Mathematics

Research article

Empirical likelihood based heteroscedasticity diagnostics for varying coefficient partially nonlinear models

Cuiping Wang[1,2]**, Xiaoshuang Zhou**[2,*] **and Peixin Zhao**[3]

[1] School of Mathematics and Statistics, Shandong University of Technology, Zibo 255022, China

[2] College of Mathematics and Big Data, Dezhou University, Dezhou 253023, China

[3] College of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing 400067, China

* **Correspondence:** Email: xiaoshuang1227@126.com.

Abstract: Heteroscedasticity diagnostics of error variance is essential before performing some statistical inference work. This paper is concerned with the statistical diagnostics for the varying coefficient partially nonlinear model. We propose a novel diagnostic approach for heteroscedasticity of error variance in the model by combining it with the empirical likelihood method. Under some mild conditions, the nonparametric version of the Wilks theorem is obtained. Furthermore, simulation studies and a real data analysis are implemented to evaluate the performances of our proposed approaches.

Keywords: empirical likelihood; heteroscedasticity diagnostics; hypothesis test; varying coefficient partially nonlinear model

Mathematics Subject Classification: 62G05, 62G20, 62H15

1. Introduction

As one of the most important semiparametric regression models, the varying coefficient partially linear model (VCPLM), is an integration of the varying-coefficient model and the classic linear model. It can be described as follows:

$$Y = X^T \alpha(U) + Z^T \beta + \varepsilon, \tag{1.1}$$

where $Y \in R$ represents the response variable, $X = (X_1, X_2, \cdots, X_q)^T$, $Z = (Z_1, Z_2, \cdots, Z_P)^T$ and $U \in R$ represent the associated covariates, ε is the model error, $\beta = (\beta_1, \cdots, \beta_P)^T$ is an unknown p-dimensional parameter vector, and $\alpha(\cdot) = (\alpha_1(\cdot), \cdots, \alpha_q(\cdot))^T$ is an unknown q-dimensional varying

coefficient function vector. It has been extensively researched due to the fact that model (1.1) has both the interpretability of parametric structure and the flexibility of nonparametric structure.

It is observed that the existence form of the unknown β is linear through the function $Z^T\beta$ in model (1.1). Actually, the absolute linear relationship may be inappropriate in practical applications. To further explore more accurate information between Y and some certain covariates, the VCPNLM, which was introduced by Li and Mei [1], inherits the following form

$$Y = g(Z, \beta) + X^T\alpha(U) + \varepsilon. \tag{1.2}$$

The difference between these two models is that model (1.2) extends $Z^T\beta$ in (1.1) to $g(Z, \beta)$, and $g(\cdot, \cdot)$ is set to be a pre-known function. It is notable that the dimensions of the parameter vector β and the covariate Z in $g(Z, \beta)$ are not necessarily consistent. Taking the generalized linear model as an example, we can write $\exp(a + Z^T\beta)$ as $g(Z, \tilde{\beta})$ with $g(\cdot) = \exp(\cdot)$ and $\tilde{\beta} = (a, \beta^T)^T$, where both a and β are parameters.

Compared to the VCPLM (1.1), VCPNLM (1.2) has a stronger adaptability. Therefore, it is a of great significance to conduct some statistical studies related to model (1.2). Fortunately, it has been extensively researched since its introduction. In the work of Li and Mei [1], they presented the profile nonlinear least square estimators for both the unknown β and $\alpha(\cdot)$. Zhou et al. [2] developed the construction of confidence regions for the unknown quantities by employing the empirical likelihood technique. Jiang et al. [3] put forward a robust estimation approach based on a novel loss function related to the exponential squared in the case where variables have measurement errors. Xiao and Chen [4] promoted the local bias-corrected empirical likelihood procedures to deal with the additive errors for the nonparametric component. Dai and Huang [5] were dedicated to treating the distorted measurement errors in both the response and the covariates. Qian and Huang [6] proposed the corrected profile least squares estimation procedure with measurement errors in the nonparametric part. For model (1.2) with data missing, Wang et al. [7] used the inverse probability weighted profile nonlinear least squares approach and the empirical likelihood technique to deal with the missing covariates. Xia et al. [8] developed the statistical inferences with missing responses. Furthermore, Xiao and Liang [9] performed a robust two-stage estimator method from the aspect of a modal regression. Xiao and Shi [10] studied the robust estimation for model (1.2) with a nonignorable missing response. Zhou and Zhao [11] studied model (1.2) under the framework of a quantile regression with the censored response variable and a missing censoring indicator.

It is noted that all the aforementioned studies were conducted under the assumption of the equal variances of $\varepsilon_i, i = 1, \cdots, n$, that is $E\varepsilon_i = 0$ and $Var\varepsilon_i = \sigma^2, i = 1, \cdots, n$. However, using procedures for homoscedastic models in the case of heteroscedastic errors may lead to the loss of efficiency. Therefore, it is crucial and meaningful to ensure the absence of heteroscedasticity before we perform some statistical inference work. The diagnosis of heteroscedasticity has received sufficient attention from many scholars. We can refer to [12–14] for partial linear models, and refer to [15–17] for VCPLMs with measurement errors. As we have seen, some statistical works based on the empirical likelihood technique ([18–20]) inherit many advantages. A significant advantage is that there is no need for a variance estimation. Recently, it has also been shown to work well in the issue of testing the underlying heteroscedastic errors. Readers can be referred to but not limited to [12–14]. To our knowledge, research related to the diagnosis of heteroscedasticity for VCPNLM (1.2) by means of the empirical likelihood technique has not yet emerged.

Taking all the above statements into account, in this paper, we plan to perform a diagnostic method for heteroscedasticity based on the empirical likelihood technique in model (1.2). Assuming that the variance of ε_i satisfies $Var\varepsilon_i = \sigma_i^2$, then the hypotheses testing problem can be defined as follows:

$$H_0 : \sigma_i^2 = \sigma^2, \quad VS \quad H_1 : \sigma_i^2 \neq \sigma^2, \tag{1.3}$$

where σ denotes an ordinary constant. We are concerned in constructing a test for heteroscedasticity by invoking the empirical likelihood technique, which does not specify the distribution of the errors. Under several regularity conditions, we attempt to derive the corresponding Wilk's theorem. Finally, we expect to verify the feasibility of our proposed method through some simulation studies.

The remainder is described as follows: the methodology and the main results of the empirical likelihood based diagnostics method are introduced in Section 2; some simulation studies are implemented to exhibit the finite sample performances of our proposed test statistics in Section 3; we use the Boston housing price data to illustrate our proposed method in Section 4; and the conclusions and ongoing works are presented in Section 5. The proofs of the main results are presented in the Appendix.

2. Methodology

Denote $\{(Y_i, Z_i, X_i, U_i), i = 1, \ldots, n\}$ as the i.i.d. copies of $\{(Y, Z, X, U)\}$; then, the individual form of the VCPNLM is as follows:

$$Y_i = g(Z_i, \beta) + X_i^T \alpha(U_i) + \varepsilon_i, \quad i = 1, \ldots, n. \tag{2.1}$$

Suppose that β is known beforehand; then, (2.1) can be reexpressed as the following varying coefficient model:

$$Y_i - g(Z_i, \beta) = \sum_{j=1}^{q} \alpha_j(U_i) X_{ij} + \varepsilon_i, \quad i = 1, \cdots, n. \tag{2.2}$$

First, we employ the classic local linear smooth technique to derive the estimator of $\{\alpha_j(\cdot), j = 1, \cdots, q\}$ in model (2.2). Based on Taylor's expansion and for u in a small neighborhood of u_0, $\alpha_j(u)$ can be locally approximated via the following linear form:

$$\alpha_j(u) \approx \alpha_j(u_0) + \alpha'_j(u_0)(u - u_0), \quad j = 1, \cdots, q. \tag{2.3}$$

According to (2.3), the estimator of $\{(\alpha_j(u_0), \alpha'_j(u_0)), j = 1, \ldots, q\}$ can be deduced by minimizing the weighted local least-squares problem as follows:

$$\sum_{i=1}^{n} \{Y_i - g(Z_i, \beta) - \sum_{j=1}^{q} [\alpha_j(u_0) + \alpha'_j(u_0)(U_i - u_0)] X_{ij}\}^2 K_h(U_i - u_0), \tag{2.4}$$

where $K(\cdot)$ is the kernel function, $K_h(\cdot)$ has the form $K(\cdot/h)/h$, and h is the bandwidth which can be determined by some usual methods.

We introduce the following matrix notations for simplicity in description. Let

$$\mathbf{Y} = (Y_1, \cdots, Y_n)^T, \quad \mathbf{M} = (X_1^T \alpha(U_1), \cdots, X_n^T \alpha(U_n))^T, \quad \mathbf{g}(Z, \beta) = (g(Z_1, \beta), \cdots, g(Z_n, \beta))^T,$$

$$\mathbf{K}(u_0) = \text{diag}(K_h(U_1 - u_0), \ldots, K_h(U_n - u_0)), \quad \mathbf{H}(u_0) = (\alpha(u_0)^T, h\alpha'(u_0)^T)^T,$$

and

$$\mathbf{X}(u_0) = \begin{pmatrix} X_1^T & h^{-1}(U_1 - u_0)X_1^T \\ \vdots & \vdots \\ X_n^T & h^{-1}(U_n - u_0)X_n^T \end{pmatrix}.$$

Under the above matrix representations, the estimator of $\mathbf{H}(u_0)$ can be derived by the following:

$$\hat{\mathbf{H}}(u_0) = [\mathbf{X}(u_0)^T \mathbf{K}(u_0) \mathbf{X}(u_0)]^{-1} \mathbf{X}(u_0)^T \mathbf{K}(u_0)[\mathbf{Y} - \mathbf{g}(Z, \beta)]. \tag{2.5}$$

Then, the estimator of $\alpha(\cdot)$ at u_0 can be obtained by taking only the first part, that is,

$$\hat{\alpha}(u_0) = (\mathbf{I}_q, \mathbf{0}_q)[\mathbf{X}(u_0)^T \mathbf{K}(u_0) \mathbf{X}(u_0)]^{-1} \mathbf{X}(u_0)^T \mathbf{K}(u_0)[\mathbf{Y} - \mathbf{g}(Z, \beta)], \tag{2.6}$$

where $\mathbf{I}_q$ represents the $q \times q$ identity matrix, and $\mathbf{0}_q$ represents the $q \times q$ matrix for all entries 0.

For the purpose of testing the heteroscedasticity in model (2.1), first we rewrite the expression of the error variance as follows:

$$\text{Var}\,(\varepsilon_i) = \sigma^2 m_i,$$

where $m_i > 0$. Similar to arguments in Liu et al. [16], we presume that m_i possesses the subsequent sturcture:

$$m_i = m(U_i, \gamma). \tag{2.7}$$

Here, m_i is supposed to rely on the known covariate U_i and an unknown $q \times 1$ vector γ. It is remarkable that the structure of the function $m(\cdot, \cdot)$ is usually known in advance. In addition, we assume that $m(\cdot, \cdot)$ is differentiable with respect to γ and there exists a unique γ^* that satisfies $m(U_i, \gamma^*) = 1$ for all U_i. Thus, (1.3) is converted to the following hypothesis problem:

$$H_0 : \gamma = \gamma^*, \text{VS} \quad H_1 : \gamma \neq \gamma^*. \tag{2.8}$$

In order to utilize the empirical likelihood technique, we first consider the following estimation function:

$$\begin{cases} h_{1i} = \eta_i[(Y_i - X_i^T \alpha(U_i) - g(Z_i, \beta))^2 - \sigma^2], \\ h_{2i} = \tilde{g}'(Z_i, \beta)[Y_i - X_i^T \alpha(U_i) - g(Z_i, \beta)]. \end{cases} \tag{2.9}$$

Denote $\eta_i = (\dot{m}_i^T, 1)^T$, and $\dot{m}_i$ represents the derivative of m_i with respect to δ under the null hypothesis H_0. Write $h_i = (h_{1i}^T, h_{2i}^T)^T \in R^{p+q+1}$; then, we can easily know that $E(h_i) = 0$ under H_0. Intuitively, the above heteroscedasticity test problem is converted to testing whether $E(h_i) = 0$. This can be completed by means of the empirical likelihood technique.

Denote $p_1, p_2, \ldots, p_n$ be some nonnegative numbers whose sum is 1, that is $\sum_{i=1}^n p_i = 1$. Under the null hypothesis H_0, we can construct the profile empirical likelihood ratio for γ, σ^2, β as follows:

$$R_0(\gamma; \sigma^2, \beta) = \max \left\{ \prod_{i=1}^n n p_i \mid \sum_{i=0}^n p_i h_i = 0, p_i \geq 0, \sum_{i=0}^n p_i = 1 \right\}. \tag{2.10}$$

Here, β and σ^2 are nuisance parameters. It is noteworthy that $R_0(\gamma; \sigma^2, \beta)$ cannot be used to construct a test directly, for it contains the unknown γ, β, σ^2, and $\alpha(\cdot)$. A measure to deal with this issue is to substitute $\hat{\alpha}(\cdot)$ for $\alpha(\cdot)$. Therefore, we denote $\hat{h}_{1i}$ and $\hat{h}_{2i}$ as follows:

$$\begin{cases} \hat{h}_{1i} = \eta_i[(Y_i - X_i^T \hat{\alpha}(U_i) - g(Z_i, \beta))^2 - \sigma^2], \\ \hat{h}_{2i} = \tilde{g}'(Z_i, \beta)[Y_i - X_i^T \hat{\alpha}(U_i) - g(Z_i, \beta)]. \end{cases} \tag{2.11}$$

Denote $\hat{h}_i = (\hat{h}_{1i}^T, \hat{h}_{2i}^T)^T$. Naturally, the estimated profile empirical likelihood ratio is expressed by the following:

$$R(\gamma; \sigma^2, \beta) = \max\{\prod_{i=1}^{n} np_i \mid \sum_{i=0}^{n} p_i\hat{h}_i = 0, p_i \geq 0, \sum_{i=0}^{n} p_i = 1\}. \tag{2.12}$$

Combining the method of the Lagrange multiplier, we can obtain the optimal value of p_i as follows:

$$p_i = \frac{1}{n(1 + \lambda^T\hat{h}_i)}, \tag{2.13}$$

and λ satisfies the following equation:

$$\frac{1}{n}\sum_{i=1}^{n} \frac{\hat{h}_i}{1 + \lambda^T\hat{h}_i} = 0. \tag{2.14}$$

Substituting (2.13) into (2.12), we have the following:

$$-2\log R\left(\gamma; \sigma^2, \beta\right) = 2\sum_{i=1}^{n} \log\left(1 + \lambda^T\hat{h}_i\right). \tag{2.15}$$

To establish the nonparametric Wilk's theorem for $-2\log R(\gamma; \sigma^2, \beta)$, the following regularity conditions C1–C7 are needed with references to Zhou et al. [2], and the condition C8 is needed for the proof.

C1 The density function $f(u)$ of U is Lipschitz continuous and has bounded away from zero on its bounded support $\mathbb{U}$.

C2 $\Gamma(u)$ is a $q \times q$ nonsingular matrix for u in the support. Both $\Gamma(u), \Gamma(u)^{-1}$, and $\Phi(u)$ are Lipschitz continuous.

C3 There exists an $s > 2$ that satisfies $E\|X\|^{2s} < \infty$, $E\|g'(Z,\beta)\|^{2s} < \infty$, $E\|\varepsilon\|^{2s} < \infty$, and $E\|U\|^{2s} < \infty$, where $\|\cdot\|$ denotes the Euclidean norm. Meanwhile, for some $0 < \delta < 2 - s^{-1}$, $n^{2\delta-1}h \to \infty$ holds.

C4 $\{\alpha_j''(u), j = 1, \ldots, q\}$ with respect to u is continuous in $U \in \Omega$.

C5 The Kernel function $K(\cdot)$ is a univariate symmetric density function that satisfies the Lipschitz condition. The functions $u^3K(u)$ and $u^3K'(u)$ are bounded and $\int u^4K(u)du \leq \infty$.

C6 $nh^8 \to 0$ and $nh^2/(\log n)^2 \to \infty$ hold.

C7 $g(z,\beta)$ is continuous with respect to β for any z, and $g''(z,\beta)$ with respect to β are all continuous, where $\beta \in \mathcal{B}$ and $\mathcal{B}$ is a compact set.

C8

$$\frac{1}{n}\begin{pmatrix} \sum_{i=1}^{n} \eta_i\eta_i^{\tau} & \sum_{i=1}^{n} \eta_ig'(Z_i,\beta) \\ \sum_{i=1}^{n} \eta_ig'(Z_i,\beta)^T & \sum_{i=1}^{n} g'(Z_i,\beta)g'(Z_i,\beta)^T \end{pmatrix} \to \Sigma \triangleq \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

and

$$B_{11} > \mathbf{0}, \quad B_{22} > \mathbf{0}.$$

The following Theorem 1 describes the asymptotic behavior of $-2\log R(\gamma; \sigma^2, \beta)$.

Theorem 1. *Suppose that Conditions C1–C8 hold. Under the null hypothesis, we have the following:*

$$-2\log R\left(\gamma; \sigma^2, \beta\right) \xrightarrow{\mathcal{L}} \chi^2_{p+q+1},$$

where "$\xrightarrow{\mathcal{L}}$" represents the convergence in distribution and χ^2_{p+q+1} represents the chi-square distribution with $p + q + 1$ degrees of freedom.

To deal with the so-called nuisance parameters β and σ^2, under the null hypothesis H_0, we address

$$R(\gamma) \triangleq \max_{\sigma^2, \beta} R\left(\gamma; \sigma^2, \beta\right), \tag{2.16}$$

that is, maximizing (2.16) with respect to β and σ^2. Then, $R(\gamma)$ has the following asymptotic result:

$$R(\gamma) \xrightarrow{\mathcal{L}} \chi_q^2.$$

3. Simulation study

In this section, we assess the finite sample performances of our proposed work by some simulation studies. Let the data be generated from the following VCPNLM:

$$Y = g(Z, \beta) + X\alpha(U) + \varepsilon. \tag{3.1}$$

Specifically, $g(Z, \beta) = \exp\{Z\beta\}$ with $\beta = 2$ and $\alpha(U) = \sin(2\pi U)$. The model error ε is supposed to come from the normal distribution (Case 1) and the uniform distribution (Case 2), respectively, with $E(\varepsilon|X, Z, U) = 0$ and $Var(\varepsilon|X, Z, U) = \sigma^2 m(U, \gamma)$, where $\sigma^2 = 1$, $m(U, \gamma) = \exp(\gamma U)$. Obviously, $\gamma = 0$ corresponds to the null hypothesis, and $\gamma \neq 0$ corresponds to the alternative hypothesis. Moreover, in the following simulation, the covariates are generated on the base of $Z \sim N(0, 1)$, $X \sim N(0, 1)$ and $U \sim U(0, 1)$, and naturally Y is generated according to the model (3.1). Throughout the simulation studies, we select the Epanechnikov kernel $K(u) = \frac{3}{4}(1 - u^2)_+$ as the kernel function in our simulation, and the bandwidth h is taken as $h = cn^{-1/5}$, where the constant c is chosen as the standard deviation of the covariate U.

To evaluate the performance of the proposed method, the sample size in our simulation is taken as $n = 200, 400$, and 600, respectively, and the nominal level is 0.05. For each situation, we repeat 1000 simulation replications. With these replications, the power of the proposed empirical likelihood ratio test is displayed in Table 1 and Figure 1. Then, we can make the following observations:

(i) The power declines rapidly when the null hypothesis holds, and it converges to the correct nominal level when the sample increases. This result declares that our proposed testing method can control the probability of making the Type I error.

(ii) For any given n, the simulation performances under different error distribution cases are very similar. This result also indicates that our proposed test method is efficient for different model errors.

(iii) The power quickly tends to 1 when the sample increases and when the alternative hypothesis holds. In this respect, we can also demonstrate that our proposed heteroscedasticity test for the VCPNLM is effective.

Next, we compare the proposed empirical likelihood ratio test method with the profile likelihood ratio (PLR) test method used in [21]. In this simulation, the nominal level is taken as $\alpha = 0.05$, the sample size is taken $n = 400$, and the experiments are repeated 1000 times for each case. The simulation results of the 1000 replicates are shown in Figure 2, where the dashed line is the empirical power function based on the empirical likelihood ratio (ELR) test method proposed by this paper, and the dotted line is the empirical power function based on the PLR method.

From Figure 2, we can see that the empirical power functions obtained by the ELR test method and the PLR test method both rapidly increase as the value of γ increases. In addition, the empirical power derived with the ELR test method, is superior to that obtained by the PLR test method.

Table 1. The power of empirical likelihood ratio test under different model errors.

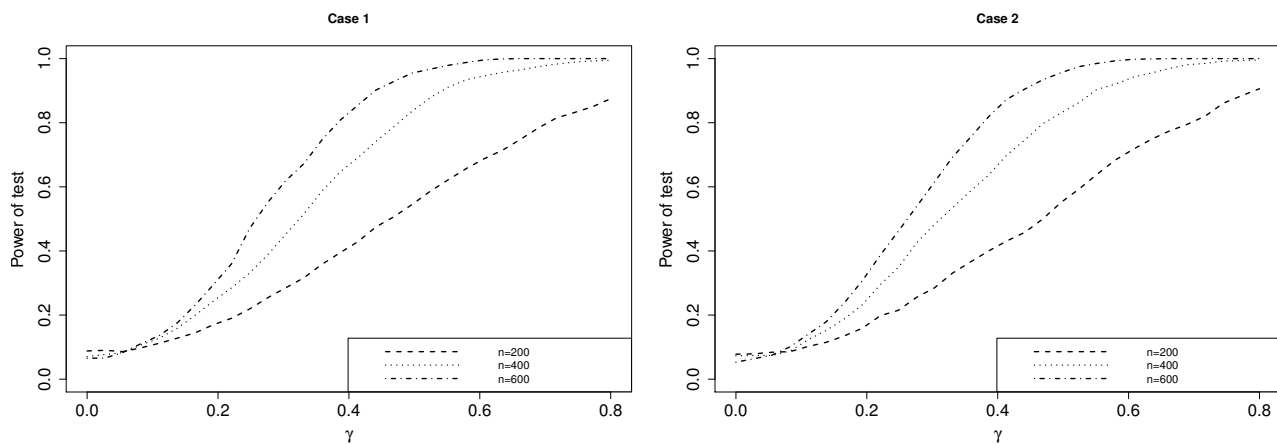| Model error	n	$\gamma = 0$	$\gamma = 0.2$	$\gamma = 0.4$	$\gamma = 0.6$	$\gamma = 0.8$
Case 1	200	0.099	0.166	0.404	0.689	0.893
	400	0.074	0.258	0.647	0.944	0.994
	600	0.051	0.320	0.859	0.992	1.000
Case 2	200	0.091	0.156	0.433	0.712	0.906
	400	0.079	0.250	0.683	0.946	0.998
	600	0.050	0.362	0.842	0.993	1.000

Figure 1. The power curves of empirical likelihood ratio test under model errors.
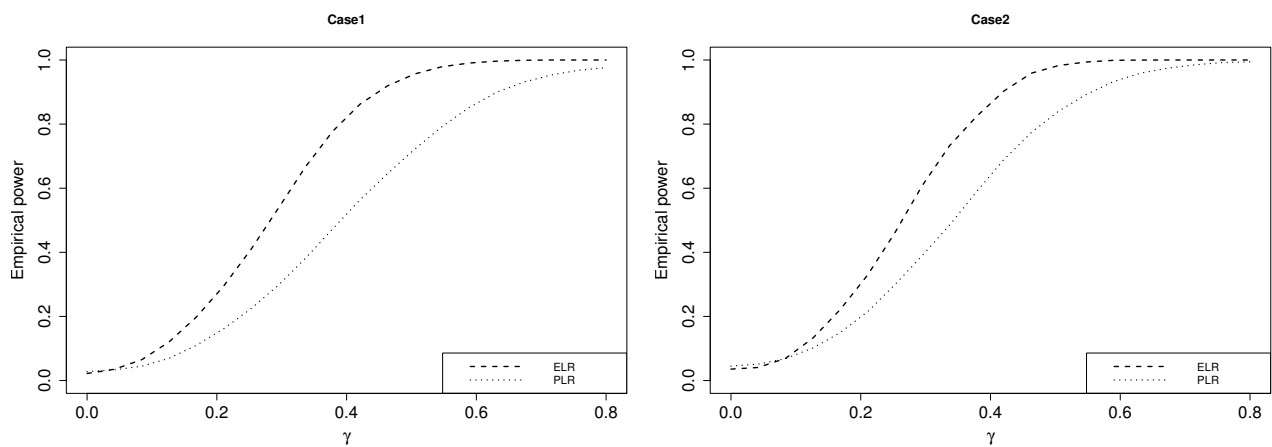

Figure 2. The testing power of the ELR and PLR methods under $n = 400$.

4. Application to Boston housing price data

In this section, we analyse the Boston housing price data to illustrate the model testing procedure proposed by this paper. The data set contains information of 506 different houses from a variety of locations in Boston Standard Metropolitan Statistical Area in 1970. Many researchers have analyzed this data set by using the partially linear additive model, the partially linear additive spatial autoregressive model, the partially linear single-index model, and other semiparametric models (see in [22–24]). The objective of these studies is to evaluate the influencing factors of the price of owner-occupied homes such as the the per capita crime rate by town, the weighted distances to five Boston employment centres, the average number of rooms per dwelling, and other factors. Hence, for the purpose of our demonstration, we take the indexes as the pupil-teacher ratio by town (denoted by PTRATIO), the index of accessibility to radial highways(denoted by RAD), the percentage of lower status of the population (denoted by LSTAT), the per capita crime rate by town (denoted by CRIME), and the median value of owner-occupied homes in USD 1000's (denoted by MEDV).

In addition, [25] pointed that the covariate CRIME has a nonlinear effect on the response. Hence, we fit this data set by using the following model:

$$Y_i = \exp(Z_{i1}\beta_1 + Z_{i2}\beta_2) + X_i\alpha(U_i) + \varepsilon_i, \quad i = 1, \cdots, 506,$$

where Y_i is the response MEDV, U_i is the covariate CRIME, X_i is the covariate log(LSTAT), and Z_{i1} and Z_{i2} are covariates RAD and PTRATIO, respectively. The logarithmic transformation for the covariate LSTAT is taken to ease off the trouble caused by big gaps in the domain.

Here, we consider the null hypothesises $H_0 : Var(\varepsilon_i|Z_{i1}, Z_{i2}, X_i, U_i) \equiv \sigma^2$. By using the ELR testing procedure proposed by this paper, we find that the p-value of this testing problem is 0.3484. This means that the null hypothesises can not be rejected under the nominal level 0.05, which also implies that the model error ε does not have significant effect on the covariates.

5. Conclusions

In this paper, we were concerned with the statistical inferences for the VCPNLM. Combining the empirical likelihood method, we proposed a diagnostic technique for heteroscedasticity in the semiparametric varying-coefficient partially nonlinear models. Under some mild conditions, the nonparametric version of Wilks theorem was derived and proven. Furthermore, simulation studies were performed to illustrate the performances of our proposed methods. As we have known, missing data is common in many fields. Ignoring the missing data will result in the reduction of effective information. Therefore, our forthcoming work is to implement the the statistical inferences for the VCPNLM in the case of missing data.

Author contributions

Cuiping Wang: Writing-original draft, Conceptualization, Formal analysis, Methodology; Xiaoshuang Zhou: Funding acquisition, Validation and data analysis, Methodology, Supervision, Writing-review and editing; Peixin Zhao: Funding acquisition, Software, Supervision, Writing-review and editing. All authors have read and approved the final version of the manuscript for publication

Use of Generative-AI tools declaration

The authors declares they have used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

Conflict of interest

The author declares no conflicts of interest in this paper.

References

1. T. Li, C. Mei, Estimation and inference for varying coefficient partially nonlinear models, *J. Stat. Plan. Infer.*, **143** (2013), 2023–2037. http://dx.doi.org/10.1016/j.jspi.2013.05.011

2. X. Zhou, P. Zhao, X. Wang, Empirical likelihood inferences for varying coefficient partially nonlinear models, *J. Appl. Stat.*, **44** (2017), 474–492. http://dx.doi.org/10.1080/02664763.2016.1177496

3. Y. Jiang, Q. Ji, B. Xie, Robust estimation for the varying coefficient partially nonlinear models, *J. Comput. Appl. Math.*, **326** (2017), 31–43. http://dx.doi.org/10.1016/j.cam.2017.04.028

4. Y. Xiao, Z. Chen, Bias-corrected estimations in varying-coefficient partially nonlinear models with measurement error in the nonparametric part, *J. Appl. Stat.*, **45** (2018), 586–603. http://dx.doi.org/10.1080/02664763.2017.1288201

5. S. Dai, Z. Huang, Estimation for varying coefficient partially nonlinear models with distorted measurement errors, *J. Korean Stat. Soc.*, **48** (2019), 117–133. http://dx.doi.org/10.1016/j.jkss.2018.09.001

6. Y. Qian, Z. Huang, Statistical inference for a varying-coefficient partially nonlinear model with measurement errors, *Stat. Methodol.*, **32** (2016), 122130. http://dx.doi.org/10.1016/j.stamet.2016.05.004

7. X. Wang, P. Zhao, H. Du, Statistical inferences for varying coefficient partially nonlinear model with missing covariates, *Commun. Stat.-Theor. M.*, **50** (2021), 2599–2618. http://dx.doi.org/10.1080/03610926.2019.1674870

8. L. Xia, X. Wang, P. Zhao, Y. Song, Empirical likelihood for varying coefficient partially nonlinear model with missing responses, *AIMS Mathematics*, **6** (2021), 7125–7152. http://dx.doi.org/10.3934/math.2021418

9. Y. Xiao, L. Liang, Robust estimation and variable selection for varying-coefficient partially nonlinear models based on modal regression, *J. Korean Stat. Soc.*, **51** (2020), 692–715. http://dx.doi.org/10.1007/s42952-021-00158-w

10. Y. Xiao, Y. Shi, Robust estimation for varying-coefficient partially nonlinear model with nonignorable missing response, *AIMS Mathematics*, **8** (2023), 29849–29871. http://dx.doi.org/10.3934/math.20231526

11. X. Zhou, P. Zhao, Estimation and inferences for varying coefficient partially nonlinear quantile models with censoring indicators missing at random, *Comput. Stat.*, **37** (2022), 1727–1750. http://dx.doi.org/10.1007/s00180-021-01192-2

12. H. Wong, F. Liu, M. Chen, W. Ip, Empirical likelihood based diagnostics for heteroscedasticity in partial linear models, *Comput. Stat. Data Anal.*, **53** (2009), 3466–3477. http://dx.doi.org/10.1016/j.csda.2009.02.029

13. H. Wong, F. Liu, M. Chen, W. Ip, Empirical likelihood based diagnostics for heteroscedasticity in partially linear errors-in-variables models, *J. Stat. Plan. Infer.*, **139** (2009), 916–929. http://dx.doi.org/10.1016/j.jspi.2008.05.049

14. G. Fan, H. Liang, J. Wang, Empirical likelihood for a heteroscedastic partial linear errors-in-variables model, *Commun. Stat.-Theor. M.*, **41** (2012), 108–127. http://dx.doi.org/10.1080/03610926.2010.517357

15. J. Lin, Y. Zhao, H. Wang, Heteroscedasticity diagnostics in varying-coefficient partially linear regression models and applications in analyzing Boston housing data, *J. Applied Statistics*, **42** (2015), 2432–2448. http://dx.doi.org/10.1080/02664763.2015.1043623

16. F. Liu, C. Li, P. Wang, X. Kang, Empirical likelihood based diagnostics for heteroscedasticity in semiparametric varying-coefficient partially linear errors-in-variables models, *Commun. Stat.-Theor. M.*, **47** (2018), 5485–5496. http://dx.doi.org/10.1080/03610926.2017.1395050

17. F. Liu, W. Gao, J. He, X. Fu, X. Kang, Empirical likelihood based diagnostics for heteroscedasticity in semiparametric varying-coefficient partially linear models with missing responses, *J. Syst. Sci. Complex.*, **34** (2021), 1175–1188. http://dx.doi.org/10.1007/s11424-020-9240-7

18. A. Owen, Empirical likelihood ratio confidence intervals for single functional, *Biometrika*, **75** (1988), 237–249. http://dx.doi.org/10.1093/biomet/75.2.237

19. A. Owen, Empirical likelihood ratio confidence regions, *Ann. Stat.*, **18** (1990), 90–120. http://dx.doi.org/10.1214/aos/1176347494

20. A. Owen, Empirical likelihood for linear models, *Ann. Stat.*, **19** (1991), 1725–1747. http://dx.doi.org/10.1214/aos/1176348368

21. J. Fan, T. Huang, Profile likelihood inferences on semiparametric varying-coefficient partially linear models, *Bernoulli*, **11** (2005), 1031–1057. http://dx.doi.org/10.3150/bj/1137421639

22. Y. Sun, H. Yan, W. Zhang, Z. Lu, A semiparametric spatial dynamic model, *Ann. Stat.*, **42** (2014), 700–727. http://dx.doi.org/10.1214/13-AOS1201

23. J. Du, X. Sun, R. Cao, Z. Zhang, Statistical inference for partially linear additive spatial autoregressive models, *Spat. Stat.*, **25** (2018), 52–67. http://dx.doi.org/10.1016/j.spasta.2018.04.008

24. S. Cheng, J. Chen, Estimation of partially linear single-index spatialautoregressive model, *Stat. Papers*, **62** (2021), 495–531. http://dx.doi.org/10.1007/s00362-019-01105-y

25. H. Zhang, G. Cheng, Y. Liu, Linear or nonlinear? Automatic structure discovery for partially linear models, *J. Am. Stat. Assoc.*, **106** (2011), 1099–1112. http://dx.doi.org/10.1198/jasa.2011.tm10281

26. S. Wang, M. Wu, Z. Jia, *Inequalities in matrix theory (Chinese)*, 2 Eds., Beijing: Science Press, 2006.

27. J. Gao, Asymptotic theory for partly linear models, *Commun. Stat.-Theor. M.*, **24** (1995), 1985–2009. http://dx.doi.org/10.1080/03610929508831598

Appendix

Several Lemmas are needed to prove the main result.

Lemma 1. *Assuming that Conditions C1–C8 hold. Then we get the following conclusions:*

$$\max_{1 \le j \le q} \sup_{u \in U} \left| \hat{\alpha}_j(u) - \alpha_j(u) \right| = O(d_n) \quad a.s.,$$

and $d_n = h^2 + (\log n / nh)^{1/2}$. If $h = dn^{-1/5}$ with a constant d, then we have

$$\max_{1 \le j \le q} \sup_{u \in U} \left| \hat{\alpha}_j(u) - \alpha_j(u) \right| = O\left(n^{-2/5 (\log n)^{1/2}} \right) \quad a.s.$$

Proof. The proof can be derived in [2]. □

Lemma 2. $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ *is a real symmetric matrix, if $B_{22} > 0$, write $B_{11.2} \triangleq B_{11} - B_{12} B_{22}^{-1} B_{21}$, then we have*

    (a) $B > 0 \Leftrightarrow B_{22} > 0, B_{11.2} > 0$.
    (b) If $B_{22} > 0$, then $B \ge 0 \Leftrightarrow B_{11.2} \ge 0$.

Proof. The proof can be seen in [26]. □

Lemma 3. *Let $\theta_i, i = 1, \cdots, n$ be i.i.d. random variables with $E(\theta_i) = 0$ and $\mathrm{Var}(\theta_i) = \sigma^2 < \infty$, then for any permutation $(l_1, l_2, \cdots, l_n)$ of $(1, 2, \cdots, n)$, we have*

$$\max_{1 \le j \le n} \left| \sum_{i=1}^{j} \theta_{l_i} \right| = O\left(n^{1/2} \log n \right) \quad a.s.$$

Proof. The proof of Lemma 3 can be referred to [27]. □

Lemma 4. *Assuming that Conditions C1–C8 and H_0 hold, then we have*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{h}_i = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} h_i + o_p(1).$$

Proof. Firstly, we prove
    (1). $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{h}_{1i} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} h_{1i} + o_p(1)$,

(2). $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{h}_{2i} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} h_{2i} + o_p(1)$.

Firstly, we consider the component $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{h}_{1i}$,

$$
\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{h}_{1i} &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \eta_i \left[(Y_i - X_i^T \hat{\alpha}(U_i) - g(Z_i, \beta))^2 - \sigma^2 \right] \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \eta_i \left(\varepsilon_i^2 - \sigma^2 \right) - \frac{2}{\sqrt{n}} \sum_{i=1}^{n} \eta_i \varepsilon_i [X_i^T \hat{\alpha}(U_i) - X_i^T \alpha(U_i)] \quad + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \eta_i [X_i^T \hat{\alpha}(U_i) - X_i^T \alpha(U_i)]^2 \\
&\triangleq \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \eta_i \left(\varepsilon_i^2 - \sigma^2 \right) + R_1 + R_2,
\end{aligned}
$$

where $R_1 = \frac{2}{\sqrt{n}} \sum_{i=1}^{n} \eta_i \varepsilon_i [X_i^T \hat{\alpha}(U_i) - X_i^T \alpha(U_i)]$ and $R_2 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \eta_i [X_i^T \hat{\alpha}(U_i) - X_i^T \alpha(U_i)]^2$. Therefore, we can derive that

$$
\begin{aligned}
|R_1| &= \frac{2}{\sqrt{n}} \left| \sum_{i=1}^{n} \eta_i \varepsilon_i (X_i^T \hat{\alpha}(U_i) - X_i^T \alpha(U_i)) \right| \\
&\leq \frac{2}{\sqrt{n}} \sup_{1 \leq i \leq n} \|\eta_i\| \cdot \max_{1 \leq k \leq n} \left\| \sum_{i=1}^{k} \varepsilon_i \right\| \cdot \sup_{1 \leq i \leq n} \|(\hat{\alpha}(U_i) - \alpha(U_i))\| \cdot \max_{1 \leq i \leq n} \|X_i\| \\
&= o_p(1).
\end{aligned}
$$

Similar to the discussion of R_1, $R_2 = o_p(1)$ holds. Then

$$
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{h}_{1i} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \eta_i \left(\varepsilon_i^2 - \sigma^2 \right) + o_p(1).
$$

Next, we consider the component $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{h}_{2i}$,

$$
\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{h}_{2i} &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \tilde{g}'(Z_i, \beta) \left[Y_i - X_i^T \hat{\alpha}(U_i) - g(Z_i, \beta) \right] \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \tilde{g}'(Z_i, \beta) \varepsilon_i - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \tilde{g}'(Z_i, \beta) [X_i^T \hat{\alpha}(U_i) - X_i^T \alpha(U_i)],
\end{aligned}
$$

where

$$
\begin{aligned}
&\left\| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \tilde{g}'(Z_i, \beta) [X_i^T \hat{\alpha}(U_i) - X_i^T \alpha(U_i)] \right\| \\
&\leq \sqrt{n} \max_{1 \leq i \leq n} \|\tilde{g}'(Z_i, \beta)\| \cdot \max_{1 \leq i \leq n} \|X_i\| \cdot \sup_{1 \leq i \leq n} \|\hat{\alpha}(U_i) - \alpha(U_i)\| \\
&= o_p(1).
\end{aligned}
$$

So $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{h}_{2i} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \tilde{g}'(Z_i, \beta) \varepsilon_i + o_p(1)$. $\qquad \square$

Lemma 5. *Assuming that Conditions C1–C8 and H_0 hold. Then*

$$
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{h}_i \xrightarrow{\mathcal{L}} N(\mathbf{0}, \Sigma').
$$

Proof. Denote $b = (b_1^T, b_2^T)^T$, where $b_1 \in R^{q+1}$ and $b_1 \in R^p$, Therefore, b can be regarded as a $(p+q+1)$-dimensional vector.

$$(b_1^T, b_2^T) \begin{pmatrix} \sum_{i=1}^n \eta_i(\varepsilon_i^2 - \sigma^2) \\ \sum_{i=1}^n g'(Z_i, \beta)\varepsilon_i \end{pmatrix} = \sum_{i=1}^n \left(b_1^T \eta_i \left(\varepsilon_i^2 - \sigma^2 \right) + b_2^T g'(Z_i, \beta)\varepsilon_i \right) \triangleq \sum_{i=1}^n \xi_i.$$

Denote $\mu_i = E\varepsilon^i$, then we have $\mu_1 = E\varepsilon = 0$ and $\mu_2 = E\varepsilon^2 = \sigma^2$,

$$\text{Var}(\xi_i) = (b_1^T, b_2^T) \begin{pmatrix} \eta_i \eta_i^T (\mu_4 - \mu_2^2) & \eta_i g'(Z_i, \beta)^T \mu_3 \\ z_i \eta_i^T \mu_3 & g'(Z_i, \beta)g'(Z_i, \beta)^T \mu_2 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

According to Condition C8, matrix $\Sigma = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ is nonnegative definite.

We know matrix $B_{22} > 0$ and $B_{11} - B_{12}B_{22}^{-1}B_{21} \geq 0$ from Lemma 3(b), this together with Cauchy-Schwartz inequality show that

$$\mu_3^2 = \left(E\left(\varepsilon_i^2 - \sigma^2 \right) \varepsilon_i \right)^2 < E\left(\varepsilon_i^2 - \sigma^2 \right)^2 E\varepsilon_i^2 = \left(\mu_4 - \mu_2^2 \right) \mu_2.$$

Since the relationship between the ε_i and $\varepsilon_i^2 - \sigma^2$ is nonlinear, the above inequality holds strictly. Then,

$$B_{11} \left(\mu_4 - \mu_2^2 \right) - B_{12}B_{22}^{-1}B_{21}\mu_3^2/\mu_2 > \left(B_{11} - B_{12}B_{22}^{-1}B_{21} \right) \mu_3^2/\mu_2 \geq 0.$$

Naturally, we get $B_{11} \left(\mu_4 - \mu_2^2 \right) - B_{12}B_{22}^{-1}B_{21}\mu_3^2/\mu_2 > 0, B_{22}\mu_2 > 0$, it follows from Lemma 3 that

$$\begin{pmatrix} B_{11} \left(\mu_4 - \mu_2^2 \right) & B_{12}\mu_3 \\ B_{21}\mu_3 & B_{22}\mu_2 \end{pmatrix} \triangleq \Sigma'$$

is a positive definite matrix. This indicates that the Lindeberg Condition is met. So by means of the Lindeberg-Feller central limit theorem, we obtain

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n b^\tau \hat{h}_i \overset{\mathcal{L}}{\to} N\left(\mathbf{0}, b^\tau \Sigma' b \right).$$

This together with Cramer-Wold method, we deduce that $\frac{1}{\sqrt{n}} \sum_{i=1}^n \hat{h}_i \overset{\mathcal{L}}{\to} N\left(\mathbf{0}, \Sigma' \right)$. The proof is finished. $\square$

Lemma 6. *Under null hypothesis and conditions C1–C6, we have*

$$\frac{1}{n} \sum_{i=1}^n \hat{h}_i \hat{h}_i^T \overset{\mathcal{P}}{\to} \Sigma'.$$

Proof. It is easy to obtain that

$$\frac{1}{n} \sum_{i=1}^n \hat{h}_i \hat{h}_i^T = \frac{1}{n} \begin{pmatrix} \sum_{i=1}^n \hat{h}_{1i}\hat{h}_{1i}^\tau & \sum_{i=1}^n \hat{h}_{1i}\hat{h}_{2i}^T \\ \sum_{i=1}^n \hat{h}_{2i}\hat{h}_{1i}^\tau & \sum_{i=1}^n \hat{h}_{2i}\hat{h}_{2i}^T \end{pmatrix},$$

$$\frac{1}{n}\sum_{i=1}^{n}\hat{h}_{1i}\hat{h}_{1i}^{T} = \frac{1}{n}\sum_{i=1}^{n}\eta_{i}\eta_{i}^{T}[(Y_{i} - X_{i}^{T}\hat{\alpha}(U_{i}) - g(Z_{i},\beta))^{2} - \sigma^{2}]^{2}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\eta_{i}\eta_{i}^{T}\left[\varepsilon_{i}^{2} - \sigma^{2} + \left(X_{i}^{T}\hat{\alpha}(U_{i}) - X_{i}^{T}\alpha(U_{i})\right)^{2} + 2\varepsilon_{i}\left(X_{i}^{T}\hat{\alpha}(U_{i}) - X_{i}^{T}\alpha(U_{i})\right)\right]^{2}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\eta_{i}\eta_{i}^{T}\left(\varepsilon_{i}^{2} - \sigma^{2}\right)^{2} + \frac{1}{n}\sum_{i=1}^{n}\eta_{i}\eta_{i}^{\tau}\left(X_{i}^{T}\hat{\alpha}(U_{i}) - X_{i}^{T}\alpha(U_{i})\right)^{4}$$

$$+ \frac{4}{n}\sum_{i=1}^{n}\eta_{i}\eta_{i}^{T}\varepsilon_{i}^{2}\left(X_{i}^{T}\hat{\alpha}(U_{i}) - X_{i}^{T}\alpha(U_{i})\right)^{2}$$

$$+ \frac{2}{n}\sum_{i=1}^{n}\eta_{i}\eta_{i}^{T}\left(\varepsilon_{i}^{2} - \sigma^{2}\right)\left(X_{i}^{T}\hat{\alpha}(U_{i}) - X_{i}^{T}\alpha(U_{i})\right)^{2}$$

$$+ \frac{4}{n}\sum_{i=1}^{n}\eta_{i}\eta_{i}^{T}\left(\varepsilon_{i}^{2} - \sigma^{2}\right)\varepsilon_{i}\left(X_{i}^{T}\hat{\alpha}(U_{i}) - X_{i}^{T}\alpha(U_{i})\right)$$

$$+ \frac{4}{n}\sum_{i=1}^{n}\eta_{i}\eta_{i}^{\tau}\varepsilon_{i}\left(X_{i}^{T}\hat{\alpha}(U_{i}) - X_{i}^{T}\alpha(U_{i})\right)^{3}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\eta_{i}\eta_{i}^{T}\left(\varepsilon_{i}^{2} - \sigma^{2}\right)^{2} + \frac{1}{n}\sum_{i=1}^{n}\eta_{i}\eta_{i}^{T}\left(X_{i}^{T}\hat{\alpha}(U_{i}) - X_{i}^{T}\alpha(U_{i})\right)^{4}$$

$$+ \frac{6}{n}\sum_{i=1}^{n}\eta_{i}\eta_{i}^{\tau}\varepsilon_{i}^{2}\left(X_{i}^{T}\hat{\alpha}(U_{i}) - X_{i}^{T}\alpha(U_{i})\right)^{2}$$

$$- \frac{2}{n}\sigma^{2}\sum_{i=1}^{n}\eta_{i}\eta_{i}^{\tau}\left(X_{i}^{T}\hat{\alpha}(U_{i}) - X_{i}^{T}\alpha(U_{i})\right)^{2}$$

$$+ \frac{4}{n}\sum_{i=1}^{n}\eta_{i}\eta_{i}^{\tau}\varepsilon_{i}^{3}\left(X_{i}^{T}\hat{\alpha}(U_{i}) - X_{i}^{T}\alpha(U_{i})\right)$$

$$- \frac{4}{n}\sigma^{2}\sum_{i=1}^{n}\eta_{i}\eta_{i}^{\tau}\varepsilon_{i}\left(X_{i}^{T}\hat{\alpha}(U_{i}) - X_{i}^{T}\alpha(U_{i})\right)$$

$$+ \frac{4}{n}\sum_{i=1}^{n}\eta_{i}\eta_{i}^{\tau}\varepsilon_{i}\left(X_{i}^{T}\hat{\alpha}(U_{i}) - X_{i}^{T}\alpha(U_{i})\right)^{3}$$

$$= \frac{1}{n}\sum_{i=1}^{n}h_{1i}h_{1i}^{T} + N_{1} + N_{2} + N_{3} + N_{4} + N_{5} + N_{6}.$$

Using Condition C6 and Lemma 1, we get

$$|N_{1}| = \left|\frac{1}{n}\sum_{i=1}^{n}\left(X_{i}^{T}\hat{\alpha}(U_{i}) - X_{i}^{T}\alpha(U_{i})\right)^{4}\right| \leq \sup_{1\leq i\leq n}\left\|X_{i}^{T}\hat{\alpha}(U_{i}) - X_{i}^{T}\alpha(U_{i})\right\|^{4} = O\left(C_{n}^{4}\right)O_{p}(1) = o_{p}(1).$$

Using the similar derivation method, we arrive at the following conclusion

$$N_{i} = o_{p}(1), \quad i = 2,\cdots,6.$$

Moreover, we have

$$\frac{1}{n} \sum_{i=1}^{n} \hat{h}_i \hat{h}_i^{\tau} = \frac{1}{n} \sum_{i=1}^{n} h_i h_i^{\tau} + o_p(1).$$

Invoking the law of large number, it is easy to get

$$\frac{1}{n} \sum_{i=1}^{n} \hat{h}_i \hat{h}_i^{T} \xrightarrow{\mathcal{P}} \Sigma'.$$

$\square$

Lemma 7. *Denote $\hat{h}_{max} = \max\{\hat{h}_1, \cdots, \hat{h}_n\}$, then under the null hypothesis and Conditions C1–C8, it holds*

$$\hat{h}_{max} = o_p(n^{1/2}).$$

Proof. The proof can be inspired from [18]. $\square$

Lemma 8. *The conclusion about the the Lagrange multiplier λ is as follows:*

$$\|\lambda\| = O_p\left(n^{-1/2}\right).$$

Proof. It can be get from [18], thus we omit here. $\square$

Proof of Theorem 1. Based on the above Lemmas 7 and 8 and the Taylor expansion of (2.15), we deduce that

$$-2 \log R\left(\gamma; \sigma^2, \beta\right) = 2 \sum_{i=1}^{n} \left(\lambda^T \hat{h}_i - \frac{1}{2}\left(\lambda^T \hat{h}_i\right)^2\right) + o_p(1).$$

By Lemmas 5–8, we have

$$-2 \log R(\gamma; \sigma^2, \beta) = \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{h}_i\right)^T \left(\frac{1}{n} \sum_{i=1}^{n} \hat{h}_i \hat{h}_i^T\right)^{-1} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{h}_i\right) + o_p(1).$$

Similar to [18], $-2 \log R\left(\gamma; \sigma^2, \beta\right) \xrightarrow{\mathcal{L}} \chi^2_{p+q+1}$ can be derived. Theorem 1 follows clearly. $\square$