

Research Article

Block Empirical Likelihood for Longitudinal Single-Index Varying-Coefficient Model

Yunquan Song,^{1,2} Ling Jian,² and Lu Lin¹

¹ Shandong University Qilu Securities Institute for Financial Studies and School of Mathematics, Shandong University, Jinan 250100, China

² College of Science, China University of Petroleum, Qingdao 266580, China

Correspondence should be addressed to Lu Lin; linlu@sdu.edu.cn

Received 26 May 2013; Revised 15 August 2013; Accepted 15 August 2013

Academic Editor: Zhiwei Gao

Copyright © 2013 Yunquan Song et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we consider a single-index varying-coefficient model with application to longitudinal data. In order to accommodate the within-group correlation, we apply the block empirical likelihood procedure to longitudinal single-index varying-coefficient model, and prove a nonparametric version of Wilks' theorem which can be used to construct the block empirical likelihood confidence region with asymptotically correct coverage probability for the parametric component. In comparison with normal approximations, the proposed method does not require a consistent estimator for the asymptotic covariance matrix, making it easier to conduct inference for the model's parametric component. Simulations demonstrate how the proposed method works.

1. Introduction

The single-index varying-coefficient model which was proposed by Huang and Zhenheng [1] is a very important tool to explore the dynamic pattern in many complex dynamic systems, such as economics, finance, politics, epidemiology, medical science, and ecology. As mentioned in Gao et al. [2], the concept of complex dynamic systems arises in many varieties. Such systems are often concurrent and distributed, because they have to react to various kinds of events, signals, and conditions. They may be characterized by a system with uncertainties, time delays, stochastic perturbations, hybrid dynamics, distributed dynamics, chaotic dynamics, and a large number of algebraic loops. Moreover, many related literatures, such as Jian et al. [3] and Hu et al. [4], have been proposed. The single-index varying-coefficient models is one method that can be used to describe the complex dynamic systems. They are natural extensions of classical parametric models with good interpretability and are becoming more and more popular in data analysis.

Longitudinal data arise frequently in many scientific studies. For longitudinal data, we know that the data that are collected from the same subject at different times are

correlated and that the observations from different subjects are often independent. Therefore, it is of great interest to estimate the regression function incorporating the within-subject correlation to improve the efficiency of estimation. The single-index varying-coefficient model is a popular nonparametric fitting technique; it is easily interpreted in real applications because it has the features of the single-index model and the varying-coefficient model. In addition, the single-index varying-coefficient model may include cross-product terms of some components of covariates. Hence, it has considerable flexibility to cater for a complex multivariate nonlinear structure.

Without loss of generality, we consider a longitudinal study with N subjects and n_i observations over time for the i th subject ($i = 1, \dots, N$) with a total of $n = \sum_{i=1}^N n_i$ observations. In this article, we apply longitudinal data to a single-index varying-coefficient model, and propose a single-index varying-coefficient longitudinal data model of the form

$$y_{ij} = g^T(\beta^T x_{ij}) z_{ij} + \varepsilon_{ij}, \quad i = 1, \dots, N; \quad j = 1, \dots, n_i, \quad (1)$$

where $(x_{ij}, z_{ij}) \in R^p \times R^q$ is a vector of covariates, y_{ij} is the j th measurement on the i th unit, β is a $p \times 1$ vector of unknown

parameters, $g(\cdot)$ is an $q \times 1$ vector of unknown functions, and ε_{ij} is a random error with mean 0 and finite variance σ^2 , assuming that ε_{ij} and (x_{ij}, z_{ij}) are independent. For the sake of identifiability, it is often assumed that $\|\beta\| = 1$ and the first nonzero element is positive, where $\|\cdot\|$ denotes the Euclidean metric.

Obviously, model (1) includes a class of important statistical models. For example, if $q = 1$ and $z_{ij} = 1$, model (1) reduces to the single-index longitudinal data model which was proposed by Bai et al. [5] to estimate the index coefficient and unknown link function in a single-index model for longitudinal data by combining penalized splines and quadratic inference functions. If $p = 1$ and $\beta = 1$, (1) is the varying-coefficient longitudinal data model studied by Chiang et al. [6], Huang et al. [7], and Qu and Li [8], among others. So model (1) is easily interpreted in real applications because it has the features of the single-index longitudinal data model and the varying-coefficient longitudinal data model. In addition, model (1) may include cross-product terms of some components of x_{ij} and z_{ij} . Hence, it has considerable flexibility to cater for complex multivariate nonlinear structure.

When $n_i = 1$, model (1) reduces to the nonlongitudinal single-index varying-coefficient model. Some authors have studied the estimation and application of the model. Recently, empirical likelihood methods have been applied to nonlongitudinal single-index varying-coefficient model. For example, Xue and Wang [9] developed statistical techniques for the unknown coefficient functions and single-index parameters in the single-index varying-coefficient models. They first estimate the nonparametric component via the local linear fitting, then construct an estimated empirical likelihood ratio function, and hence obtain a maximum empirical likelihood estimator for the parametric component. The motivation is that empirical likelihood based inference has many desirable statistical properties. For example, this method does not involve any variance estimation which is rather complicated in nonparametric or semiparametric regression settings and hence are robust against the heteroscedasticity; confidence region based on the empirical likelihood method does not have predetermined symmetry so that it can better correspond to the true shape of the underlying distribution, and so on. Owen [10, 11] and many others developed this into a general methodology. For example, Wang and Jing [12], Chen and Qin [13], Shi and Lau [14], and Xue and Zhu [15–17], among others. A recent survey on empirical likelihood can be found in the monograph of Owen [18]. More methods about the single-index varying-coefficient model have been proposed, such as Huang and Zhang [19] and Feng and Xue [20]. When $n_i > 1$, model (1) is the single-index longitudinal data model. The usual empirical likelihood method cannot be applied, however, to the single-index longitudinal data model (1) due to correlation within groups. In this paper, we propose a block empirical likelihood procedure to accommodate this correlation. A nonparametric version of the Wilks' theorem is derived, which can be used to construct confidence regions with asymptotically correct coverage probabilities for the

parametric component in the model. Compared with normal approximations, our method has the appealing feature that it does not require one to construct a consistent estimator for the asymptotic covariance matrix. Furthermore, the block empirical likelihood method avoids intensive Monte Carlo simulations usually required by the bootstrap method.

The rest of the paper is organized as follows. Section 2 introduces the estimated block empirical likelihood method. Section 3 derives the nonparametric version of Wilks' theorem. Section 4 provides a data-driven procedure to choose the tuning parameters. A simulation study is given in Section 5. Proof of the main result is relegated to Section 6.

2. Block Empirical Likelihood Method

In this section, we are to extend the results of You et al. [21] and Xue and Wang [9] to the single-index varying-coefficient longitudinal data model.

To apply the block empirical likelihood method to model (1), we introduce an auxiliary random vector

$$\eta_{ij}(\beta) = \{y_{ij} - g^T(\beta^T x_{ij})z_{ij}\} \dot{g}^T(\beta^T x_{ij})z_{ij}x_{ij}\omega(\beta^T x_{ij}), \quad (2)$$

where $\dot{g}(\cdot)$ stands for the derivative of the function vector $g(\cdot)$, and $\omega(\cdot)$ is a bounded weight function with a bounded support \mathcal{U}_ω , which is introduced to control the boundary effect in the estimations of $g(\cdot)$ and $\dot{g}(\cdot)$. For convenience, we pointed that $\omega(\cdot)$ is the indicator function of the set \mathcal{U}_ω . Note that $E\{\eta_{ij}(\beta)\} = 0$ if $\beta = \beta_0$. Hence, the problem of testing whether β is the true parameter is equivalent to testing whether $E\{\eta_{ij}(\beta)\} = 0$, for $i = 1, \dots, N$; $j = 1, \dots, n_i$. Because of the unknowns $g(\cdot)$ and $\dot{g}(\cdot)$, we cannot directly use the block empirical likelihood method to make statistical inference on β . A natural way is to replace $g(\cdot)$ and $\dot{g}(\cdot)$ by their estimators. In this paper, we estimate the vector functions $g(\cdot)$ and $\dot{g}(\cdot)$ via the local linear regression technique (see, e.g., Fan and Gijbels [22]). The local linear estimators for $g(u)$ and $\dot{g}(u)$ are defined as $\hat{g}(u; \beta) = \hat{a}$ and $\hat{\dot{g}}(u; \beta) = \hat{b}$ at the fixed point β_0 , where \hat{a} and \hat{b} minimize the sum of weighted squares:

$$\sum_{i=1}^N \sum_{j=1}^{n_i} \left[y_{ij} - \{a + b(\beta^T x_{ij} - u)\} z_{ij} \right]^2 K_h(\beta^T x_{ij} - u), \quad (3)$$

where $K_h(\cdot) = h^{-1}K(\cdot/h)$, $K(\cdot)$ is a kernel function, and $h = h_n$ is a bandwidth sequence that decreases to 0 as n increase to ∞ . It follows from the least squares theory that

$$\left(\hat{g}^T(u; \beta), \hat{\dot{g}}^T(u; \beta) \right)^T = S_n^{-1}(u; \beta) \xi_n(u; \beta), \quad (4)$$

where

$$S_n(u; \beta) = \begin{pmatrix} S_{n,0}(u; \beta) & S_{n,1}(u; \beta) \\ S_{n,1}(u; \beta) & S_{n,2}(u; \beta) \end{pmatrix}, \quad (5)$$

$$\xi_n(u; \beta) = \begin{pmatrix} \xi_{n,0}(u; \beta) & \xi_{n,1}(u; \beta) \\ \xi_{n,1}(u; \beta) & \xi_{n,2}(u; \beta) \end{pmatrix}$$

with

$$S_{n,k}(u; \beta) = \frac{1}{n} \sum_{i=1}^N \sum_{j=1}^{n_i} z_{ij} z_{ij}^T \left(\frac{\beta^T x_{ij} - u}{h} \right)^k K_h(\beta^T x_{ij} - u),$$

$$\xi_{n,k}(u; \beta) = \frac{1}{n} \sum_{i=1}^N \sum_{j=1}^{n_i} z_{ij} y_{ij} \left(\frac{\beta^T x_{ij} - u}{h} \right)^k K_h(\beta^T x_{ij} - u). \quad (6)$$

Remark 1. Since the convergence rate of the estimator of $\hat{g}'_0(u)$ is slower than that of the estimator of $g_0(u)$ if the same bandwidth is used, this leads to a slower convergence rate for the estimator $\hat{\beta}$ of β than \sqrt{n} . To increase the convergence rate of the estimator of $\hat{g}'_0(u)$, we introduce the another bandwidth h_1 to replace h in $\hat{g}(u; \beta)$ and define it as $\hat{g}_{h_1}(u; \beta)$.

Similar to Owen [11] and Shi and Lau [14], $\{\hat{r}_{ij}(\beta) = y_{ij} - \hat{g}^T(\beta^T x_{ij}) z_{ij}, i = 1, \dots, N; j = 1, \dots, n_i\}$ can be treated as a random sieve approximation of the random error sequence $\{\varepsilon_{ij}, i = 1, \dots, N; j = 1, \dots, n_i\}$. In order to deal with the correlation within groups, we use the block empirical likelihood procedure proposed by You et al. [21]. Unlike the usual empirical likelihood method, the block empirical likelihood procedure takes the "data" $\hat{r}_{ij}(\beta)$ for $j = 1, \dots, n_i$ into account as a whole. Let $\hat{\eta}_{ij}(\beta) = \hat{r}_{ij}(\beta) \hat{g}^T(\beta^T x_{ij}) z_{ij} x_{ij} \omega(\beta^T x_{ij})$ be $\eta_{ij}(\beta)$, with $g(\beta^T x_{ij})$ and $\dot{g}(\beta^T x_{ij})$ replaced by $\hat{g}(\beta^T x_{ij}; \beta)$ and $\hat{\dot{g}}(\beta^T x_{ij}; \beta)$, respectively, for $i = 1, \dots, N; j = 1, \dots, n_i$. Then an estimated block empirical likelihood function for β is defined as

$$\hat{L}(\beta) = \max \left\{ \prod_{i=1}^n p_i \mid p_i \geq 0, \sum_{i=1}^k p_i = 1, \sum_{i=1}^k p_i \sum_{j=1}^{n_i} \hat{\eta}_{ij}(\beta) = 0 \right\}. \quad (7)$$

For a given β a unique maximum exists, provided that 0 is inside the convex hull of the points $\sum_{j=1}^{n_i} \hat{\eta}_{ij}(\beta)$ for $i = 1, \dots, N$. The maximum of (7) may be found via the method of Lagrange multipliers. The optimal value for p_i satisfying (7) may be shown to be

$$p_i = \frac{1}{N} \times \frac{1}{1 + \lambda^T \sum_{j=1}^{n_i} \hat{\eta}_{ij}(\beta)}, \quad (8)$$

where the Lagrange multiplier $\lambda = (\lambda_1, \dots, \lambda_p)^T$ is the solution of the following equation:

$$0 = \frac{1}{N} \sum_{i=1}^N \frac{\sum_{j=1}^{n_i} \hat{\eta}_{ij}(\beta)}{1 + \lambda^T \sum_{j=1}^{n_i} \hat{\eta}_{ij}(\beta)}. \quad (9)$$

Since $p_1 \times \dots \times p_N$ is maximized for $p_i = 1/N$ in the absence of parametric constraints, we define the corresponding estimated profile block empirical log-likelihood ratio as

$$\hat{l}(\beta) = - \sum_{i=1}^N \log \left[1 + \lambda^T \sum_{j=1}^{n_i} \hat{\eta}_{ij}(\beta) \right]. \quad (10)$$

We will show in the next section that if β_0 is the true parameter vector, $\hat{l}(\beta_0)$ is asymptotically chi-square distributed.

3. Theoretical Properties

Throughout this article, we assume that N increases to push up the total sample size $n = n_1 + \dots + n_N$, while the n_i is fixed. To establish the nonparametric Wilks' theorem for $LR(\beta_0)$, we first make the following assumptions.

- (A.1) The density function of $\beta^T x_{ij}$, $f(u)$, is bounded away from zero for $u \in \mathcal{U}_\omega$ and β near β_0 and satisfies the Lipschitz condition of order 1 on \mathcal{U}_ω , where \mathcal{U}_ω is the support of $\omega(u)$.
- (A.2) The function $g_k(u)$, $1 \leq k \leq q$, have continuous second derivatives on \mathcal{U}_ω , where $g_k(u)$ are the k th components of $g(u)$.
- (A.3) $E(\|x_{ij}\|^6) < \infty$, $E(\|z_{ij}\|^6) < \infty$, and $E(|\varepsilon_{ij}|^6) < \infty$.
- (A.4) $Nh^2/(\log N)^2 \rightarrow \infty$, $Nh^4 \log N \rightarrow 0$; $Nhh_1^3/(\log N)^2 \rightarrow \infty$, $Nh_1^5 = O(1)$.
- (A.5) The kernel $K(\cdot)$ is a symmetric probability density function with a bounded support and satisfies the Lipschitz condition of order 1 and $\int u^2 K(u) du \neq 0$.
- (A.6) The matrix $D(u) = E(z_{ij} z_{ij}^T \mid \beta_0^T x_{ij} = u)$ is positive definite, and each entry of $D(u)$ and $C(u) = E(v_{ij} z_{ij}^T \mid \beta_0^T x_{ij} = u)$ satisfies the Lipschitz condition of order 1 on \mathcal{U}_ω , where $v_{ij} = x_{ij} \dot{g}_0^T(\beta_0^T x_{ij}) z_{ij} \omega(\beta_0^T x_{ij})$, and \mathcal{U}_ω is defined in (A.1).
- (A.7) The matrices $B(\beta_0) = E(v_{ij} v_{ij}^T)$ and $B_*(\beta_0) = B(\beta_0) - E\{C(\beta_0^T x_{ij}) \dot{g}_0(\beta_0^T x_{ij}) E(x_{ij}^T \mid \beta_0^T x_{ij})\}$ are positive definite, where v_{ij} is defined in (A.6).

Remark 2. Condition (A.1) is used to bound the density function of $\beta^T x_{ij}$ away from zero. This ensures that the denominators of $\hat{g}(u; \beta)$ and $\hat{\dot{g}}(u; \beta)$ are, in probability one, bounded away from 0 for $u \in \mathcal{U}_\omega$. The second derivatives in (A.2) are standard smoothness conditions. (A.3)–(A.5) are necessary conditions for the asymptotic normality or the uniform consistency of the estimators. It should be pointed out that the condition can be replaced by $E(\|x_{ij}\|^{6+\delta}) < \infty$, $E(\|z_{ij}\|^{6+\delta}) < \infty$, and $E(|\varepsilon_{ij}|^{6+\delta}) < \infty$ for some $\delta > 0$. In the current work, the exponential index of the norm is set as 6 for it is the minimum value to meet the asymptotic normality or the uniform consistency of the estimators. Conditions (A.6) and (A.7) ensure that the asymptotic variance for the estimator of β_0 exists.

Let $\mathcal{B} = \{\beta \in R^p : \|\beta\| = 1, \text{ and the first nonzero element is positive}\}$. Then β_0 is an inner point of set \mathcal{B} . The following theorem shows that $-2\hat{l}(\beta_0)$ is asymptotically distributed as a weighted sum of independent χ_1^2 variables.

Theorem 3. Suppose that (A.1)–(A.7) hold, then as $N \rightarrow \infty$,

$$-2\widehat{l}(\beta_0) \xrightarrow{D} \omega_1 \chi_{1,1}^2 + \cdots + \omega_p \chi_{1,p}^2, \quad (11)$$

where \xrightarrow{D} represents convergence in distribution, $\chi_{1,1}^2, \dots, \chi_{1,p}^2$ are independent χ_1^2 variables, and the weights ω_j , for $1 \leq j \leq p$, are the eigenvalues of $G(\beta_0) = B^{-1}(\beta_0)A(\beta_0)$. Here $B(\beta_0)$ is defined in condition (A.7),

$$A(\beta_0) = B(\beta_0) - E \left\{ C(\beta_0^T x_{ij}) D^{-1}(\beta_0^T x_{ij}) C^T(\beta_0^T x_{ij}) \right\}, \quad (12)$$

and $C(u)$ and $D(u)$ are defined in condition (A.6).

To apply Theorem 3 to construct a confidence region or interval for β_0 , we need to consistently estimate the unknown weights ω_j . By the plug-in method, $A(\beta_0)$ and $B(\beta_0)$ can be consistently estimated by

$$\begin{aligned} \widehat{A}(\widehat{\beta}) &= \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{n_i} \left\{ \widehat{v}_{ij} \widehat{v}_{ij}^T - \widehat{C}(\widehat{\beta}^T x_{ij}) \widehat{D}^{-1}(\widehat{\beta}^T x_{ij}) \widehat{C}^T(\widehat{\beta}^T x_{ij}) \right\}, \\ & \quad (13) \end{aligned}$$

$$\widehat{B}(\widehat{\beta}) = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{n_i} \widehat{v}_{ij} \widehat{v}_{ij}^T, \quad (14)$$

respectively, where $\widehat{\beta}$ is the maximum empirical likelihood estimator of β_0 defined by (9), $\widehat{v}_{ij} = x_{ij} \widehat{g}^T(\widehat{\beta}^T x_{ij}; \widehat{\beta}) z_{ij} \omega(\widehat{\beta}^T x_{ij})$, $\widehat{C}(\cdot) = \sum_{i=1}^n W_{ni}(\cdot) \widehat{v}_{ij} z_{ij}^T$, and $\widehat{D}(\cdot) = \sum_{i=1}^n W_{ni}(\cdot) z_{ij} z_{ij}^T$ with

$$W_{nij}(\cdot) = \frac{K_1(\widehat{\beta}^T x_{ij} - \cdot / b_n)}{\sum_{k=1}^N \sum_{l=1}^{n_k} K_1(\widehat{\beta}^T x_{kl} - \cdot / b_n)}, \quad (15)$$

where $K_1(\cdot)$ is a kernel function, and b_n is bandwidth with $0 < b_n \rightarrow 0$.

This implies that the eigenvalues of $\widehat{G}(\widehat{\beta}) = \widehat{B}^{-1}(\widehat{\beta})\widehat{A}(\widehat{\beta})$, say $\widehat{\omega}_j$, consistently estimate ω_j for $j = 1, \dots, p$. Let $\widehat{c}_{1-\alpha}$ be the $1 - \alpha$ quantile of the conditional distribution of the weighted sum $\widehat{s} = \widehat{\omega}_1 \chi_{1,1}^2 + \cdots + \widehat{\omega}_p \chi_{1,p}^2$ given the data. Then an approximate $1 - \alpha$ confidence region for β_0 can be defined as follows:

$$\mathcal{R}(\alpha) = \left\{ \beta \in \mathcal{B} : -2\widehat{l}(\beta) \leq \widehat{c}_{1-\alpha} \right\}. \quad (16)$$

In practice, the conditional distribution of the weighted sum \widehat{s} , given the sample $\{(y_{ij}, x_{ij}, z_{ij}), 1 \leq i \leq N; 1 \leq j \leq n_i\}$, can be calculated using Monte Carlo simulations by repeatedly generating independent samples $\chi_{1,1}^2, \dots, \chi_{1,p}^2$ from the χ_1^2 distribution.

In addition to the above, direct way of approximating the asymptotic distributions, we can also consider the following alternative. The alternative is motivated by the results of Rao

and Scott [24]. Now, we propose another adjusted empirical log-likelihood, whose asymptotic distribution is chi-squared with p degrees of freedom. The adjustment technique is developed by Wang and Rao [25] by using an approximate result in Rao and Scott [24]. Note that $\widehat{\rho}(\beta)$ can be written as

$$\widehat{\rho}(\beta) = \frac{\text{tr} \left\{ \widehat{A}^{-1}(\beta) \widehat{A}(\beta) \right\}}{\text{tr} \left\{ \widehat{B}^{-1}(\beta) \widehat{A}(\beta) \right\}}. \quad (17)$$

By examining the asymptotic expansion of $-2\widehat{l}(\beta)$, which is specified in the proof of Theorem 4 below, we define an adjustment factor

$$\widehat{r}(\beta) = \frac{\text{tr} \left\{ \widehat{A}^{-1}(\beta) \widehat{\Sigma}(\beta) \right\}}{\text{tr} \left\{ \widehat{B}^{-1}(\beta) \widehat{\Sigma}(\beta) \right\}}, \quad (18)$$

by replacing $\widehat{A}(\beta)$ in $\widehat{\rho}(\beta)$ by $\widehat{\Sigma}(\beta)$, where $\widehat{\Sigma}(\beta) = \left\{ \sum_{i=1}^N \sum_{j=1}^{n_i} \widehat{\eta}_{ij}(\beta) \right\} \left\{ \sum_{i=1}^N \sum_{j=1}^{n_i} \widehat{\eta}_{ij}(\beta) \right\}^T$. The adjusted empirical log-likelihood ration is defined by

$$\widehat{l}_a(\beta) = \widehat{r}(\beta) \left\{ -2\widehat{l}(\beta) \right\}, \quad (19)$$

where $\widehat{l}(\beta)$ is defined in (10).

Theorem 4. Suppose that conditions (A.1)–(A.6) hold. Then, $\widehat{l}_a(\beta) \xrightarrow{D} \chi_p^2$.

According to Theorem 4, $\widehat{l}_a(\beta)$ can be used to construct an approximate confidence region for β_0 . Let

$$\mathcal{R}_a(\alpha) = \left\{ \beta \in \mathcal{B} : \widehat{l}_a(\beta) \leq \chi_p^2(1 - \alpha) \right\}. \quad (20)$$

Then, $\mathcal{R}_a(\alpha)$ gives a confidence region for β_0 with asymptotically correct coverage probability $1 - \alpha$.

4. Bandwidth Selection

For practical implementation, the tuning parameters need to be decided. We employ a data-driven procedure to choose the tuning parameter h , where h controls the smoothness of $\widehat{g}(\cdot)$ and $\widehat{\widehat{g}}(\cdot)$. We all know that various existing bandwidth selection techniques for nonparametric regression, such as the cross-validation, generalized cross-validation, and the modified multifold cross-validation criterion, can be adapted for the estimation $\widehat{g}(\cdot)$ and $\widehat{\widehat{g}}(\cdot)$. Because the algorithm of the modified multifold cross-validation criterion proposed by Cai et al. [26] to select the optimal bandwidth is simple and quick, throughout the empirical studies in this paper, we consider the modified multifold cross-validation criterion. Specifically, let m and M be two given positive integers and $n > mM$. The basic idea is first to use M subseries of lengths $n - km$ ($k = 1, \dots, M$) to estimate the unknown coefficient functions and then to compute the one-step forecasting error of the next section of the sample of length m based on

the estimated models. More precisely, we choose h which minimizes

$$\text{AMS}(h) = \sum_{k=1}^M \frac{1}{m} \sum_{i=n-km+1}^{n-km+m} \sum_{j=1}^{n_i} \left\{ y_{ij} - \sum_{l=1}^q \hat{g}_{l,k}(\beta^T x_{ij}) z_{ij}^{(l)} \right\}^2, \quad (21)$$

where $\hat{g}_{l,k}(\cdot)$ are computed from the sample $\{(y_{ij}, x_{ij}, z_{ij}), 1 \leq i \leq n - km; 1 \leq j \leq n_i\}$ with bandwidth equal to $h/(n - km)^{1/5}$. Note that for different sample size, we rescale bandwidth according to its optimal rate, that is, $h \propto n^{-1/5}$. Since the selected bandwidth does not depend critically on the choice of m and M , to computation expediency, we take $m = [0.1n]$ and $M = 5$ in our simulation.

Let h_{opt} be the bandwidth obtained by minimizing (21) with respect to $h > 0$; that is, $h_{\text{opt}} = \inf_{h>0} \text{AMS}(h)$. Then h_{opt} is the optimal bandwidth for estimating $\hat{g}(\cdot)$. When calculating the block empirical likelihood ratios and estimator of β_0 , we use the approximation bandwidth

$$h = h_{\text{opt}} n^{-1/20} (\log n)^{-1/2}, \quad h_1 = h_{\text{opt}}, \quad (22)$$

because this insures that the required bandwidth has correct order of magnitude for the optimal asymptotic performance (see, e.g., Carroll et al. [27]), and the bandwidth \hat{h} satisfies condition (A.4).

5. A Simulation Study

In this section, we carry out some simulations to study the finite sample performance of the estimated block empirical likelihood method.

Example 5. The data are generated from

$$Y_{ij} = g_0(\beta X_{ij}) + g_1(\beta X_{ij}) Z_{ij} + \varepsilon_{ij}, \quad (23)$$

$$i = 1, \dots, N, \quad j = 1, \dots, n_i,$$

where $X_{ij} \sim N(0, 1)$, $Z_{ij} \sim N(0, 1)$, $g(t) = \sin(2\pi t)$, $\varepsilon_{ij} = a\varepsilon_{i,j-1} + e_{ij}$, and e'_{ij} s are i.i.d $N(0, 1)$. For each combination of N , n_i , and a , 1000 samples are generated from the above model in all simulations. For each sample, a 95% confidence interval for $\beta = 2$ are computed using our estimated block empirical likelihood method. For the smoother, we used a local linear smoother with the Gaussian kernel $K_h(t) = (1/h\sqrt{2\pi}) \exp(-t^2/2h^2)$ with a modified multifold cross-validation criterion bandwidth throughout all smoothing steps. Some representative coverage probabilities and coverage confidence intervals are reported in Table 1. Simulation results show that our estimated block empirical likelihood confidence regions have high coverage probabilities and short average confidence interval lengths.

Example 6. Consider the regression model

$$Y_{ij} = g_0(\beta_0^T X_{ij}) + g_1(\beta_0^T X_{ij}) Z_{ij}^{(1)} + g_2(\beta_0^T X_{ij}) Z_{ij}^{(2)} + \varepsilon_{ij}, \quad (24)$$

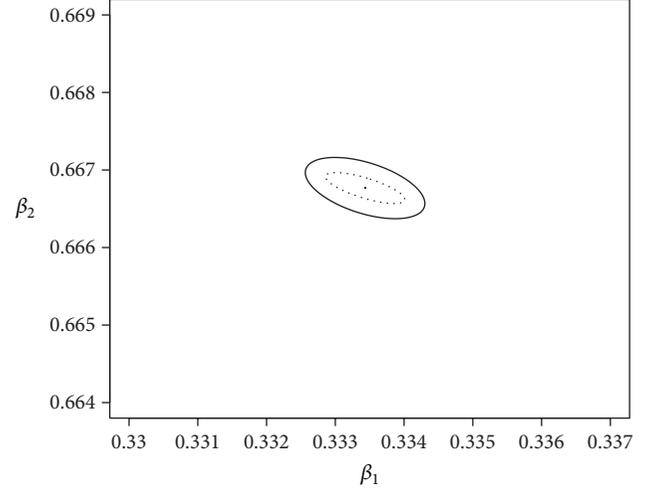


FIGURE 1: Averages of 95% confidence regions of (β_1, β_2) , based on EEL (solid curve) and AEL (dashed curve) when $n = 100$ in the cases of Example 6.

where $\beta_0 = (1/3, 2/3)$ and the ε_{ij} s are independent $N(0, 1^2)$ random variables. The sample $\{X_{ij} = (X_{ij}^{(1)}, X_{ij}^{(2)})^T; 1 \leq i \leq N, 1 \leq j \leq n_i\}$ was generated from a bivariate uniform distribution on $[-1, 1]^2$ with independent components, $\{Z_{ij} = (Z_{ij}^{(1)}, Z_{ij}^{(2)})^T; 1 \leq i \leq N, 1 \leq j \leq n_i\}$ was generated from a bivariate normal distribution $N(0, \Sigma)$ with $\text{var}(Z_{ij}^{(1)}) = \text{var}(Z_{ij}^{(2)}) = 1$, and correlation coefficient between $Z_{ij}^{(1)}$ and $Z_{ij}^{(2)}$ is $\rho = 0.5$. In model (24), the coefficient functions are $g_0(u) = 8 \exp(-2u^2)$, $g_1(u) = 6u^2$ and $g_2(u) = 4 \sin(\pi u)$.

For the smoother, we use a local linear smoother with a Gaussian kernel $K_h(u) = (1/h\sqrt{2\pi}) \exp(-u^2/2h^2)$, and use the modified multifold cross-validation criterion proposed by Cai et al. [26] to select the optimal bandwidth throughout all smoothing steps because the algorithm is simple and quick. We take the weight function $\omega(u) = I_{[-1,1]}$. The sample size for the simulated data is 100, and the run is 1000 times in all simulations.

The confidence regions of β_0 and their coverage probabilities, with nominal level $1 - \alpha = 0.95$, were computed from 1000 runs. The estimated block empirical likelihood was used to construct the confidence regions. The simulated results are given in Figure 1. Simulation results show that our block empirical likelihood confidence regions have high coverage probabilities and short average confidence interval lengths.

The histograms of the 1000 estimators of the parameters β_1 and β_2 are in Figures 2(a) and 2(b), respectively. The Q-Q plots of the 1000 estimators of the parameters β_1 and β_2 are in Figures 3(a) and 3(b), respectively. Figures 2 and 3 show empirically that these estimators are asymptotically normal. The means of the estimates of the unknown parameters β_1 and β_2 are 0.33342 and 0.66673, respectively, and their biases (standard deviations) are 0.000128 (0.00308) and 0.000603 (0.00352), respectively.

TABLE 1: The selection probabilities of adaptive EL shrinkage estimation.

N	Number of replicates	CI	CP
$a = 0.2$			
50	$n_1 = \dots = n_{25} = 4, n_{26} = \dots = n_{50} = 4$	[1.645015, 2.349996]	0.9372
50	$n_1 = \dots = n_{25} = 4, n_{26} = \dots = n_{50} = 5$	[1.634619, 2.355068]	0.9270
100	$n_1 = \dots = n_{25} = 4, n_{26} = \dots = n_{100} = 4$	[1.634619, 2.355068]	0.9343
100	$n_1 = \dots = n_{25} = 4, n_{26} = \dots = n_{100} = 5$	[1.634619, 2.355068]	0.9424
$a = 0.4$			
50	$n_1 = \dots = n_{25} = 4, n_{26} = \dots = n_{50} = 4$	[1.634619, 2.355068]	0.9428
50	$n_1 = \dots = n_{25} = 4, n_{26} = \dots = n_{50} = 5$	[1.634619, 2.355068]	0.9334
100	$n_1 = \dots = n_{25} = 4, n_{26} = \dots = n_{100} = 4$	[1.634619, 2.355068]	0.9427
100	$n_1 = \dots = n_{25} = 4, n_{26} = \dots = n_{100} = 5$	[1.634619, 2.355068]	0.9351

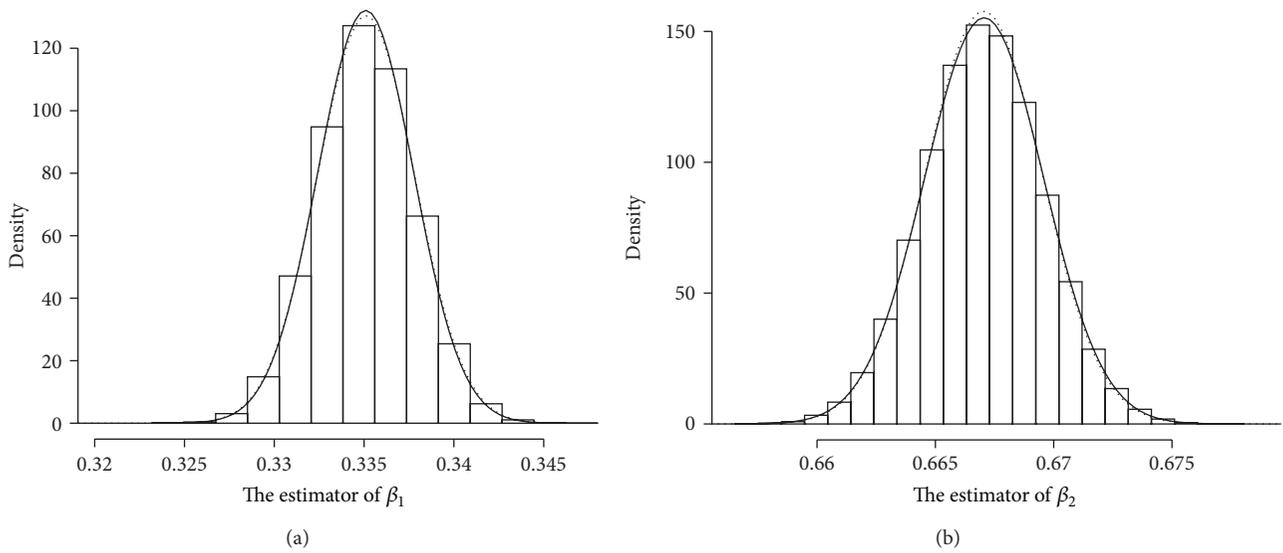


FIGURE 2: The histograms of the 1000 estimators of every parameter, the estimated curve of density (solid curve) and the curve of normal density (dashed curve): (a) for β_1 and (b) for β_2 .

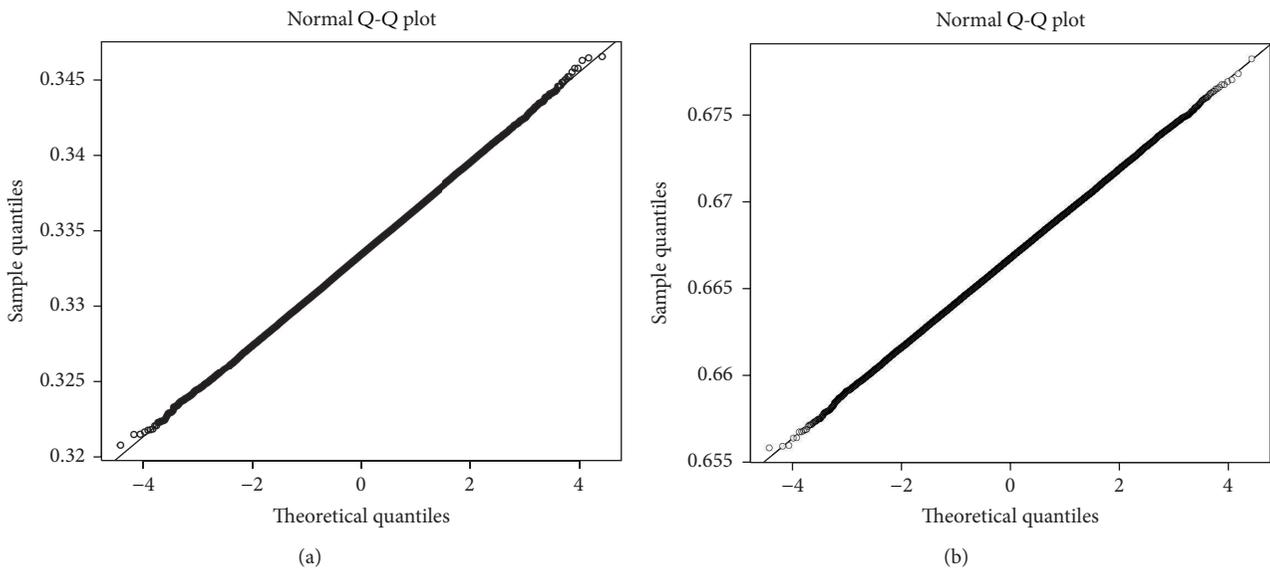


FIGURE 3: (a) for β_1 and (b) for β_2 : the Q-Q plot of the 1000 estimators of every parameter.

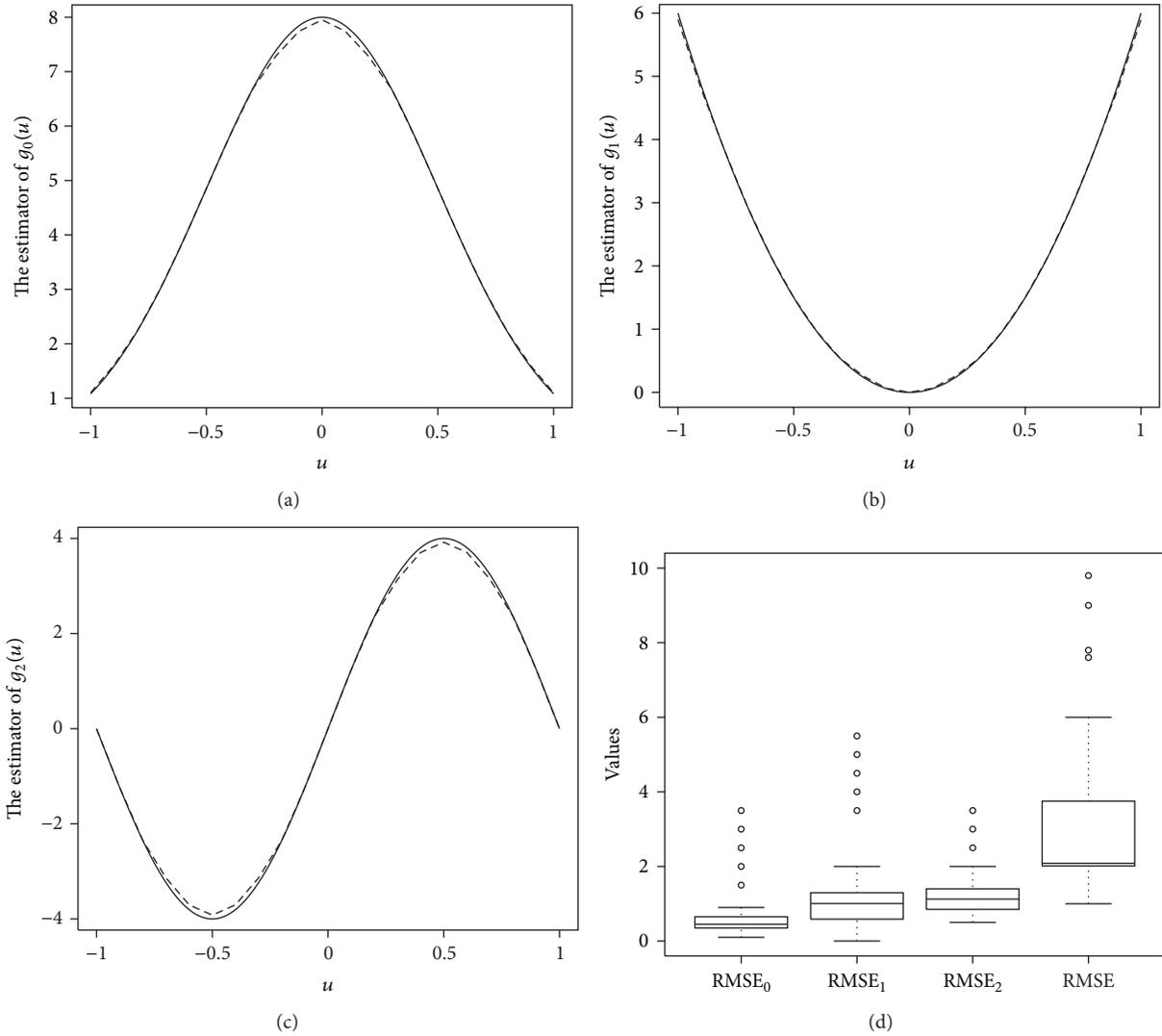


FIGURE 4: The true curve (solid curve) and the estimated curve (dashed curve); (d) the boxplots of the 1000 RMSE values in estimations of $g_0(\cdot)$, $g_1(\cdot)$, and $g_2(\cdot)$ and the sum of the three RMSEs.

We also consider the average estimates of the coefficient functions $g_0(u)$, $g_1(u)$, and $g_2(u)$ over the 1000 replicates. The estimators $\hat{g}_j(\cdot)$ are assessed via the root mean squared errors (RMSE); that is, $RMSE = \sum_{j=0}^2 RMSE_j$, where

$$RMSE_j = \left[n_{\text{grid}}^{-1} \sum_{k=1}^{n_{\text{grid}}} \{ \hat{g}_j(u_k) - g_j(u_k) \}^2 \right]^{1/2}, \quad (25)$$

and $\{u_k, k = 1, \dots, n_{\text{grid}}\}$ are regular grid points. The boxplot for the 1000 RMSEs is given in Figure 4. From Figures 4(a)–4(c) we see that every estimated curve agrees with the true function curve very closely. Figure 4(d) shows that all RMSEs of estimates for the unknown functions are very small.

Example 7. We now apply the block empirical likelihood method to analyze the data from a longitudinal hormone study [28]. The study involved 34 women whose urine samples were collected in one menstrual cycle and whose urinary

progesterone was assayed on alternate days. A total of 492 observations were obtained, with each woman contributing from 11 to 28 observations over time. Each woman’s cycle length was standardized uniformly to a reference 28-day cycle since the change of the progesterone level for each woman depends on time during a menstrual cycle. In the following, we consider the following model:

$$Y_{ij} = g_0(\beta_1 AGE_{ij} + \beta_2 BMI_{ij}) + g_1(\beta_1 AGE_{ij} + \beta_2 BMI_{ij}) t_{ij} + \varepsilon_{ij}, \quad (26)$$

where Y_{ij} is the j th log-transformed progesterone value measured at standardized day t_{ij} since menstruation for i th woman, and AGE_{ij} and BMI_{ij} are age and body mass index for the i th individual at day t_{ij} , respectively.

We apply the block empirical likelihood method to fit the data. Because we focus on the estimators of β_1 and β_2 , we only

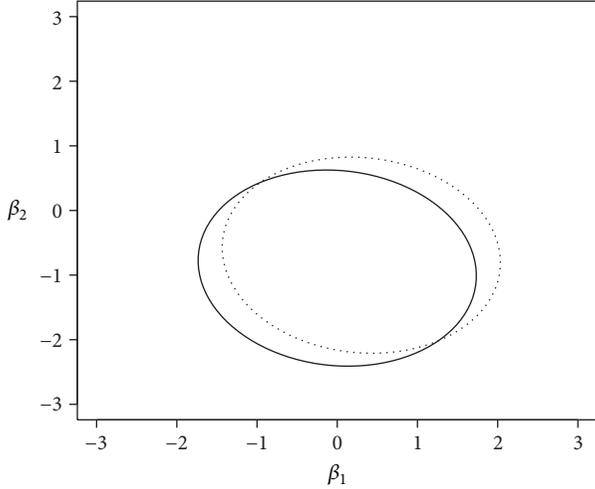


FIGURE 5: The 0.95 confidence regions for the regression coefficients β_1 and β_2 with correlation structures being independence (dotted curve) and first-autoregressive (sold curve).

summarize the estimators of β_1 and β_2 in Figure 5. Next, we denote $\hat{\beta}_{\text{ind}}$ and $\hat{\beta}_{\text{AR}}$ as the estimators of $\beta = (\beta_1, \beta_2)$, when the correlation structures are specified as independence and first-order autoregressive, respectively. We see from Figure 5 that both $\hat{\beta}_{\text{ind}}$ and $\hat{\beta}_{\text{AR}}$ are significant for neither of confidence regions for the two estimators including (0,0). Therefore, we conclude that the parameters β_1 and β_2 are not significant, which is consistent with the conclusion of Zhang et al. [28].

6. Proof of the Theorem

In order to prove Theorem 3, we introduce the following several lemmas. The following lemma gives uniformly convergent rates of $\hat{g}(u; \beta)$ and $\hat{g}'(u; \beta)$. This lemma is straightforward extension of known results in nonparametric function estimation. Moreover, the proofs of Lemma 9 and Lemma 10 is similar with the corresponding Lemma 9 and Lemma 10 of Xue and Wang [9]. We hence omit these proofs.

Lemma 8. Let $\mathcal{B}_n = \{\beta \in \mathcal{B} : \|\beta - \beta_0\| \leq c_0 n^{-1/2}\}$ for some positive constant c_0 . Suppose that conditions (A.1)–(A.3), (A.5), and (A.6) hold. Then

$$\begin{aligned} & \sup_{u \in \mathcal{U}_\omega, \beta \in \mathcal{B}_n} \|\hat{g}(u; \beta) - g_0(u)\| \\ &= O_p \left(\left\{ \frac{\log(1/h)}{nh} \right\}^{1/2} + h^2 \right), \\ & \sup_{u \in \mathcal{U}_\omega, \beta \in \mathcal{B}_n} \|\hat{g}'(u; \beta) - \dot{g}_0(u)\| \\ &= O_p \left(\left\{ \frac{\log(1/h)}{nh^3} \right\}^{1/2} + h \right). \end{aligned} \quad (27)$$

In order to describe Lemma 9, we use the following notations. Denote $\mathcal{G} = \{g : \mathcal{U}_\omega \times \mathcal{B} \mapsto R^q\}$, $\|g\|_{\mathcal{G}} =$

$\sup_{u \in \mathcal{U}_\omega, \beta \in \mathcal{B}_n} \|g(u; \beta)\|$. From Lemma 8, we have $\|\hat{g} - g_0\|_{\mathcal{G}} = o_p(1)$ and $\|\hat{g}' - \dot{g}_0\|_{\mathcal{G}} = o_p(1)$; hence, we can assume that g lies in \mathcal{G}_δ with $\delta = \delta_n \rightarrow 0$ and $\delta > 0$, where

$$\mathcal{G}_\delta = \{g \in \mathcal{G} : \|g - g_0\|_{\mathcal{G}} \leq \delta, \|\dot{g} - \dot{g}_0\|_{\mathcal{G}} \leq \delta\}. \quad (28)$$

Let $g_0(\beta^T \mathcal{X}; \beta) = E\{g_0(\beta_0^T \mathcal{X}) | \beta^T \mathcal{X}\}$ and $\dot{g}_0(\beta^T \mathcal{X}; \beta) = E\{\dot{g}_0(\beta_0^T \mathcal{X}) | \beta^T \mathcal{X}\}$,

$$\begin{aligned} & Q(g, \beta) \\ &= E \left[\{ \mathcal{Y} - g^T(\beta^T \mathcal{X}; \beta) \mathcal{X} \} \dot{g}^T(\beta^T \mathcal{X}; \beta) \mathcal{X} \omega(\beta^T \mathcal{X}) \right], \end{aligned} \quad (29)$$

$$\begin{aligned} & Q_n(g, \beta) \\ &= \frac{1}{n} \sum_{i=1}^N \sum_{j=1}^{n_i} \{ y_{ij} - g^T(\beta^T x_{ij}; \beta) z_{ij} \} \dot{g}^T \\ & \quad \times (\beta^T x_{ij}; \beta) z_{ij} x_{ij} \omega(\beta^T x_{ij}). \end{aligned} \quad (30)$$

Lemma 9. Suppose that conditions (A.1)–(A.6) hold. Let

$$\begin{aligned} & J_1(g, \beta) = Q_n(g, \beta) - Q(g, \beta) - Q_n(g_0, \beta_0), \\ & J_2(\hat{g}, \beta) = Q(g, \beta) - Q(g_0, \beta) \\ & \quad - \omega(g_0(\beta^T \mathcal{X}; \beta); \beta) \\ & \quad \times \{g(\beta^T \mathcal{X}; \beta) - g_0(\beta^T \mathcal{X})\}, \\ & J_3(g, \beta) = \omega(g_0(\beta^T \mathcal{X}), \beta) \{g(\beta^T \mathcal{X}; \beta) - g_0(\beta^T \mathcal{X})\} \\ & \quad - \omega(g_0(\beta^T \mathcal{X}; \beta_0), \beta_0) \\ & \quad \times \{g(\beta_0^T \mathcal{X}; \beta_0) - g_0(\beta_0^T \mathcal{X}; \beta)\}, \\ & J_4(\hat{g}, \beta_0) = Q_n(g_0, \beta_0) \\ & \quad + \omega(g_0(\beta^T \mathcal{X}), \beta) \{g(\beta^T \mathcal{X}; \beta) - g_0(\beta^T \mathcal{X})\}. \end{aligned} \quad (31)$$

Then

$$\sup_{(g, \beta) \in \mathcal{G} \times \mathcal{B}_n} \|J_1(g, \beta)\| = o_p(n^{1/2}), \quad (32)$$

$$\sup_{\beta \in \mathcal{B}_n} \|J_2(\hat{g}, \beta)\| = o_p(n^{1/2}), \quad (33)$$

$$\sup_{(g, \beta) \in \mathcal{G} \times \mathcal{B}_n} \|J_3(g, \beta)\| = o(n^{1/2}), \quad (34)$$

$$\sqrt{n} J_4(\hat{g}, \beta_0) \xrightarrow{D} N(0, \sigma^2 A(\beta_0)), \quad (35)$$

where $A(\beta_0)$ is defined in (12).

Lemma 10. Suppose that conditions (A.1)–(A.6) hold. Then

$$\sup_{\beta \in \mathcal{B}_n} \|Q_n(\hat{g}, \beta)\| = O_p(n^{-1/2}), \quad (36)$$

$$\sup_{\beta \in \mathcal{B}_n} \|R_n(\beta) - \sigma^2 B(\beta_0)\| = o_p(1), \quad (37)$$

$$\sup_{\beta \in \mathcal{B}_n} \max_{\substack{1 \leq i \leq N \\ 1 \leq j \leq n_i}} \|\hat{\eta}_{ij}(\beta)\| = o_p(n^{1/2}), \quad (38)$$

$$\sup_{\beta \in \mathcal{B}_n} \|\lambda(\beta)\| = o_p(n^{-1/2}), \quad (39)$$

where $Q_n(\hat{g}, \beta)$ is defined in (30), $R_n(\beta) = n^{-1} \sum_{i=1}^N \sum_{j=1}^{n_i} \hat{\eta}_{ij}(\beta) \hat{\eta}_{ij}^T(\beta)$, $B(\beta_0)$ is defined in condition (A.7), and $\hat{\eta}_{ij}(\beta)$ is defined in (2).

Proof of Theorem 3. Note that, when $\beta = \beta_0$, Lemma 10 also holds. Applying the Taylor expansion to (7) and invoking Lemma 10, we can obtain

$$-2\hat{l}(\beta_0) = -\sum_{i=1}^N \sum_{j=1}^{n_i} \left[\lambda^T \hat{\eta}_{ij}(\beta_0) - \frac{1}{2} \{ \lambda^T \hat{\eta}_{ij}(\beta_0) \}^2 \right] + o_p(1). \quad (40)$$

By (9) and Lemma 10, we have

$$\begin{aligned} \sum_{i=1}^N \sum_{j=1}^{n_i} \{ \lambda^T \hat{\eta}_{ij}(\beta_0) \}^2 &= \sum_{i=1}^N \sum_{j=1}^{n_i} \lambda^T \hat{\eta}_{ij}(\beta_0) + o_p(1), \\ \lambda &= \left\{ \sum_{i=1}^N \sum_{j=1}^{n_i} \hat{\eta}_{ij}(\beta_0) \hat{\eta}_{ij}^T(\beta_0) \right\}^{-1} \sum_{i=1}^N \sum_{j=1}^{n_i} \hat{\eta}_{ij}(\beta_0) + o_p(n^{-(1/2)}). \end{aligned} \quad (41)$$

This together with (40) proves that

$$-2\hat{l}(\beta_0) = nQ_n^T(\hat{g}, \beta_0) R_n^{-1}(\beta_0) Q_n(\hat{g}, \beta_0) + o_p(1), \quad (42)$$

where $Q_n(\hat{g}, \beta_0)$ and $R_n(\beta_0)$ are defined in (30) and (37), respectively. From (37) of Lemma 10 and (42), we obtain

$$\begin{aligned} -2\hat{l}(\beta_0) &= \left\{ (\sigma^2 A)^{-1/2} \sqrt{n} Q_n(\hat{g}, \beta_0) \right\}^T G(\beta_0) \\ &\quad \times \left\{ (\sigma^2 A)^{-1/2} \sqrt{n} Q_n(\hat{g}, \beta_0) \right\} + o_p(1), \end{aligned} \quad (43)$$

where $G(\beta_0) = A^{1/2}(\beta_0) B^{-1}(\beta_0) A^{1/2}(\beta_0)$. Let $G_0 = \text{diag}(\omega_1, \dots, \omega_p)$, where ω_i , $1 \leq i \leq p$, are the eigenvalues of $G(\beta_0)$. Then there exists an orthogonal matrix H such that $H^T G_0 H$. Using the notations of Lemma 9, we have

$$\begin{aligned} Q_n(\hat{g}, \beta) &= J_1(\hat{g}, \beta) + J_2(\hat{g}, \beta) + J_3(\hat{g}, \beta) \\ &\quad + J_4(\hat{g}, \beta) + Q(g_0, \beta). \end{aligned} \quad (44)$$

Noting that $Q(g_0, \beta_0)$, from the above equation and Lemma 9, we have

$$Q_n(\hat{g}, \beta_0) = J_4(\hat{g}, \beta_0) + o_p(n^{-1/2}). \quad (45)$$

Hence, by (35) of Lemma 9, we have

$$H \{ \sigma^{-2} A^{-1}(\beta_0) \}^{1/2} \sqrt{n} Q_n(\hat{g}, \beta_0) \xrightarrow{D} N(0, I_p), \quad (46)$$

where I_p is the $p \times p$ identity matrix. This together with (43) proves Theorem 3. \square

Proof of Theorem 4. By Lemma 10 and, similarly to the proof of (42), we can obtain

$$\hat{l}(\beta) = -\frac{n}{2} Q_n^T(\hat{g}, \beta) \{ \sigma^2 B(\beta) \}^{-1} Q_n(\hat{g}, \beta) + O_p(1), \quad (47)$$

uniformly for $\beta \in \mathcal{B}_n$, where $o_p(1)$ tends to 0 in probability uniformly for $\beta \in \mathcal{B}_n$. Note that $\hat{A}(\beta_0) \xrightarrow{p} A(\beta_0)$ and $\hat{B}(\beta_0) \xrightarrow{p} B(\beta_0)$. By the expansion of $\hat{l}_a(\beta_0)$, defined in (19) and (47), we get

$$\hat{l}_a(\beta_0) = Q_n^T(\hat{g}, \beta) \{ \sigma^{-2} A^{-1}(\beta_0) \} Q_n(\hat{g}, \beta_0) + O_p(1). \quad (48)$$

This together with (44) and (48) proves Theorem 4. Then we complete the proof. \square

Acknowledgments

This research was supported by NNSF project (11171188 and 11231005) of China, Mathematical Finance-Backward Stochastic Analysis and Computations in Financial Risk Control of China (11221061), NSF and SRRF projects (ZR2010AZ001 and BS2011SF006) of Shandong Province of China, K C Wong-HKBU Fellowship Programme for Mainland China Scholars 2010-11, and the Fundamental Research Funds for the Central Universities (27R1310008A).

References

- [1] Z. Huang, "Empirical likelihood for single-index varying-coefficient models with right-censored data," *Journal of the Korean Statistical Society*, vol. 39, no. 4, pp. 533–544, 2010.
- [2] Z. Gao, D. Kong, and C. Gao, "Modeling and control of complex dynamic systems: applied mathematical aspects," *Journal of Applied Mathematics*, vol. 2012, Article ID 869792, 5 pages, 2012.
- [3] L. Jian, S. Shen, and Y. Song, "Improving the solution of least squares support vector machines with application to a blast furnace system," *Journal of Applied Mathematics*, vol. 2012, Article ID 949654, 12 pages, 2012.
- [4] J. Hu and Z. Gao, "Modules identification in gene positive networks of hepatocellular carcinoma using pearson agglomerative method and pearson cohesion coupling modularity," *Journal of Applied Mathematics*, vol. 2012, Article ID 248658, 21 pages, 2012.
- [5] Y. Bai, W. K. Fung, and Z. Y. Zhu, "Penalized quadratic inference functions for single-index models with longitudinal data," *Journal of Multivariate Analysis*, vol. 100, no. 1, pp. 152–161, 2009.
- [6] C. T. Chiang, J. A. Rice, and C. O. Wu, "Smoothing spline estimation for varying coefficient models with repeatedly measured dependent variables," *Journal of the American Statistical Association*, vol. 96, no. 454, pp. 605–619, 2001.

- [7] J. Z. Huang, C. O. Wu, and L. Zhou, "Varying-coefficient models and basis function approximations for the analysis of repeated measurements," *Biometrika*, vol. 89, no. 1, pp. 111–128, 2002.
- [8] A. Qu and R. Li, "Quadratic inference functions for varying-coefficient models with longitudinal data," *Biometrics*, vol. 62, no. 2, pp. 379–391, 2006.
- [9] L. Xue and Q. Wang, "Empirical likelihood for single-index varying-coefficient models," *Bernoulli*, vol. 18, no. 3, pp. 836–856, 2012.
- [10] A. B. Owen, "Empirical likelihood ratio confidence intervals for a single functional," *Biometrika*, vol. 75, no. 2, pp. 237–249, 1988.
- [11] A. Owen, "Empirical likelihood ratio confidence regions," *The Annals of Statistics*, vol. 18, no. 1, pp. 90–120, 1990.
- [12] Q. H. Wang and B. Y. Jing, "Empirical likelihood for partial linear models with fixed designs," *Statistics & Probability Letters*, vol. 41, no. 4, pp. 425–433, 1999.
- [13] S. X. Chen and Y. S. Qin, "Empirical likelihood confidence intervals for local linear smoothers," *Biometrika*, vol. 87, no. 4, pp. 946–953, 2000.
- [14] J. Shi and T. S. Lau, "Empirical likelihood for partially linear models," *Journal of Multivariate Analysis*, vol. 72, no. 1, pp. 132–148, 2000.
- [15] L. Xue and L. Zhu, "Empirical likelihood for a varying coefficient model with longitudinal data," *Journal of the American Statistical Association*, vol. 102, no. 478, pp. 642–654, 2007.
- [16] L. Xue and L. Zhu, "Empirical likelihood semiparametric regression analysis for longitudinal data," *Biometrika*, vol. 94, no. 4, pp. 921–937, 2007.
- [17] L. G. Xue and L. Zhu, "Empirical likelihood for single-index models," *Journal of Multivariate Analysis*, vol. 97, no. 6, pp. 1295–1312, 2006.
- [18] A. B. Owen, *Empirical Likelihood*, Chapman & Hall, London, UK, 2001.
- [19] Z. Huang and R. Zhang, "Testing for the parametric parts in a single-index varying-coefficient model," *Science China: Mathematics*, vol. 55, no. 5, pp. 1017–1028, 2012.
- [20] S. Feng and L. Xue, "Variable selection for single-index varying-coefficient model," *Frontiers of Mathematics in China*, vol. 8, no. 3, pp. 541–565, 2013.
- [21] J. H. You, G. M. Chen, and Y. Zhou, "Block empirical likelihood for longitudinal partially linear regression models," *The Canadian Journal of Statistics*, vol. 34, no. 1, pp. 79–96, 2006.
- [22] J. Fan and I. Gijbels, *Local Polynomial Modelling and Its Applications*, vol. 66 of *Monographs on Statistics and Applied Probability*, Chapman & Hall, London, UK, 1996.
- [23] J. You and Y. Zhou, "Empirical likelihood for semiparametric varying-coefficient partially linear regression models," *Statistics & Probability Letters*, vol. 76, no. 4, pp. 412–422, 2006.
- [24] J. N. K. Rao and A. J. Scott, "The analysis of categorical data from complex sample surveys: chi-squared tests for goodness of fit and independence in two-way tables," *Journal of the American Statistical Association*, vol. 76, no. 374, pp. 221–230, 1981.
- [25] Q. Wang and J. N. K. Rao, "Empirical likelihood-based inference in linear errors-in-covariables models with validation data," *Biometrika*, vol. 89, no. 2, pp. 345–358, 2002.
- [26] Z. Cai, J. Fan, and Q. Yao, "Functional-coefficient regression models for nonlinear time series," *Journal of the American Statistical Association*, vol. 95, no. 451, pp. 941–956, 2000.
- [27] R. J. Carroll, J. Fan, I. Gijbels, and M. P. Wand, "Generalized partially linear single-index models," *Journal of the American Statistical Association*, vol. 92, no. 438, pp. 477–489, 1997.
- [28] D. Zhang, X. Lin, J. Raz, and M. Sowers, "Semiparametric stochastic mixed models for longitudinal data," *Journal of the American Statistical Association*, vol. 93, no. 442, pp. 710–719, 1998.