right) T_{1 i j}$, we have

$$
\begin{aligned}
& \mathbb{E}\left[2 \Delta_f\left(x_i\right) T_{1 i j} \mid z_i\right] \\
& =\int_{-\infty}^{+\infty} \int_{x_j} 2 \Delta_f\left(x_i\right) g\left(x_j, q_j\right) \frac{1}{h^d} K^x\left(\frac{x_j-x_i}{h}, x_i\right)\left(k^{-}\left(\frac{q_j-\gamma_0^v}{h}\right)-k^{-}\left(\frac{q_j-\gamma_0}{h}\right)\right) f\left(x_j, q_j\right) d x_j d q_j \\
& =2 \Delta_f\left(x_i\right) \int_{-1}^0 \int_{u_x} g\left(x_i+u_x h, \gamma_0^v+u_q h\right) K^x\left(u_x, x_i\right) k^{-}\left(u_q\right) f\left(x_i+u_x h, \gamma_0^v+u_q h\right) d u_x d u_q \\
& \quad-2 \Delta_f\left(x_i\right) \int_{-1}^0 \int_{u_x} g\left(x_i+u_x h, \gamma_0+u_q h\right) K^x\left(u_x, x_i\right) k^{-}\left(u_q\right) f\left(x_i+u_x h, \gamma_0+u_q h\right) d u_x d u_q \\
& =2 \Delta_f\left(x_i\right) \int_{-1}^0 \int_{u_x}\left[\bar{g}\left(x_i, \gamma_0^v+u_q h\right)+\bar{g}_1\left(x_i, \gamma_0^v+u_q h\right) u_x h+o\left(u_x h\right)\right] K^x\left(u_x, x_i\right) k^{-}\left(u_q\right) d u_x d u_q \\
& \quad-2 \Delta_f\left(x_i\right) \int_{-1}^0 \int_{u_x}\left[\bar{g}\left(x_i, \gamma_0+u_q h\right)+\bar{g}_1\left(x_i, \gamma_0+u_q h\right) u_x h+o\left(u_x h\right)\right] K^x\left(u_x, x_i\right) k^{-}\left(u_q\right) d u_x d u_q \\
& =2 \Delta_f\left(x_i\right) \int_{-1}^0\left[\bar{g}\left(x_i, \gamma_0^v+u_q h\right)-\bar{g}\left(x_i, \gamma_0+u_q h\right)\right] k^{-}\left(u_q\right) d u_q \\
& =2 \Delta_f\left(x_i\right) \bar{g}_2\left(x_i, \gamma_0\right) \frac{v}{\sqrt{n / h}}+O\left(\frac{v^2}{n / h}\right),
\end{aligned}
$$

where $\bar{g}(x_j, q_j)=g(x_j, q_j) f(x_j, q_j)$, $\bar{g}_1(x_j, q_j)=\frac{\partial \bar{g}(x_j, q_j)}{\partial x_j}$ and $\bar{g}_2(x_j, q_j)=\frac{\partial \bar{g}(x_j, q_j)}{\partial q_j}$. Since

$$
\begin{aligned}
& \mathbb{E}\left[2 \Delta_f\left(x_i\right) T_{1 i j} \mid z_j\right] \\
& =\int_{x_i} 2 \Delta_f\left(x_i\right) g\left(x_j, q_j\right) \frac{1}{h^d} K^x\left(\frac{x_j-x_i}{h}, x_i\right)\left(k^{-}\left(\frac{q_j-\gamma_0^v}{h}\right)-k^{-}\left(\frac{q_j-\gamma_0}{h}\right)\right) f\left(x_i\right) d x_i \\
& =\frac{2}{h} g\left(x_j, q_j\right)\left(k^{-}\left(\frac{q_j-\gamma_0^v}{h}\right)-k^{-}\left(\frac{q_j-\gamma_0}{h}\right)\right) \int_{u_x} \Delta_f\left(x_j+u_x h\right) K^x\left(u_x, x_j-u_x h\right) f\left(x_j-u_x h\right) d u_x \\
& =\frac{2}{h} g\left(x_j, q_j\right)\left(k^{-}\left(\frac{q_j-\gamma_0^v}{h}\right)-k^{-}\left(\frac{q_j-\gamma_0}{h}\right)\right) \Delta_f\left(x_j\right) f\left(x_j\right)
\end{aligned}
$$

and

$$
\begin{aligned}
& \mathbb{E}\left[4 \Delta_f^2\left(x_i\right) T_{1 i j}^2\right] \\
& =\int_{x_i} \int_{q_j} \int_{x_j} \frac{4}{h^{2 d}} \Delta_f^2\left(x_i\right) g\left(x_j, q_j\right)^2 K^x\left(\frac{x_j-x_i}{h}, x_i\right)^2\left(k^{-}\left(\frac{q_j-\gamma_0^v}{h}\right)-k^{-}\left(\frac{q_j-\gamma_0}{h}\right)\right)^2 f\left(x_j, q_j\right) d x_j d q_j f\left(x_i\right) d x_i \\
& =\int_{x_i} \int_{u_q} \int_{u_x}\left\{\begin{array}{c}
\frac{4}{h^d} \Delta_f^2\left(x_i\right) g\left(x_i+u_x h, \gamma_0+u_q h\right)^2 K^x\left(u_x, x_i\right)^2 \\
\cdot\left(k^{-}\left(u_q+\frac{\gamma_0-\gamma_0^v}{h}\right)-k^{-}\left(u_q\right)\right)^2 f\left(x_i+u_x h, \gamma_0+u_q h\right)
\end{array}\right\} d u_x d u_q f\left(x_i\right) d x_i \\
& =O\left(\frac{1}{h^d} \frac{v}{\sqrt{n h}}\right),
\end{aligned}
$$

we have

$$
\begin{aligned}
& \frac{1}{n} \sum_{i=1}^n \mathbb{E}\left[2 \Delta_f\left(x_i\right) T_{1 i j} \mid z_i\right]=2 \Delta_f\left(x_i\right) \bar{g}_2\left(x_i, \gamma_0\right) \frac{v}{\sqrt{n / h}}+O\left(\frac{v^2}{n / h}\right), \\
& \frac{1}{n} \sum_{i=1}^n \mathbb{E}\left[2 \Delta_f\left(x_i\right) T_{1 i j} \mid z_j\right]=\frac{2}{h} g\left(x_j, q_j\right)\left(k^{-}\left(\frac{q_j-\gamma_0^v}{h}\right)-k^{-}\left(\frac{q_j-\gamma_0}{h}\right)\right) \Delta_f\left(x_j\right) f\left(x_j\right), \\
& \frac{1}{n} \mathbb{E}\left[4 \Delta_f^2\left(x_i\right) T_{1 i j}^2\right]^{1 / 2}=\frac{1}{n} O\left(\frac{1}{h^{d / 2}} \frac{1}{(n h)^{1 / 4}}\right)=o(1).
\end{aligned}
$$

By Lemma 8.4 of Newey and Mcfadden(1994), we have

$$
\frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j=1, j \neq i}^n \mathbb{E}\left[2 \Delta_f\left(x_i\right) T_{1 i j}\right] \approx 2 \Delta_f\left(x_i\right) \bar{g}_2\left(x_i, \gamma_0\right) \frac{v}{\sqrt{n / h}},
$$

where the extra terms $\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[2\Delta_f(x_i)T_{1ij}|z_j]-\mathbb{E}[2\Delta_f(x_i)T_{1ij}]$ and $\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[2\Delta_f(x_i)T_{1ij}|z_j]-\mathbb{E}[2\Delta_f(x_i)T_{1ij}]$ are $o_p(1)$. Similarly for $2\Delta_f(x_i)T_{2ij}$, we have

$$\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\mathbb{E}[2\Delta_f(x_i)T_{2ij}]\approx -2\Delta_f\left(x_i\right)\overline{g}_2\left(x_i,\gamma_0\right)\tfrac{v}{\sqrt{n/h}}.$$

Hence,

$$\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\mathbb{E}\left[2\Delta_f(x_i)\left(T_{1ij}+T_{2ij}\right)\right]\approx 0$$

and

$$\begin{aligned}
&Var\left(\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\mathbb{E}\left[2\Delta_f(x_i)\left(T_{1ij}+T_{2ij}\right)\right]\right)\\
&\leq Var\left(\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\mathbb{E}\left[2\Delta_f(x_i)T_{1ij}\right]\right)+Var\left(\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\mathbb{E}\left[2\Delta_f(x_i)T_{2ij}\right]\right)\\
&=o\left(1\right).
\end{aligned}$$

So the result of interest is derived. ■

Lemma 2

$$\begin{aligned}
&\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}2\Delta_f\left(x_i\right)\left(T_{3ij}+T_{4ij}\right)\\
&\approx\tfrac{2}{nh}\sum_{i=1}^{n}\Delta_f\left(x_i\right)f\left(x_i\right)u_i\left[\left(k^-\left(\tfrac{q_i-\gamma_0^v}{h}\right)-k^-\left(\tfrac{q_i-\gamma_0}{h}\right)\right)-\left(k^+\left(\tfrac{q_i-\gamma_0^v}{h}\right)-k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right)\right].
\end{aligned}$$

Proof. Since

$$\mathbb{E}\left[2\Delta_f\left(x_i\right)T_{3ij}|z_i\right]=0,$$

$$\begin{aligned}
&\mathbb{E}\left[2\Delta_f\left(x_i\right)T_{3ij}|z_j\right]\\
&=\tfrac{2}{h^d}u_j\left(k^-\left(\tfrac{q_j-\gamma_0^v}{h}\right)-k^-\left(\tfrac{q_j-\gamma_0}{h}\right)\right)\int_{x_i}\Delta_f\left(x_i\right)K^x\left(\tfrac{x_j-x_i}{h},x_i\right)f\left(x_i\right)dx_i\\
&=\tfrac{2}{h}u_j\left(k^-\left(\tfrac{q_j-\gamma_0^v}{h}\right)-k^-\left(\tfrac{q_j-\gamma_0}{h}\right)\right)\int_{x_i}\Delta_f\left(x_j-u_xh_j\right)K^x\left(u_x,x_j-u_xh\right)f\left(x_j+-u_xh\right)dx_i\\
&\approx\tfrac{2}{h}u_j\left(k^-\left(\tfrac{q_j-\gamma_0^v}{h}\right)-k^-\left(\tfrac{q_j-\gamma_0}{h}\right)\right)\Delta_f\left(x_j\right)f\left(x_j\right)
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\left[4\Delta_f^2(x_i)T_{3ij}^2\right]\\
&=\int_{x_i}\int_{q_j}\int_{x_j}\tfrac{4}{h^{2d}}\Delta_f^2\left(x_i\right)\sigma^2\left(x_j,q_j\right)K^x\left(\tfrac{x_j-x_i}{h},x_i\right)^2\left(k^-\left(\tfrac{q_j-\gamma_0^v}{h}\right)-k^-\left(\tfrac{q_j-\gamma_0}{h}\right)\right)^2f\left(x_j,q_j\right)dx_jdq_jf\left(x_i\right)dx_i\\
&=\int_{x_i}\int_{u_q}\int_{u_x}\left\{\begin{array}{c}\tfrac{4}{h^d}\Delta_f^2\left(x_i\right)\sigma^2\left(x_i+u_xh,\gamma_0+u_qh\right)K^x\left(u_x,x_i\right)^2\\ \cdot\left(k^-\left(u_q-\tfrac{v}{\sqrt{nh}}\right)-k^-\left(u_q\right)\right)^2f\left(x_i+u_xh,\gamma_0+u_qh\right)\end{array}\right\}du_xdu_qf\left(x_i\right)dx_i\\
&=O\left(\tfrac{1}{h^d}\tfrac{v}{\sqrt{nh}}\right),
\end{aligned}$$

we have

$$\begin{aligned}
&\tfrac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[2\Delta_f\left(x_i\right)T_{3ij}\right]\approx\tfrac{2}{nh}\sum_{i=1}^{n}u_j\left(k^-\left(\tfrac{q_j-\gamma_0^v}{h}\right)-k^-\left(\tfrac{q_j-\gamma_0}{h}\right)\right)\Delta_f\left(x_j\right)f\left(x_j\right),\\
&\tfrac{1}{n}\mathbb{E}\left[4\Delta_f^2(x_i)T_{3ij}^2\right]^{1/2}=\tfrac{1}{n}O\left(\tfrac{1}{h^{d/2}}\tfrac{1}{(nh)^{1/4}}\right)=o\left(1\right).
\end{aligned}$$

By Lemma 8.4 of Newey and Mcfadden(1994), we have

$$\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}2\Delta_f\left(x_i\right)T_{3ij}\approx\tfrac{2}{nh}\sum_{i=1}^{n}\Delta_f\left(x_i\right)f\left(x_i\right)u_i\left(k^-\left(\tfrac{q_i-\gamma_0^v}{h}\right)-k^-\left(\tfrac{q_i-\gamma_0}{h}\right)\right).$$

A similar result can be derived for $\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}2\Delta_f(x_i)T_{4ij}$ and the result of interest is then obtained. ■

Lemma 3

$$\begin{aligned}&\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}2\Delta_f(x_i)(T_{5ij}+T_{6ij})\\ \approx&\frac{2}{nh}\sum_{i=1}^{n}\Delta_f(x_i)f(x_i)(1,x_i',q_i)\delta_0\left[\left(k^{-}\left(\frac{q_i-\gamma_0^v}{h}\right)1(\gamma_0^v-h\leq q_i\leq\gamma_0)-k^{-}\left(\frac{q_i-\gamma_0}{h}\right)1(\gamma_0-h\leq q_i\leq\gamma_0)\right)\right].\end{aligned}$$

Proof. Taking T_5 to illustrate, we have

$$\begin{aligned}&\mathbb{E}\left[2\Delta_f(x_i)T_{5ij}|z_j\right]\\ =&-\frac{2}{h^d}\int_{x_i}\Delta_f(x_i)\left(1,x_j',q_j\right)\delta_0K^x\left(\frac{x_j-x_i}{h},x_i\right)k^{-}\left(\frac{q_j-\gamma_0}{h}\right)1\left(\gamma_0-h\leq q_j\leq\gamma_0\right)f(x_i)\,dx_i\\ =&-\frac{2}{h}\left(1,x_j',q_j\right)\delta_0k^{-}\left(\frac{q_j-\gamma_0}{h}\right)1\left(\gamma_0-h\leq q_j\leq\gamma_0\right)\int_{u_x}\Delta_f(x_j-u_xh)K^x(u_x,x_j-u_xh)f(x_j-u_xh)\,du_x\\ \approx&-\frac{2}{h}\left(1,x_j',q_j\right)\delta_0k^{-}\left(\frac{q_j-\gamma_0}{h}\right)1\left(\gamma_0-h\leq q_j\leq\gamma_0\right)\Delta_f(x_j)f(x_j),\end{aligned}$$

$$\begin{aligned}&\mathbb{E}\left[2\Delta_f(x_i)T_{5ij}|z_i\right]\\ =&-\frac{2}{h^d}\int_{q_i}\int_{x_i}\Delta_f(x_i)\left(1,x_j',q_j\right)\delta_0K^x\left(\frac{x_j-x_i}{h},x_i\right)k^{-}\left(\frac{q_j-\gamma_0}{h}\right)1\left(\gamma_0-h\leq q_j\leq\gamma_0\right)f(x_j,q_j)\,dx_idq_j\\ =&-2\Delta_f(x_i)\int_{-1}^{0}\int_{u_x}(1,x_i'+u_x'h,\gamma_0+u_qh)\delta_0K^x(u_x,x_i)k^{-}(u_q)f(x_i+u_xh,\gamma_0+u_qh)\,du_x\\ \approx&-2\Delta_f(x_i)(1,x_i',\gamma_0)\delta_0f(x_i,\gamma_0)\approx-2\Delta_f(x_i)^2\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}\left[4\Delta_f^2(x_i)T_{5ij}^2\right]\\ =&\int_{x_i}\int_{q_j}\int_{x_j}\frac{4}{h^{2d}}\Delta_f^2(x_i)\left[\left(1,x_j',q_j\right)\delta_0\right]^2K^x\left(\frac{x_j-x_i}{h},x_i\right)^2\left(k^{-}\left(\frac{q_j-\gamma_0}{h}\right)\right)^2f(x_j,q_j)\,dx_jdq_jf(x_i)\,dx_i\\ =&\int_{x_i}\int_{-1}^{0}\int_{u_x}\left\{\begin{array}{c}\frac{4}{h^d}\Delta_f^2(x_i)\left[(1,x_i'+u_x'h,\gamma_0+u_qh)\delta_0\right]^2\\ \cdot K^x(u_x,x_i)^2\left(k^{-}(u_q)\right)^2f(x_i+u_xh,\gamma_0+u_qh)\end{array}\right\}du_xdu_qf(x_i)\,dx_i\\ =&O\left(\frac{1}{h^d}\right).\end{aligned}$$

Hence,

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[2\Delta_f(x_i)T_{5ij}|z_i\right]\approx-2\Delta_f(x_i)^2,\\ &\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[2\Delta_f(x_i)T_{5ij}|z_j\right]\approx-\frac{2}{h}\left(1,x_j',q_j\right)\delta_0k^{-}\left(\frac{q_j-\gamma_0}{h}\right)1\left(\gamma_0-h\leq q_j\leq\gamma_0\right)\Delta_f(x_j)f(x_j),\\ &\frac{1}{n}\mathbb{E}\left[4\Delta_f^2(x_i)T_{5ij}^2\right]^{1/2}=\frac{1}{n}O\left(\frac{1}{h^{d/2}}\right)=o(1).\end{aligned}$$

A similar result can be derived for $\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[2\Delta_f(x_i)T_{6ij}|z_i\right]$. Then by Lemma 8.4 of Newey and Mcfadden(1994), we have

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[2\Delta_f(x_i)\left(T_{5ij}+T_{5ij}\right)\right]\\ \approx&\frac{2}{nh}\sum_{i=1}^{n}\left(1,x_j',q_j\right)\delta_0\left[k^{-}\left(\frac{q_j-\gamma_0^v}{h}\right)1\left(\gamma_0^v-h\leq q_j\leq\gamma_0\right)-k^{-}\left(\frac{q_j-\gamma_0}{h}\right)1\left(\gamma_0-h\leq q_j\leq\gamma_0\right)\right]\Delta_f(x_j)f(x_j),\end{aligned}$$

where the extra terms $\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[2\Delta_f(x_i)T_{5ij}|z_i]-\mathbb{E}[2\Delta_f(x_i)T_{5ij}]$ and $\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[2\Delta_f(x_i)T_{6ij}|z_i]-\mathbb{E}[2\Delta_f(x_i)T_{6ij}]$ are $o_p(1)$. ■

Lemma 4 $Cov\left(S_1(v_1),S_1(v_2)\right)\approx4\mathbb{E}\left[\Delta_f^2(x_i)f^2(x_i)(\sigma_+^2(x_i)+\sigma_-^2(x_i))|q_i=\gamma_0\right]f_q(\gamma_0)v_1v_2\xi_{(1)}$.

Proof. Without loss of generality, assume $v_1 \geq v_2 \geq 0$. Then

$$\begin{aligned}
&4nCov\left(\Delta_f(x_i) f(x_i) u_i \left(k^+\left(\tfrac{q_i-\gamma_0^{v_1}}{h}\right) - k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right), \Delta_f(x_i) f(x_i) u_i \left(k^+\left(\tfrac{q_i-\gamma_0^{v_2}}{h}\right) - k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right)\right) \\
&= 4n\mathbb{E}\left[\Delta_f^2(x_i) f(x_i)^2 u_i^2 \left(k^+\left(\tfrac{q_i-\gamma_0^{v_1}}{h}\right) - k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right)\left(k^+\left(\tfrac{q_i-\gamma_0^{v_2}}{h}\right) - k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right)\right] f(u_i|x_i,q_i) f(x_i,q_i)\, du_i dq_i dx_i \\
&= 4n\int_{x_i}\int_{\gamma_0}^{\gamma_0^{v_2}}\int_{u_i} \Delta_f^2(x_i) f(x_i)^2 u_i^2 \left(k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right)^2 f(u_i|x_i,q_i) f(x_i,q_i)\, du_i dq_i dx_i \\
&\quad - 4n\int_{x_i}\int_{\gamma_0^{v_2}}^{\gamma_0^{v_1}}\int_{u_i} \Delta_f^2(x_i) f(x_i)^2 u_i^2 k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\left(k^+\left(\tfrac{q_i-\gamma_0^{v_2}}{h}\right) - k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right) f(u_i|x_i,q_i) f(x_i,q_i)\, du_i dq_i dx_i \\
&\quad + 4n\int_{x_i}\int_{\gamma_0^{v_1}}^{\gamma_0+h}\int_{u_i} \left\{\begin{array}{c} \Delta_f^2(x_i) f(x_i)^2 u_i^2 \left(k^+\left(\tfrac{q_i-\gamma_0^{v_1}}{h}\right) - k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right) \\ \cdot\left(k^+\left(\tfrac{q_i-\gamma_0^{v_2}}{h}\right) - k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right) f(u_i|x_i,q_i) f(x_i,q_i) \end{array}\right\} du_i dq_i dx_i \ (*) \\
&\quad + 4n\int_{x_i}\int_{\gamma_0+h}^{\gamma_0^{v_2}+h}\int_{u_i} \Delta_f^2(x_i) f(x_i)^2 u_i^2 k^+\left(\tfrac{q_i-\gamma_0^{v_1}}{h}\right) k^+\left(\tfrac{q_i-\gamma_0^{v_2}}{h}\right) f(u_i|x_i,q_i) f(x_i,q_i)\, du_i dq_i dx_i \\
&\approx 4\mathbb{E}\left[\Delta_f^2(x_i) f(x_i)^2 \sigma_+^2(x_i) | q_i = \gamma_0\right] f_q(\gamma_0) v_1 v_2 \xi_{(1)},
\end{aligned}$$

where $\sigma_+^2(x_i) = \mathbb{E}[u_i^2|x_i, q_i = \gamma_0+]$ and $\xi_{(1)} = \int_0^1 k_+'(t)^2 dt$. For a more detailed proof, we refer to that of Lemma C.4 of DH. Similarly,

$$\begin{aligned}
&4nCov\left(\Delta_f(x_i) f(x_i) u_i \left(k^-\left(\tfrac{q_i-\gamma_0^{v_1}}{h}\right) - k^-\left(\tfrac{q_i-\gamma_0}{h}\right)\right), \Delta_f(x_i) f(x_i) u_i \left(k^-\left(\tfrac{q_i-\gamma_0^{v_2}}{h}\right) - k^-\left(\tfrac{q_i-\gamma_0}{h}\right)\right)\right) \\
&\approx 4\mathbb{E}\left[\Delta_f^2(x_i) f(x_i)^2 \sigma_-^2(x_i) | q_i = \gamma_0\right] f_q(\gamma_0) v_1 v_2 \xi_{(1)}
\end{aligned}$$

and

$$\begin{aligned}
&4nCov\left(\Delta_f(x_i) f(x_i) u_i \left(k^\pm\left(\tfrac{q_i-\gamma_0^{v_1}}{h}\right) - k^\pm\left(\tfrac{q_i-\gamma_0}{h}\right)\right), \Delta_f(x_i) f(x_i) u_i \left(k^\mp\left(\tfrac{q_i-\gamma_0^{v_2}}{h}\right) - k^\mp\left(\tfrac{q_i-\gamma_0}{h}\right)\right)\right) \\
&\approx 4\mathbb{E}\left[\Delta_f^2(x_i) f(x_i)^2 \sigma_-^2(x_i) | q_i = \gamma_0\right] f_q(\gamma_0) v_1 v_2 \xi_{(1)} \tfrac{v}{\sqrt{nh}} = o_p(1),
\end{aligned}$$

where $\sigma_-^2(x_i) = \mathbb{E}[u_i^2|x_i, q_i = \gamma_0-]$. So the result of interest is obtained by summing up terms. ■

Lemma 5 *$Cov(S_2(v_1), S_2(v_2)) = o(v_1 v_2)$, $Cov(S_1(v_1), S_2(v_2)) = o(v_1 v_2)$ and $Cov(S_2(v_1), S_1(v_2)) = o(v_1 v_2)$.*

Proof. From Lemma 3, by tedious calculation, we can obtain

$$Var(S_2(v_2)) = o(v_2).$$

Hence,

$$\begin{aligned}
Cov(S_2(v_1), S_2(v_2)) &\leq \sqrt{Var(S_2(v_1))\, Var(S_2(v_2))} = o(v_1 v_2), \\
Cov(S_1(v_1), S_2(v_2)) &\leq \sqrt{Var(S_1(v_1))\, Var(S_2(v_2))} = o(v_1 v_2), \\
Cov(S_2(v_1), S_1(v_2)) &\leq \sqrt{Var(S_2(v_1))\, Var(S_1(v_2))} = o(v_1 v_2).
\end{aligned}$$

■

Lemma 6 *For any $\phi_1, \phi_2 > 0$, there exists $\eta > 0$ such that*

$$P\left\{\sup_{|v_1-v_2|<\eta}\left|\widehat{nh}\left(\widehat{Q}_n(\gamma_0^{v_1}) - \widehat{Q}_n(\gamma_0)\right) - \widehat{nh}\left(\widehat{Q}_n(\gamma_0^{v_2}) - \widehat{Q}_n(\gamma_0)\right)\right| > \phi_1\right\} < \phi_2,$$

where

$$\widehat{nh}\left(\widehat{Q}_n(\gamma_0^{v}) - \widehat{Q}_n(\gamma_0)\right) = nh\left(\widehat{Q}_n(\gamma_0^{v}) - \widehat{Q}_n(\gamma_0)\right) - \mathbb{E}\left[nh\left(\widehat{Q}_n(\gamma_0^{v}) - \widehat{Q}_n(\gamma_0)\right)\right].$$

Proof. From Proposition 3, we have

$$
\begin{aligned}
&\lim_{n\to\infty}\mathbb{E}\left[\widehat{nh}\left(\widehat{Q}_n\left(\gamma_0^{v_1}\right)-\widehat{Q}_n\left(\gamma_0\right)\right)-\widehat{nh}\left(\widehat{Q}_n\left(\gamma_0^{v_2}\right)-\widehat{Q}_n\left(\gamma_0\right)\right)\right]^2\\
&=\lim_{n\to\infty}\mathbb{E}\left[\widehat{nh}\left(\widehat{Q}_n\left(\gamma_0^{v_1}\right)-\widehat{Q}_n\left(\gamma_0\right)\right)\right]^2+\lim_{n\to\infty}\mathbb{E}\left[\widehat{nh}\left(\widehat{Q}_n\left(\gamma_0^{v_2}\right)-\widehat{Q}_n\left(\gamma_0\right)\right)\right]^2\\
&\quad-2\lim_{n\to\infty}\mathbb{E}\left[\widehat{nh}\left(\widehat{Q}_n\left(\gamma_0^{v_1}\right)-\widehat{Q}_n\left(\gamma_0\right)\right)\right]\left[\widehat{nh}\left(\widehat{Q}_n\left(\gamma_0^{v_2}\right)-\widehat{Q}_n\left(\gamma_0\right)\right)\right]\\
&=4\left(v_1^2+v_2^2-2v_1v_2\right)\mathbb{E}\left[\Delta_f^2(x_i)f^2(x_i)(\sigma_+^2(x_i)+\sigma_-^2(x_i))|q_i=\gamma_0\right]f_q(\gamma_0)\xi_{(1)}+o\left(\left|v_1-v_2\right|^2\right)\\
&\le C\left|v_1-v_2\right|^2.
\end{aligned}
$$

By Markov's inequality, the result follows. ■

Lemma 7 $\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}2\Delta_f(x_i)\left(T_{1ij}+T_{2ij}\right)\approx 0.$

Proof. The proof is similar to that of Lemma 1. For $2\Delta_f(x_i)T_{1ij}$, we now have

$$
\begin{aligned}
&\mathbb{E}\left[2\Delta_f(x_i)T_{1ij}|X_i\right]\\
&=\int_{-\infty}^{+\infty}\int_{x_j}2\Delta_f\left(x_i\right)g\left(x_j,q_j\right)\frac{1}{h^d}K^x\left(\frac{x_j-x_i}{h},x_i\right)\left(k^-\left(\frac{q_j-\gamma_0^v}{h}\right)-k^-\left(\frac{q_j-\gamma_0}{h}\right)\right)f\left(x_j,q_j\right)dx_jdq_j\\
&=2\Delta_f\left(x_i\right)\int_{-1}^{0}\int_{u_x}g\left(x_i+u_xh,\gamma_0^v+u_qh\right)K^x\left(u_x,x_i\right)k^-\left(u_q\right)f\left(x_i+u_xh,\gamma_0^v+u_qh\right)du_xdu_q\\
&\quad-2\Delta_f\left(x_i\right)\int_{-1}^{0}\int_{u_x}g\left(x_i+u_xh,\gamma_0+u_qh\right)K^x\left(u_x,x_i\right)k^-\left(u_q\right)f\left(x_i+u_xh,\gamma_0+u_qh\right)du_xdu_q\\
&=2\Delta_f\left(x_i\right)\int_{-1}^{0}\int_{u_x}\left[\begin{array}{c}\overline{g}\left(x_i,\gamma_0^v+u_qh\right)+\overline{g}_1^{(1)}\left(x_i,\gamma_0^v+u_qh\right)u_xh^1\\+\cdots+\overline{g}_1^{(s)}\left(x_i,\gamma_0^v+u_qh\right)u_xh^s+o\left(u_x^sh^s\right)\end{array}\right]K^x\left(u_x,x_i\right)k^-\left(u_q\right)du_xdu_q\\
&\quad-2\Delta_f\left(x_i\right)\int_{-1}^{0}\int_{u_x}\left[\begin{array}{c}\overline{g}\left(x_i,\gamma_0+u_qh\right)+\overline{g}_1^{(1)}\left(x_i,\gamma_0+u_qh\right)u_xh^1\\+\cdots+\overline{g}_1^{(s)}\left(x_i,\gamma_0+u_qh\right)u_xh^s+o\left(u_x^sh^s\right)\end{array}\right]K^x\left(u_x,x_i\right)k^-\left(u_q\right)du_xdu_q\\
&\approx 2\Delta_f\left(x_i\right)\int_{-1}^{0}\left[\begin{array}{c}\left(\overline{g}\left(x_i,\gamma_0^v+u_qh\right)-\overline{g}\left(x_i,\gamma_0+u_qh\right)\right)+h^1\left(\overline{g}_1^{(1)}\left(x_i,\gamma_0^v+u_qh\right)-\overline{g}_1^{(1)}\left(x_i,\gamma_0+u_qh\right)\right)\\+\cdots+h^s\left(\overline{g}_1^{(s)}\left(x_i,\gamma_0^v+u_qh\right)-\overline{g}_1^{(s)}\left(x_i,\gamma_0+u_qh\right)\right)\end{array}\right]k^-\left(u_q\right)du_q\\
&\approx 2\Delta_f\left(x_i\right)\frac{v}{n\rho_n^2}\left[\overline{g}_2\left(x_i,\gamma_0\right)+\cdots+h^s\overline{g}_2^{(s)}\left(x_i,\gamma_0\right)\right],
\end{aligned}
$$

where $\overline{g}(x_j,q_j)=g(x_j,q_j)f(x_j,q_j)$, $\overline{g}_1^{(n)}(x_j,q_j)=\frac{\partial^n\overline{g}(x_j,q_j)}{(\partial x_j)^n}$ and $\overline{g}_2^{(n)}(x_j,q_j)=\frac{\partial\overline{g}_1^{(n)}(x_j,q_j)}{\partial q_j}$. Since

$$
\begin{aligned}
&\mathbb{E}\left[2\Delta_f(x_i)T_{1ij}|X_j\right]\\
&=\int_{x_j}2\Delta_f\left(x_i\right)g\left(x_j,q_j\right)\frac{1}{h^d}K^x\left(\frac{x_j-x_i}{h},x_i\right)\left(k^-\left(\frac{q_j-\gamma_0^v}{h}\right)-k^-\left(\frac{q_j-\gamma_0}{h}\right)\right)f\left(x_i\right)dx_i\\
&=\frac{2}{h}g\left(x_j,q_j\right)\left(k^-\left(\frac{q_j-\gamma_0^v}{h}\right)-k^-\left(\frac{q_j-\gamma_0}{h}\right)\right)\int_{u_x}\Delta_f\left(x_j+u_xh\right)K^x\left(u_x,x_j-u_xh\right)f\left(x_j-u_xh\right)du_x\\
&=\frac{2}{h}g\left(x_j,q_j\right)\left(k^-\left(\frac{q_j-\gamma_0^v}{h}\right)-k^-\left(\frac{q_j-\gamma_0}{h}\right)\right)\Delta_f\left(x_j\right)f\left(x_j\right)
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E}\left[4\Delta_f^2(x_i)T_{1ij}^2\right]\\
&=\int_{x_i}\int_{q_j}\int_{x_j}\frac{4}{h^{2d}}\Delta_f^2\left(x_i\right)g\left(x_j,q_j\right)^2K^x\left(\frac{x_j-x_i}{h},x_i\right)^2\left(k^-\left(\frac{q_j-\gamma_0^v}{h}\right)-k^-\left(\frac{q_j-\gamma_0}{h}\right)\right)^2f\left(x_j,q_j\right)dx_jdq_jf\left(x_i\right)dx_i\\
&=\int_{x_i}\int_{u_q}\int_{u_x}\left\{\begin{array}{c}\frac{4}{h^d}\Delta_f^2\left(x_i\right)g\left(x_i+u_xh,\gamma_0+u_qh\right)^2K^x\left(u_x,x_i\right)^2\\\cdot\left(k^-\left(u_q+\frac{\gamma_0-\gamma_0^v}{h}\right)-k^-\left(u_q\right)\right)^2f\left(x_i+u_xh,\gamma_0+u_qh\right)\end{array}\right\}du_xdu_qf\left(x_i\right)dx_i\\
&=O\left(\frac{\rho_n^2}{h^d}\frac{v}{nh\rho_n^2}\right)=\left(\frac{1}{nh^{d+1}}\right),
\end{aligned}
$$

we have

$$
\begin{aligned}
&\tfrac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[2\Delta_f(x_i)T_{1ij}|X_i\right]\approx 2\Delta_f\left(x_i\right)\tfrac{v}{n\rho_n^2}\left[\overline{g}_2\left(x_i,\gamma_0\right)+\cdots+h^s\overline{g}_2^{(s)}\left(x_i,\gamma_0\right)\right],\\
&\tfrac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[2\Delta_f(x_i)T_{1ij}|X_j\right]=\tfrac{2}{h}g\left(x_j,q_j\right)\left(k^-\left(\tfrac{q_j-\gamma_0^v}{h}\right)-k^-\left(\tfrac{q_j-\gamma_0}{h}\right)\right)\Delta_f\left(x_j\right)f\left(x_j\right),\\
&\tfrac{1}{n}\mathbb{E}\left[4\Delta_f^2(x_i)T_{1ij}^2\right]^{1/2}=\tfrac{1}{n}O\left(\tfrac{1}{\sqrt{nh^{d+1}}}\right)=o\left(1\right).
\end{aligned}
$$

By Lemma 8.4 of Newey and Mcfadden(1994), we have

$$
\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}2\Delta_f(x_i)T_{1ij}\approx 2\Delta_f\left(x_i\right)\tfrac{v}{n\rho_n^2}\left[\overline{g}_2\left(x_i,\gamma_0\right)+\cdots+h^s\overline{g}_2^{(s)}\left(x_i,\gamma_0\right)\right],
$$

where the extra terms $\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[2\Delta_f(x_i)T_{1ij}|X_j]-\mathbb{E}[2\Delta_f(x_i)T_{1ij}]$ and $\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[2\Delta_f(x_i)T_{1ij}|X_j]-\mathbb{E}[2\Delta_f(x_i)T_{1ij}]$ are $o_p(1)$. Similarly for $2\Delta_f(x_i)T_{2ij}$, we have

$$
\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}2\Delta_f(x_i)T_{2ij}\approx -2\Delta_f\left(x_i\right)\tfrac{v}{n\rho_n^2}\left[\overline{g}_2\left(x_i,\gamma_0\right)+\cdots+h^s\overline{g}_2^{(s)}\left(x_i,\gamma_0\right)\right].
$$

Hence,

$$
\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}2\Delta_f(x_i)\left(T_{1ij}+T_{2ij}\right)\approx 0
$$

and the result of interest is derived. ■

Lemma 8

$$
\begin{aligned}
&\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}2\Delta_f\left(x_i\right)\left(T_{3ij}+T_{4ij}\right)\\
\approx\ &\tfrac{2}{nh}\sum_{i=1}^{n}\Delta_f\left(x_i\right)f\left(x_i\right)u_i\left[\left(k^-\left(\tfrac{q_i-\gamma_0^v}{h}\right)-k^-\left(\tfrac{q_i-\gamma_0}{h}\right)\right)-\left(k^+\left(\tfrac{q_i-\gamma_0^v}{h}\right)-k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right)\right].
\end{aligned}
$$

Proof. The proof is the same as that of Lemma 2. ■

Lemma 9

$$
\begin{aligned}
&\tfrac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}2\Delta_f\left(x_i\right)\left(T_{5ij}+T_{6ij}\right)\\
\approx\ &\tfrac{2}{nh}\sum_{i=1}^{n}\Delta_f\left(x_i\right)f\left(x_i\right)\left(1,x_i',q_i\right)\delta_n\left[\left(k^-\left(\tfrac{q_i-\gamma_0^v}{h}\right)1(\gamma_0^v-h\leq q_i\leq\gamma_0)-k^-\left(\tfrac{q_i-\gamma_0}{h}\right)1(\gamma_0-h\leq q_i\leq\gamma_0)\right)\right].
\end{aligned}
$$

Proof. The proof is the same as that of Lemma 3. ■

Lemma 10 $Cov\left(S_1(v_1),S_1(v_2)\right)\approx\Sigma_n v_2$.

Proof. Without loss of generality, assume $v_1\geq v_2\geq 0$. Then

$$
\begin{aligned}
&4nCov\left(\Delta_f\left(x_i\right)f\left(x_i\right)u_i\left(k^+\left(\tfrac{q_i-\gamma_0^{v_1}}{h}\right)-k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right),\Delta_f\left(x_i\right)f\left(x_i\right)u_i\left(k^+\left(\tfrac{q_i-\gamma_0^{v_2}}{h}\right)-k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right)\right)\\
&=4n\mathbb{E}\left[\Delta_f^2\left(x_i\right)f\left(x_i\right)^2u_i^2\left(k^+\left(\tfrac{q_i-\gamma_0^{v_1}}{h}\right)-k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right)\left(k^+\left(\tfrac{q_i-\gamma_0^{v_2}}{h}\right)-k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right)\right]f\left(u_i|x_i,q_i\right)f\left(x_i,q_i\right)du_idq_idx_i\\
&=4n\int_{x_i}\int_{\gamma_0}^{\gamma_0^{v_2}}\int_{u_i}\Delta_f^2\left(x_i\right)f\left(x_i\right)^2u_i^2\left(k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right)^2f\left(u_i|x_i,q_i\right)f\left(x_i,q_i\right)du_idq_idx_i\ (*)\\
&\quad-4n\int_{x_i}\int_{\gamma_0^{v_2}}^{\gamma_0^{v_1}}\int_{u_i}\Delta_f^2\left(x_i\right)f\left(x_i\right)^2u_i^2k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\left(k^+\left(\tfrac{q_i-\gamma_0^{v_2}}{h}\right)-k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right)f\left(u_i|x_i,q_i\right)f\left(x_i,q_i\right)du_idq_idx_i\\
&\quad+4n\int_{x_i}\int_{\gamma_0^{v_1}}^{\gamma_0+h}\int_{u_i}\left\{\begin{array}{c}\Delta_f^2\left(x_i\right)f\left(x_i\right)^2u_i^2\left(k^+\left(\tfrac{q_i-\gamma_0^{v_1}}{h}\right)-k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right)\\ \cdot\left(k^+\left(\tfrac{q_i-\gamma_0^{v_2}}{h}\right)-k^+\left(\tfrac{q_i-\gamma_0}{h}\right)\right)f\left(u_i|x_i,q_i\right)f\left(x_i,q_i\right)\end{array}\right\}du_idq_idx_i\\
&\quad+4n\int_{x_i}\int_{\gamma_0+h}^{\gamma_0^{v_2}+h}\int_{u_i}\Delta_f^2\left(x_i\right)f\left(x_i\right)^2u_i^2k^+\left(\tfrac{q_i-\gamma_0^{v_1}}{h}\right)k^+\left(\tfrac{q_i-\gamma_0^{v_2}}{h}\right)f\left(u_i|x_i,q_i\right)f\left(x_i,q_i\right)du_idq_idx_i\\
&\approx 4k_+\left(0\right)^2\mathbb{E}\left[\Delta_f^2\left(x_i\right)f\left(x_i\right)^2\sigma_+^2\left(x_i\right)|q_i=\gamma_0\right]f_q\left(\gamma_0\right)/\rho_n^2v_2,
\end{aligned}
$$

where $\sigma_+^2(x_i) = \mathbb{E}[u_i^2|x_i, q_i = \gamma_0+]$. Similarly,

$$\begin{aligned}&4nCov\left[\Delta_f(x_i) f(x_i) u_i \left(k^-\left(\tfrac{q_i-\gamma_0^{v_1}}{h}\right) - k^-\left(\tfrac{q_i-\gamma_0}{h}\right)\right), \Delta_f(x_i) f(x_i) u_i \left(k^-\left(\tfrac{q_i-\gamma_0^{v_2}}{h}\right) - k^-\left(\tfrac{q_i-\gamma_0}{h}\right)\right)\right] \\ &\approx 4k_+(0)^2 \mathbb{E}\left[\Delta_f^2(x_i) f(x_i)^2 \sigma_+^2(x_i) | q_i = \gamma_0\right] f_q(\gamma_0)/\rho_n^2 v_2,\end{aligned}$$

and

$$\begin{aligned}&4nCov\left[\Delta_f(x_i) f(x_i) u_i \left(k^\pm\left(\tfrac{q_i-\gamma_0^{v_1}}{h}\right) - k^\pm\left(\tfrac{q_i-\gamma_0}{h}\right)\right), \Delta_f(x_i) f(x_i) u_i \left(k^\mp\left(\tfrac{q_i-\gamma_0^{v_2}}{h}\right) - k^\mp\left(\tfrac{q_i-\gamma_0}{h}\right)\right)\right] \\ &\approx 4k_+(0)^2 \mathbb{E}\left[\Delta_f^2(x_i) f(x_i)^2 \sigma_+^2(x_i) | q_i = \gamma_0\right] f_q(\gamma_0)/\rho_n^2 v_2.\end{aligned}$$

If $v_1 \le v_2 \le 0$, the result is similar except that the term $\sigma_+^2(x_i)$ is replaced by $\sigma_-^2(x_i)$.

If $v_1 v_2 < 0$, then the four terms are all $o(v_2)$. ■

Lemma 11 *$Cov(S_2(v_1), S_2(v_2)) = o(v_2)$, $Cov(S_1(v_1), S_2(v_2)) = o(v_2)$ and $Cov(S_2(v_1), S_1(v_2)) = o(v_2)$.*

Proof. The proof idea is the same as that in Lemma 5. ■

Lemma 12 *For any $\phi_1, \phi_2 > 0$, there exists $\eta > 0$ such that*

$$P\left\{\sup_{|v_1-v_2|<\eta}\left|\widehat{nh}\left(\widehat{Q}_n(\gamma_0^{v_1}) - \widehat{Q}_n(\gamma_0)\right) - \widehat{nh}\left(\widehat{Q}_n(\gamma_0^{v_2}) - \widehat{Q}_n(\gamma_0)\right)\right| > \phi_1\right\} < \phi_2,$$

where

$$\widehat{nh}\left(\widehat{Q}_n(\gamma_0^{v}) - \widehat{Q}_n(\gamma_0)\right) = nh\left(\widehat{Q}_n(\gamma_0^{v}) - \widehat{Q}_n(\gamma_0)\right) - \mathbb{E}\left[nh\left(\widehat{Q}_n(\gamma_0^{v}) - \widehat{Q}_n(\gamma_0)\right)\right].$$

Proof. Without loss of generality, assume $v_1 \ge v_2 \ge 0$. From Proposition 8, we have

$$\begin{aligned}&\lim_{n\to\infty}\mathbb{E}\left[\widehat{nh}\left(\widehat{Q}_n(\gamma_0^{v_1}) - \widehat{Q}_n(\gamma_0)\right) - \widehat{nh}\left(\widehat{Q}_n(\gamma_0^{v_2}) - \widehat{Q}_n(\gamma_0)\right)\right]^2 \\ &= \lim_{n\to\infty}\mathbb{E}\left[\widehat{nh}\left(\widehat{Q}_n(\gamma_0^{v_1}) - \widehat{Q}_n(\gamma_0)\right)\right]^2 + \lim_{n\to\infty}\mathbb{E}\left[\widehat{nh}\left(\widehat{Q}_n(\gamma_0^{v_2}) - \widehat{Q}_n(\gamma_0)\right)\right]^2 \\ &\quad - 2\lim_{n\to\infty}\mathbb{E}\left[\widehat{nh}\left(\widehat{Q}_n(\gamma_0^{v_1}) - \widehat{Q}_n(\gamma_0)\right)\right]\left[\widehat{nh}\left(\widehat{Q}_n(\gamma_0^{v_2}) - \widehat{Q}_n(\gamma_0)\right)\right] \\ &= 16k^2(0)(v_1 + v_2 - 2v_2) V_1 f_q(\gamma_0) \\ &\le C|v_1 - v_2|.\end{aligned}$$

By Markov's inequality, the result follows. ■

Lemma 13 $\frac{1}{n}\sum_{j=1}^n 2\Delta_f(x_o)(T_{1j} + T_{2j}) \approx 0$.

Proof. First,

$$\begin{aligned}&\mathbb{E}[2\Delta_f(x_o)T_{1j}] \\ &= \int_{-\infty}^{+\infty}\int_{x_j} 2\Delta_f(x_o) g(x_j, q_j) \tfrac{1}{h^d} K^x\left(\tfrac{x_j - x_o}{h}, x_o\right)\left(k^-\left(\tfrac{q_j-\gamma_0^v}{h}\right) - k^-\left(\tfrac{q_j-\gamma_0}{h}\right)\right) f(x_j, q_j)\, dx_j dq_j \\ &= 2\Delta_f(x_o)\int_{-1}^0\int_{u_x} g(x_o + u_x h, \gamma_0^v + u_q h) K^x(u_x, x_o) k^-(u_q) f(x_o + u_x h, \gamma_0^v + u_q h)\, du_x du_q \\ &\quad - 2\Delta_f(x_o)\int_{-1}^0\int_{u_x} g(x_o + u_x h, \gamma_0 + u_q h) K^x(u_x, x_o) k^-(u_q) f(x_o + u_x h, \gamma_0 + u_q h)\, du_x du_q \\ &= 2\Delta_f(x_o)\int_{-1}^0\int_{u_x}\left[\overline{g}(x_o, \gamma_0^v + u_q h) + \cdots + \overline{g}_1^{(s)}(x_o, \gamma_0^v + u_q h) u_x h + o(u_x h)\right] K^x(u_x, x_o) k^-(u_q)\, du_x du_q \\ &\quad - 2\Delta_f(x_o)\int_{-1}^0\int_{u_x}\left[\overline{g}(x_o, \gamma_0 + u_q h)\cdots + \overline{g}_1^{(s)}(x_o, \gamma_0 + u_q h) u_x h + o(u_x h)\right] K^x(u_x, x_o) k^-(u_q)\, du_x du_q \\ &\approx 2\Delta_f(x_o)\int_{-1}^0\left[\left(\overline{g}(x_o, \gamma_0^v) + \cdots + \overline{g}_1^{(s)}(x_o, \gamma_0^v) h^s\right) - \left(\overline{g}(x_o, \gamma_0) + \cdots + \overline{g}_1^{(s)}(x_o, \gamma_0) h^s\right)\right] k^-(u_q)\, du_q \\ &\approx 2\Delta_f(x_o)\left(\overline{g}_2(x_o, \gamma_0) + \cdots + \overline{g}_{12}^{(s)}(x_o, \gamma_0) h^s\right)\tfrac{v}{nh^{d-1}\Delta_o^2},\end{aligned}$$

where $\overline{g}(x_j,q_j)=g(x_j,q_j)f(x_j,q_j)$, $\overline{g}_1^{(s)}(x_j,q_j)=\frac{\partial \overline{g}(x_j,q_j)}{(\partial x_j)^s}$, $\overline{g}_2(x_j,q_j)=\frac{\partial \overline{g}(x_j,q_j)}{\partial q_j}$ and $\overline{g}_{12}^{(s)}(x_j,q_j)=\frac{\partial \overline{g}_1^{(s)}(x_j,q_j)}{\partial q_j}$. Similarly, we have

$$\mathbb{E}\left[2\Delta_f(x_o)T_{2j}\right]\approx -2\Delta_f\left(x_o\right)\left(\overline{g}_2\left(x_o,\gamma_0\right)+\cdots+\overline{g}_{12}^{(s)}\left(x_o,\gamma_0\right)h^s\right)\frac{v}{nh^{d-1}\Delta_o^2}.$$

Hence,

$$\mathbb{E}\left[\frac{1}{n}\sum_{j=1}^n 2\Delta_f(x_o)\Delta_f(x_o)\left(T_{1j}+T_{2j}\right)\right]\approx 0.$$

Next,

$$\begin{aligned}
&\frac{1}{n}\mathbb{E}\left[4\Delta_f^2(x_o)T_{1j}^2\right]\\
&=\frac{1}{n}\int_{q_j}\int_{x_j}\frac{4}{h^{2d}}\Delta_f^2\left(x_o\right)g\left(x_j,q_j\right)^2K^x\left(\frac{x_j-x_o}{h},x_o\right)^2\left(k^-\left(\frac{q_j-\gamma_0^v}{h}\right)-k^-\left(\frac{q_j-\gamma_0}{h}\right)\right)^2 f\left(x_j,q_j\right)dx_jdq_j\\
&=\frac{1}{n}\int_{u_q}\int_{u_x}\left\{\begin{array}{c}\frac{4}{h^d}\Delta_f^2\left(x_o\right)g\left(x_o+u_xh,\gamma_0+u_qh\right)^2K^x\left(u_x,x_o\right)^2\\ \cdot\left(k^-\left(u_q+\frac{\gamma_0-\gamma_0^v}{h}\right)-k^-\left(u_q\right)\right)^2 f\left(x_o+u_xh,\gamma_0+u_qh\right)\end{array}\right\}du_xdu_q\\
&=O\left(\frac{\Delta_o^2}{nh^d}\frac{v^2}{nh^2\Delta_o^2}\right)=o(1),
\end{aligned}$$

so

$$Var\left[\frac{1}{n}\sum_{j=1}^n 2\Delta_f(x_o)T_{1j}\right]=\frac{1}{n}Var\left[2\Delta_f(x_o)T_{1j}\right]=\frac{1}{n}\mathbb{E}\left[4\Delta_f^2(x_o)T_{1j}^2\right]-\frac{1}{n}\mathbb{E}\left[2\Delta_f(x_o)T_{1j}\right]^2=o(1).$$

With similar results derived for $\frac{1}{n}\sum_{j=1}^n 2\Delta_f(x_o)T_{2j}$ and

$$\begin{aligned}
&\frac{1}{n}\mathbb{E}\left[4\Delta_f(x_o)^2T_{1j}T_{2j}\right]\\
&=\frac{1}{n}\int_{u_q}\int_{u_x}\left\{\begin{array}{c}\frac{4}{h^d}\Delta_f^2\left(x_o\right)g\left(x_0+u_xh,\gamma_0+u_qh\right)^2\left(k^-\left(u_q-\frac{v}{nh^d\Delta_o^2}\right)-k^-\left(u_q\right)\right)\\ \cdot\left(k^+\left(u_q-\frac{v}{nh^d\Delta_o^2}\right)-k^+\left(u_q\right)\right)K^x\left(u_x,x_o\right)^2 f\left(x_0+u_xh,\gamma_0+u_qh\right)\end{array}\right\}du_xdu_q\\
&=O\left(\frac{\Delta_o^2}{nh^d}\frac{v}{nh^d\Delta_o^2}\right),
\end{aligned}$$

we have

$$\frac{1}{n}Cov\left[2\Delta_f(x_o)T_{1j},2\Delta_f(x_o)T_{2j}\right]=O\left(\frac{1}{nh^d}\frac{v}{nh^d}\right)-O\left(\frac{1}{nh^d}\frac{v}{nh^2}\right)=o(1).$$

As a result,

$$Var\left[\frac{1}{n}\sum_{j=1}^n 2\Delta_f(x_o)(T_{1j}+T_{2j})\right]=o(1).$$

By Markov's inequality, the result of interest is obtained. ■

Lemma 14 $\frac{1}{n}\sum_{j=1}^n 2\Delta_f\left(x_o\right)\left(T_{5j}+T_{6j}\right)\approx -2k_-(0)\frac{v}{nh^d}f(x_o,\gamma_0)^2.$

Proof. Since

$$\begin{aligned}
&\mathbb{E}[2\Delta_f(x_o)T_{5j}]\\
&=-\int_{q_j}\int_{x_j}\frac{2}{h^d}\Delta_f(x_o)(1,x_j',q_j)\delta_nK^x\left(\frac{x_j-x_o}{h},x_o\right)k^-\left(\frac{q_j-\gamma_0}{h}\right)1(\gamma_0-h\le q_j\le\gamma_0)f(x_j,q_j)dx_jdq_j\\
&=-2\Delta_f(x_o)\int_{-1}^0\int_{u_x}(1,x_o'+u_x'h,\gamma_0+u_qh)\delta_nK^x(u_x,x_o)k^-(u_q)f(x_o+u_xh,\gamma_0+u_qh)du_xdu_q\\
&=-2\Delta_f(x_o)(1,x_o',\gamma_0)\delta_nf\left(x_o,\gamma_0\right)\left(1+O\left(h\right)\right)\approx -2\Delta_f^2(x_o)
\end{aligned}$$

and, similarly,

$$\begin{aligned}
&\mathbb{E}[2\Delta_f(x_o)T_{6j}] \\
&= \int_{q_j}\int_{x_j}\tfrac{2}{h^d}\Delta_f(x_o)(1,x_o',q_j)\delta_n K^x\left(\tfrac{x_j-x_o}{h},x_o\right)k^-\left(\tfrac{q_j-\gamma_0^v}{h}\right)1(\gamma_0^v-h\le q_j\le\gamma_0)f(x_j,q_j)dx_jdq_j \\
&= 2\Delta_f(x_o)\int_{-1+\frac{v}{nh^d\Delta_o^2}}^{0}\int_{u_x}(1,x_o'+u_x'h,\gamma_0+u_qh)\delta_nK^x(u_x,x_o)k^-\left(u_q-\tfrac{v}{nh^d\Delta_o^2}\right)f(x_o+u_xh,\gamma_0+u_qh)du_xdu_q \\
&= 2\Delta_f(x_o)(1,x_o',\gamma_0)\delta_nf(x_o,\gamma_0)\left(1-k_+(0)\tfrac{v}{nh^d\Delta_o^2}+O(h)\right)\approx 2\Delta_f^2(x_o)(1-k_-(0)\tfrac{v}{nh^d\Delta_o^2}),
\end{aligned}$$

the result of interest follows. ■

Lemma 15 *$Cov\,(S_1(v_1),S_1(v_2))\approx\Sigma_o v_2$.*

Proof. Without loss of generality, assume $v_1\ge v_2\ge 0$. Then

$$\begin{aligned}
&4nh^{2d}Cov\left(\Delta_f(x_o)u_j\left(K_{h,j}^{\gamma_0^{v_1}+}-K_{h,j}^{\gamma_0+}\right),\Delta_f(x_i)u_i\left(K_{h,j}^{\gamma_0^{v_2}+}-K_{h,j}^{\gamma_0+}\right)\right) \\
&= 4nh^d\mathbb{E}\left[\Delta_f^2(x_o)u_j^2K^x\left(\tfrac{x_j-x_o}{h},x_o\right)^2\left(k^+\left(\tfrac{q_j-\gamma_0^{v_1}}{h}\right)-k^+\left(\tfrac{q_j-\gamma_0}{h}\right)\right)\left(k^+\left(\tfrac{q_j-\gamma_0^{v_2}}{h}\right)-k^+\left(\tfrac{q_j-\gamma_0}{h}\right)\right)\right] \\
&= 4nh^d\int_{x_j}\int_{\gamma_0}^{\gamma_0^{v_2}}\int_{u_i}\Delta_f^2(x_o)u_j^2K^x\left(\tfrac{x_j-x_o}{h},x_o\right)^2\left(k^+\left(\tfrac{q_j-\gamma_0}{h}\right)\right)^2f(u_j|x_j,q_j)f(x_j,q_j)du_jdq_jdx_j\ (*) \\
&\quad -4nh^d\int_{x_i}\int_{\gamma_0^{v_2}}^{\gamma_0^{v_1}}\int_{u_i}\Delta_f^2(x_o)u_j^2K^x\left(\tfrac{x_j-x_o}{h},x_o\right)^2k^+\left(\tfrac{q_j-\gamma_0}{h}\right)\left(k^+\left(\tfrac{q_j-\gamma_0^{v_2}}{h}\right)-k^+\left(\tfrac{q_j-\gamma_0}{h}\right)\right)f(u_j|x_j,q_j)f(x_j,q_j)du_jdq_jdx_j \\
&\quad +4nh^d\int_{x_i}\int_{\gamma_0^{v_1}}^{\gamma_0+h}\int_{u_i}\left\{\begin{array}{c}\Delta_f^2(x_o)u_j^2K^x\left(\tfrac{x_j-x_o}{h},x_o\right)^2\left(k^+\left(\tfrac{q_j-\gamma_0^{v_1}}{h}\right)-k^+\left(\tfrac{q_j-\gamma_0}{h}\right)\right)\\ \cdot\left(k^+\left(\tfrac{q_j-\gamma_0^{v_2}}{h}\right)-k^+\left(\tfrac{q_j-\gamma_0}{h}\right)\right)f(u_j|x_j,q_j)f(x_j,q_j)\end{array}\right\}du_jdq_jdx_j \\
&\quad +4nh^d\int_{x_i}\int_{\gamma_0+h}^{\gamma_0^{v_2}+h}\int_{u_i}\Delta_f^2(x_o)u_j^2K^x\left(\tfrac{x_j-x_o}{h},x_o\right)^2k^+\left(\tfrac{q_j-\gamma_0^{v_1}}{h}\right)k^+\left(\tfrac{q_j-\gamma_0^{v_2}}{h}\right)f(u_j|x_j,q_j)f(x_j,q_j)du_jdq_jdx_j \\
&\approx 4k_+(0)^2\sigma_+^2(x_o)f_q(x_o,\gamma_0)^3\kappa^2v_2
\end{aligned}$$

and

$$\begin{aligned}
&4nh^{2d}Cov\left(\Delta_f(x_o)u_j\left(K_{h,j}^{\gamma_0^{v_1}-}-K_{h,j}^{\gamma_0-}\right),\Delta_f(x_o)u_j\left(K_{h,j}^{\gamma_0^{v_2}-}-K_{h,j}^{\gamma_0-}\right)\right) \\
\approx\,&4k_-^2(0)\sigma_+^2(x_o)f(x_o,\gamma_0)^3\kappa^2v_2,
\end{aligned}$$

and

$$\begin{aligned}
&-4nh^{2d}Cov\left(\Delta_f(x_o)e_j\left(K_{h,j}^{\gamma_0^{v_1}\pm}-K_{h,j}^{\gamma_0\pm}\right),\Delta_f(x_o)e_j\left(K_{h,j}^{\gamma_0^{v_2}\mp}-K_{h,j}^{\gamma_0\mp}\right)\right) \\
\approx\,&4k_+(0)k_-(0)\sigma_+^2(x_o)f(x_o,\gamma_0)^3\kappa^2v_2,
\end{aligned}$$

where $\sigma_\pm^2(x_o)=\mathbb{E}[e_j^2|x_j=x_o,q_j=\gamma_0^\pm]$, $\kappa^2=\int K(u_x)^2du_x$.

If $v_1\le v_2\le 0$, the result is similar except that the term $\sigma_+^2(x_o)$ is replaced by $\sigma_-^2(x_o)$.

If $v_1v_2<0$, then the four terms are all $o(v_2)$. ■

Lemma 16 *$Cov\,(S_2(v_1),S_2(v_2))=o\,(v_2)\,,Cov\,(S_1(v_1),S_2(v_2))=o\,(v_2)$ and $Cov\,(S_2(v_1),S_1(v_2))=o\,(v_2)$.*

Proof. The proof idea is the same as that in Lemma 11. ■

Lemma 17 *For any $\phi_1,\phi_2>0$, there exists $\eta>0$ such that*

$$P\left\{\sup_{|v_1-v_2|<\eta}\left|\widehat{nh^d}\left(\widetilde{Q}_n\left(\gamma_0^{v_1}\right)-\widetilde{Q}_n\left(\gamma_0\right)\right)-\widehat{nh^d}\left(\widetilde{Q}_n\left(\gamma_0^{v_2}\right)-\widetilde{Q}_n\left(\gamma_0\right)\right)\right|>\phi_1\right\}<\phi_2,$$

where

$$\widehat{nh^d}\left(\widetilde{Q}_n\left(\gamma_0^v\right)-\widetilde{Q}_n\left(\gamma_0\right)\right)=nh^d\left(\widetilde{Q}_n\left(\gamma_0^v\right)-\widetilde{Q}_n\left(\gamma_0\right)\right)-\mathbb{E}\left[nh^d\left(\widetilde{Q}_n\left(\gamma_0^v\right)-\widetilde{Q}_n\left(\gamma_0\right)\right)\right].$$

Proof. Without loss of generality, assume $v_1 \geq v_2 \geq 0$. From Proposition 11, we have

$$\begin{aligned}
&\lim_{n\to\infty}\mathbb{E}\left[\widehat{nh^d}\left(\widetilde{Q}_n\left(\gamma_0^{v_1}\right)-\widetilde{Q}_n\left(\gamma_0\right)\right)-\widehat{nh^d}\left(\widetilde{Q}_n\left(\gamma_0^{v_2}\right)-\widetilde{Q}_n\left(\gamma_0\right)\right)\right]^2\\
&=\lim_{n\to\infty}\mathbb{E}\left[\widehat{nh^d}\left(\widetilde{Q}_n\left(\gamma_0^{v_1}\right)-\widetilde{Q}_n\left(\gamma_0\right)\right)\right]^2+\lim_{n\to\infty}\mathbb{E}\left[\widehat{nh^d}\left(\widetilde{Q}_n\left(\gamma_0^{v_2}\right)-\widetilde{Q}_n\left(\gamma_0\right)\right)\right]^2\\
&\quad-2\lim_{n\to\infty}\mathbb{E}\left[\widehat{nh^d}\left(\widetilde{Q}_n\left(\gamma_0^{v_1}\right)-\widetilde{Q}_n\left(\gamma_0\right)\right)\right]\left[\widehat{nh^d}\left(\widetilde{Q}_n\left(\gamma_0^{v_2}\right)-\widetilde{Q}_n\left(\gamma_0\right)\right)\right]\\
&=16k^2(0)\,f\left(x_o,\gamma_0\right)^3\kappa^2\sigma_+^2\left(x_o\right)\left(v_1+v_2-2v_2\right)\\
&\leq C\left|v_1-v_2\right|.
\end{aligned}$$

By Markov's inequality, the result follows. ∎

Lemma 18 $\left|\mathbb{E}_1\left[\left(m_2-m_1\right)L_{b,21}\right]\right|=O_p\left(b^\eta\right)$.

Proof. We have

$$\begin{aligned}
&\left|\mathbb{E}\left[\left(m(x_2,q_2)-m(x_1,q_1)\right)L_{b,21}|x_1,q_1\right]\right|\\
&=\left|\int\left(m(x_2,q_2)-m(x_1,q_1)\right)f(x_2,q_2)\tfrac{1}{b^d}L^x\left(\tfrac{x_2-x_1}{b},x_1\right)l\left(\tfrac{q_2-q_1}{b}\right)dx_2dq_2\right|\\
&=\left|\int\left\{\begin{array}{c}\left(Q_m\left(\left(x_2,q_2\right),\left(x_1,q_1\right)\right)+R_m\left(\left(x_2,q_2\right),\left(x_1,q_1\right)\right)\right)\\ \cdot\left(f(x_1,q_1)+Q_f\left(\left(x_2,q_2\right),\left(x_1,q_1\right)\right)+R_f\left(\left(x_2,q_2\right),\left(x_1,q_1\right)\right)\right)\\ \cdot\frac{1}{b^d}L^x\left(\frac{x_2-x_1}{b},x_1\right)l\left(\frac{q_2-q_1}{b}\right)\end{array}\right\}dx_2dq_2\right|,
\end{aligned}$$

where $Q_m\left(\left(x_2,q_2\right),\left(x_1,q_1\right)\right)$ is the $(s-1)$th-order Taylor expansion of $m(x_2,q_2)$ at $m(x_1,q_1)$, $R_m\left(\left(x_2,q_2\right),\left(x_1,q_1\right)\right)$ is the remainder term, $Q_f\left(\left(x_2,q_2\right),\left(x_1,q_1\right)\right)$ is $(\lambda-1)$th-order Taylor expansion of $f(x_2,q_2)$ at $f(x_1,q_1)$, and $R_f\left(\left(x_2,q_2\right),\left(x_1,q_1\right)\right)$ is the remainder term. From Assumption L,

$$\int Q_m\left(\left(x_2,q_2\right),\left(x_1,q_1\right)\right)\left(f(x_1,q_1)+Q_f\left(\left(x_2,q_2\right),\left(x_1,q_1\right)\right)\right)\tfrac{1}{b^d}L^x\left(\tfrac{x_2-x_1}{b},x_1\right)l\left(\tfrac{q_2-q_1}{b}\right)dx_2dq=0,$$

so $\left|\mathbb{E}\left[\left(m(x_2,q_2)-m(x_1,q_1)\right)L_{b,21}|x_1\right]\right|$ is bounded by

$$\begin{aligned}
&\left|\int R_m\left(\left(x_2,q_2\right),\left(x_1,q_1\right)\right)f(x_1,q_1)\tfrac{1}{b^d}L^x\left(\tfrac{x_2-x_1}{b},x_1\right)l\left(\tfrac{q_2-q_1}{b}\right)dx_2dq_2\right|\\
&+\left|\int\left(m(x_2,q_2)-m(x_1,q_1)\right)R_f\left(\left(x_2,q_2\right),\left(x_1,q_1\right)\right)\tfrac{1}{b^d}L^x\left(\tfrac{x_2-x_1}{b},x_1\right)l\left(\tfrac{q_2-q_1}{b}\right)dx_2dq_2\right|\\
&\leq Cb^s+Cb^{\lambda+1}\leq Cb^\eta,
\end{aligned}$$

where $\eta=\min\left(\lambda+1,s\right)$. ∎

Supplement D: Parametric Tests for Threshold Effects when Instruments are present

This supplement discusses the asymptotic distribution of the Wald-type and score-type test statistics under the null and local alternatives when instruments are available. We also provide implementation details for the use of Hansen's (1996) simulation method in the current context.

For $\gamma \in \Gamma$, define

$$\begin{aligned}
&\Omega_1(\gamma) = \mathbb{E}\left[\mathbf{z}_i\mathbf{z}_i'\varepsilon_i^2 1(q_i \le \gamma)\right], \Omega_2(\gamma) = \mathbb{E}\left[\mathbf{z}_i\mathbf{z}_i'\varepsilon_i^2 1(q_i > \gamma)\right], \\
&Q_1(\gamma) = \mathbb{E}\left[\mathbf{z}_i\mathbf{x}_i' 1(q_i \le \gamma)\right], Q_2(\gamma) = \mathbb{E}\left[\mathbf{z}_i\mathbf{x}_i' 1(q_i > \gamma)\right], \\
&V_1(\gamma) = \left[Q_1(\gamma)'\,\Omega_1(\gamma)^{-1} Q_1(\gamma)\right]^{-1}, V_2(\gamma) = \left[Q_2(\gamma)'\,\Omega_2(\gamma)^{-1} Q_2(\gamma)\right]^{-1}, \\
&\Omega = \mathbb{E}\left[\mathbf{z}_i\mathbf{z}_i'\varepsilon_i^2\right], Q = \mathbb{E}\left[\mathbf{z}_i\mathbf{x}_i'\right], V = \left[Q'\Omega^{-1}Q\right]^{-1}.
\end{aligned}$$

$S_1(\gamma)$ is a mean-zero Gaussian process with covariance kernel $\mathbb{E}\left[S_1(\gamma_1) S_1(\gamma_2)'\right] = \Omega_1(\gamma_1 \wedge \gamma_2)$, $S = \lim_{\gamma \to \infty} S_1(\gamma)$, and $S_2(\gamma) = S - S_1(\gamma)$. $S(\gamma)$ is a mean zero Gaussian process with covariance kernel

$$H(\gamma_1, \gamma_2) = \mathbb{E}\left[\left(\mathbf{z}_i 1(q_i \le \gamma_1) - Q_1(\gamma_1) V Q'\Omega^{-1}\mathbf{z}_i\right)\left(\mathbf{z}_i 1(q_i \le \gamma_2) - Q_1(\gamma_2) V Q'\Omega^{-1}\mathbf{z}_i\right)' \varepsilon_i^2\right].$$

Given the threshold point γ, the 2SLS estimators for β_1 and β_2 are

$$\begin{aligned}
\widetilde{\beta}_1(\gamma) &= \left(\widehat{Q}_1(\gamma)'\left(\tfrac{1}{n}Z_{\le\gamma}'Z_{\le\gamma}\right)^{-1}\widehat{Q}_1(\gamma)\right)^{-1}\left(\widehat{Q}_1(\gamma)'\left(\tfrac{1}{n}Z_{\le\gamma}'Z_{\le\gamma}\right)^{-1}\tfrac{1}{n}Z_{\le\gamma}'Y\right), \\
\widetilde{\beta}_2(\gamma) &= \left(\widehat{Q}_2(\gamma)'\left(\tfrac{1}{n}Z_{>\gamma}'Z_{>\gamma}\right)^{-1}\widehat{Q}_2(\gamma)\right)^{-1}\left(\widehat{Q}_2(\gamma)'\left(\tfrac{1}{n}Z_{>\gamma}'Z_{>\gamma}\right)^{-1}\tfrac{1}{n}Z_{>\gamma}'Y\right),
\end{aligned}$$

where $\widehat{Q}_1(\gamma) = n^{-1}\sum_{i=1}^n \mathbf{z}_i\mathbf{x}_i' 1(q_i \le \gamma)$ and $\widehat{Q}_2(\gamma) = n^{-1}\sum_{i=1}^n \mathbf{z}_i\mathbf{x}_i' 1(q_i > \gamma)$. The residual from this equation is

$$\widetilde{\varepsilon}_i(\gamma) = y_i - \mathbf{x}_i'\widetilde{\beta}_1(\gamma) 1(q_i \le \gamma) - \mathbf{x}_i'\widetilde{\beta}_2(\gamma) 1(q_i > \gamma).$$

The GMM estimators for β_1 and β_2 are

$$\begin{aligned}
\widehat{\beta}_1(\gamma) &= \left(\widehat{Q}_1(\gamma)'\,\widetilde{\Omega}_1^{-1}(\gamma)\,\widehat{Q}_1(\gamma)\right)^{-1}\left(\widehat{Q}_1(\gamma)'\,\widetilde{\Omega}_1^{-1}(\gamma)\,\tfrac{1}{n}Z_{\le\gamma}'Y\right), \\
\widehat{\beta}_2(\gamma) &= \left(\widehat{Q}_2(\gamma)'\,\widetilde{\Omega}_2^{-1}(\gamma)\,\widehat{Q}_2(\gamma)\right)^{-1}\left(\widehat{Q}_2(\gamma)'\,\widetilde{\Omega}_2^{-1}(\gamma)\,\tfrac{1}{n}Z_{>\gamma}'Y\right),
\end{aligned}$$

where the weight matrices

$$\widetilde{\Omega}_1(\gamma) = \tfrac{1}{n}\textstyle\sum_{i=1}^n \mathbf{z}_i\mathbf{z}_i'\widetilde{\varepsilon}_i^2(\gamma) 1(q_i \le \gamma),\ \widetilde{\Omega}_2(\gamma) = \tfrac{1}{n}\sum_{i=1}^n \mathbf{z}_i\mathbf{z}_i'\widetilde{\varepsilon}_i^2(\gamma) 1(q_i > \gamma).$$

The estimated covariance matrices for the GMM estimators are

$$\widehat{V}_1(\gamma) = \left(\widehat{Q}_1(\gamma)'\,\widetilde{\Omega}_1^{-1}(\gamma)\,\widehat{Q}_1(\gamma)\right)^{-1},\ \widehat{V}_2(\gamma) = \left(\widehat{Q}_2(\gamma)'\,\widetilde{\Omega}_2^{-1}(\gamma)\,\widehat{Q}_2(\gamma)\right)^{-1}.$$

When H_0 holds, $\delta = 0$, and then the 2SLS estimator for β is

$$\widetilde{\beta} = \left(\widehat{Q}'\left(\tfrac{1}{n}Z'Z\right)^{-1}\widehat{Q}\right)^{-1}\left(\widehat{Q}'\left(\tfrac{1}{n}Z'Z\right)^{-1}\tfrac{1}{n}Z'Y\right),$$

where $\widehat{Q} = n^{-1}\sum_{i=1}^n \mathbf{z}_i\mathbf{x}_i'$. Note here that the underlying assumption in this specification testing context is $\mathbb{E}[\varepsilon|\mathbf{z}] = 0$, so that the 2SLS estimator can be applied. Correspondingly, the GMM estimator for β is

$$\widehat{\beta} = \left(\widehat{Q}'\widetilde{\Omega}^{-1}\widehat{Q}\right)^{-1}\left(\widehat{Q}'\widetilde{\Omega}^{-1}\tfrac{1}{n}Z'Y\right),$$

and the residual is

$$\widehat{\varepsilon}_i = y_i - \mathbf{x}_i'\widehat{\beta},$$

where the weight matrix is

$$\widetilde{\Omega} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_i\mathbf{z}_i'\widetilde{\varepsilon}_i^2$$

with $\widetilde{\varepsilon}_i = y_i - \mathbf{x}_i'\widetilde{\beta}$. The estimated covariance matrix for the GMM estimator is

$$\widehat{V} = \left(\widehat{Q}'\widetilde{\Omega}^{-1}\widehat{Q}\right)^{-1}.$$

Wald-type Tests

Define

$$W_n(\gamma) = \left(\widehat{V}_1(\gamma) + \widehat{V}_2(\gamma)\right)^{-1/2}\sqrt{n}\left(\widehat{\beta}_1(\gamma) - \widehat{\beta}_2(\gamma)\right), \gamma \in \Gamma.$$

The Wald-type test statistic is a functional of $W_n(\cdot)$. Two test statistics are the most popular. The first is the Kolmogorov-Smirnov sup-type statistic

$$K_n^{\omega} = \sup_{\gamma\in\Gamma}\left\|W_n(\gamma)\right\|,$$

and the second is the Cramér–von Mises average-type statistic

$$C_n^{\omega} = \int_{\Gamma}\left\|W_n(\gamma)\right\| w(\gamma)d\gamma,$$

where $w(\gamma)$ in C_n^{ω} is a known positive weight function with $\int_{\Gamma} w(\gamma)d\gamma = 1$. For example, $w(\tau) = 1/|\Gamma|$ with $|\Gamma|$ being the length of Γ. But if we have some information on the locations where threshold effects are most likely to occur, we can impose larger weights on the neighborhoods of such locations. The choice of the norm $\|\cdot\|$ is also an issue. The Euclidean norm $\|\cdot\|_2$ is obviously natural, e.g., CH use (the square of) this norm. Yu (2013b) suggests using the ℓ_1 norm in testing quantile threshold effects, and Bai (1996) suggests using the ℓ_∞ norm in structural change tests.

The following theorem states the asymptotic distribution of a general continuous functional $g(\cdot)$ of $W_n(\cdot)$ under the local alternative $\delta_n = n^{-1/2}c$. The corresponding test statistic is denoted as g_n^{ω}.

Theorem 10 *If $\delta_n = n^{-1/2}c$, $\mathbb{E}\left[\|x\|^4\right] < \infty$, $\mathbb{E}[q^4] < \infty$, $\mathbb{E}[\varepsilon^4]$ and $\mathbb{E}\left[\|\mathbf{z}\|^4\right] < \infty$, then*

$$g_n^{\omega} \xrightarrow{d} g_c^{\omega} = g(W^c),$$

where

$$\begin{aligned} W^c(\gamma) = {} & (V_1(\gamma) + V_2(\gamma))^{-1/2}\left[V_1(\gamma)Q_1(\gamma)'\Omega_1(\gamma)^{-1}Q_1(\gamma\wedge\gamma_0) - V_2(\gamma)Q_2(\gamma)'\Omega_2(\gamma)^{-1}Q_2(\gamma\vee\gamma_0)\right]c \\ & + (V_1(\gamma) + V_2(\gamma))^{-1/2}\left[V_1(\gamma)Q_1(\gamma)'\Omega_1(\gamma)^{-1}S_1(\gamma) - V_2(\gamma)Q_2(\gamma)'\Omega_2(\gamma)^{-1}S_2(\gamma)\right]. \end{aligned}$$

Proof. Under the local alternative $\delta_n = n^{-1/2}c$, $Y = X_{\le\gamma_0}(\beta + \delta_n) + X_{>\gamma_0}\beta + \boldsymbol{\varepsilon} = X\beta + X_{\le\gamma_0}\delta_n + \boldsymbol{\varepsilon}$, so that

$$\begin{aligned} \widetilde{\beta}_1(\gamma) &= \left(X_{\le\gamma}'Z_{\le\gamma}\left(Z_{\le\gamma}'Z_{\le\gamma}\right)^{-1}Z_{\le\gamma}'X_{\le\gamma}\right)^{-1}\left(X_{\le\gamma}'Z_{\le\gamma}\left(Z_{\le\gamma}'Z_{\le\gamma}\right)^{-1}Z_{\le\gamma}'Y\right) \\ &= \beta + O_p(1)\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\left[\mathbf{x}_i'\delta_n 1\left(q_i \le \gamma_0\wedge\gamma\right) + \varepsilon_i 1\left(q_i \le \gamma\right)\right] \\ &= \beta + o_p(1) \text{ uniformly in } \gamma \in \Gamma. \end{aligned}$$

Similarly, $\widetilde{\beta}_2(\gamma)$ is uniformly consistent for β. As a result,

$$\begin{aligned}\widetilde{\varepsilon}_i(\gamma) &= y_i - \mathbf{x}_i'\widetilde{\beta}_1(\gamma)1(q_i \le \gamma) - \mathbf{x}_i'\widetilde{\beta}_2(\gamma)1(q_i > \gamma) \\ &= \mathbf{x}_i'\beta + \mathbf{x}_i'\delta_n 1(q_i \le \gamma_0) + \varepsilon_i - \mathbf{x}_i'(\beta + o_p(1)) \\ &= \varepsilon_i + o_p(\|\mathbf{x}_i\|) \text{ uniformly in } \gamma \in \Gamma,\end{aligned}$$

so that

$$\begin{aligned}\widetilde{\Omega}_1(\gamma) &= \frac{1}{n}\sum_{i=1}^n \mathbf{z}_i\mathbf{z}_i'\widetilde{\varepsilon}_i^2(\gamma)1(q_i \le \gamma) \\ &= \frac{1}{n}\sum_{i=1}^n \mathbf{z}_i\mathbf{z}_i'(\varepsilon_i + o_p(\|\mathbf{x}_i\|))^2 1(q_i \le \gamma) \xrightarrow{p} \Omega_1(\gamma)\end{aligned}$$

uniformly in $\gamma \in \Gamma$ by a standard argument. Similarly, $\widetilde{\Omega}_2(\gamma) \xrightarrow{p} \Omega_2(\gamma)$ uniformly in $\gamma \in \Gamma$. Now,

$$\sqrt{n}\left(\widehat{\beta}_1(\gamma) - \beta\right) = \left[\widehat{Q}_1(\gamma)'\widetilde{\Omega}_1(\gamma)^{-1}\widehat{Q}_1(\gamma)\right]^{-1}\left[\widehat{Q}_1(\gamma)'\widetilde{\Omega}_1(\gamma)^{-1}\frac{1}{\sqrt{n}}Z_{\le\gamma}'\left(X_{\le\gamma_0}\delta_n + \boldsymbol{\varepsilon}\right)\right],$$

where $\widehat{Q}_1(\gamma) \xrightarrow{p} Q_1(\gamma)$, $\frac{1}{\sqrt{n}}Z_{\le\gamma}'X_{\le\gamma_0}\delta_n \xrightarrow{p} Q_1(\gamma \wedge \gamma_0)c$ uniformly in $\gamma \in \Gamma$, and $\frac{1}{\sqrt{n}}Z_{\le\gamma}'\boldsymbol{\varepsilon} \rightsquigarrow S_1(\gamma)$. Hence

$$\sqrt{n}\left(\widehat{\beta}_1(\gamma) - \beta\right) \rightsquigarrow V_1(\gamma)Q_1(\gamma)'\Omega_1(\gamma)^{-1}\left[Q_1(\gamma \wedge \gamma_0)c + S_1(\gamma)\right].$$

Similarly,

$$\sqrt{n}\left(\widehat{\beta}_2(\gamma) - \beta\right) \rightsquigarrow V_2(\gamma)Q_2(\gamma)\Omega_2(\gamma)^{-1}\left[Q_2(\gamma \vee \gamma_0)c + S_2(\gamma)\right].$$

From the arguments above and by the CMT, $\widehat{V}_1(\gamma) \xrightarrow{p} V_1(\gamma)$ and $\widehat{V}_2(\gamma) \xrightarrow{p} V_2(\gamma)$ uniformly in $\gamma \in \Gamma$. Finally, $W_n(\gamma) \rightsquigarrow W^c(\gamma)$ as specified in the theorem, where the second part of $W^c(\gamma)$ is the process in Theorem 4 of CH. ■

Score-type Tests

The score-type tests are based on

$$\begin{aligned}T_n(\gamma) = &\left[n^{-1}\sum_{i=1}^n \left(\mathbf{z}_i 1(q_i \le \gamma) - \widehat{Q}_1(\gamma)\widehat{V}\widehat{Q}'\widetilde{\Omega}^{-1}\mathbf{z}_i\right)\left(\mathbf{z}_i 1(q_i \le \gamma) - \widehat{Q}_1(\gamma)\widehat{V}\widehat{Q}'\widetilde{\Omega}^{-1}\mathbf{z}_i\right)'\widehat{\varepsilon}_i^2\right]^{-1/2} \\ &\cdot n^{-1/2}\sum_{i=1}^n \left[\mathbf{z}_i 1(q_i \le \gamma) - \widehat{Q}_1(\gamma)\widehat{V}\widehat{Q}'\widetilde{\Omega}^{-1}\mathbf{z}_i\right]\widehat{\varepsilon}_i, \gamma \in \Gamma.\end{aligned}$$

Note here that although $\widehat{Q}_1(\gamma)\widehat{V}\widehat{Q}'\widetilde{\Omega}^{-1}n^{-1/2}\sum_{i=1}^n \mathbf{z}_i\widehat{\varepsilon}_i = o_p(1)$, $\mathbf{z}_i 1(q_i \le \gamma)$ is recentered by $\widehat{Q}_1(\gamma)\widehat{V}\widehat{Q}'\widetilde{\Omega}^{-1}\mathbf{z}_i$. This is because the effect of $\widehat{\beta}$ will not disappear asymptotically so the asymptotic distribution of $n^{-1/2}\sum_{i=1}^n \mathbf{z}_i 1(q_i \le \gamma)\widehat{\varepsilon}_i$ differs from $n^{-1/2}\sum_{i=1}^n \mathbf{z}_i 1(q_i \le \gamma)\varepsilon_i$ under H_0. Recentering is to offset the effect of $\widehat{\beta}$. Since only $\widehat{\beta}$ is used in the construction of $T_n(\cdot)$, this type of tests is constructed under H_0 and only one GMM estimator needs to be constructed. This significantly lightens the computation burden. Given $T_n(\cdot)$, we can similarly construct the Kolmogorov-Smirnov sup-type statistic K_n^s and the Cramér–von Mises average-type statistic C_n^s.

The following theorem states the asymptotic distribution of a general continuous functional $g(\cdot)$ of $T_n(\cdot)$ under the local alternative $\delta_n = n^{-1/2}c$. The corresponding test statistic is denoted as g_n^s.

Theorem 11 *If* $\delta_n = n^{-1/2}c$, $\mathbb{E}\left[\|x\|^4\right] < \infty$, $\mathbb{E}\left[q^4\right] < \infty$, $\mathbb{E}\left[\varepsilon^4\right]$ *and* $\mathbb{E}\left[\|\mathbf{z}\|^4\right] < \infty$, *then*

$$g_n^s \xrightarrow{d} g_c^s = g(T^c),$$

where

$$T^c(\gamma) = H(\gamma,\gamma)^{-1/2}\left\{S(\gamma) + \left[Q_1(\gamma\wedge\gamma_0) - Q_1(\gamma)VQ'\Omega^{-1}Q_1(\gamma_0)\right]c\right\}.$$

Proof. As in the last theorem, we can show $\widehat{\beta} \xrightarrow{p} \beta$, $\widetilde{\Omega} \xrightarrow{p} \Omega$, and $\widehat{V} \xrightarrow{p} V$ under the local alternative.

$$\begin{aligned}
& n^{-1/2}\textstyle\sum_{i=1}^n \mathbf{z}_i 1(q_i \le \gamma)\widehat{\varepsilon}_i \\
&= n^{-1/2}\textstyle\sum_{i=1}^n \mathbf{z}_i 1(q_i \le \gamma)\left(y_i - \mathbf{x}_i'\widehat{\beta}\right) \\
&= n^{-1/2}\textstyle\sum_{i=1}^n \mathbf{z}_i 1(q_i \le \gamma)\left(y_i - \mathbf{x}_i'\beta\right) - n^{-1}\textstyle\sum_{i=1}^n \mathbf{z}_i\mathbf{x}_i' 1(q_i \le \gamma)\sqrt{n}\left(\widehat{\beta}-\beta\right) \\
&= n^{-1/2}\textstyle\sum_{i=1}^n \mathbf{z}_i 1(q_i \le \gamma)\left(\mathbf{x}_i'\delta_n 1(q_i \le \gamma_0) + \varepsilon_i\right) - \widehat{Q}_1(\gamma)\sqrt{n}\left(\widehat{\beta}-\beta\right),
\end{aligned}$$

where $n^{-1/2}\sum_{i=1}^n \mathbf{z}_i 1(q_i \le \gamma)\mathbf{x}_i'\delta_n 1(q_i \le \gamma_0) \xrightarrow{p} Q_1(\gamma\wedge\gamma_0)c$, $\widehat{Q}_1(\gamma) \xrightarrow{p} Q_1(\gamma)$ uniformly in $\gamma \in \Gamma$, and $n^{-1/2}\sum_{i=1}^n \mathbf{z}_i 1(q_i \le \gamma)\varepsilon_i \rightsquigarrow S_1(\gamma)$. Next,

$$\begin{aligned}
& n^{-1/2}\sum_{i=1}^n \widehat{Q}_1(\gamma)\widehat{V}\widehat{Q}'\widetilde{\Omega}^{-1}\mathbf{z}_i\widehat{\varepsilon}_i \\
=\; & \widehat{Q}_1(\gamma)\left(\widehat{Q}'\widetilde{\Omega}^{-1}\widehat{Q}\right)^{-1}\widehat{Q}'\widetilde{\Omega}^{-1}n^{-1/2}\sum_{i=1}^n \mathbf{z}_i\left(-\mathbf{x}_i'(\widehat{\beta}-\beta) + \mathbf{x}_i'\delta_n 1(q_i \le \gamma_0) + \varepsilon_i\right) \\
=\; & -\widehat{Q}_1(\gamma)\sqrt{n}\left(\widehat{\beta}-\beta\right) + \widehat{Q}_1(\gamma)\widehat{V}\widehat{Q}'\widetilde{\Omega}^{-1}\left(n^{-1}\sum_{i=1}^n \mathbf{z}_i\mathbf{x}_i' 1(q_i \le \gamma_0)\right)c \\
& + \widehat{Q}_1(\gamma)\widehat{V}\widehat{Q}'\widetilde{\Omega}^{-1}\left(n^{-1/2}\sum_{i=1}^n \mathbf{z}_i\varepsilon_i\right),
\end{aligned}$$

where the second term in the last equality converges in probability to $Q_1(\gamma)VQ'\Omega^{-1}Q_1(\gamma_0)c$ uniformly in $\gamma \in \Gamma$, and $n^{-1/2}\sum_{i=1}^n \mathbf{z}_i\varepsilon_i \xrightarrow{d} N(0,\Omega)$. In summary,

$$\begin{aligned}
& n^{-1/2}\sum_{i=1}^n\left[\mathbf{z}_i 1(q_i \le \gamma) - \widehat{Q}_1(\gamma)\widehat{V}\widehat{Q}'\widetilde{\Omega}^{-1}\mathbf{z}_i\right]\widehat{\varepsilon}_i \\
=\; & n^{-1/2}\sum_{i=1}^n\left[\mathbf{z}_i 1(q_i \le \gamma) - \widehat{Q}_1(\gamma)\widehat{V}\widehat{Q}'\widetilde{\Omega}^{-1}\mathbf{z}_i\right]\varepsilon_i \\
& + \left[Q_1(\gamma\wedge\gamma_0) - Q_1(\gamma)VQ'\Omega^{-1}Q_1(\gamma_0)\right]c + o_p(1) \\
\rightsquigarrow\; & S(\gamma) + \left[Q_1(\gamma\wedge\gamma_0) - Q_1(\gamma)VQ'\Omega^{-1}Q_1(\gamma_0)\right]c,
\end{aligned}$$

and it is not hard to show $n^{-1}\sum_{i=1}^n\left(\mathbf{z}_i 1(q_i \le \gamma) - \widehat{Q}_1(\gamma)\widehat{V}\widehat{Q}'\widetilde{\Omega}^{-1}\mathbf{z}_i\right)\left(\mathbf{z}_i 1(q_i \le \gamma) - \widehat{Q}_1(\gamma)\widehat{V}\widehat{Q}'\widetilde{\Omega}^{-1}\mathbf{z}_i\right)'\widehat{\varepsilon}_i^2 \xrightarrow{p} H(\gamma,\gamma)$ uniformly in $\gamma \in \Gamma$, so the results of the theorem follow.

$$H(\gamma_1,\gamma_2) = \mathbb{E}\left[\left(\mathbf{z}_i 1(q_i \le \gamma_1) - Q_1(\gamma_1)VQ'\Omega^{-1}\mathbf{z}_i\right)\left(\mathbf{z}_i 1(q_i \le \gamma_2) - Q_1(\gamma_2)VQ'\Omega^{-1}\mathbf{z}_i\right)'\varepsilon_i^2\right].$$

■

To understand $S(\gamma)$ in $T^c(\gamma)$, consider a simple case where $\mathbf{x} = (1, x')'$, q follows a uniform distribution on $[0,1]$ and is independent of $(\mathbf{z}', x', \varepsilon)'$. In this case,

$$H(\gamma_1,\gamma_2) = (\gamma_1\wedge\gamma_2)\,\Omega - \gamma_1\gamma_2 QVQ'.$$

If $d_{\mathbf{z}} = \overline{d}$, i.e., the model is just-identified, then

$$\begin{aligned}H(\gamma_1, \gamma_2) &= \mathbb{E}\left[\left(\mathbf{z}_i 1(q_i \leq \gamma_1) - Q_1(\gamma_1) Q^{-1}\mathbf{z}_i\right)\left(\mathbf{z}_i 1(q_i \leq \gamma_2) - Q_1(\gamma_2) Q^{-1}\mathbf{z}_i\right)' \varepsilon_i^2\right] \\ &= \Omega_1(\gamma_1 \wedge \gamma_2) - Q_1(\gamma_1) Q^{-1}\Omega_1(\gamma_2) - \Omega_1(\gamma_1) Q'^{-1}Q_1(\gamma_2)' + Q_1(\gamma_1) Q^{-1}\Omega Q'^{-1}Q_1(\gamma_2)',\end{aligned}$$

and we can let, for $\gamma \in \Gamma$,

$$\begin{aligned}T_n(\gamma) = &\left[n^{-1}\sum_{i=1}^{n}\left(z_i 1(q_i \leq \gamma) - \widehat{Q}_1(\gamma)\widehat{Q}^{-1}z_i\right)\left(z_i 1(q_i \leq \gamma) - \widehat{Q}_1(\gamma)\widehat{Q}^{-1}z_i\right)'\widehat{\varepsilon}_i^2\right]^{-1/2} \\ &\cdot n^{-1/2}\sum_{i=1}^{n}\left[z_i 1(q_i \leq \gamma) - \widehat{Q}_1(\gamma)\widehat{Q}^{-1}z_i\right]\widehat{\varepsilon}_i, \gamma \in \Gamma.\end{aligned} \tag{36}$$

Combining these two cases, $H(\gamma_1, \gamma_2)$ reduces to $(\gamma_1 \wedge \gamma_2 - \gamma_1\gamma_2)\,\Omega$, where $d_z = d$. In other words, $\Omega^{-1/2}S(\gamma)$ is a standard d-dimensional Brownian Bridge. Now, the local power is generated by $[Q_1(\gamma \wedge \gamma_0) - Q_1(\gamma) V Q'\Omega^{-1}Q_1(\gamma_0)]c = (\gamma \wedge \gamma_0 - \gamma\gamma_0)\, Qc$. Of course, the construction of $T_n(\gamma)$ can be greatly simplified in this simple case, e.g., let

$$T_n(\gamma) = \widetilde{\Omega}^{-1/2} \cdot n^{-1/2}\sum_{i=1}^{n}[\mathbf{z}_i 1(q_i \leq \gamma) - \gamma\mathbf{z}_i]\widehat{\varepsilon}_i,$$

which converges to the standard d-dimensional Brownian Bridge. In linear regression, we need only replace $\mathbf{z}_i$ in all formula of (36) by $\mathbf{x}_i$.

Simulating Critical Values

The asymptotic distributions in the above two theorems are nonpivotal, but the simulation method in Hansen (1996) can be extended to the present case. More specifically, let $\{\xi_i^*\}_{i=1}^n$ be i.i.d. $N(0,1)$ random variables, and set

$$W_n^*(\gamma) = \left(\widehat{V}_1(\gamma) + \widehat{V}_2(\gamma)\right)^{-1/2}\sqrt{n}\left(\widehat{\beta}_1^*(\gamma) - \widehat{\beta}_2^*(\gamma)\right), \gamma \in \Gamma,$$

and, for $\gamma \in \Gamma$,

$$\begin{aligned}T_n^*(\gamma) = &\left[n^{-1}\sum_{i=1}^{n}\left(\mathbf{z}_i 1(q_i \leq \gamma) - \widehat{Q}_1(\gamma)\widehat{V}\widehat{Q}'\widetilde{\Omega}^{-1}\mathbf{z}_i\right)\left(\mathbf{z}_i 1(q_i \leq \gamma) - \widehat{Q}_1(\gamma)\widehat{V}\widehat{Q}'\widetilde{\Omega}^{-1}\mathbf{z}_i\right)'\widehat{\varepsilon}_i^2\right]^{-1/2} \\ &\cdot n^{-1/2}\sum_{i=1}^{n}\left[\mathbf{z}_i 1(q_i \leq \gamma) - \widehat{Q}_1(\gamma)\widehat{V}\widehat{Q}'\widetilde{\Omega}^{-1}\mathbf{z}_i\right]\widehat{\varepsilon}_i\xi_i^*, \gamma \in \Gamma,\end{aligned} \tag{37}$$

where $\widehat{\beta}_1^*(\gamma)$ and $\widehat{\beta}_2^*(\gamma)$ are similarly defined as $\widehat{\beta}_1(\gamma)$ and $\widehat{\beta}_2(\gamma)$ with the only difference being that y_i is replaced by $\widetilde{\varepsilon}_i(\gamma)\xi_i^*$; more specifically,

$$\begin{aligned}\widehat{\beta}_1^*(\gamma) &= \left(\widehat{Q}_1(\gamma)'\widetilde{\Omega}_1^{-1}(\gamma)\widehat{Q}_1(\gamma)\right)^{-1}\left(\widehat{Q}_1(\gamma)'\widetilde{\Omega}_1^{-1}(\gamma)\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i 1(q_i \leq \gamma)\widetilde{\varepsilon}_i(\gamma)\xi_i^*\right), \\ \widehat{\beta}_2^*(\gamma) &= \left(\widehat{Q}_2(\gamma)'\widetilde{\Omega}_2^{-1}(\gamma)\widehat{Q}_2(\gamma)\right)^{-1}\left(\widehat{Q}_2(\gamma)'\widetilde{\Omega}_2^{-1}(\gamma)\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i 1(q_i > \gamma)\widetilde{\varepsilon}_i(\gamma)\xi_i^*\right).\end{aligned}$$

Our test rejects H_0 if g_n^{ω} (g_n^s) is greater than the $(1-\alpha)$th conditional quantile of $g(W_n^*(\gamma))$ $(g(T_n^*(\gamma)))$. Equivalently, the p-value transformation can be employed. Take the score test as an example. Define $p_n^* = 1 - F_n^*(g_n^s)$, and $p_n = 1 - F_0(g_n^s)$, where F_n^* is the conditional distribution of $g(T_n^*(\gamma))$ given the original data, and F_0 is the asymptotic distribution of $g(T_n(\gamma))$ under the null. Our test rejects H_0 if $p_n^* \leq \alpha$. By stochastic equicontinuity of the $T_n(\gamma)$ process, we can replace Γ by finite grids with the distance between adjacent grid points going to zero as $n \to \infty$. A natural choice of the grids for Γ is the q_i's in Γ. Also, the conditional distribution can be approximated by standard simulation techniques. More specifically, the following procedure is used.

Step 1: generate $\left\{\xi_{ij}^*\right\}_{i=1}^n$ be i.i.d. $N(0,1)$ random variables.
Step 2: set $T_n^{j*}(\gamma_l)$ as in (37), where $\{\gamma_l\}_{l=1}^L$ is a grid approximation of Γ. Note here that the same $\left\{\xi_{ij}^*\right\}_{i=1}^n$ are used for all γ_l, $l = 1, \cdots, L$.
Step 3: set $g_n^{j*} = g\left(T_n^{j*}\right)$.
Step 4: repeat Step 1-3 J times to generate $\left\{g_n^{j*}\right\}_{j=1}^J$.
Step 5: if $p_n^{J*} = J^{-1}\sum_{j=1}^J 1\left(g_n^{j*} \geq g_n^s\right) \leq \alpha$, we reject H_0; otherwise, accept H_0.

It can be shown that $p_n^* = p_n + o_p(1)$ under both the null and local alternative. Hence $p_n^* \xrightarrow{d} p_c = 1 - F_0\left(g_c^s\right)$ under the local alternative, and $p_n^* \xrightarrow{d} U$, the uniform distribution on $[0,1]$, under the null. The proof is similar to that of Yu (2013b, 2016) and so it is omitted here.

Additional References

Bai, J., 1996, Testing for Parameter Constancy in Linear Regressions: An Empirical Distribution Function Approach, *Econometrica*, 64, 597-622.

Bhattacharya, P.K. and P.J. Brockwell, 1976, The Minimum of an Additive Process with Applications to Signal Estimation and Storage Theory, *Z. Wahrschein. Verw. Gebiete*, 37, 51-75.

Fan, J., 1993, Local Linear Regression Smoothers and Their Minimax Efficiency, *Annals of Statistics*, 21, 196-216.

De Jong, P., 1987, A Central Limit Theorem for Generalized Quadratic Forms, *Probability Theory and Related Fields*, 75, 261-277.

Hall, P., 1984, Central Limit Theorem for Integrated Square Error of Multivariate Nonparametric Density Estimators, *Journal of Multivariate Analysis*, 14, 1-16.

Hsiao, C., Q. Li and J.S. Racine, 2007, A Consistent Model Specification Test with Mixed Discrete and Continuous Data, *Journal of Econometrics*, 140, 802-826.

Kim, J. and D. Pollard, 1990, Cube Root Asymptotics, *Annals of Statistics*, 18, 191-219.

Newey, W.K., 1994, Kernel Estimation of Partial Means and a General Variance Estimator, *Econometric Theory*, 10, 233-253.

Pollard, D., 1993, Asymptotics for a Binary Choice Model, Preprint, Department of Statistics, Yale University.

Sun, Y.X., 2005, Adaptive Estimation of the Regression Discontinuity Model, mimio, Department of Economics, University of California, San Diego.

Van der Vaart, A.W. and J. A. Wellner, 1996, *Weak Convergence and Empirical Processes: With Applications to Statistics*, New York: Springer.