Supplementary Material to "NONPARAMETRIC TESTING IN REGRESSION MODELS WITH WILCOXON-TYPE GENERALIZED LIKELIHOOD RATIO"

Long Feng[1], Zhaojun Wang[1], Chunming Zhang[2], Changliang Zou[1],

[1]*Nankai University and* [2]*University of Wisconsin–Madison*

This supplemental file contains some more discussions on the proposed WGLR test, all the technical proofs and some additional simulation results. The following materials are included:

- Appendix A: Some discussions on two modifications of WGLR tests
- Appendix B: Proof of Theorem 6
- Appendix C: Proofs of other theorems
- Appendix D: Simulation results in heteroscedasticity cases
- Appendix E: Some additional simulation results in Section 3
- Appendix F: Some additional figures in Section 4

Appendix

Appendix A: Some discussions on two modifications of WGLR tests

In constructing λ_{na}, both the "likelihood" function and local smoother are of Wilcoxon-type and come from (2.2). If we replace the local linear estimators in (1.3) by the local linear Walsh-average estimators and denote the resulting test statistic as ω_{na}, i.e.,

$$\omega_{na} = \frac{n}{2}\left(\log\sum_{i=1}^{n}\widehat{\varepsilon}_i^{a2} - \log\sum_{i=1}^{n}\widehat{\varepsilon}_i^{2}\right) \tag{A.1}$$

Similar to Theorems 1-2, we can establish the asymptotic normality of ω_{na}.

Theorem 7 *(i) Suppose the conditions in Theorem 2 hold. Under H_0, we have*

$$(\omega_{na} - \mu_{\omega na})/\sigma_{\omega na} \xrightarrow{d} N(0,1),$$

where

$$\begin{aligned}\mu_{\omega na} =& \frac{1}{h}(p+1)|\Omega|\left(K(0)\sigma^{-2}\zeta^2 - \frac{1}{2}\int K^2(t)dt\right), \zeta^2 = \varphi^{-1/2}\tau\int xG(-x)dG(x),\\ \sigma^2_{\omega na} =& 2h^{-1}\sigma^{-2}\tau^2(p+1)|\Omega|\Bigg(\int K^2(t)dt - \sigma^{-2}\zeta^2\int K(t)K*K(t)dt\\ &+\frac{1}{4}\sigma^{-2}\tau^2\int(K*K)^2(t)dt\Bigg);\end{aligned}$$

(ii) Suppose the conditions in Theorem 2 hold. under $H_{na}^{'}$, we have

$$[\omega_{na} - \mu_{\omega na} - \sigma^{-2}\tau^2 d_{2na}]/\sigma^{*}_{\omega na} \xrightarrow{d} N(0,1),$$

where $\sigma^{*2}_{\omega na} = \sigma^2_{\omega na} + \sigma^{-2}h^{-1}E[\boldsymbol{G}(U)^T\mathbf{W}\mathbf{W}^T\boldsymbol{G}(U)]$.

This theorem implies that the power of ω_{na} is

$$\beta_{\omega_{na}} = \Phi\left(-\frac{\sigma_{\omega na}}{\sqrt{\sigma^2_{\omega na} + n\sigma^{-2}B(G)}}z_\alpha + \frac{2^{-1}n\sigma^{-2}B(G)}{\sqrt{\sigma^2_{\omega na}+n\sigma^{-2}B(G)}}\right).$$

It is difficult to calculate the ARE of ω_{na} with respect to the GLR test under the general cases. For convenience, we choose the same bandwidth for ω_{na} and GLR and then

$$\begin{aligned}&\mathrm{ARE}(\omega_{na}, \mathrm{GLR})\\ &= \frac{\sigma\tau^{-1}\left(\int\left\{K(t) - \frac{1}{2}K*K(t)\right\}^2 dt\right)^{1/2}}{\left(\int K^2(t)dt - \sigma^{-2}\zeta^2\int K(t)K*K(t)dt + \frac{1}{4}\sigma^{-2}\tau^2\int(K*K)^2(t)dt\right)^{1/2}}.\end{aligned}$$

Table A.1 shows the ARE of ω_{na}, LOSS (Hong and Lee 2013) and WGLR with respect to GLR for a number of distributions and kernel functions. We observe that ARE(ω_{na}, GLR)'s are similar for different kernels and generally much smaller than ARE(WGLR, GLR) for heavy-tailed distribution. To a certain extent, ω_{na} can be viewed as some compromise between the GLR and WGLR tests. Moreover, if the local linear Walsh-average estimators in (2.4) are replaced by the local linear estimators, similar results to those in Table A.1 can be obtained.

Table A.1: The asymptotic efficiency comparisons of ω_{na} and WGLR. $t(d)$: student's t-distribution with d degrees of freedom. $T(\rho,\sigma)$:Tukey contaminated normal with CDF $F(x) = (1-\rho)\Phi(x) + \rho\Phi(x/\sigma)$ where $\rho \in [0,1]$ is the contamination proportion.

	Epanechnikov	Biweight	Triweight	Gaussian	
		ARE(LOSS, GLR)			
	1.46	1.46	1.46	1.49	
Errors		ARE(ω_{na}, GLR)			ARE(WGLR, GLR)
$N(0,1)$	0.99	0.99	0.99	0.99	0.96
$t(3)$	1.16	1.16	1.16	1.16	1.90
$t(4)$	1.07	1.07	1.07	1.07	1.40
$T(0.05, 10)$	1.59	1.58	1.58	1.59	4.77
$T(0.10, 10)$	1.88	1.87	1.87	1.89	7.19

Appendix B: Proof of Theorem 6

Here we only provide the proof of Theorem 6. Obviously, Theorem 1 and 2 are the special cases of Theorem 6 with $\varrho^2(x)=1$. Let $r_n=1/\sqrt{nh}$. For ease of illustration, we need some notations:

$$\begin{aligned}
\xi_i &= \varphi^{-1/2}\rho(U_i)\{G(-\varepsilon_i)-1/2\},\ w_0=\int\int t^2(s+t)^2K(t)K(s+t)dtds,\\
R_{n10} &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \xi_i(\alpha_0''(U_i)+\boldsymbol{A}_0''(U_i)^T\boldsymbol{X}_i)\int t^2K(t)dt(1+O(h)+O(n^{-1/2})),\\
R_{n20} &= \frac{1}{2}\frac{1}{\sqrt{n}}\sum_{i=1}^n \xi_i\alpha_0''(U_i)w_0.\\
R_{n30} &= \frac{1}{8}E[\alpha_0''(U_i)^2+\boldsymbol{A}_0''(U_i)^T\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{A}_0''(U_i)]w_0(1+O(n^{-1/2})),\\
\boldsymbol{\alpha}_n(u_0) &= r_n^2\boldsymbol{\Gamma}(u_0)^{-1}\sum_{i=1}^n\eta_i\boldsymbol{W}_iK((U_i-u_0)/h), \boldsymbol{c}=(c,\boldsymbol{0}_{1\times p})^T, \boldsymbol{\Gamma}(u)=\boldsymbol{\Sigma}(u)f(u),\\
\boldsymbol{R}_n(u_0) &= r_n^2\sum_{i=1}^n\boldsymbol{\Gamma}(u_0)^{-1}\Big(\alpha(U_i)-\boldsymbol{\beta}(u_0)^T\boldsymbol{V}_i(u_0)+\boldsymbol{A}(U_i)^T\boldsymbol{X}_i-\boldsymbol{\gamma}(u_0)^T\boldsymbol{Z}_i(u_0)\Big)\boldsymbol{W}_iK((U_i-u_0)/h),\\
R_{1n} &= \sum_{k=1}^n\xi_k\boldsymbol{R}_n(U_k)^T\mathbf{W}_k/\rho^2(U_k),\ R_{2n}=\sum_{k=1}^n\boldsymbol{\alpha}_n(U_k)^T\mathbf{W}_k\mathbf{W}_k^T\boldsymbol{R}_n(U_k)/\rho^2(U_k),\\
R_{3n} &= \frac{1}{2}\sum_{k=1}^n\boldsymbol{R}_n(U_k)^T\mathbf{W}_k\mathbf{W}_k^T\boldsymbol{R}_n(U_k)/\rho^2(U_k).
\end{aligned}$$

Lemma 1 *Let $\widehat{\boldsymbol{A}}$ be the local linear Walsh-average estimator. Then, under conditions (A1)-(A5), uniformly for $u_0\in\Omega$,*

$$(\widehat{\alpha}(u_0),\widehat{\boldsymbol{A}}(u_0)^T)^T-(\alpha(u_0),\boldsymbol{A}(u_0)^T)^T=\boldsymbol{c}+(\boldsymbol{\alpha}_n(u_0)+\boldsymbol{R}_n(u_0))(1+o_p(1)),$$

and under condition (A4′), $c=0$ and $\eta_i=\xi_i$, $i=1,\ldots,n$.

Proof. From Shang *et al.* (2012), we can easily obtain the result. □

Proof of Theorem 6 By Proposition 1, the estimators $\widehat{\rho}_k(x)$ is consistent. Note that under H_{0a}, $\widehat{\varepsilon}_i^a=\varepsilon_i$. Thus, by Slutsky's theorem, we only need to consider the asymptotic property of

$$\begin{aligned}
\lambda_{na} &= \frac{\varphi^{-1/2}}{n+1}\sum\sum_{i\le j}(|\tau^{-1}\varepsilon_i+\tau^{-1}\varepsilon_j|-|\rho^{-1}(U_i)\widehat{\varepsilon}_i+\rho^{-1}(U_j)\widehat{\varepsilon}_j|)\\
&= \frac{\varphi^{-1/2}}{n+1}\sum\sum_{i\le j}\Big(|\tau^{-1}\varepsilon_i+\tau^{-1}\varepsilon_j|-|\tau^{-1}\varepsilon_i+\tau^{-1}\varepsilon_j+\phi_{ij})|\Big),
\end{aligned}$$

where $\phi_{ij} = \rho^{-1}(U_i)(\alpha(U_i) - \widehat{\alpha}(U_i) + \boldsymbol{A}(U_i)^T\boldsymbol{X}_i - \widehat{\boldsymbol{A}}(U_i)^T\boldsymbol{X}_i) + \rho^{-1}(U_j)(\alpha(U_j) - \widehat{\alpha}(U_j) + \boldsymbol{A}(U_j)^T\boldsymbol{X}_j - \widehat{\boldsymbol{A}}(U_j)^T\boldsymbol{X}_j)$. By using the identity

$$|z| - |z-y| = y\text{sgn}(z) + 2(z-y)\{I(0<z<y) - I(y<z<0)\}$$

which holds for $z \neq 0$, we have

$$\begin{aligned}\lambda_{na} =& \Bigg(-\frac{\varphi^{-1/2}}{n+1}\sum\sum_{i\leq j}\phi_{ij}\text{sgn}(\varepsilon_i+\varepsilon_j) + \frac{2\varphi^{-1/2}\tau^{-1}}{n+1}\sum\sum_{i\leq j}(\tau^{-1}(\varepsilon_i+\varepsilon_j)+\phi_{ij}) \\ &\qquad \times \Big\{I(0<\tau^{-1}(\varepsilon_i+\varepsilon_j)<-\phi_{ij})) - I(-\phi_{ij}<\tau^{-1}(\varepsilon_i+\varepsilon_j)<0)\Big\}\Bigg) \\ \doteq& A_h - B_h.\end{aligned}$$

Firstly, we consider A_h. By Lemma 1, we have

$$\begin{aligned}A_h =& \Bigg(-\frac{\varphi^{-1/2}}{n+1}\sum\sum_{i\leq j}\left(\boldsymbol{R}_n(U_i)^T\boldsymbol{W}_i/\rho(U_i) + \boldsymbol{R}_n(U_j)^T\boldsymbol{W}_j/\rho(U_j)\right)\text{sgn}(\varepsilon_i+\varepsilon_j) \\ & -\frac{\varphi^{-1/2}}{n+1}\sum\sum_{i\leq j}\left(\boldsymbol{\alpha}_n(U_i)^T\boldsymbol{W}_i/\rho(U_i) + \boldsymbol{\alpha}_n(U_j)^T\boldsymbol{W}_j/\rho(U_j)\right)\text{sgn}(\varepsilon_i+\varepsilon_j)\Bigg)(1+o_p(1)) \\ \doteq& (C_h+D_h)(1+o_p(1)).\end{aligned}$$

Firstly, note that D_h

$$\begin{aligned}&\frac{\varphi^{-1/2}}{n+1}\sum\sum_{i\leq j}\left(\boldsymbol{\alpha}_n(U_i)^T\boldsymbol{W}_i/\rho(U_i) + \boldsymbol{\alpha}_n(U_j)^T\boldsymbol{W}_j/\rho(U_j)\right)\text{sgn}(\varepsilon_i+\varepsilon_j) \\ =& \sum_{i=1}^n \boldsymbol{\alpha}_n(U_i)^T\boldsymbol{W}_i/\rho(U_i)\left(\sum\sum_{j\neq i}\frac{\varphi^{-1/2}}{n+1}\text{sgn}(\varepsilon_i+\varepsilon_j)\right) + \frac{\varphi^{-1/2}}{n+1}\sum_{i=1}^n \boldsymbol{\alpha}_n(U_i)^T\boldsymbol{W}_i\text{sgn}(\varepsilon_i)/\rho(U_i) \\ =& \sum_{i=1}^n \boldsymbol{\alpha}_n(U_i)^T\boldsymbol{W}_i\xi_i/\rho^2(U_i) + o_p(h^{-1/2}),\end{aligned}$$

where the last equality holds because of the facts that

$$\begin{aligned}&\sum\sum_{j\neq i}\frac{\varphi^{-1/2}}{n+1}\text{sgn}(\varepsilon_i+\varepsilon_j) = \varphi^{-1/2}(G(\varepsilon_i - 0.5)) + o_p(n^{-1/2}), \\ &E\left(\frac{\varphi^{-1/2}}{n+1}\sum_{i=1}^n \boldsymbol{\alpha}_n(U_i)^T\boldsymbol{W}_i\text{sgn}(\varepsilon_i)/\rho(U_i)\right) = O\left(\frac{1}{nh}\right) = o(h^{-1/2}), \\ &\text{var}\left(\frac{\varphi^{-1/2}}{n+1}\sum_{i=1}^n \boldsymbol{\alpha}_n(U_i)^T\boldsymbol{W}_i\text{sgn}(\varepsilon_i)/\rho(U_i)\right) = O\left(\frac{1}{n}+\frac{1}{n^2h}\right) = o(h^{-1}).\end{aligned}$$

Similarly, we can prove that

$$
\begin{aligned}
&\frac{\varphi^{-1/2}}{n+1}\sum\sum_{i\le j}\left(\boldsymbol{R}_n(U_i)^T\boldsymbol{W}_i/\rho(U_i)+\boldsymbol{R}_n(U_j)^T\boldsymbol{W}_j/\rho(U_j))\right)\mathrm{sgn}(\varepsilon_i+\varepsilon_j)\\
=&\sum_{i=1}^n\boldsymbol{R}_n(U_i)^T\boldsymbol{W}_i\xi_i/\rho^2(U_i)+o_p(h^{-1/2}).
\end{aligned}
$$

Thus,

$$
A_h=\sum_{i=1}^n\boldsymbol{\alpha}_n(U_i)^T\boldsymbol{W}_i\xi_i/\rho^2(U_i)+\sum_{i=1}^n\boldsymbol{R}_n(U_i)^T\boldsymbol{W}_i\xi_i/\rho^2(U_i)+o_p(h^{-1/2}).
$$

Next, we consider B_h which can be written as

$$
\begin{aligned}
B_h=&-\frac{2\varphi^{-1/2}}{n+1}\sum\sum_{i\le j}(\tau^{-1}(\varepsilon_i+\varepsilon_j)+\phi_{ij})I(0<\tau^{-1}(\varepsilon_i+\varepsilon_j)<-\phi_{ij}))\\
&+\frac{2\varphi^{-1/2}}{n+1}\sum\sum_{i\le j}(\tau^{-1}(\varepsilon_i+\varepsilon_j)+\phi_{ij})I(-\phi_{ij}<\tau^{-1}(\varepsilon_i+\varepsilon_j)<0)\\
\doteq&E_h+F_h.
\end{aligned}
$$

On the set $\{\phi_{ij}<0\}$ and conditional on $\{\boldsymbol{X}_i,U_i\}$,

$$
\begin{aligned}
E(E_h)=&-\frac{2\varphi^{-1/2}}{n+1}\sum\sum_{i\le j}\int\int_{-y}^{-\phi_{ij}-y}(\tau^{-1}(x+y)+\phi_{ij})g(x)g(y)dxdy\\
=&\frac{2\varphi^{-1/2}}{n+1}\sum\sum_{i\le j}\int\frac{1}{2}\tau^2\phi_{ij}^2g^2(y)dy+O(n^{-1/2}h^{-3/2})\\
=&\frac{1}{n+1}\sum\sum_{i\le j}\phi_{ij}^2+o(h^{-1/2})
\end{aligned}
$$

and $\mathrm{var}(E_h) = O(n^{-1}h^{-2}) = o(h^{-1})$. Similarly, on the set $\{\phi_{ij} > 0\}$, $E(F_h) = \frac{1}{n+1}\sum\sum_{i\le j}\phi_{ij}^2 + o(h^{-1/2})$ and $\mathrm{var}(E_h) = O(n^{-1}h^{-2}) = o(h^{-1})$. Thus,

$$
\begin{aligned}
B_h =& \frac{1}{n+1}\sum\sum_{i\le j}\Bigg(\left((\boldsymbol{\alpha}_n(U_i) + \boldsymbol{R}_n(U_i))^T\boldsymbol{W}_i\right)/\rho(U_i) \\
&+ ((\boldsymbol{\alpha}_n(U_j) + \boldsymbol{R}_n(U_j))^T\boldsymbol{W}_j)/\rho(U_j)\Bigg)^2 + o_p(h^{-1/2}) \\
=& \frac{1}{2}\sum_{i=1}^n(\boldsymbol{\alpha}_n(U_i)^T\boldsymbol{W}_i)^2/\rho^2(U_i) + \frac{1}{2}\sum_{i=1}^n(\boldsymbol{R}_n(U_i)^T\boldsymbol{W}_i)^2/\rho^2(U_i) \\
&+ \sum_{i=1}^n \boldsymbol{R}_n(U_i)^T\boldsymbol{W}_i\boldsymbol{W}_i^T\boldsymbol{\alpha}_n(U_i)/\rho^2(U_i) \\
&+ \frac{1}{n+1}\sum\sum_{j\neq i}\boldsymbol{\alpha}_n(U_i)^T\boldsymbol{W}_i\boldsymbol{\alpha}_n(U_j)^T\boldsymbol{W}_j/\rho(U_i)/\rho(U_j) \\
&+ \frac{1}{n+1}\sum\sum_{j\neq i}\boldsymbol{R}_n(U_i)^T\mathbf{W}_iR_n(U_j)^T\boldsymbol{W}_j/\rho(U_i)/\rho(U_j) \\
&+ \frac{1}{n+1}\sum\sum_{j\neq i}\boldsymbol{\alpha}_n(U_i)^T\boldsymbol{W}_i\boldsymbol{R}_n(U_j)^T\boldsymbol{W}_j/\rho(U_i)/\rho(U_j) + o_p(h^{-1/2}).
\end{aligned}
$$

After calculating the expectation and variance of the last three sums, we can prove that

$$
\begin{aligned}
&\frac{1}{n+1}\sum\sum_{j\neq i}\boldsymbol{\alpha}_n(U_i)^T\boldsymbol{W}_i\boldsymbol{\alpha}_n(U_j)^T\boldsymbol{W}_j/\rho(U_i)/\rho(U_j) = O_p(1 + (nh)^{-1/2} + (nh)^{-1}), \\
&\frac{1}{n+1}\sum\sum_{j\neq i}\boldsymbol{R}_n(U_i)^T\mathbf{W}_iR_n(U_j)^T\boldsymbol{W}_j/\rho(U_i)/\rho(U_j) = O_p(h^4 + n^{-1/2}h^{7/2} + n^{-1}h^3), \\
&\frac{1}{n+1}\sum\sum_{j\neq i}\boldsymbol{\alpha}_n(U_i)^T\boldsymbol{W}_i\boldsymbol{R}_n(U_j)^T\boldsymbol{W}_j/\rho(U_i)/\rho(U_j) = O_p(h^2 + n^{-1/2}h^{3/2} + n^{-1}h),
\end{aligned}
$$

and accordingly,

$$
\begin{aligned}
B_h =& \frac{1}{2}\sum_{i=1}^n(\boldsymbol{\alpha}_n(U_i)^T\boldsymbol{W}_i)^2/\rho^2(U_i) + \frac{1}{2}\sum_{i=1}^n(\boldsymbol{R}_n(U_i)^T\boldsymbol{W}_i)^2/\rho^2(U_i) \\
&+ \sum_{i=1}^n \boldsymbol{R}_n(U_i)^T\boldsymbol{W}_i\boldsymbol{W}_i^T\boldsymbol{\alpha}_n(U_i)/\rho^2(U_i) + o_p(h^{-1/2}).
\end{aligned}
$$

This leads to

$$
\begin{aligned}
\lambda_{na} =& \sum_{i=1}^n \boldsymbol{\alpha}_n(U_i)^T\boldsymbol{W}_i\xi_i/\rho^2(U_i) - \frac{1}{2}\sum_{i=1}^n(\boldsymbol{\alpha}_n(U_i)^T\boldsymbol{W}_i)^2/\rho^2(U_i) \\
&+ R_{n1} - R_{n2} - R_{n3} + o_p(h^{-1/2}).
\end{aligned}
$$

Taking the same procedure as Lemma 7.2 in Fan *et al.* (2001), we can show that

$$\begin{aligned} R_{n1} &= n^{1/2}h^2 R_{n10} + O(n^{-1/2}h), \\ R_{n2} &= n^{1/2}h^2 R_{n20} + O(n^{-1/2}h), \\ R_{n3} &= nh^4 R_{n30} + O(h^3). \end{aligned}$$

Also, similar to Lemma 7.4 in Fan *et al.* (2001), it can be verified that

$$\begin{aligned} \sum_{i=1}^{n} \boldsymbol{\alpha}_n(U_i)^T \boldsymbol{W}_i \xi_i/\rho^2(U_i) &= \frac{1}{h}(p+1)K(0)Ef(U)^{-1} \\ &\quad + \frac{1}{n}\sum_{j\neq i}\sum \rho^{-1}(U_i)\rho^{-1}(U_j)\xi_i\xi_j \boldsymbol{W}_i^T \boldsymbol{\Gamma}(U_j)^{-1}\boldsymbol{W}_j K_n(U_i-U_j) + o_p(h^{-1/2}), \\ \sum_{i=1}^{n} (\boldsymbol{\alpha}_n(U_i)\boldsymbol{W}_i^T)^2/\rho^2(U_i) &= \frac{1}{h}(p+1)Ef(U)^{-1}\int K^2(t)dt \\ &\quad + \frac{2}{nh}\sum_{i<j}\sum \rho^{-1}(U_i)\rho^{-1}(U_j)\xi_i\xi_j \boldsymbol{W}_i^T \boldsymbol{\Gamma}^{-1}(U_i) K*K((U_i-U_j)/h)\boldsymbol{W}_j + o_p(h^{-1/2}). \end{aligned}$$

Thus, $\lambda_{na} = \mu_{na} - d_{1na} + W(n)h^{-1/2}/2 + o_p(h^{-1/2})$, where $d_{1na} = \tau^{-2}[nh^4 R_{30} - n^{1/2}h^2(R_{n10} - R_{n20})] = O_p(nh^4 + n^{1/2}h^2) = o_p(h^{-1/2})$ and

$$\begin{aligned} W(n) &= \frac{\sqrt{h}}{n}\sum_{j\neq i}\sum \rho^{-1}(U_i)\rho^{-1}(U_j)\xi_i\xi_j[2K_h(U_i-U_j) - K_h*K_h(U_i-U_j)]\boldsymbol{W}_i^T\boldsymbol{\Gamma}(U_j)^{-1}\boldsymbol{W}_j \\ &= \frac{\sqrt{h}}{n}\sum_{j\neq i}\sum \zeta_i\zeta_j[2K_h(U_i-U_j) - K_h*K_h(U_i-U_j)]\boldsymbol{W}_i^T\boldsymbol{\Gamma}(U_j)^{-1}\boldsymbol{W}_j \end{aligned}$$

where $\zeta_i = \rho^{-1}(U_i)\xi_i = \varphi^{-1/2}\tau(G(\varepsilon_i) - 0.5)$. It remains to show that

$$W(n) \xrightarrow{d} N(0, v)$$

with $v = 2||2K - K*K||_2^2(p+1)Ef(U)^{-1}$. Similar to Fan *et al.* (2001), by applying Theorem 2 in De Jong (1987), we can easily obtain the result.

Under H'_{1a} and by similar arguments as above, it can be checked that

$$\begin{aligned} \lambda_{na} =& \mu_{na} + d_{2na} - W(n)h^{-1/2}/2 \\ &- \sum_{i=1}^{n} h^{-1/2}\boldsymbol{G}^T(U_i)\boldsymbol{W}_i\xi_i/\rho^2(U_i) + o_p(h^{-1/2}), \end{aligned}$$

Then we can obtain the assertion. $\square$

Appendix C: Proofs of other theorems

Proof of Proposition 1 By Fan and Gijbels (1996), we can easily show that $\widehat{f}(u) = f(u)(1+o_p(1))$. Thus, we only need to show that

$$\widetilde{\rho}^{-1}(u) = \frac{\varphi^{-1/2}}{2n(n-1)t_n f^2(u)} \sum_{i=1}^n \sum_{j=1}^n I(|\widehat{\varepsilon}_i + \widehat{\varepsilon}_j| \le t_n) K_h(U_i - u) K_h(U_j - u)$$

is a ratio-consistent estimator of $\rho^{-1}(u)$.

$$\begin{aligned}
\widetilde{\rho}^{-1}(u) =& \frac{\varphi^{-1/2}}{2n^2 t_n f^2(u)} \sum_{i=1}^n \sum_{j=1}^n I(|\varrho(U_i)\varepsilon_i + \varrho(U_j)\varepsilon_j| \le t_n) K_h(U_i - u) K_h(U_j - u) \\
&+ \frac{\varphi^{-1/2}}{2n^2 t_n f^2(u)} \sum_{i=1}^n \sum_{j=1}^n (I(|\widehat{\varepsilon}_i + \widehat{\varepsilon}_j| \le t_n) - I(|\varrho(U_i)\varepsilon_i - \varrho(U_j)\varepsilon_j| \le t_n)) \\
&\times K_h(U_i - u) K_h(U_j - u) \\
&\doteq U_{n1} + U_{n2}
\end{aligned}$$

Clearly, $U_{n1} = \frac{1}{n^2}\sum_{i=1}^n \sum_{j=1}^n W_n(i,j)$ is of the form of U-statistic since U_{n1} is symmetric in this arguments. Note that

$$\begin{aligned}
E(W_n^2(i,j)) =& \frac{3}{t_n^2 f^4(u)} E(I(|\varrho(U_i)\varepsilon_i + \varrho(U_j)\varepsilon_j| \le t_n) K_h^2(U_i - u) K_h^2(U_j - u)) \\
=& \frac{3}{t_n^2 h^4 f^4(u)} E\Bigg(\int \left(G\left(\frac{t_n}{\varrho(U_i)} - \frac{\varrho(U_j)}{\varrho(U_i)}\epsilon \right) - G\left(-\frac{t_n}{\varrho(U_i)} - \frac{\varrho(U_j)}{\varrho(U_i)}\epsilon \right) \right) \\
&\times g(\epsilon) d\epsilon K^2\left(\frac{U_i - u}{h}\right) K^2\left(\frac{U_j - u}{h}\right) \Bigg) \\
=& \frac{\sqrt{3} v_0^2}{t_n \tau \varrho(u) h^2 f^2(u)} (1+o(1)) = O(t_n^{-1} h^{-2}) = o(n)
\end{aligned}$$

where the antepenultimate equality is followed by a simple calculation similar to Parzen (1962). Thus, $U_{n1} = E(W_n(i,j)) + o_p(1)$. Similarly,

$$\begin{aligned}
E(W_n(i,j)) =& \frac{\varphi^{-1/2}}{2t_n f^2(u)} E(I(|\varrho(U_i)\varepsilon_i + \varrho(U_j)\varepsilon_j| \le t_n) K_h(U_i - u) K_h(U_j - u)) \\
=& \frac{\varphi^{-1/2}}{2t_n h^2 f^2(u)} E\Bigg(\int \left(G\left(\frac{t_n}{\varrho(U_i)} - \frac{\varrho(U_j)}{\varrho(U_i)}\epsilon \right) - G\left(-\frac{t_n}{\varrho(U_i)} - \frac{\varrho(U_j)}{\varrho(U_i)}\epsilon \right) \right) \\
&\times g(\epsilon) d\epsilon K\left(\frac{U_i - u}{h}\right) K\left(\frac{U_j - u}{h}\right) \Bigg) \\
=& \frac{1}{\varrho(u)\tau}(1+o(1))
\end{aligned}$$

Thus, $U_{n1} = \rho^{-1}(u) + o_p(1)$. Similarly, we can show that $U_{n2} = O(h^2 + (nh)^{-1/2}) = o(1)$ by Lemma 1. Thus, $\widehat{\rho}^{-1}(u)$ is a ratio-consistent estimator of $\rho^{-1}(u)$. □

Proof of Theorem 7 Taking the same procedure as Fan et al. (2001), under H_0, we have

$$\begin{aligned}\omega_{na} =& \sum_{i=1}^{n} \boldsymbol{\alpha}_n(U_i)^T \boldsymbol{W}_i \varepsilon_i/\sigma^2 - \frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{\alpha}_n(U_i)^T \boldsymbol{W}_i)^2/\sigma^2 \\ &+ \sum_{i=1}^{n} \boldsymbol{R}_n(U_i)^T \boldsymbol{W}_i \varepsilon_i/\sigma^2 - R_{n2}/\sigma^2 - R_{n3}/\sigma^2 + o_p(h^{-1/2}).\end{aligned}$$

Also, we can verify that

$$\begin{aligned}\sum_{i=1}^{n} \boldsymbol{\alpha}_n(U_i)^T \boldsymbol{W}_i \varepsilon_i =& \frac{1}{h}(p+1)K(0)\zeta^2 Ef(U)^{-1} \\ &+ \frac{1}{n}\sum\sum_{j\neq i} \varepsilon_i \xi_j \boldsymbol{W}_i^T \boldsymbol{\Gamma}(U_j)^{-1} \boldsymbol{W}_j K_n(U_i - U_j) + o_p(h^{-1/2}), \\ \sum_{i=1}^{n} \boldsymbol{R}_n(U_i)^T \boldsymbol{W}_i =& n^{1/2}h^2 R_{\omega n10} + O(n^{-1/2}h).\end{aligned}$$

Thus,

$$\omega_{na} = \mu_{\omega na} + W_\omega(n)h^{-1/2}/2 + o_p(h^{-1/2}),$$

where

$$\begin{aligned}W_\omega(n) =& \frac{2h^{1/2}}{n}\sum\sum_{j\neq i} \sigma^{-2}\varepsilon_i \xi_j \boldsymbol{W}_i^T \boldsymbol{\Gamma}(U_j)^{-1} \boldsymbol{W}_j K_n(U_i - U_j) \\ &- \frac{1}{nh^{1/2}}\sum\sum_{j\neq i} \sigma^{-2}\xi_i \xi_j \boldsymbol{W}_i^T \boldsymbol{\Gamma}^{-1}(U_i) K * K((U_i - U_j)/h) \boldsymbol{W}_j.\end{aligned}$$

Applying the martingale central limit theorem (Hall and Heyde 1980), we can verify that

$$W_\omega(n) \xrightarrow{d} N(0, \varsigma),$$

where

$$\begin{aligned}\varsigma =& 2(p+1)Ef(U)^{-1}\Big(4\sigma^{-2}\tau^2 \int K^2(t)dt - 4\sigma^{-4}\zeta^2\tau^2 \int K(t)K * K(t)dt \\ &+ \sigma^{-4}\tau^4 \int (K * K)^2(t)dt\Big).\end{aligned}$$

Under H_{1a}' and by the similar arguments as above, it can be verified that

$$\begin{aligned}\omega_{na} =& \mu_{\omega na} + \sigma^{-2}\tau^2 d_{2na} - W_\omega(n)h^{-1/2}/2 \\ &- \sum_{i=1}^{n} \sqrt{n}\boldsymbol{G}_n^T(U_i)\boldsymbol{W}_i \varepsilon_i/\sigma^2 + o_p(h^{-1/2}),\end{aligned}$$

from which the assertion follows immediately.

Proof of Theorem 3 Denote

$$\boldsymbol{\Gamma} = \begin{pmatrix} \boldsymbol{\Gamma}_{11} & \boldsymbol{\Gamma}_{12} \\ \boldsymbol{\Gamma}_{21} & \boldsymbol{\Gamma}_{22}, \end{pmatrix}, \ \boldsymbol{\Gamma}_{1,2} = \boldsymbol{\Gamma}_{11} - \boldsymbol{\Gamma}_{12}\boldsymbol{\Gamma}_{22}^{-1}\boldsymbol{\Gamma}_{21},$$

where $\boldsymbol{\Gamma}_{11}, \boldsymbol{\Gamma}_{12}, \boldsymbol{\Gamma}_{21}, \boldsymbol{\Gamma}_{22}$ are $(p_1+1)\times(p_1+1), (p_1+1)\times p_2, p_2\times(p_1+1), p_2\times p_2$ matrices and $p_2 = p - p_1$. Taking the same procedure as for $\widehat{\boldsymbol{A}}$, we have

$$\begin{aligned}\widehat{\boldsymbol{A}}_2^b(u_0) - \boldsymbol{A}_2(u_0) =& r_n^2\boldsymbol{\Gamma}_{22}^{-1}(u_0)\sum_{i=1}^n \Big(\xi_i + \boldsymbol{A}_2(U_i)^T\boldsymbol{X}_i^{(2)} - \bar{\eta}_2(u_0, \boldsymbol{X}_i^{(2)}, U_i)\Big) \\ & \times \boldsymbol{X}_i^{(2)} K((U_i - u_0)/h)(1 + o_p(1)),\end{aligned}$$

where $\bar{\eta}_2(u_0, \boldsymbol{X}_i^{(2)}, U_i) = \boldsymbol{A}_2(u_0)^T\boldsymbol{X}_i^{(2)} + \boldsymbol{A}_2'(u_0)^T\boldsymbol{X}_i^{(2)}(U_i - u_0)$. Note that $\lambda_{nb} = \lambda_{na} - \lambda'_{nb}$ where

$$\lambda'_{nb} = \frac{\varphi^{-1/2}\tau^{-1}}{n+1}\sum\sum_{i\le j}\left(|\widehat{\varepsilon}_i^a + \widehat{\varepsilon}_j^a| - |\widehat{\varepsilon}_i^b + \widehat{\varepsilon}_j^b|\right).$$

Similar to the proof of Theorem 6, under H_{0b}, we have

$$\begin{aligned}\lambda'_{nb}\tau^2 =& r_n^2\sum_{i=1}^n\sum_{j=1}^n \xi_i\xi_j {\boldsymbol{X}_i^{(2)}}^T\boldsymbol{\Gamma}_{22}^{-1}(U_i)\boldsymbol{X}_j^{(2)}K((U_i - U_j)/h) \\ & - \frac{1}{2}r_n^4\sum_{i=1}^n\left(\sum_{j=1}^n \xi_j K((U_i - U_j)/h){\boldsymbol{X}_i^{(2)}}^T\right)\boldsymbol{\Gamma}_{22}^{-1}(U_i)\boldsymbol{X}_i^{(2)}{\boldsymbol{X}_i^{(2)}}^T\boldsymbol{\Gamma}_{22}^{-1}(U_i) \\ & \times \left(\sum_{j=1}^n \xi_j K((U_i - U_j)/h)\boldsymbol{X}_i^{(2)}\right) + o_p(h^{-1/2}),\end{aligned}$$

Consequently,

$$\begin{aligned}& -\lambda_{nb}\tau^2 \\ &= -r_n^2\sum_{i=1}^n\sum_{j=1}^n \xi_i\xi_j\left(\boldsymbol{W}_j^{(1)} - \boldsymbol{\Gamma}_{12}(U_i)\boldsymbol{\Gamma}_{22}^{-1}(U_i)\boldsymbol{X}_j^{(2)}\right)^T\boldsymbol{\Gamma}_{1,2}^{-1}(U_i) \\ &\quad \times \left(\boldsymbol{W}_i^{(1)} - \boldsymbol{\Gamma}_{12}(U_i)\boldsymbol{\Gamma}_{22}^{-1}(U_i)\boldsymbol{X}_j^{(2)}\right)K((U_i - U_j)/h) \\ &\quad + \frac{r_n^4}{2}\sum_{i=1}^n\sum_{j=1}^n \xi_i\xi_j\sum_{k=1}^n(\boldsymbol{W}_i^{(1)} - \boldsymbol{\Gamma}_{12}(U_k)\boldsymbol{\Gamma}_{22}^{-1}(U_k)\boldsymbol{X}_i^{(2)})^T\boldsymbol{\Gamma}_{1,2}^{-1}(U_k) \\ &\quad \times (\boldsymbol{W}_k^{(1)} - \boldsymbol{\Gamma}_{12}(U_k)\boldsymbol{\Gamma}_{22}^{-1}(U_k)\boldsymbol{X}_k^{(2)})(\boldsymbol{W}_k^{(1)} - \boldsymbol{\Gamma}_{12}(U_k)\boldsymbol{\Gamma}_{22}^{-1}(U_k)\boldsymbol{X}_k^{(2)})^T \\ &\quad \times \boldsymbol{\Gamma}_{1,2}^{-1}(U_k)(\boldsymbol{W}_j^{(1)} - \boldsymbol{\Gamma}_{12}(U_k)\boldsymbol{\Gamma}_{22}^{-1}(U_k)\boldsymbol{X}_j^{(2)}) \\ &\quad + R_{n4} + R_{n5} + o_p(h^{-1/2}),\end{aligned}$$

where $\boldsymbol{W}_i^{(1)} = (1, \boldsymbol{X}_i^{(1)^T})^T$ and

$$
\begin{aligned}
R_{n4} =& \frac{r_n^4}{2} \sum_{i=1}^n \sum_{j=1}^n \xi_i \xi_j \sum_{k=1}^n (\boldsymbol{W}_i^{(1)} - \boldsymbol{\Gamma}_{12}(U_k)\boldsymbol{\Gamma}_{22}^{-1}(U_k)\boldsymbol{X}_i^{(2)})^T \boldsymbol{\Gamma}_{1,2}^{-1}(U_k) \\
& \times (\boldsymbol{W}_k^{(1)} - \boldsymbol{\Gamma}_{12}(U_k)\boldsymbol{\Gamma}_{22}^{-1}(U_k)\boldsymbol{X}_k^{(2)})\boldsymbol{X}_k^{(2)^T} \boldsymbol{\Gamma}_{22}^{-1}(U_k)\boldsymbol{X}_j^{(2)} \\
& \times K((U_i - U_k)/h)K((U_j - U_k)/h), \\
R_{n5} =& \frac{r_n^4}{2} \sum_{i=1}^n \sum_{j=1}^n \xi_i \xi_j \sum_{k=1}^n (\boldsymbol{W}_j^{(1)} - \boldsymbol{\Gamma}_{12}(U_k)\boldsymbol{\Gamma}_{22}^{-1}(U_k)\boldsymbol{X}_j^{(2)})^T \boldsymbol{\Gamma}_{1,2}^{-1}(U_k) \\
& \times (\boldsymbol{W}_k^{(1)} - \boldsymbol{\Gamma}_{12}(U_k)\boldsymbol{\Gamma}_{22}^{-1}(U_k)\boldsymbol{X}_k^{(2)})\boldsymbol{X}_k^{(2)^T} \boldsymbol{\Gamma}_{22}^{-1}(U_k)\boldsymbol{X}_j^{(2)} \\
& \times K((U_i - U_k)/h)K((U_j - U_k)/h).
\end{aligned}
$$

After some tedious calculation, as $nh^{3/2} \to \infty$, $E(R_{n4}^2) = O(n^{-2}h^{-4}) = o(h^{-1})$ and thus $R_{n4} = o_p(h^{-1/2})$. Similarly, we can show $R_{n5} = o_p(h^{-1/2})$. As a consequence,

$$
\begin{aligned}
& -\lambda_{nb}\tau^2 \\
& = -r_n^2 \sum_{i=1}^n \sum_{j=1}^n \xi_i \xi_j (\boldsymbol{W}_j^{(1)} - \boldsymbol{\Gamma}_{12}(U_i)\boldsymbol{\Gamma}_{22}^{-1}(U_i)\boldsymbol{X}_j^{(2)})^T \boldsymbol{\Gamma}_{1,2}^{-1}(U_i) \\
& \quad \times (\boldsymbol{W}_i^{(1)} - \boldsymbol{\Gamma}_{12}(U_i)\boldsymbol{\Gamma}_{22}^{-1}(U_i)\boldsymbol{X}_j^{(2)})K((U_i - U_j)/h) \\
& \quad + \frac{r_n^4}{2} \sum_{i=1}^n \sum_{j=1}^n \xi_i \xi_j \sum_{k=1}^n (\boldsymbol{W}_i^{(1)} - \boldsymbol{\Gamma}_{12}(U_k)\boldsymbol{\Gamma}_{22}^{-1}(U_k)\boldsymbol{X}_i^{(2)})^T \boldsymbol{\Gamma}_{1,2}^{-1}(U_k) \\
& \quad \times (\boldsymbol{W}_k^{(1)} - \boldsymbol{\Gamma}_{12}(U_k)\boldsymbol{\Gamma}_{22}^{-1}(U_k)\boldsymbol{X}_k^{(2)})(\boldsymbol{W}_k^{(1)} - \boldsymbol{\Gamma}_{12}(U_k)\boldsymbol{\Gamma}_{22}^{-1}(U_k)\boldsymbol{X}_k^{(2)})^T \\
& \quad \times \boldsymbol{\Gamma}_{1,2}^{-1}(U_k)(\boldsymbol{W}_j^{(1)} - \boldsymbol{\Gamma}_{12}(U_k)\boldsymbol{\Gamma}_{22}^{-1}(U_k)\boldsymbol{X}_j^{(2)}) + o_p(h^{-1/2})
\end{aligned}
$$

The remaining proof follows the same lines as those in the proof of Theorem 6. □

Proof of Theorem 4 Let $\eta_i = \delta^{-1}(G(2c - \varepsilon_i) - 1/2)$ and $\boldsymbol{\Gamma}_{2,1} = \boldsymbol{\Gamma}_{22} - \boldsymbol{\Gamma}_{21}\boldsymbol{\Gamma}_{11}^{-1}\boldsymbol{\Gamma}_{12}$. Analogously to the arguments for $\widehat{\boldsymbol{A}}$, we get

$$
(\widehat{\alpha}^c(u_0), \widehat{\boldsymbol{A}}_1^c(u_0)^T)^T - (\alpha(u_0), \boldsymbol{A}_1(u_0)^T)^T = (c, \boldsymbol{0}_{1\times p_1})^T + (\widetilde{\alpha}_n(u_0) + \widetilde{R}_n(u_0))(1 + o_p(1)),
$$

where

$$
\begin{aligned}
\widetilde{\alpha}_n(u_0) &= r_n^2 \begin{pmatrix} 1 & 0 \\ 0 & \boldsymbol{\Gamma}_{11} \end{pmatrix}^{-1} \sum_{i=1}^n \eta_i (1, \boldsymbol{X}_i^{(1)^T})^T K_h((U_i - u_0)/h), \\
\widetilde{R}_n(u_0) &= r_n^2 \begin{pmatrix} 1 & 0 \\ 0 & \boldsymbol{\Gamma}_{11} \end{pmatrix}^{-1} \sum_{i=1}^n \Big(\alpha(U_i) - \boldsymbol{\beta}(u_0)^T \boldsymbol{V}_i(u_0) + \boldsymbol{A}_1(U_i)^T \boldsymbol{X}_i^{(1)} \\
& \qquad - \boldsymbol{A}_1(u_0)^T \boldsymbol{X}_i^{(1)} - \boldsymbol{A}_1''(u_0)^T \boldsymbol{X}_i^{(1)} \Big)(1, \boldsymbol{X}_i^{(1)^T})^T K((U_i - u_0)/h).
\end{aligned}
$$

Define $\bar{\phi}_{ij} = \alpha(U_i) - \widehat{\alpha}^c(U_i) + \boldsymbol{A}(U_i)^T\boldsymbol{X}_i - \widehat{\boldsymbol{A}}^c(U_i)^T\boldsymbol{X}_i + \alpha(U_j) - \widehat{\alpha}^c(U_j) + \boldsymbol{A}(U_j)^T\boldsymbol{X}_j - \widehat{\boldsymbol{A}}^c(U_j)^T\boldsymbol{X}_j$ and $\boldsymbol{W}_i^{(1)} = (1, \boldsymbol{X}_i^{(1)^T})^T$. Thus,

$$
\begin{aligned}
\lambda_{n2u} =& \frac{\psi^{-1}\delta}{n+1}\sum\sum_{i\le j}(|\widehat{\varepsilon}_i^c + \widehat{\varepsilon}_j^c| - |\widehat{\varepsilon}_i + \widehat{\varepsilon}_j|) \\
=& \frac{\psi^{-1}\delta}{n+1}\sum\sum_{i\le j}(|\varepsilon_i + \varepsilon_j - 2c + \bar{\phi}_{ij}| - |\varepsilon_i + \varepsilon_j - 2c + \phi_{ij}|) \\
=& \frac{\psi^{-1}\delta}{n+1}\sum\sum_{i\le j}(|\varepsilon_i + \varepsilon_j - 2c + \bar{\phi}_{ij}| - |\varepsilon_i + \varepsilon_j - 2c|) \\
&- \frac{\psi^{-1}\delta}{n+1}\sum\sum_{i\le j}(|\varepsilon_i + \varepsilon_j - 2c + \phi_{ij}| - |\varepsilon_i + \varepsilon_j - 2c|) \\
\doteq& G_h - H_h.
\end{aligned}
$$

Taking the same procedure as in the proof of Theorem 6, we can show that

$$
\begin{aligned}
G_h =& \psi^{-1}\delta^2\sum_{i=1}^n(\widetilde{\boldsymbol{\alpha}}_n(U_i) + \widetilde{\boldsymbol{R}}_n(U_i))^T\boldsymbol{W}_i^{(1)}\eta_i + ((\widetilde{\boldsymbol{\alpha}}_n(U_i) + \widetilde{\boldsymbol{R}}_n(U_i))^T\boldsymbol{W}_i^{(1)})^2 + o_p(h^{-1/2}), \\
H_h =& \psi^{-1}\delta^2\sum_{i=1}^n(\boldsymbol{\alpha}_n(U_i) + \boldsymbol{R}_n(U_i))^T\boldsymbol{W}_i\eta_i + ((\boldsymbol{\alpha}_n(U_i) + \boldsymbol{R}_n(U_i))^T\boldsymbol{W}_i)^2 + o_p(h^{-1/2}).
\end{aligned}
$$

Finally, similar to the proof of Theorem 3, we can obtain the result. □

Proof of Theorem 5 Denote $\xi_i^* = \sum\sum_{j\ne i}\frac{\varphi^{-1/2}\tau}{2(n+1)}\mathrm{sgn}(\varepsilon_i^* + \varepsilon_j^*)$. We will show that $E^*(\xi_i^*) = o_p(1)$ and $E^*(\xi_i^{*2}) = \tau^2(1 + o_p(1))$ where E^* denotes the conditional expectation given $\{\boldsymbol{X}_i, U_i, Y_i\}_{i=1}^n$.

$$
\begin{aligned}
E^*(\xi_i^*) =& E^*\left(\sum\sum_{j\ne i}\frac{\varphi^{-1/2}\tau}{2(n+1)}\mathrm{sgn}(\varepsilon_i^* + \varepsilon_j^*)\right) = \frac{\varphi^{-1/2}n\tau}{2(n+1)}E^*(\mathrm{sgn}(\varepsilon_i^* + \varepsilon_j^*)) \\
=& \frac{\varphi^{-1/2}\tau}{2n(n+1)}\sum_{i=1}^n\sum_{j=1}^n\mathrm{sgn}(\widehat{\varepsilon}_i + \widehat{\varepsilon}_j) = \frac{\varphi^{-1/2}\tau}{2n(n+1)}\sum_{1\le i,j\le n}\mathrm{sgn}(\varepsilon_i + \varepsilon_j + \phi_{ij}) \\
=& \frac{\varphi^{-1/2}\tau}{2n(n+1)}\sum_{i=1}^n\sum_{j=1}^n[\mathrm{sgn}(\varepsilon_i + \varepsilon_j + \phi_{ij}) - \mathrm{sgn}(\varepsilon_i + \varepsilon_j)] \\
&+ \frac{\varphi^{-1/2}\tau}{2n(n+1)}\sum_{i=1}^n\sum_{j=1}^n\mathrm{sgn}(\varepsilon_i + \varepsilon_j) \\
\doteq& I_{1h} + I_{2h},
\end{aligned}
$$

where ϕ_{ij} is defined in the proof of Theorem 6. Taking the similar procedure as for dealing with B_h, we obtain that $E(I_{1h}^2) = O((nh)^{-1})$ and $E(I_{2h}^2) = O(n^{-1})$. Thus,

$E^*(\xi_i^*) = o_p((nh)^{-1/2} + n^{-1/2}) = o_p(1)$. Next, we consider the second moment.

$$
\begin{aligned}
E^*(\xi_i^{*2}) = & E^*\left(\sum_{j\neq i}\sum \frac{\varphi^{-1/2}\tau}{2(n+1)}\mathrm{sgn}(\varepsilon_i^* + \varepsilon_j^*)\right)^2 \\
= & \frac{\varphi^{-1}\tau^2}{4n(n+1)^2}\sum_{i=1}^n\sum_{j=1}^n(\mathrm{sgn}(\varepsilon_i + \varepsilon_j + \phi_{ij})\mathrm{sgn}(\varepsilon_i + \varepsilon_l + \phi_{il})) \\
= & \frac{\varphi^{-1}\tau^2}{4n(n+1)^2}\sum_{i=1}^n\sum_{j=1}^n(\mathrm{sgn}(\varepsilon_i + \varepsilon_j)\mathrm{sgn}(\varepsilon_i + \varepsilon_l)) \\
& + \frac{\varphi^{-1}\tau^2}{4n(n+1)^2}\sum_{i=1}^n\sum_{j=1}^n\sum_{l=1}^n([\mathrm{sgn}(\varepsilon_i + \varepsilon_j + \phi_{ij}) - \mathrm{sgn}(\varepsilon_i + \varepsilon_j)]\mathrm{sgn}(\varepsilon_i + \varepsilon_l + \phi_{il})) \\
& + \frac{\varphi^{-1}\tau^2}{4n(n+1)^2}\sum_{i=1}^n\sum_{j=1}^n\sum_{l=1}^n(\mathrm{sgn}(\varepsilon_i + \varepsilon_j)[\mathrm{sgn}(\varepsilon_i + \varepsilon_l + \phi_{il}) - \mathrm{sgn}(\varepsilon_i + \varepsilon_l)]) \\
\doteq & J_{1h} + J_{2h} + J_{3h}.
\end{aligned}
$$

Using the similar arguments as for B_h, it can be verified that $E(J_{1h}) = \tau^2$, $\mathrm{var}(J_{1h}) = O(n^{-1})$, $E(J_{2h}^2) = O((nh)^{-1})$ and $E(J_{3h}^2) = O((nh)^{-1})$. Hence, $E^*(\xi_i^{*2}) = \tau^2 + O_p((nh)^{-1/2} + n^{-1/2})$. With these results, Theorem 5 can be established by mimicking the proof of Theorem 6. □

Proof of Corollaries 1-2 The proof of these two results are similar and thus we only elaborate on the first one. We decompose this problem as the following two simple null hypothesis

$$H_{0g1} : (\alpha_0, \alpha_1) = (\beta_0 + c, \beta_1) \quad \text{versus} \quad H_{1g1} : m(x) = \alpha_0 + \alpha_1 x \tag{A.2}$$

and

$$H_{0g2} : (\alpha_0, \alpha_1) = (\beta_0 + c, \beta_1) \quad \text{versus} \quad H_{1g2} : m(x) \neq \alpha_0 + \alpha_1 x, \tag{A.3}$$

where β_0, β_1 are the true value of parameters. The WGLR test statistics for the hypotheses (A.2) and (A.3) are denoted as λ_{ng2} and λ_{ng2}, respectively. It can be easily seen that $\lambda_{ng} = \lambda_{ng2} - \lambda_{ng1}$. According to Theorem 3, we have $\sigma_{ng}^{-1}(\lambda_{ng2} - \mu_{ng}) \xrightarrow{d} N(0, 1)$. Furthermore, by Theorem 3.6.1 in Hettmansperger and McKean (2010), we have $\lambda_{ng1} = O_p(1) = o_p(h^{-1/2})$, from which the result immediately follows. □

References

De Jong, P. (1987). A central limit theorem for generalized quadratic forms. *Probability Theory Related Field* **75**, 261–277.

Hall, P., and Heyde, C. C. (1980), *Martingale Limit Theory and Its Application*, New York: Academic Press.

Kou, H. L., Sievers, G. L. and McKean, J. (1987). An estimator of the scale parameter for the rank analysis of linear models under general score functions. *Scandinavian Journal of Statistics* **14**, 131–141.

Appendix D: Simulation results in heteroscedasticity cases

In this subsection, we conduct a simulation study in the heteroscedasticity case. All the settings are the same as the above subsection except that the variance function is

$$\varrho^2(u) = \frac{e^u}{\int_0^1 e^t dt}.$$

Similar to Koul *et al.*(1987), we choose the bandwidth t_n in (2.14) as $t_n = h\gamma_\alpha$ where γ_α is the α-th quantile of the empirical distribution function of $\{|\widehat{\varepsilon}_i - \widehat{\varepsilon}_j|\}_{1\le i<j\le n}$. Here we choose $\alpha = 0.8$.

Tables A.2 and A.3 report the simulated level of our test and power comparison with other tests, respectively. The simulated results are similar to the homoscedasticity case. We can control the empirical sizes in most cases. Under the normal cases, WGLR performs a litter worse than GLR, Zheng and LOSS tests. However, under the non-normal cases, WGLR is significantly powerful than the other tests. And WGLR still performs better than WQ in the model (I) and (VI). Thus, our WGLR procedure is also robust in the heteroscedasticity case.

Table A.2: Simulated level (%) of test on testing (3.1) with heteroscedastic error

	$n=25$			$n=50$			$n=100$		
h	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
0.12	7.5	7.0	7.1	5.4	4.5	5.8	5.6	5.9	4.8
0.15	7.2	6.7	7.3	5.7	4.9	4.8	5.3	5.2	4.9
0.18	6.3	5.1	5.3	4.3	4.0	4.7	4.1	5.4	5.3
0.21	4.5	5.0	4.2	5.5	4.4	3.8	4.2	4.1	5.7

Table A.3: Empirical power (%) of tests on testing (3.1) with heteroscedastic error.

		$n = 25$					$n = 50$				
	Models	WGLR	GLR	WQ	Zheng	LOSS	WGLR	GLR	WQ	Zheng	LOSS
$N(0, 1)$	(I)	13	13	4.8	25	23	37	27	4.0	43	39
	(II)	17	22	13	33	37	45	46	19	66	61
	(III)	11	24	20	17	14	40	51	47	41	26
	(IV)	16	32	48	46	21	50	65	82	77	49
	(V)	22	30	18	40	41	58	62	37	74	70
	(VI)	28	28	5.8	45	42	62	57	09	76	71
$t(3)$	(I)	40	16	5.5	37	27	75	39	4.6	50	56
	(II)	49	33	19	50	46	84	64	39	68	75
	(III)	43	32	35	23	20	87	68	83	41	47
	(IV)	55	43	71	57	35	90	77	97	79	68
	(V)	61	42	35	51	56	92	77	70	77	80
	(VI)	65	41	09	62	53	96	72	18	79	81
$T(0.05, 10)$	(I)	66	35	5.6	57	44	96	54	5.5	65	56
	(II)	78	50	44	72	60	99	70	71	77	73
	(III)	69	48	72	38	31	99	69	99	67	53
	(IV)	81	57	96	71	52	99	74	100	87	65
	(V)	86	60	69	69	66	100	74	95	80	79
	(VI)	89	58	19	74	65	100	76	32	83	80

Appendix E: Some additional simulation results in Section 3

Table A.4: Simulated level (%) of test on testing linearity

	$n = 60$				$n = 100$			
h	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(d)
0.06	7.2	6.8	7.0	7.3	6.2	6.4	6.5	6.6
0.09	6.4	6.0	6.3	6.6	6.1	5.8	6.2	5.6
0.12	5.3	5.2	4.7	4.6	5.8	5.5	4.4	4.9
0.15	5.0	5.4	4.5	5.3	5.1	4.7	5.3	5.0

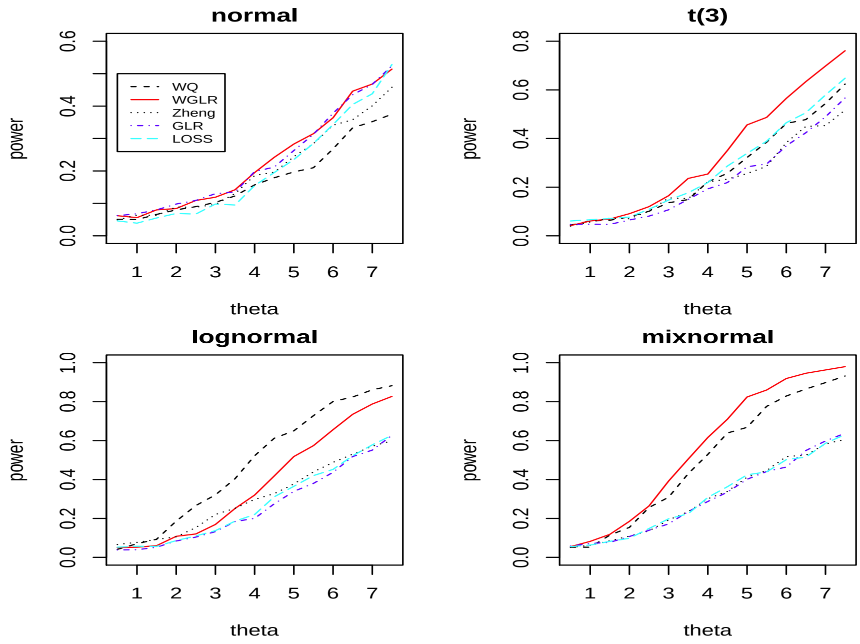

Figure A.1: Simulated power curves of square alternative on testing linearity.

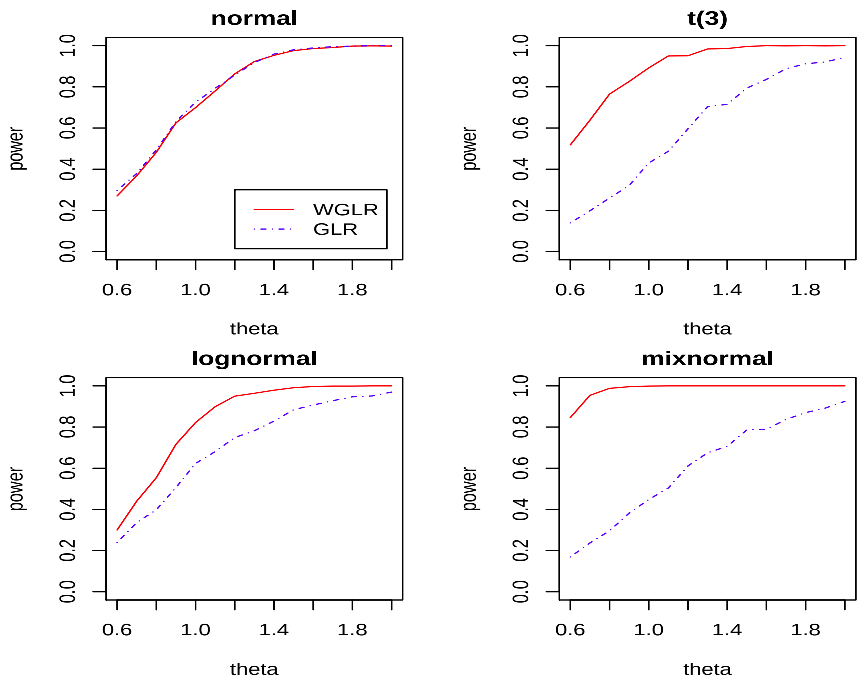

Figure A.2: Simulated power curves on testing homogeneity with $n = 100, h = 0.09$.

Table A.5: Simulated level (%) of test on testing homogeneity

	$n = 60$				$n = 100$			
h	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(d)
0.06	6.9	7.3	7.2	7.3	6.2	6.3	6.0	6.9
0.09	5.9	6.2	5.5	6.8	6.3	5.7	6.2	5.9
0.12	5.3	5.8	5.7	4.4	5.8	5.8	4.7	5.7
0.15	4.4	5.7	4.9	5.9	5.0	4.2	5.6	5.2

Appendix F: Some additional figures in Section 4

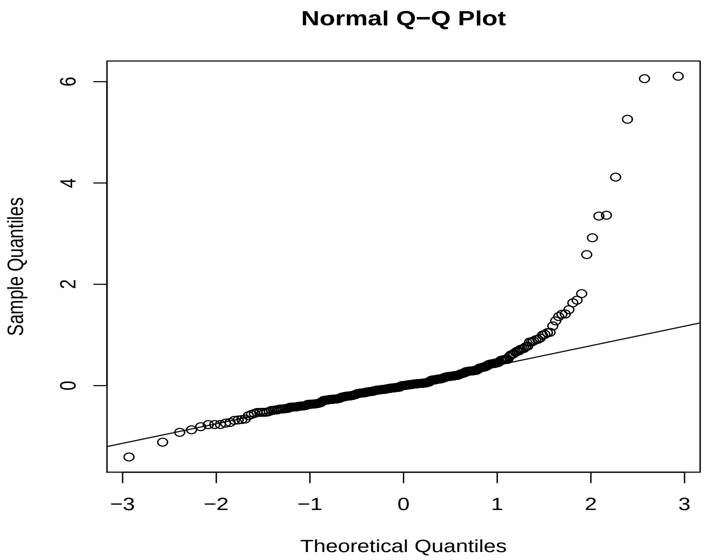

Figure A.3: The normal QQ-plot for the residuals

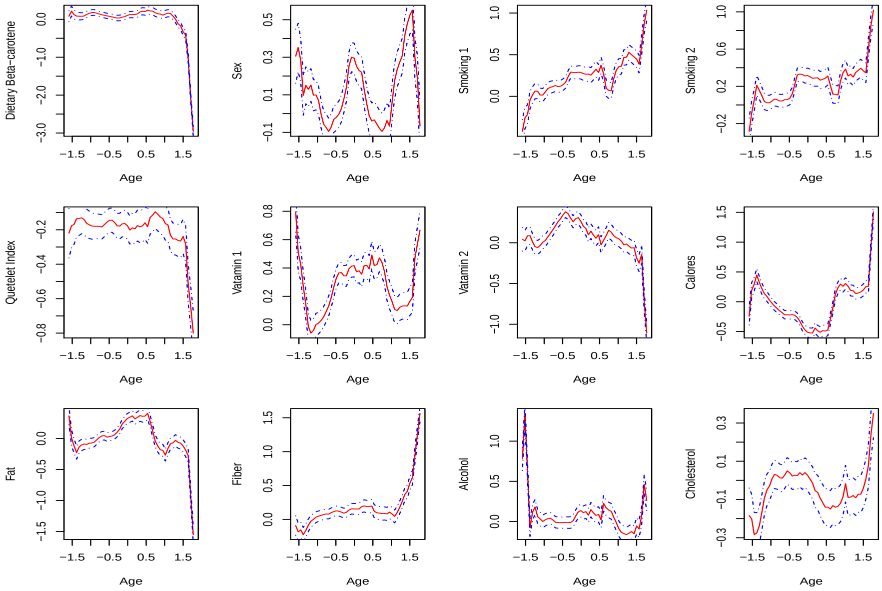

Figure A.4: Fitted coefficient functions and corresponding pointwise 95% confidence interval.