

TESTS FOR VARYING-COEFFICIENT PARTS ON VARYING-COEFFICIENT SINGLE-INDEX MODEL

ZHENSHENG HUANG AND RIQUAN ZHANG

ABSTRACT. To study the relationship between the levels of chemical pollutants and the number of daily total hospital admissions for respiratory diseases and to find the effect of temperature/relative humidity on the admission number, Wong *et al.* [17] introduced the varying-coefficient single-index model (VCSIM). As pointed out, it is a popular multivariate nonparametric fitting technique. However, the tests of the model have not been very well developed. In this paper, based on the estimators obtained by the local linear technique, the average method and the one-step back-fitting technique in the VCSIM, the generalized likelihood ratio (GLR) tests for varying-coefficient parts on the VCSIM are established. Under the null hypotheses the new proposed GLR tests follow the χ^2 -distribution asymptotically with scale constant and degree of freedom independent of the nuisance parameters, known as Wilks phenomenon. Simulations are conducted to evaluate the test procedure empirically. A real example is used to illustrate the performance of the testing approach.

1. Introduction

To study the relationship between the levels of chemical pollutants and the number of daily total hospital admissions for respiratory diseases in Hong Kong, to find the effect of temperature/relative humidity on the admission number, and to examine how the effect of weather variables varies with time, Wong *et al.* [17] introduced the VCSIM,

$$(1) \quad Y = a_0(\alpha_0^T \mathbf{X}) + \sum_{j=1}^p a_j(U) Z_j + \varepsilon,$$

where $a_0(\cdot)$ is an unknown univariable function, $\alpha_0 \in \mathbb{R}^q$ is an unknown parametric vector with $\|\alpha_0\| = 1$ for identifiability, $\mathbf{X} \in \mathbb{R}^q$, $Y \in \mathbb{R}$ and

Received June 5, 2008; Revised March 8, 2009.

2000 *Mathematics Subject Classification.* Primary 62G08; Secondary 62G20.

Key words and phrases. averaged method, back-fitting algorithms, generalized likelihood ratio test, local linear method, varying-coefficient single-index model, Wilks phenomenon.

The research was supported in partly the Natural Science Foundation of China (10871072), Shanxi's Natural Science Foundation of China (2007011014) and Ph. D. Program Scholarship Fund of ECNU 2009.

$\mathbf{Z} = (Z_1, \dots, Z_p)^T \in \mathbb{R}^p$; $\mathbf{a}(\cdot) = (a_1(\cdot), \dots, a_p(\cdot))^T$ is an unknown function vector from \mathbb{R} to \mathbb{R}^p , U is a covariate, which may be one entry of \mathbf{Z} or \mathbf{X} , or some other covariate; the error ε is independent of $(\mathbf{X}, U, \mathbf{Z})$ with $E(\varepsilon) = 0$ and $\text{Var}(\varepsilon) = \sigma^2$.

The VCSIM avoids the so-called ‘‘curse of dimensionality’’ and is flexible enough to cover a variety of situations. When $\mathbf{a}(\cdot) = 0$, or equivalently, there are no predictors \mathbf{Z} and U , it is the nonparametric single-index model (Härdle, Hall and Ichimura [11]). When $q = 1$, i.e., \mathbf{X} is a scalar, the model becomes the varying coefficient model which has a varying intercept with a different smoothing variable from that of the coefficient functions (Hastie and Tibshirani [13]). Zhang and Li [20] and Ip, Wong, and Zhang [15] discussed the estimation and inference for the model. When there is no predictors \mathbf{X} , The VCSIM is the varying coefficient model with a single smoothing variable (Hastie and Tibshirani [13]). The varying coefficient model has extremely wide applications. For example, see Hoover *et al.* [14] for novel applications of the model to longitudinal data; Fan, Yao, and Cai [1] for applications in ecologic data analysis; Chen and Tsay [3] and Cai, Fan, and Yao [1] for nonlinear time series applications. If $\mathbf{X} = \mathbf{Z}$ and $\mathbf{a}(u) = \mathbf{a}$, then the VCSIM is the case of the extended partially linear single index model considered by Xia, Tong, and Li [18]. Note that in their paper the theory covers the case of nonlinear time series as well.

Wong *et al.* [17] employed the local linear method, the average method and the back-fitting technique to obtain the estimates of the unknown parameters and the unknown functions of the model (1), and gave their asymptotic distribution. They thought that in many practical situations the partially linear single-index (PLSI) model, which was discussed by Carroll *et al.* [2], is not complex enough to capture the underlying relationship between the response variable Y and its covariates (\mathbf{X}, \mathbf{Z}) . They gave the VCSIM that allows the effect of \mathbf{Z} on Y to be varying with some covariates. After fitting the model, another important issue arises regarding whether the varying coefficients are in fact not varying, or whether the coefficient functions are not those given ones. Zhang [19] discussed the testing problem for nonparametric parts on the PLSI models by the GLR test, and the Wilks phenomenon was unveiled. However, there has been less attention focused on the test problem for the varying-coefficient parts on the VCSIM, i.e.,

$$(2) \quad H_0 : a_j(u) = a_{j0}(u) \longleftrightarrow H_1 : a_j(u) \neq a_{j0}(u), \quad j = 1, \dots, p.$$

If $\{a_{j0}(\cdot)\}_{j=1}^p$ are some constants, the testing problem (2) answers the question that whether the coefficient functions are not varying; when $\{a_{j0}(\cdot)\}_{j=1}^p$ are some functions, it tests whether the coefficient functions are not these given functions.

Obviously the tools used in the inferences of the PLSI models cannot be used directly in the testing problem (2) because of the difference in the process of estimating the unknown parameters and the unknown functions in two models. In this paper, we extend the generalized likelihood ratio (GLR) tests to the

testing problem (2) for the VCSIM (1) by using the estimation method proposed by Wong *et al.* [17]. This not only provides a useful tool to the frequently asked question, but also enriches the GLR test theory. The GLR test was proposed by Fan *et al.* [9] for the inference of nonparametric models. It is constructed by replacing the maximum likelihood estimator in the maximum likelihood ratio test by a reasonable nonparametric estimator of the coefficient function. The GLR test is extended in many fields because of its some nice properties. Fan and J. Zhang [7] proposed the sieve empirical likelihood ratio test to handle the case that the distribution of the stochastic error in the model is completely unspecified. Fan and W. Zhang [8] employed the GLR test for the spectral density. Fan and Jiang [6] extended the GLR tests to additive models, while Ip *et al.* [15] also extended the GLR test to varying-coefficient models with different smoothing variables. Therefore, one naturally ask if the GLR test can be applicable to the VCSIM (1). This question is solved in this paper. And it is showed that the proposed GLR test shares the same nice properties as that of Fan *et al.* [9].

The rest of this paper is organized as follows. In Section 2 we introduce the estimating process. In Section 3 the GLR test statistic is constructed and its asymptotic null distribution, power and minimax rate are developed. Simulations are given in Section 4 and we provide an example of testing on a real data set in Section 5. The proofs are given in Appendix.

2. Estimation

Suppose that $(Y_i, \mathbf{X}_i, U_i, \mathbf{Z}_i)_{i=1}^n$ is an i.i.d sample from the VCSIM (1), and that for every $j = 0, \dots, p$, $a_j(\cdot)$ has a continuous second derivative. For v in a small neighborhood of u , $a_j(\cdot)$ can be approximated locally by a linear function, i.e., $a_j(v) \approx a_j(u) + a'_j(u)(v - u) \equiv a_j + b_j(v - u)$, $j = 0, \dots, p$. Throughout the article, we write $K_{j,h_j}(\cdot) = K_j(\cdot/h_j)h_j^{-1}$, where $K_j(\cdot)$ is a bounded, non-negative, compactly supported symmetric about zero and Lipschitz continuous density function, and h_j ($h_j > 0$) is a bandwidth, $j = 1, 2, 3, 4$; $\{t_i\}_{i=1}^n$ denotes a vector $(t_1, \dots, t_n)^T$. The estimation procedure from Wong *et al.* [17] is as follows.

Step 1. Define the quasi-initial estimators of $a_j(\cdot)$.

With given α , by using the local linear method in model (1), we define the quasi-initial estimators of $\{a_j(\cdot)\}_{j=0}^p$ as

$$(3) \quad \check{a}_0(t; \alpha) = \frac{1}{n} \sum_{k=1}^n \tilde{a}_0(t, U_k; \alpha),$$

$$(4) \quad \check{a}_j(u; \alpha) = \frac{1}{n} \sum_{k=1}^n \tilde{a}_j(\alpha^T \mathbf{X}_k, u; \alpha),$$

where

$$\begin{aligned}\tilde{a}_0(t, u; \alpha) &= \sum_{i=1}^n e_{1,2p+2}^T (V^T W V)^{-1} V_i W_{ii} Y_i, \\ \tilde{a}_j(t, u; \alpha) &= \sum_{i=1}^n e_{j+1,2p+2}^T (V^T W V)^{-1} V_i W_{ii} Y_i\end{aligned}$$

for every $j = 1, \dots, p$. Here $e_{j,2p+2}$ is the $(2p + 2) \times 1$ unit vector with 1 at the j -th position, $V = V(x_0, u_0)$ denotes an $n \times (2p + 2)$ matrix with $V_i^T = (1, \mathbf{Z}_i^T, (\alpha^T \mathbf{X}_i - t)/h_1, ((U_i - u)/h_2)Z_{i1}, \dots, ((U_i - u_0)/h_2)Z_{ip})$ as its i -th row, where $\mathbf{Z}_i = (Z_{i1}, \dots, Z_{ip})^T$ and $W = \text{diag}(W_{11}, \dots, W_{nn})$ with $W_{ii} = K_{1,h_1}(\alpha^T \mathbf{X}_i - t)K_{2,h_2}(U_i - u)$.

Step 2. Obtain the quasi-efficient estimators of $a_j(\cdot)$.

To decrease the variances of the quasi-initial estimators, a one-step back-fitting technique is employed to obtain the quasi-efficient estimators of $\{a_j(\cdot)\}_{j=0}^p$, we have

$$(5) \quad \hat{a}_0(t; \alpha) = \sum_{i=1}^n e_{1,2}^T (\tilde{X}^T W_1 \tilde{X})^{-1} \tilde{X}_i K_{3,h_3}(\alpha^T \mathbf{X}_i - t) \tilde{Y}_i,$$

$$(6) \quad \hat{a}_j(u; \alpha) = \sum_{i=1}^n e_{j,2p}^T (\tilde{U}^T W_2 \tilde{U})^{-1} \tilde{U}_i K_{4,h_4}(U_i - u) \tilde{Y}_i,$$

where $e_{1,2} = (1, 0)^T$, \tilde{X} denotes an $n \times 2$ matrix with $\tilde{X}_i^T = (1, (\alpha^T \mathbf{X}_i - t)/h_3)$ as its i -th row, $W_1 = \text{diag}(W_{11}, \dots, W_{nn})$ with $W_{ii} = K_{3,h_3}(\alpha^T \mathbf{X}_i - t)$, $\tilde{Y}_i = Y_i - \sum_{j=1}^p \hat{a}_j(U_i; \alpha) Z_{ij}$, $e_{j,2p}$ is the $2p \times 1$ unit vector with 1 at the j -th position, \tilde{U} denotes an $n \times 2p$ matrix with $\tilde{U}_i^T = (\mathbf{Z}_i^T, ((U_i - u)/h_4)\mathbf{Z}_i^T)$ as its i -th row, where $\mathbf{Z}_i = (Z_{i1}, \dots, Z_{ip})^T$ and $W_2 = \text{diag}(W_{11}, \dots, W_{nn})$ with $W_{ii} = K_{4,h_4}(U_i - u)$, $\tilde{Y}_i = Y_i - \hat{a}_0(\alpha^T \mathbf{X}_i)$.

Step 3. Obtain the estimator of α_0 . The estimator $\hat{\alpha}_0$ of α_0 is defined as the minimizer of the sum of squares

$$\sum_{i=1}^n \left(Y_i - \hat{a}_0(\alpha^T \mathbf{X}_i; \alpha) - \sum_{j=1}^p \hat{a}_j(U_i; \alpha) Z_{ij} \right)^2,$$

subject to $\|\alpha\| = 1$.

Step 4. Obtain the estimators of all unknown functions. With obtained $\hat{\alpha}$ from step 3, we define the estimates of $\{a_j(\cdot)\}_{j=0}^p$ as

$$\hat{a}_0(t) = \hat{a}_0(t; \hat{\alpha}_0), \quad \hat{a}_j(u) = \hat{a}_j(u; \hat{\alpha}_0), \quad j = 1, \dots, p.$$

Remark 1. When α is a known constant α_0 or estimated to the order $O_p(n^{-1/2})$, Wong *et al.* [17] discussed the asymptotic properties of the proposed quasi-efficient estimators. They showed that the quasi-efficient estimators of the unknown function $a_0(\cdot)$ has the same asymptotic normality properties as the local

linear estimator in univariate model $Y = a_0(X) + \varepsilon$, which was discussed by Fan and Gijbels [5]; and the quasi-efficient estimators of the coefficient functions $\{a_j(\cdot)\}_{j=1}^p$ have the same asymptotic normality distribution as the local linear estimators in functional-coefficient regression model $Y = \sum_{j=1}^p a_j(U)Z_j + \varepsilon$ (see Cai *et al.* [1]).

3. Generalized likelihood ratio tests

3.1. The generalized likelihood ratio tests

In this section, based on the estimation procedure in Section 2, we define the GLR statistic and develop its asymptotic theory under the model (1), the Wilks phenomenon is unveiled.

Assume that $\varepsilon \sim N(0, \sigma^2)$ for simplicity. In fact, the normality assumptions are not needed, which is demonstrated in Fan *et al.* [9] and our proofs in Appendix.

The log-likelihood under the model (1) is

$$l = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n \left(Y_i - a_0(\alpha_0^T \mathbf{X}_i) - \sum_{j=1}^p a_j(U_i)Z_{ij} \right)^2.$$

Replacing the unknown functions by the estimators given in Section 2, we have

$$l = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n \left(Y_i - \hat{a}_0(\hat{\alpha}_0^T \mathbf{X}_i) - \sum_{j=1}^p \hat{a}_j(U_i)Z_{ij} \right)^2.$$

Maximizing l over σ^2 leads to the generalized likelihood under the model (1),

$$\ell_n(H_1) = -(n/2) \log(2\pi/n) - (n/2) \log(RSS_1) - n/2,$$

where

$$RSS_1 = \sum_{i=1}^n \left(Y_i - \hat{a}_0(\hat{\alpha}_0^T \mathbf{X}_i) - \sum_{j=1}^p \hat{a}_j(U_i)Z_{ij} \right)^2.$$

Similarly, the maximum log-likelihood under H_0 can be expressed as

$$\ell_n(H_0) = -(n/2) \log(2\pi/n) - (n/2) \log(RSS_0) - n/2,$$

where

$$RSS_0 = \sum_{i=1}^n \left(Y_i - \bar{a}_0(\bar{\alpha}_0^T \mathbf{X}_i) - \sum_{j=1}^p a_{j0}(U_i)Z_{ij} \right)^2,$$

and $\bar{a}_0(\cdot)$, $\bar{\alpha}_0$ as the estimators of $a_0(\cdot)$ and α_0 respectively under H_0 , where $\bar{\alpha}_0$ is the minimizer of the sum of squares

$$\sum_{i=1}^n \left(Y_i - \sum_{j=1}^p a_{j0}(U_i; \alpha)Z_{ij} - \hat{a}_0^{(0)}(\alpha^T \mathbf{X}_i; \alpha) \right)^2,$$

subject to $\|\alpha\| = 1$ with

$$\hat{a}_0^{(0)}(t; \alpha) = \sum_{i=1}^n e_{1,2}^T (\tilde{X}^T W_1 \tilde{X})^{-1} \tilde{X}_i K_{3,h_3}(\alpha^T \mathbf{X}_i - t) (Y_i - \sum_{j=1}^p a_{j0}(U_i; \alpha) Z_{ij}),$$

and $\bar{a}_0(t) = \hat{a}_0^{(0)}(t; \bar{\alpha}_0)$.

Following Fan *et al.* [9], the GLR statistics is defined as

$$(7) \quad \lambda_n = \ell_n(H_1) - \ell_n(H_0) = \frac{n}{2} \log \frac{RSS_0}{RSS_1} \approx \frac{n}{2} \frac{RSS_0 - RSS_1}{RSS_1},$$

which compares the nearly best fitting likelihood in the full model with that under the null model. Intuitively, the null hypothesis is rejected when λ_n is too large.

3.2. Asymptotic null distribution

To derive the asymptotic distribution of λ_n under H_0 , the following assumptions will be used. Condition A:

- (A₁) For every $j = 0, 1, \dots, p$, $a_j(\cdot)$ has a Lipschitz continuous second derivative.
- (A₂) For every $j = 1, 2, 3, 4$, $K_j(\cdot)$ are some bounded, nonnegative, compactly supported, symmetric about zero and Lipschitz continuous density functions.
- (A₃) The joint density $p(t, u)$ of $\alpha^T \mathbf{X}$ and U , the marginal density $p_1(t)$ of $\alpha^T \mathbf{X}$, the marginal density $p_2(u)$ of U are compactly supported, bounded, Lipschitz continuous and bounded away from zero by a constant, and U has a bounded support Ω , $\alpha^T \mathbf{X}$ has a bounded support L for α in a neighborhood of α_0 .
- (A₄) The function $E(\mathbf{Z} | \alpha^T \mathbf{X} = t, U = u)$, $E(\mathbf{Z}\mathbf{Z}^T | \alpha^T \mathbf{X} = t, U = u)$ and $E((\mathbf{Z}\mathbf{Z}^T) * (\mathbf{Z}\mathbf{Z}^T) | \alpha^T \mathbf{X} = t, U = u)$ are Lipschitz continuous, where $A * B$ is the Hadamard product of matrix A and B .
- (A₅) $nh_1 h_2 \rightarrow \infty$, $h_1/h_j \rightarrow 0$, $h_2/h_j \rightarrow 0$, $nh_j^{3/2} \rightarrow \infty$, $nh_j^5 \rightarrow 0$, $j = 3, 4$.
- (A₆) $\{\varepsilon_i\}_{i=1}^p$ are i.i.d and $E|\varepsilon_i|^4 < \infty$.

Theorem 1. *Suppose Condition A holds. Then, under H_0 ,*

$$(8) \quad \sigma_n^{-1}(\lambda_n - \mu_n + d_n) \xrightarrow{d} N(0, 1).$$

Furthermore, if $a_{j0}(\cdot)$ is linear or $nh_4^{9/2} \rightarrow 0$, then, under H_0 ,

$$(9) \quad r_K \lambda_n \sim_a \chi^2(r_K \mu_n),$$

where $r_K = 2\mu_n/\sigma_n^2$, $\mu_n = \frac{p|\Omega|}{h_4} \left(K_4(0) - \frac{1}{2} \int K_4^2(t) dt \right)$, $\sigma_n^2 = \frac{2p|\Omega|}{h_4} \int (K_4(t) - \frac{1}{2} K_4 * K_4(t))^2 dt$, $d_n = O_p(nh_4^4 + \sqrt{nh_4^2})$, $K_4 * K_4(t)$ denotes the convolution of K_4 .

Remark 2. The scale constant r_K and the degree of freedom $r_K\mu_n$ of the asymptotic chi-square distribution are independent of the nuisance parameters or the functions σ^2 , α_0 , $p(t, u)$, $p_1(t)$, $p_2(u)$. This phenomenon is called the Wilks phenomenon.

Remark 3. When $a_0(\cdot) = 0$, the model (2) reduces to the varying-coefficient model proposed by Hastie and Tibshirani [13]. The new GLR statistic becomes the GLR statistic of Fan *et al.* [9]. Thus the two statistics share the same asymptotic properties.

3.3. Power of generalized likelihood ratio tests

Based on the estimators in Section 2, we now consider the power of the GLR tests.

Assume that $h_4 = o(n^{-1/5})$, so that the second term in the definition of d_n is of smaller order than σ_n (see Fan *et al.* [9]). As is stated later in Theorem 3, the optimal bandwidth for the testing problem (2) is $h_4 = O(n^{-2/9})$, which satisfies the condition $h_4 = o(n^{-1/5})$. Under these assumptions, Theorem 1 leads to an approximate level- γ test based on the GLR statistic,

$$\phi_{h_4} = I\{\lambda_n(H_0) - \mu_n \geq Z_\gamma \sigma_n\}.$$

If we consider the contiguous alternative of form

$$H_{1n} : a_j(u) = a_{j0}(u) + g_{jn}(u), \quad j = 1, \dots, p,$$

where $g_{jn}(u)$ has a Lipschitz continuous second derivative for every j , the power of the new GLR tests can be approximated by using the following theorem.

Theorem 2. *Suppose that Condition A holds and that for every $j = 1, \dots, p$, $nh_4 E(g_{jn}^2(U)|Z_j) \rightarrow C_1$, $nh_4 E((g_{jn}''(U))^2|Z_j) \rightarrow C_2$, $E((g_{jn}(U)Z_j)^4|Z_j) = O((nh_4)^{-3/2})$, $E((g_{jn}''(U)Z_j)^4|Z_j) = O((nh_4)^{-3/2})$ for some C_1 and C_2 which are dependent of $G_n = (g_{1n}(u), \dots, g_{pn}(u))$. Then, under H_{1n} ,*

$$(10) \quad \sigma_{1n}^{*-1}(\lambda_n - \mu_n - d_{1n}) \xrightarrow{d} N(0, 1),$$

where $\sigma_{1n}^{*2} = \sigma_n^2 + n\sigma^{-2} E(\sum_{j=1}^p g_{jn}(U)Z_j)^2$ and $d_{1n} = (n/2\sigma^2) E(\sum_{j=1}^p g_{jn}(U)Z_j)^2$.

To study the optimal property of the GLR tests, we consider the class of functions $\mathbb{G}_n = \{G_n = (g_{1n}(u), \dots, g_{pn}(u))\}$, satisfying the regularity conditions:

$$\begin{aligned} \text{Var}\left(\left(\sum_{j=1}^p g_{jn}(U)Z_j\right)^2\right) &\leq M\left(E\left(\sum_{j=1}^p g_{jn}(U)Z_j\right)^2\right)^2, \\ nE\left(\sum_{j=1}^p g_{jn}(U)Z_j\right)^2 &> M_n \rightarrow \infty \end{aligned}$$

for some constants $M > 0$ and $M_n \rightarrow \infty$. For a given $\rho > 0$, let

$$\mathbb{G}_n(\rho) = \left\{ G_n = (g_{1n}(u), \dots, g_{pn}(u)) \in \mathbb{G}_n : E \left(\sum_{j=1}^p g_{jn}(U) Z_j \right)^2 \geq \rho^2 \right\}.$$

The maximum of the probabilities of type II error is the given by

$$\beta(\gamma, \rho) = \sup_{G_n \in \mathbb{G}_n(\rho)} \beta(\gamma, G_n),$$

where $\beta(\gamma, G_n) = P(\phi_{h_4} = 0 \mid a_j(u) = a_{j0}(u) + g_{jn}(u), j = 1, \dots, p)$ is the probability of a type II error at the alternative $H_{1n} : a_j(u) = a_{j0}(u) + g_{jn}(u), j = 1, \dots, p$. The minimax rate of ϕ_{h_4} is defined as the smallest ρ_n such that:

- (i) for every $\rho > \rho_n, \gamma > 0$, and for every $\beta > 0$, there exists a constant c , such that $\beta(\gamma, c\beta) \leq \beta + o(1)$; and
- (ii) for any sequence $\rho_n^* = o(\rho_n)$, there exist $\gamma > 0, \beta > 0$ such that for any $c > 0, P(\phi_{h_4} = 1 \mid a_j(u) = a_{j0}(u), j = 1, \dots, p) = \gamma + o(1)$, and $\liminf_n \beta(\gamma, c\rho_n^*) > \beta$.

This measures how close are the alternative that can be detected by the new GLR tests ϕ_{h_4} . The rate depends on the bandwidth h_4 . To stress its dependence, we write it as $\rho_n(h_4)$.

Theorem 3. *Under Condition (A), the generalized likelihood can detect alternatives with rate $\rho_n(h_4) = n^{-4/9}$, when $h_4 = C_* n^{-2/9}$ for some constant C_* .*

Remark 4. The GLR tests are asymptotically optimal in terms of rates of convergence for an ordinary nonparametric hypothesis testing according to the formulations of Lepski and Spokoiny [16]. When $p = 1, Z_1 = 1$ and $a_0(\cdot) = 0$, the model (1) becomes an ordinary nonparametric model, in which Lepski and Spokoiny [16] proved the optimal rate to be $n^{-4/9}$ for testing H_0 . Therefore, the new GLR test is optimal in this sense that it achieves the optimal rate of convergence.

4. Simulations

In this section, the purpose of the simulations is twofold: demonstrating the Wilks phenomenon, and the power of the proposed GLR test. The effect of the error ε distribution on the performance of the test is also investigated. Numerical results show that the new GLR test works well. Throughout this section, the Epanechnikov kernel is used.

We consider the following model:

$$(11) \quad Y = a_0(\alpha_1 X_1 + \alpha_2 X_2) + a_1(U) Z_1 + a_2(U) Z_2 + \varepsilon,$$

and testing the null hypothesis

$$H_0 : a_1(u) = \sin(6\pi u), \quad a_2(u) = \cos(6\pi u),$$

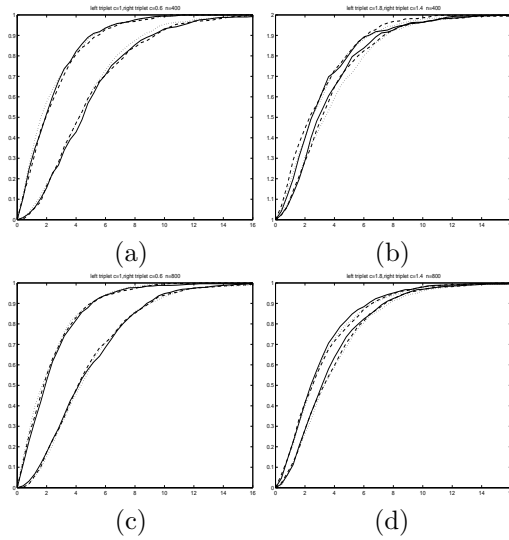


FIGURE 1. The empirical distributions of the GLR statistic under H_0 in several cases. (a) $n = 400$, $c=0.5$ (left triplet) and 1 (right triplet); (b) $n = 400$, $c=1.5$ (left triplet) and 2 (right triplet); (c) $n = 800$, $c=0.5$ (left triplet) and 1 (right triplet); (d) $n = 800$, $c=1.5$ (left triplet) and 2 (right triplet); where $F(x) = N(0, 1)$ (solid curve), $t(6)$ (dashed curve) and $U[-0.5, 0.5]$ (dotted curve).

where $a_0(x) = 2 \exp(-x^2)$, $(X_1, X_2, U, Z_1, Z_2)^T \sim U([-0.5, 0.5]^5)$, $\alpha_1 = 1/\sqrt{5}$, $\alpha_2 = 2/\sqrt{5}$, $\varepsilon \sim F(x)$, which is a distribution function, and ε is independent of (X_1, X_2, U, Z_1, Z_2) .

In this section, in order to satisfy the order requirements of the bandwidths in Condition (A_5) and Theorem 3, and for simplicity, we take $h_1 = h_2 = cn^{-1/3}$, $h_3 = h_4 = cn^{-2/9}$, where c is some positive constant and n is the size of the sample.

To show the Wilks phenomenon, the error distribution $F(x)$ is taken as $N(0, 1)$, $t(5)$ and $U[-0.5, 0.5]$ respectively. Sample sizes $n = 400$ and 800 from the model (11) are used under H_0 with every $F(x)$. To study the effect of the bandwidth on the new GLR test, c takes on the values 0.5, 1, 1.5 and 2. For each combination of (c, n, F) , we simulate 300 replications for the corresponding model at each run.

The empirical distributions of the GLR statistics λ_n are shown in Figure 1(a-d) in several cases under H_0 .

From Figure 1, we can see that the nuisance function $F(x)$ has less effect on the empirical distribution of the GLR statistic under H_0 , since the lines under

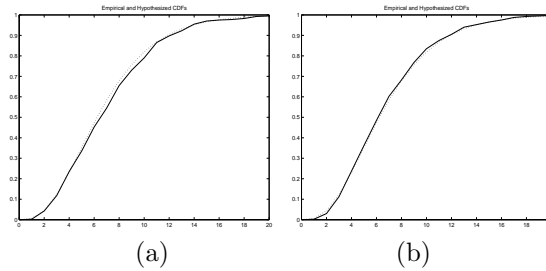


FIGURE 2. The empirical distributions (solid curve) of the scaled GLR statistic under H_0 and the chi-square distributions (dotted curve) with degrees of freedom of d_0 : (a) $n = 400, c = 0.5, F(x) = N(0, 1)$, (b) $n = 800, c = 2, F(x) = t(5)$.

different $F(x)$'s are very close to one another. We also see that the empirical distribution of the GLR statistic under H_0 are dependent of the bandwidth.

To check whether the normalized statistic $r_K \lambda_n$ follows asymptotically the chi-square distribution, we equate the mean and variance of the scaled statistic $r_K \lambda_n$ to the corresponding mean and variance of a chi-square random variable, say $\chi^2(d_0)$, with degrees of freedom d_0 . This results in $r_K = 2\mu_0/\sigma_0^2$ and $d_0 = 2\mu_0^2/\sigma_0^2$ with μ_0 and σ_0^2 the simulated mean and variance of λ_n . We calculate further the empirical distribution of the scaled GLR statistic, and compare it with the $\chi^2(d_0)$ distribution based on 300 replications from the model (11) under H_0 with every $F(x)$ and every c . Here we only show two simulated results in Figure 2, as the empirical distribution do not depend sensitively on the error distribution $F(x)$.

From Figure 2, we can see that the empirical distribution of the scaled GLR statistic and the $\chi^2(d_0)$ distribution are nearly the same. This demonstrates empirically that the null distribution of the GLR statistic can be satisfactorily approximated by the χ^2 -distribution.

For the power assessment, we evaluate the power for a sequence of alternative models indexed by θ ,

$$(12) \quad \begin{aligned} H_{1\theta} : a_1(u) &= (1 + \theta) \sin(6\pi u), \\ a_2(u) &= (1 + \theta) \cos(6\pi u), \quad \theta = 0.0, 0.1, 0.2, \dots, \end{aligned}$$

ranging from the null model to reasonably far away from it. The above functional forms were also used in Fan and Jiang [6], Fan and W. Zhang [8] and Ip *et al.* [15]. Figure 3 reports the differences between the null and the alternatives in (12).

For each given value of γ (0.05 or 0.10), the critical values of our test are calculated by the corresponding $\chi^2(r_K \mu_n)$, where r_K and μ_n are given by Theorem 1 with $h_4 = n^{-2/9}$ and $|\Omega| = 1$. The sample size is $n = 200$. The

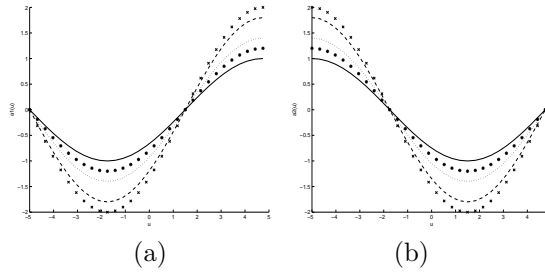


FIGURE 3. The difference between the null and the alternatives: (a) $a_1(u) = (1+\theta) \sin(6\pi u)$; (b) $a_1(u) = (1+\theta) \cos(6\pi u)$. ($- \theta = 0$; $* \theta = 0.2$; $\dots \theta = 0.4$; $- - \theta = 0.8$; $\times \theta = 1.0$).

parameter θ is related to the separation distance between the null and the alternative hypotheses. Note that when $\theta = 0$, the alternative is the same as the null hypothesis, so that the power should approximately be 0.05 (or 0.10) at the 0.05 (or 0.10) significance level γ . When θ increases, the alternative moves further away from the null hypothesis. One would expect the rejection rates of the null hypothesis to get higher and higher, which is evidenced in Table 1.

From Table 1 we see a surprisingly stable performance of the tests for different error distributions with the characteristics of heavy tail, symmetric and asymmetric densities. The numerical results here suggest that the GLR tests not only have high power for differentiating the null and the smoothing alternatives (12), but also have robustness against error distributions to some extent.

Table 1. Power of the GLR tests under different error distributions

α	Error distribution \ θ	0.0	0.4	0.8	1.2	1.4
0.05	N(0,1)	0.052	0.563	0.946	0.991	1.000
	t(5)	0.056	0.509	0.965	0.999	1.000
	$\chi^2(5)$	0.053	0.587	0.977	0.983	1.000
0.10	N(0,1)	0.096	0.784	0.996	0.999	1.000
	t(5)	0.092	0.606	0.976	0.999	1.000
	$\chi^2(5)$	0.099	0.677	0.970	0.982	1.000

To provide more information on the power, we also consider a simulation study for the behaviour of power as the sample size n changes. No loss of generalization and expositional purpose, here the error distribution $F(x)$ and the given significance level γ are taken as $N(0, 1)$ and 0.05, respectively. The simulation results are reported in Table 2, which show that with the sample size n increasing the power of the GLR tests for the alternative sequence in (12) becomes higher and higher. This in turn suggests that the proposed GLR test is very efficient.

Table 2. Power of the GLR tests under different sample sizes

$n \setminus \theta$	0.0	0.4	0.8	1.2	1.4
100	0.043	0.482	0.896	0.964	1.000
150	0.053	0.554	0.925	0.980	1.000
200	0.052	0.563	0.946	0.991	1.000
250	0.052	0.594	0.956	1.000	1.000
300	0.050	0.665	0.986	1.000	1.000

5. Real data example

We use the proposed GLR tests on an environmental data set. The data set used here consists of daily measurements of pollutants and other environmental factors in New Territories East in Hong Kong between January 1, 2000 and August 31, 2001. We consider the five pollutants and two environmental factors, i.e., sulphur dioxide (in g/m^3) X_1 , nitrogen dioxide (in g/m^3) X_2 , nitrogen oxide (in g/m^3) X_3 , respirable suspended particulate (in g/m^3) X_4 , and ozone (in g/m^3) X_5 ; temperature (in Celsius) Z_1 , relative humidity (in percent ratio) Z_2 . Wong *et al.* [17] studied the relationship between the levels of the above pollutants and the number of daily total hospital admissions (Y) for respiratory diseases in the data set. To avoid the curse of dimensionality, they analyzed the data set via the VCSIM,

$$(13) \quad Y = a_0(\alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4 + \alpha_5 X_5) + a_1(t)Z_1 + a_2(t)Z_2 + \varepsilon.$$

They suggested that the effect of two environmental factors (Z_1, Z_2) on the respiratory disease patients (Y) has a varying rate, and pointed out further that the rate is varying with the time t . The VCSIM performs fairly well in the environmental data set.

We now focus on the model diagnostic problems. Specifically we check whether the coefficient functions $a_1(\cdot), a_2(\cdot)$ in the model (13) are not varying with the time t in the environmental data set, namely whether the partially linear single-index model (PLSIM) is consistent with the data. The PLSIM has the form

$$(14) \quad Y = a_0(\alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4 + \alpha_5 X_5) + a_1 Z_1 + a_2 Z_2 + \varepsilon.$$

To answer this question, we now use our GLR statistic to test whether the null PLSIM (14) holds against the alternative VCSIM (13). To compute the p value of the test statistic, we need to find the null distribution of the GLR statistic λ_n . This can be estimated by the *conditional bootstrap method*, which was used in Fan *et al.* [6]. In this section we will use the bandwidths $\hat{\mathbf{h}}_{opt} = (0.1180, 0.1180, 0.2405, 0.2405)^T$, and take the Epanechnikov kernel. Our interest is to test whether the PLSIM (14) is adequate for the environmental data set. To see this, the P-value for the tests are summarized in Table 3, which provides stark evidence that the VCSIM (13) is appropriate for this environmental data set at the 0.01 significance level.

Table 3. P-value of the GLR tests under four different bandwidths

Bandwidth	$\frac{4}{5}\hat{\mathbf{h}}_{opt}$	$\hat{\mathbf{h}}_{opt}$	$\frac{5}{4}\hat{\mathbf{h}}_{opt}$	$2\hat{\mathbf{h}}_{opt}$
GLR statistic	32.8046	30.3271	23.8011	18.3422
P-value	0.0040	0.0030	0.0000	0.0080

Based on the above analysis, we now use our testing approach to check whether there is any covariate that can be deleted from the VCSIM (13). Firstly, to examine whether the temperature (Z_1) affects the number of daily total hospital admissions (Y) for respiratory diseases in the environmental data set, we consider the testing hypothesis $H_0 : Y = a_0(\alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4 + \alpha_5 X_5) + a_2(t)Z_2 + \varepsilon$ against $H_1 : Y = a_0(\alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4 + \alpha_5 X_5) + a_1(t)Z_1 + a_2(t)Z_2 + \varepsilon$. As a result, under the bandwidth $2\hat{\mathbf{h}}_{opt}$ the GLR statistic is $\lambda_n = 38.7325$ and the p value is close to 0, which strongly rejects the null hypothesis. Therefore, the variable temperature (Z_1) is significant at the significant level 0.01. Finally by the same token, we can show that the variable relative humidity (Z_2) plays a significant role in determining the respiratory disease patients (Y) as well.

Appendix

Lemma. *Suppose Conditions (A₁ – A₆) in Section 3.2 hold. Then uniformly in $u \in \Omega$ and $t \in L$,*

$$(15) \quad \begin{aligned} \hat{a}_0(t) - a_0(t) &= (R_0(t) + e_0(t))(1 + o_p(1)); \\ \hat{a}_j(u) - a_j(u) &= (R_j(u) + e_j(u))(1 + o_p(1)), \quad j = 1, \dots, p; \\ \bar{a}_0(t) - a_0(t) &= (R_0(t) + e_0(t))(1 + o_p(1)); \\ \hat{\alpha}_0 - \alpha_0 &= O_p(n^{-1/2}); \quad \bar{\alpha}_0 - \alpha_0 = O_p(n^{-1/2}), \end{aligned}$$

where $R_0(t) = \frac{1}{2}\mu_2(K_3)a_0''(t)h_3^2$, $e_0(t) = \frac{1}{n}\sum_{k=1}^n p_1^{-1}(t)K_{3,h_3}(\alpha_0^T \mathbf{X}_k - t)\varepsilon_k$, $R_j(u) = \frac{1}{2}\mu_2(K_4) \times a_j''(u)h_4^2$, $e_j(u) = s_j^{-1}(u)\frac{1}{n}\sum_{k=1}^n Z_{kj}K_{4,h_4}(U_k - u)\varepsilon_k$, $\mu_2(K_j) = \int t^2 K_j(t)$, $s_j(u) = E(Z_j^2|U = u)p_2(u)$.

The proof of Lemma can be completed using the methods similar to that used in Wong *et al.* [17].

Proof of Theorem 1. Note that, under H_0 ,

$$(16) \quad \begin{aligned} &\frac{1}{2}(RSS_0 - RSS_1) \\ &= \frac{1}{2}\sum_{i=1}^n \left(Y_i - a_0(\alpha_0^T \mathbf{X}_i) - \sum_{j=1}^p a_{j0}(U_i)Z_{ij} + a_0(\alpha_0^T \mathbf{X}_i) \right. \\ &\quad \left. + \sum_{j=1}^p a_{j0}(U_i)Z_{ij} - \hat{a}_0(\hat{\alpha}_0^T \mathbf{X}_i) - \sum_{j=1}^p \hat{a}_{j0}(U_i)Z_{ij} \right)^2 \end{aligned}$$

$$\begin{aligned}
& -\frac{1}{2} \sum_{i=1}^n \left(Y_i - \bar{a}_0(\bar{\alpha}_0^T \mathbf{X}_i) - \sum_{j=1}^p a_{j0}(U_i) Z_{ij} \right)^2 \\
& = \frac{1}{2} \sum_{i=1}^n \left(\varepsilon_i + a_0(\alpha_0^T \mathbf{X}_i) - \hat{a}_0(\alpha_0^T \mathbf{X}_i) + \hat{a}_0(\alpha_0^T \mathbf{X}_i) \right. \\
& \quad \left. - \hat{a}_0(\hat{\alpha}_0^T \mathbf{X}_i) + \sum_{j=1}^p (a_{j0}(U_i) - \hat{a}_{j0}(U_i)) Z_{ij} \right)^2 \\
& \quad - \frac{1}{2} \sum_{i=1}^n \left(\varepsilon_i + a_0(\alpha_0^T \mathbf{X}_i) - \bar{a}_0(\alpha_0^T \mathbf{X}_i) + \bar{a}_0(\alpha_0^T \mathbf{X}_i) - \bar{a}_0(\bar{\alpha}_0^T \mathbf{X}_i) \right)^2 \\
& = I_1 + I_2 - I_3 - I_4 + I_5 - I_6 - I_7 + I_8 + I_9 - I_{10} - I_{11} - I_{12} - I_{13} + I_{14},
\end{aligned}$$

where

$$\begin{aligned}
I_1 &= \frac{1}{2} \sum_{i=1}^n \left(a_0(\alpha_0^T \mathbf{X}_i) - \hat{a}_0(\alpha_0^T \mathbf{X}_i) \right)^2 \\
&= \frac{1}{2} \sum_{i=1}^n \left(R_0(\hat{\alpha}_0^T \mathbf{X}_i) + e_0(\hat{\alpha}_0^T \mathbf{X}_i) \right)^2 (1 + o_P(1)), \\
I_2 &= \frac{1}{2} \sum_{i=1}^n \left(\hat{a}_0(\alpha_0^T \mathbf{X}_i) - \hat{a}_0(\hat{\alpha}_0^T \mathbf{X}_i) \right)^2, \quad I_3 = \sum_{i=1}^n \left(\bar{a}_0(\alpha_0^T \mathbf{X}_i) - \bar{a}_0(\bar{\alpha}_0^T \mathbf{X}_i) \right) \varepsilon_i, \\
I_4 &= \frac{1}{2} \sum_{i=1}^n \left(\bar{a}_0(\alpha_0^T \mathbf{X}_i) - \bar{a}_0(\bar{\alpha}_0^T \mathbf{X}_i) \right)^2, \quad I_5 = \sum_{i=1}^n \left(\hat{a}_0(\alpha_0^T \mathbf{X}_i) - \hat{a}_0(\hat{\alpha}_0^T \mathbf{X}_i) \right) \varepsilon_i, \\
I_6 &= - \sum_{i=1}^n \left(\sum_{j=1}^p (a_{j0}(U_i) - \hat{a}_{j0}(U_i)) Z_{ij} \right) \varepsilon_i \\
&= \sum_{i=1}^n \left(\sum_{j=1}^p (R_j(U_i) + e_j(U_i)) Z_{ij} \right) \varepsilon_i (1 + o_P(1)), \\
I_7 &= \sum_{i=1}^n \left(a_0(\alpha_0^T \mathbf{X}_i) - \hat{a}_0(\alpha_0^T \mathbf{X}_i) \right) \left(\hat{a}_0(\alpha_0^T \mathbf{X}_i) - \hat{a}_0(\hat{\alpha}_0^T \mathbf{X}_i) \right) \\
&= \sum_{i=1}^n \left(R_0(\alpha_0^T \mathbf{X}_i) + e_0(\alpha_0^T \mathbf{X}_i) \right) \left(\hat{a}_0(\alpha_0^T \mathbf{X}_i) - \hat{a}_0(\hat{\alpha}_0^T \mathbf{X}_i) \right) (1 + o_P(1)), \\
I_8 &= \sum_{i=1}^n \left(a_0(\alpha_0^T \mathbf{X}_i) - \hat{a}_0(\alpha_0^T \mathbf{X}_i) \right) \left(\sum_{j=1}^p (a_{j0}(U_i) - \hat{a}_{j0}(U_i)) Z_{ij} \right) \\
&= \sum_{i=1}^n \left(R_0(\alpha_0^T \mathbf{X}_i) + e_0(\alpha_0^T \mathbf{X}_i) \right) \left(\sum_{j=1}^p (R_j(U_i) + e_j(U_i)) Z_{ij} \right) (1 + o_P(1)),
\end{aligned}$$

$$\begin{aligned}
I_9 &= \sum_{i=1}^n \left(\hat{a}_0(\alpha_0^T \mathbf{X}_i) - \hat{a}_0(\hat{\alpha}_0^T \mathbf{X}_i) \right) \left(\sum_{j=1}^p (a_{j0}(U_i) - \hat{a}_{j0}(U_i)) Z_{ij} \right) \\
&= \sum_{i=1}^n \left(\hat{a}_0(\alpha_0^T \mathbf{X}_i) - \hat{a}_0(\hat{\alpha}_0^T \mathbf{X}_i) \right) \left(\sum_{j=1}^p (R_j(U_i) + e_j(U_i)) Z_{ij} \right) (1 + o_P(1)), \\
I_{10} &= \frac{1}{2} \sum_{i=1}^n \left(a_0(\alpha_0^T \mathbf{X}_i) - \bar{a}_0(\alpha_0^T \mathbf{X}_i) \right)^2 \\
&= \frac{1}{2} \sum_{i=1}^n \left(R_0(\alpha_0^T \mathbf{X}_i) + e_0(\alpha_0^T \mathbf{X}_i) \right)^2 (1 + o_P(1)), \\
I_{11} &= \sum_{i=1}^n \left(a_0(\alpha_0^T \mathbf{X}_i) - \hat{a}_0(\alpha_0^T \mathbf{X}_i) \right) \varepsilon_i \\
&= \sum_{i=1}^n \left(R_0(\alpha_0^T \mathbf{X}_i) + e_0(\alpha_0^T \mathbf{X}_i) \right) \varepsilon_i (1 + o_P(1)), \\
I_{12} &= \sum_{i=1}^n \left(\bar{a}_0(\alpha_0^T \mathbf{X}_i) - a_0(\alpha_0^T \mathbf{X}_i) \right) \left(\bar{a}_0(\bar{\alpha}_0^T \mathbf{X}_i) - \bar{a}_0(\alpha_0^T \mathbf{X}_i) \right) \\
&= \sum_{i=1}^n \left(R_0(\alpha_0^T \mathbf{X}_i) + e_0(\alpha_0^T \mathbf{X}_i) \right) \left(\bar{a}_0(\bar{\alpha}_0^T \mathbf{X}_i) - \bar{a}_0(\alpha_0^T \mathbf{X}_i) \right) (1 + o_P(1)), \\
I_{13} &= \sum_{i=1}^n \left(a_0(\alpha_0^T \mathbf{X}_i) - \bar{a}_0(\alpha_0^T \mathbf{X}_i) \right) \varepsilon_i \\
&= \sum_{i=1}^n \left(R_0(\alpha_0^T \mathbf{X}_i) + e_0(\alpha_0^T \mathbf{X}_i) \right) \varepsilon_i (1 + o_P(1)), \\
I_{14} &= \frac{1}{2} \sum_{i=1}^n \left(\sum_{j=1}^p (a_{j0}(U_i) - \hat{a}_{j0}(U_i)) Z_{ij} \right)^2 \\
&= \frac{1}{2} \sum_{i=1}^n \left(\sum_{j=1}^p (R_j(U_i) + e_j(U_i)) Z_{ij} \right)^2 (1 + o_P(1)).
\end{aligned}$$

By using the same methods as in the proof of (A.2-A.5) in Zhang [19], we have

$$\begin{aligned}
(17) \quad I_1 - I_{10} &= o_p(1), \quad I_2 - I_4 = o_p(h_3^{-1/2}), \quad I_5 - I_3 = o_p(h_3^{-1/2}), \\
I_7 - I_{12} &= o_p(h_3^{-1/2}), \quad I_{11} - I_{13} = O_p\left(\frac{1}{\sqrt{nh_3}} + \sqrt{nh_3^2}\right).
\end{aligned}$$

Conjoining (16) with (17), we have

$$(18) \quad \frac{1}{2}(RSS_1 - RSS_0) = I_6 + I_8 + I_9 + I_{14} + o_p(h_3^{-1/2})$$

$$(19) \quad = \left(\sum_{k=1}^7 Q_k - Q_8 - Q_9 + \sum_{k=10}^{14} Q_k \right) (1 + o_p(1)) + o_p(h_3^{-1/2}),$$

where

$$\begin{aligned} Q_1 &= \frac{1}{2} \sum_{i=1}^n \left(\sum_{j=1}^p R_j^2(U_i) Z_{ij}^2 \right), \quad Q_2 = \frac{1}{2} \sum_{i=1}^n \left(\sum_{j=1}^p e_j^2(U_i) Z_{ij}^2 \right), \\ Q_3 &= \sum_{i=1}^n \left(\sum_{j=1}^p (R_j(U_i) e_j(U_i) Z_{ij}^2) \right), \quad Q_4 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^p \sum_{j' \neq j, j'=1}^p R_j(U_i) R_{j'}(U_i) Z_{ij} Z_{ij'}, \\ Q_5 &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^p \sum_{j' \neq j, j'=1}^p R_j(U_i) e_{j'}(U_i) Z_{ij} Z_{ij'}, \\ Q_6 &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^p \sum_{j' \neq j, j'=1}^p R_{j'}(U_i) e_j(U_i) Z_{ij} Z_{ij'}, \\ Q_7 &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^p \sum_{j' \neq j, j'=1}^p e_j(U_i) e_{j'}(U_i) Z_{ij} Z_{ij'}, \\ Q_8 &= \sum_{i=1}^n \left(\sum_{j=1}^p (R_j(U_i) Z_{ij}) \right) \varepsilon_i, \quad Q_9 = \sum_{i=1}^n \left(\sum_{j=1}^p (e_j(U_i) Z_{ij}) \right) \varepsilon_i, \\ Q_{10} &= \sum_{i=1}^n \left(\sum_{j=1}^p (R_0(\alpha_0^T \mathbf{X}_i) R_j(U_i) Z_{ij}) \right), \quad Q_{11} = \sum_{i=1}^n \left(\sum_{j=1}^p (R_0(\alpha_0^T \mathbf{X}_i) e_j(U_i) Z_{ij}) \right), \\ Q_{12} &= \sum_{i=1}^n \left(\sum_{j=1}^p (e_0(\alpha_0^T \mathbf{X}_i) R_j(U_i) Z_{ij}) \right), \quad Q_{13} = \sum_{i=1}^n \left(\sum_{j=1}^p (e_0(\alpha_0^T \mathbf{X}_i) e_j(U_i) Z_{ij}) \right), \\ Q_{14} &= \sum_{i=1}^n \left(\hat{a}_0(\alpha_0^T \mathbf{X}_i) - \hat{a}_0(\hat{\alpha}_0^T \mathbf{X}_i) \right) \left(\sum_{j=1}^p (R_j(U_i) + e_j(U_i)) Z_{ij} \right). \end{aligned}$$

Next we discuss $Q_k, k = 1, \dots, 14$, respectively.

(1) By using the Strong Law of Large Number Theory, we have

$$(20) \quad Q_8 = \sum_{i=1}^n \left(\sum_{j=1}^p (R_j(U_i) Z_{ij}) \right) \varepsilon_i = \sum_{j=1}^p \frac{h_4^2 \mu_2(K_4)}{2} \sum_{i=1}^n a''_{j0}(U_i) Z_{ij} \varepsilon_i = O_p(\sqrt{n} h_4^2).$$

Similar to Q_8 , we have

$$(21) \quad Q_1 = \frac{1}{2} \sum_{i=1}^n \left(\sum_{j=1}^p (R_j(U_i) Z_{ij}^2) \right) = \sum_{j=1}^p \frac{h_4^4 \mu_2^2(K_4)}{8} \sum_{i=1}^n (a''_{j0}(U_i))^2 Z_{ij}^2 = O_p(\sqrt{n} h_4^2),$$

$$\begin{aligned}
 Q_4 &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^p \sum_{j' \neq j, j'=1}^p R_j(U_i) R_{j'}(U_i) Z_{ij} Z_{ij'} \\
 (22) \quad &= \sum_{i=1}^n \sum_{j' \neq j, j'=1}^p \frac{h_4^4 \mu_2^2(K_4)}{8} \sum_{i=1}^n a''_{j_0}(U_i) a''_{j'_0}(U_i) Z_{ij} Z_{ij'} = O_p(\sqrt{nh_4^4}).
 \end{aligned}$$

Next we discuss Q_3 . Noting that uniformly in k

$$\frac{1}{n} \sum_{i=1}^n \frac{a''_{j_0}(U_i) Z_{ij}^2}{s_j(U_i)} K_{4,h_4}(U_i - U_k) = a''_{j_0}(U_k)(1 + o_p(1)).$$

Furthermore, by interchanging summation signs and the Strong Law of Large Number Theory, it can be shown

$$(23) \quad Q_3 = \sum_{j=1}^p \sum_{k=1}^n \frac{h_4^2 \mu_2(K_4)}{2} Z_{kj} \varepsilon_k \left(\frac{1}{n} \sum_{i=1}^n \frac{a''_{j_0}(U_i) Z_{ij}^2}{s_j(U_i)} K_{4,h_4}(U_i - U_k) \right) = O_p(\sqrt{nh_4^2}).$$

Similar to Q_3 , we have

$$(24) \quad Q_5 = O_p(\sqrt{nh_4^2}), \quad Q_6 = O_p(\sqrt{nh_4^2}).$$

(2) As to Q_7, Q_2, Q_9 , by using a similar method as in the proof of T_9, T_5 and T_2 in Ip *et al.* [15], we have

$$(25) \quad Q_7 = O_p\left(1 + (nh_4^{3/2})^{-1} + (nh_4^2)^{-1/2}\right),$$

$$\begin{aligned}
 Q_2 &= \sum_{j=1}^p \frac{1}{2n} \sum_{i,k=1, i \neq k}^n \frac{Z_{kj} Z_{ij} \varepsilon_k \varepsilon_i}{s_j(U_k) h_4} \int K_4(t) K_4\left(t - \frac{U_i - U_k}{h_4}\right) dt \\
 (26) \quad &+ \sum_{j=1}^p \frac{\sigma^2 |\Omega|}{2h_4} \int K_4^2(t) dt + O_p\left(1 + (nh_4^{3/2})^{-1} + (nh_4^2)^{-1} + (nh_4^2)^{-1/2}\right),
 \end{aligned}$$

$$\begin{aligned}
 Q_9 &= \sigma^2 \frac{p K_4(0) |\Omega|}{h_4} + O_p\left((nh_4^2)^{-1/2}\right) \\
 (27) \quad &+ \sum_{j=1}^p \frac{1}{n} \sum_{i=1}^n \sum_{k=1, k \neq i}^n \frac{Z_{kj} Z_{ij}}{s_j(U_i)} K_{4,h_4}(U_k - U_i) \varepsilon_k \varepsilon_i.
 \end{aligned}$$

(3) Now we consider $Q_{10}, Q_{11}, Q_{12}, Q_{13}$. By using the Strong Law of Large Number Theory, we have

$$(28) \quad Q_{10} = \sum_{j=1}^p \frac{\mu_2(K_3) \mu_2(K_4) (h_3 h_4)^2}{4} \sum_{i=1}^n a''_0(\alpha_0^T \mathbf{X}_i) a''_{j_0}(U_i) Z_{ij} = O_p\left(n(h_3 h_4)^2\right).$$

As to Q_{11} , using a similar method as in the proof of Q_3 , we have

$$\begin{aligned}
 (29) \quad Q_{11} &= \sum_{i=1}^n \sum_{j=1}^p \frac{\frac{1}{2}\mu_2(K_3)a_0''(\alpha_0^T \mathbf{X}_i)h_3^2}{s_j(U_i)} \frac{1}{n} \sum_{k=1}^n K_{4,h_4}(U_i - U_k) Z_{kj} Z_{ij} \varepsilon_k \\
 &= \sum_{j=1}^p \sum_{k=1}^n \frac{\mu_2(K_3)h_3^2}{2} Z_{kj} \varepsilon_k \frac{1}{n} \sum_{i=1}^n \frac{a_0''(\alpha_0^T \mathbf{X}_i) Z_{ij}}{s_j(U_i)} K_{4,h_4}(U_i - U_k) \\
 &= O_p(\sqrt{nh_3^2}).
 \end{aligned}$$

Similar to Q_{11} , we have

$$(30) \quad Q_{12} = O_p(\sqrt{nh_4^2}).$$

As to Q_{13} , we only discuss Q_{131} , one of the p terms in the sum of Q_{13} , i.e., $j = 1$. The other terms can be obtained by similar process to that for Q_{131} . Note that

$$\begin{aligned}
 (31) \quad Q_{131} &= \sum_{i=1}^n \frac{Z_{i1}}{n^2 s_1(U_i) p_1(\alpha_0^T \mathbf{X}_i)} \left(\sum_{k=1}^n K_{3,h_3}(\alpha_0^T \mathbf{X}_k - \alpha_0^T \mathbf{X}_i) \varepsilon_k \right) \\
 &\quad \times \left(\sum_{m=1}^n Z_{m1} K_{4,h_4}(U_m - U_i) \varepsilon_m \right) \\
 &= \frac{1}{n^2} \sum_{i=1}^n \sum_{k=1}^n \sum_{m=1}^n \frac{Z_{i1} Z_{m1}}{s_1(U_i) p_1(\alpha_0^T \mathbf{X}_i)} K_{3,h_3}(\alpha_0^T \mathbf{X}_k - \alpha_0^T \mathbf{X}_i) K_{4,h_4}(U_m - U_i) \varepsilon_m \varepsilon_k \\
 &= \frac{1}{n^2} \sum_{i=1}^n \sum_{k=1}^n \frac{Z_{i1} Z_{k1}}{s_1(U_i) p_1(\alpha_0^T \mathbf{X}_i)} K_{3,h_3}(\alpha_0^T \mathbf{X}_k - \alpha_0^T \mathbf{X}_i) K_{4,h_4}(U_k - U_i) \varepsilon_k^2 \\
 &\quad + \frac{1}{n^2} \sum_{i=1}^n \sum_{k=1}^n \sum_{m=1, m \neq k}^n \frac{Z_{i1} Z_{m1}}{s_1(U_i) p_1(\alpha_0^T \mathbf{X}_i)} K_{3,h_3}(\alpha_0^T \mathbf{X}_k - \alpha_0^T \mathbf{X}_i) \\
 &\quad \times K_{4,h_4}(U_m - U_i) \varepsilon_m \varepsilon_k \\
 &\equiv Q_{1311} + Q_{1312},
 \end{aligned}$$

and uniformly in k ,

$$\frac{1}{n} \sum_{i=1}^n \frac{Z_{i1}}{s_1(U_i) p_1(\alpha_0^T \mathbf{X}_i)} K_{3,h_3}(\alpha_0^T \mathbf{X}_k - \alpha_0^T \mathbf{X}_i) K_{4,h_4}(U_k - U_i) = O_p\left(1 + (nh_3 h_4)^{-1/2}\right),$$

we have, by interchanging summation signs and the Strong Law of Large Number Theory,

$$(32) \quad Q_{1311} = \frac{1}{n} \sum_{k=1}^n Z_{k1} \varepsilon_k^2 \left(\frac{1}{n} \sum_{i=1}^n \frac{Z_{i1}}{s_1(U_i) p_1(\alpha_0^T \mathbf{X}_i)} K_{3,h_3}(\alpha_0^T \mathbf{X}_k - \alpha_0^T \mathbf{X}_i) K_{4,h_4}(U_k - U_i) \right)$$

$$= O_p\left(1 + (nh_3h_4)^{-1/2}\right).$$

We may decompose Q_{1312} , it can be decomposed into three parts:

$$\begin{aligned} Q_{1312} &= \frac{1}{n^2} \sum_{i=1}^n \sum_{k=1, k \neq i}^n \frac{Z_{i1}Z_{k1}}{s_1(U_i)p_1(\alpha_0^T \mathbf{X}_i)} K_{3,h_3}(\alpha_0^T \mathbf{X}_k - \alpha_0^T \mathbf{X}_i) K_{4,h_4}(0) \varepsilon_i \varepsilon_k \\ &+ \frac{1}{n^2} \sum_{i=1}^n \sum_{m=1, m \neq i}^n \frac{Z_{i1}Z_{m1}}{s_1(U_i)p_1(\alpha_0^T \mathbf{X}_i)} K_{4,h_4}(U_m - U_i) K_{3,h_3}(0) \varepsilon_i \varepsilon_m \\ &+ \frac{1}{n^2} \sum_{i,m,k=1, i \neq m, i \neq k, m \neq k}^n \frac{Z_{i1}Z_{m1}}{s_1(U_i)p_1(\alpha_0^T \mathbf{X}_i)} K_{3,h_3}(\alpha_0^T \mathbf{X}_k - \alpha_0^T \mathbf{X}_i) \\ &\quad \times K_{4,h_4}(U_m - U_i) \varepsilon_m \varepsilon_k. \end{aligned} \tag{33}$$

$$\equiv Q_{13121} + Q_{13122} + Q_{13123}.$$

Obviously, using a similar method as in the proof of (3.54) in Fan and Gijbels [5], we have

$$\begin{aligned} Q_{13121} &= O_p\left((n^2h_3h_4^2)^{-1/2}\right), \quad Q_{13122} = O_p\left((n^2h_3^2h_4)^{-1/2}\right), \\ Q_{13123} &= O_p\left((nh_3h_4)^{-1/2}\right). \end{aligned}$$

Furthermore

$$(34) \quad Q_{1312} = O_p\left((n^2h_3h_4^2)^{-1/2} + (n^2h_3^2h_4)^{-1/2} + (nh_3h_4)^{-1/2}\right).$$

Conjoining (31),(32),(33) with (34), we have

$$(35) \quad Q_{131} = O_p\left(1 + (n^2h_3h_4^2)^{-1/2} + (n^2h_3^2h_4)^{-1/2} + (nh_3h_4)^{-1/2}\right).$$

Furthermore, we have

$$(36) \quad Q_{13} = O_p\left(1 + (n^2h_3h_4^2)^{-1/2} + (n^2h_3^2h_4)^{-1/2} + (nh_3h_4)^{-1/2}\right).$$

As to Q_{14} , using an argument similar to that for (A.3 – A.5) in Zhang [19] and direct computation, we have

$$(37) \quad Q_{14} = O_p(h_3^{-1/2}).$$

Conjoining (19)-(36) with (37), we have

$$\begin{aligned} \frac{1}{2}(RSS_0 - RSS_1) &= \frac{p\sigma^2|\Omega|}{h_4} \left(K_4(0) - \frac{1}{2} \int K_4^2(t) dt \right) \\ &+ \sum_{j=1}^p \frac{1}{n} \sum_{i,k=1, i \neq k}^n \frac{Z_{kj}Z_{ij}\varepsilon_k\varepsilon_i}{s_j(U_k)} \left(K_{4,h_4}(U_i - U_k) \right. \\ &\quad \left. - \frac{1}{2} K_{4,h_4} * K_{4,h_4}(U_i - U_k) \right) + O_p\left(nh_4^4 + \sqrt{nh_4^2}\right), \end{aligned} \tag{38}$$

$$= \sigma^2\mu_n + d_n + W_n,$$

where

$$W_n = \sum_{j=1}^p \frac{1}{n} \sum_{i,k=1, i \neq k}^n \frac{Z_{kj} Z_{ij} \varepsilon_k \varepsilon_i}{s_j(U_k)} \left(K_{4,h_4}(U_i - U_k) - \frac{1}{2} K_{4,h_4} * K_{4,h_4}(U_i - U_k) \right).$$

Write

$$w(i, k) = \frac{1}{n} \sum_{j=1}^p \frac{Z_{kj} Z_{ij} \varepsilon_k \varepsilon_i}{s_j(U_k)} \left(K_{4,h_4}(U_i - U_k) - \frac{1}{2} K_{4,h_4} * K_{4,h_4}(U_i - U_k) \right).$$

Then, W_n can be rewritten as

$$(39) \quad W_n = \sum_{i,k=1, i \neq k}^n w(i, k).$$

Obviously, $E(W_n) = 0$. Next we calculate the variance of W_n .

Noting that $\{\varepsilon_i\}_{i=1}^n$ are independent with $E(\varepsilon_i) = 0$, we have

$$\begin{aligned} \text{Var}(W_n) &= 2n(n-1)E(w(1, 2))^2 \\ &= 2\sigma^4 E \left(\sum_{j=1}^p \frac{Z_{2j} Z_{1j}}{s_j(U_2)} \left(K_{4,h_4}(U_2 - U_1) - \frac{1}{2} K_{4,h_4} * K_{4,h_4}(U_2 - U_1) \right) \right)^2 (1 + o(1)) \\ &\equiv (\Lambda_1 + \Lambda_2)(1 + o(1)), \end{aligned}$$

where

$$\begin{aligned} \Lambda_1 &= 2\sigma^4 E \left(\sum_{j=1}^p \frac{Z_{2j}^2 Z_{1j}^2}{s_j^2(U_2)} \left(K_{4,h_4}(U_2 - U_1) - \frac{1}{2} K_{4,h_4} * K_{4,h_4}(U_2 - U_1) \right)^2 \right), \\ \Lambda_2 &= 2\sigma^4 E \left(\sum_{j=1}^p \sum_{j'=1, j' \neq j}^p \frac{Z_{2j} Z_{1j} Z_{2j'} Z_{1j'}}{s_j(U_2) s_{j'}(U_2)} \left(K_{4,h_4}(U_2 - U_1) - \frac{1}{2} K_{4,h_4} * K_{4,h_4}(U_2 - U_1) \right)^2 \right), \end{aligned}$$

It is readily shown that $\Lambda_1 = \sigma^4 \sigma_n^2$, $\Lambda_2 = O(1)$. Therefore, we have $\text{Var}(W_n) = \sigma^4 \sigma_n^2 (1 + O(1))$, which implies that $W_n = O_p(h_4^{-1/2})$. Furthermore, by the definition of RSS_0 , we have, by using a similar argument as $Q_k, k = 2, 3, 8, 9$,

$$(40) \quad \begin{aligned} \frac{RSS_1}{n} &= \frac{RSS_0}{n} - \frac{2\sigma^2 \mu_n}{n} - \frac{W_n}{n} + O_P((n^{-1} h_4^4)^{1/2}) + O_p(h_4^4) \\ &= \sigma^2 (1 + o_p(1)). \end{aligned}$$

Therefore, by (7), (38) and (40), λ_n can be rewritten as

$$(41) \quad \begin{aligned} \lambda_n &= \mu_n - d_n + W_n/\sigma^2 + o_p(h_4^{-1/2}), \quad \text{i.e.,} \\ \lambda_n - \mu_n + d_n &= W_n/\sigma^2 + o_p(h_4^{-1/2}). \end{aligned}$$

As to the asymptotic distribution of W_n , using a similar argument as in Page 185 in Fan *et al.* [9] and applying Proposition 3.2 in de Jong [4], we get

$$\sigma^{-2} \sigma_n^{-1} W_n \xrightarrow{d} N(0, 1),$$

which implies that, by (41),

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

This completes the proof of Theorem 1. \square

Proof of Theorem 2. Note that, under H_{1n} , uniformly $u \in \Omega$

$$\hat{a}_j(u) - a_{j0}(u) - g_{jn}(u) = (R_j(u) + R_{jg}(u) + e_j(u))(1 + o_P(1)),$$

where $R_{jg}(u) = \frac{1}{2}\mu_2(K_4)g''_{jn}(u)h_4^2$, then, and by using the similar argument of Theorem 1, we can easily complete the rest of the proof, so we omit the details. \square

Proof of Theorem 3. The argument used here is similar to that for Theorem 8 of Fan *et al.* [9], here we only express its sketch. Under H_{1n} and under $h_4 = C_*n^{-2/9}$, we have

$$(42) \quad -\lambda_n\sigma^2 = -\mu_n\sigma^2 - W_n + d_{1n} - \sqrt{nE\left(\sum_{j=1}^p g_{jn}(U)Z_j\right)^2} + o_p(h_4^{-1/2}),$$

uniformly in $G_n \in \mathbb{G}_n$. Thus, by definition,

$$\begin{aligned} \beta(\gamma, G_n) &= P(\sigma_n^{-1}(-\lambda_n + \mu_n) \geq Z_\gamma) \\ &= P(\sigma_n^{-1}\sigma^{-2}(-W_n + d_n - d_{1n} - \sqrt{2d_{1n}}) \geq Z_\gamma) \\ &\equiv P_{1n} + P_{2n}, \end{aligned}$$

where

$$\begin{aligned} P_{1n} &= P\left(\sigma_n^{-1}\sigma^{-2}(-W_n) + n^{1/2}h_4^{5/2}b_{1n} + nh_4^{9/2}b_{2n} \right. \\ &\quad \left. - nh_4^{1/2}b_{3n} \geq Z_\gamma, |b_{1n}| \leq M, |b_{2n}| \leq M\right), \\ P_{2n} &= P\left(\sigma_n^{-1}\sigma^{-2}(-W_n) + n^{1/2}h_4^{5/2}b_{1n} + nh_4^{9/2}b_{2n} \right. \\ &\quad \left. - nh_4^{1/2}b_{3n} \geq Z_\gamma, |b_{1n}| \geq M, |b_{2n}| \geq M\right), \end{aligned}$$

and

$$\begin{aligned} b_{1n} &= (n^{1/2}h_4^{5/2}\sigma_n)^{-1}O_p(n^{1/2}h_4^2), \\ b_{2n} &= (nh_4^{9/2}\sigma_n)^{-1}O_p(nh_4^4), \\ b_{3n} &= \frac{1}{2}(h_4^{1/2}\sigma_n\sigma^2)^{-1}E\left(\sum_{j=1}^p g_{jn}(U)Z_j\right)^2(1 + o_p(1)). \end{aligned}$$

When $h_4 \leq C_0^{-1/2}n^{-1/4}$, we have $n^{1/2}h_4^{5/2} \geq C_0nh_4^{9/2}$, $n^{1/2}h_4^{5/2} \rightarrow 0$, $nh_4^{9/2} \rightarrow 0$. Thus, for $h_4 \rightarrow 0$ and $nh_4 \rightarrow \infty$, we can show that $\beta(\gamma, \rho) \rightarrow 0$ only when $nh_4^{1/2}\rho_n^2 \rightarrow \infty$. It implies that $\rho_n^2 = n^{-1}h_4^{-1/2}$ and the possible minimum value of ρ_n in this setting is $n^{-7/16}$. When $nh_4^4 \rightarrow \infty$, for any $\delta > 0$, we can

find a constant $M > 0$, such that $P_{2n} < \delta/2$, uniformly in $G_n \in \mathbb{G}_n$. Then $\beta(\gamma, \rho) \leq \delta/2 + P_{1n}$.

Note that $\sup_{\mathbb{G}_n(\rho)} P_{1n} \rightarrow 0$ only when

$$\Delta = nh_4^{9/2}M - nh_4^{9/2}\rho^2 \rightarrow -\infty.$$

Δ attains the minimum value $-\frac{8}{9}(9M)^{-1/8}n\rho^{9/4}$ at $h_4 = (\rho^2/(9M))^{1/4}$. Now simple algebra shows that in this setting the corresponding minimum value of ρ_n is $n^{-4/9}$ with $h_4 = C_*n^{-2/9}$ for some constant C_* . \square

Acknowledgements. Our sincere thanks are due to the editor and the referees for their valuable comments and helpful suggestions which have led to substantial improvements in the presentation of the manuscript.

References

- [1] Z. Cai, J. Fan, and Q. Yao, *Functional-coefficient regression models for nonlinear time series*, J. Amer. Statist. Assoc. **95** (2000), no. 451, 941–956.
- [2] R. J. Carroll, J. Fan, I. Gijbels, and M. P. Wand, *Generalized partially linear single-index models*, J. Amer. Statist. Assoc. **92** (1997), no. 438, 477–489.
- [3] R. Chen and R. S. Tsay, *Functional-coefficient autoregressive models*, J. Amer. Statist. Assoc. **88** (1993), no. 421, 298–308.
- [4] P. de Jong, *A central limit theorem for generalized quadratic forms*, Probab. Theory Related Fields **75** (1987), no. 2, 261–277.
- [5] J. Fan and I. Gijbels, *Local Polynomial Modelling and its Applications*, Monographs on Statistics and Applied Probability, 66. Chapman & Hall, London, 1996.
- [6] J. Fan and J. Jiang, *Nonparametric inferences for additive models*, J. Amer. Statist. Assoc. **100** (2005), no. 471, 890–907.
- [7] J. Fan and J. Zhang, *Sieve empirical likelihood ratio tests for nonparametric functions*, Ann. Statist. **32** (2004), no. 5, 1858–1907.
- [8] J. Fan and W. Zhang, *Generalised likelihood ratio tests for spectral density*, Biometrika **91** (2004), no. 1, 195–209.
- [9] J. Fan, C. Zhang, and J. Zhang, *Generalized likelihood ratio statistics and Wilks phenomenon*, Ann. Statist. **29** (2001), no. 1, 153–193.
- [10] J. Fan, Q. Yao, and Z. Cai, *Adaptive varying-coefficient linear models*, J. R. Stat. Soc. Ser. B Stat. Methodol. **65** (2003), no. 1, 57–80.
- [11] W. Härdle, P. Hall, and H. Ichimura, *Optimal smoothing in single-index models*, Ann. Statist. **21** (1993), no. 1, 157–178.
- [12] T. Hastie and R. Tibshirani, *Generalized Additive Models*, Monographs on Statistics and Applied Probability, 43. Chapman and Hall, Ltd., London, 1990.
- [13] ———, *Varying-coefficient models*, J. Roy. Statist. Soc. Ser. B **55** (1993), no. 4, 757–796.
- [14] D. R. Hoover, J. A. Rice, C. O. Wu, and L. P. Yang, *Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data*, Biometrika **85** (1998), no. 4, 809–822.
- [15] W. Ip, H. Wong, and R. Q. Zhang, *Generalized likelihood ratio test for varying-coefficient models with different smoothing variables*, Comput. Statist. Data Anal. **51** (2007), no. 9, 4543–4561.
- [16] O. V. Lepski and V. G. Spokoiny, *Minimax nonparametric hypothesis testing: the case of an inhomogeneous alternative*, Bernoulli **5** (1999), no. 2, 333–358.
- [17] H. Wong, W. Ip, and R. Q. Zhang, *Varying-coefficient single-index model*, Comput. Statist. Data Anal. **52** (2008), no. 3, 1458–1476.

- [18] Y. Xia, H. Tong, and W. K. Li, *On extended partially linear single-index models*, *Biometrika* **86** (1999), no. 4, 831–842.
- [19] R. Zhang, *Tests for nonparametric parts on partially linear single index models*, *Sci. China Ser. A* **50** (2007), no. 3, 439–449.
- [20] R. Q. Zhang and G. Y. Li, *Integrated estimation of functional-coefficient regression models with different smoothing variables*, *Comm. Statist. Theory Methods* **35** (2006), no. 4-6, 1093–1100.

ZHENSHENG HUANG
DEPARTMENT OF STATISTICS
EAST CHINA NORMAL UNIVERSITY
SHANGHAI, 200241, P. R. CHINA
E-mail address: stahzs@126.com

RIQUAN ZHANG
DEPARTMENT OF STATISTICS
EAST CHINA NORMAL UNIVERSITY
SHANGHAI, 200241, P. R. CHINA
AND
SHANXI DATONG UNIVERSITY
DATONG, SHANXI, 037009, P. R. CHINA
E-mail address: zhangriquan@163.com