Electronic Journal of Statistics
Vol. 6 (2012) 1040–1058
ISSN: 1935-7524
DOI: 10.1214/12-EJS701

Empirical likelihood inference for partially time-varying coefficient errors-in-variables models

Guo-Liang Fan and Hong-Xia Xu

School of Mathematics & Physics, Anhui Polytechnic University, Wuhu 241000, P. R. China
e-mail: guoliangfan@yahoo.com; hongxiaxu@ahpu.edu.cn

Han-Ying Liang

Department of Mathematics, Tongji University, Shanghai 200092, P. R. China
e-mail: hyliang83@yahoo.com

Abstract: In this paper, the empirical likelihood inferences for partially time-varying coefficient errors-in-variables model with dependent observations are investigated. We propose an empirical log-likelihood ratio function for the regression parameters and show that its limiting distribution is a mixture of central chi-squared distributions. In order that the Wilks' phenomenon holds, we construct an adjusted empirical log-likelihood ratio for the regression parameters. The adjusted empirical log-likelihood is shown to have a standard chi-squared limiting distribution. Simulations show that the proposed confidence regions have satisfactory performance.

AMS 2000 subject classifications: Primary 62G15; secondary 62E20.
Keywords and phrases: Time-varying coefficient model, errors-in-variables model, empirical likelihood, confidence region, α-mixing.

Received October 2011.

1. Introduction

In recent years, there has been increasing interest and activity in the general area of time-varying models in statistics due to they have many applications in economics and finance. For example, the market model in finance relates the return of an individual stock to the return of a market index or another individual stock and the coefficient usually is said to be beta-coefficient in the capital assets pricing model (see [28], for example). Also, [29] investigated that beta-coefficient might vary over time; [1, 20, 25] studied parametric and nonparametric time-varying coefficient models. Recently, [11] introduced the following partially time-varying coefficient model

$$Y_i = X_i^T\beta + Z_i^T\alpha(t_i) + \varepsilon_i, \quad t_i = i/n, \quad i = 1, \dots, n, \tag{1.1}$$

where Y_i is the response, $\beta = (\beta_1, \dots, \beta_p)^T$ is a vector of p-dimensional unknown parameters, $\alpha(\cdot) = (\alpha_1(\cdot), \dots, \alpha_q(\cdot))^T$ is a q-dimensional vector of unspecified smooth coefficient functions, $X_i = (X_{i1}, \ldots, X_{ip})^T$ and $Z_i = (Z_{i1}, \ldots, Z_{iq})^T$ are p-dimensional and q-dimensional random vector respectively and ε_i is the random error.

It is obvious that the partially time-varying coefficient model (1.1) covers many interesting time series models (see [11], for more details). [11] recent applied the profile least squares approach to estimate the regression parameter and the nonlinear coefficient function in the model (1.1), and established the asymptotic normality of the proposed estimators under dependent assumptions, meanwhile, they also discussed the generalized likelihood ratio test for the model (1.1). Based on the asymptotic normality of the estimator for β obtained by [11], one can construct a confidence region for parameter β. However, in many situations, there exist covariate measurement errors. For example, it has been well documented that covariates such as blood pressure, urinary sodium chloride level, and exposure to pollutants are subject to measurement errors, and these cause difficulties in conducting a statistical analysis that involves them. If the covariate variable X_i is measured with error and is not directly observable, instead, X_i is observed through $\xi_i = X_i + \eta_i$, where η_i is the measurement error with mean zero. Specifically, we consider the following partially time-varying errors-in-variables (EV) model

$$\begin{cases} Y_i = X_i^T \beta + Z_i^T \alpha(t_i) + \varepsilon_i, \\ \xi_i = X_i + \eta_i, \end{cases} \quad i = 1, \ldots, n, \tag{1.2}$$

where η_i is independent and identically distributed (i.i.d.) with mean zero and covariance matrix Σ_η, and is independent of $(X_i, Z_i, \varepsilon_i)$. In the last 20 years, a lot of investigative effort has been dedicated to the EV models in the literature. Many studies were focused on the estimation of the involved parameter in the models, see [3, 13, 27, 4] for example. Recently, [35] studied the semiparametric varying-coefficient partially linear EV model with random t_i. When $\{t_i\}$ is fixed-design, then model (1.2) contains a deterministic time trend function, which implies that $\{Y_i\}$ in model (1.2) may be nonstationary. In practice, we often deal with nonstationary components when tackling econometric and financial issues from a time perspective. For recent development of parametric and nonparametric statistical inference with nonstationary time series, we refer to [19, 9] and the references therein.

To the best of our knowledge, little is known about the asymptotic normality and the construction of empirical likelihood (EL) confidence regions of β for the partially time-varying EV model (1.2). In order to construct the confidence regions for the unknown parameter β in the partially time-varying EV model (1.2), one can use normal approximation-based method or the EL method. But it is well known that the EL method, which was introduced by [17], has many advantages over normal approximation-based method. The most appealing features of the EL method is that one can construct confidence regions without estimating the covariance of the estimator and the EL method uses only the data to determine the shape and orientation of confidence regions. In addition, the confidence region derived from the limiting normal distribution is predetermined to be symmetric which may not be adequate when the underlying distribution is typically asymmetric. More discussion on advantages of the EL method over the existing methods can be found in the recent monograph of [18]. Therefore, the EL method has been used by many authors for various regression models, which include [21, 30] for partially linear regression model, [36] for semiparametric varying-coefficient partially linear regression model, [38] for partially linear single-index model, [12] for partially linear EV model, [6] for partially linear regression model with linear process errors, [7] for heteroscedastic partially linear regression model, [8, 32] for semiparametric varying-coefficient partially linear EV model and so on.

Most of known papers related to the EL method always assumed that the data are i.i.d. However, the independence assumption for the data is not always appropriate in applications, especially for sequentially collected economic data, which often exhibit evident dependence in the data. Recently, the EL method with dependent data has attracted increasing attention from statisticians. For example, [10, 2] studied blockwise EL method with strongly dependent data. [14] used the EL method to construct confidence intervals of conditional density with strong dependent data. [15] employed the EL method to construct confidence intervals for a conditional quantile with left-truncated and strong dependent data.

In this paper, our aim is to use the EL method to construct confidence regions of β in the partially time-varying coefficient EV model (1.2) with dependent observations. Since the empirical log-likelihood ratio has not a standard chi-squared limiting distribution, we further define an adjusted empirical log-likelihood ratio, which has standard chi-squared limiting distribution. Based on the results, one can construct immediately an approximate confidence region for the regression parameter.

Throughout we assume that $\{(X_i, Z_i, \varepsilon_i)\}$ is a sequence of stationary α-mixing random variables. Recall that a sequence $\{\xi_k, k \geq 1\}$ is said to be α-mixing if the α-mixing coefficient

$$\alpha(n) :\stackrel{\text{def}}{=} \sup_{k\geq 1} \sup\{|P(\mathcal{A}\cap\mathcal{B}) - P(\mathcal{A})P(\mathcal{B})| : \mathcal{A} \in \mathcal{F}_{n+k}^{\infty}, \mathcal{B} \in \mathcal{F}_1^k\}$$

converges to zero as $n \to \infty$, where $\mathcal{F}_a^b = \sigma\{\xi_i, a \leq i \leq b\}$ denotes the σ-algebra generated by $\xi_a, \xi_{a+1}, \ldots, \xi_b$. Among various mixing conditions used in the literature, the α-mixing is reasonably weak and is known to be fulfilled by many stochastic processes including many time series models. For example, [34] derived the conditions under which a linear process is α-mixing. In fact, under very mild assumptions linear autoregressive and more generally bilinear time series models are strongly mixing with mixing coefficients decaying exponentially, i.e., $\alpha(k) = O(\rho^k)$ for some $0 < \rho < 1$; see [5], page 99, for more details.

The rest of this paper is organized as follows. Section 2 proposes the empirical log-likelihood ratio function for β. Assumptions and asymptotic distribution of the adjusted empirical log-likelihood ratio are given in Section 3. Some simulation studies and a real-data example are conducted in Section 4. Section 5 gives the proof of the main results.

2. Empirical likelihood of the parametric components

In this section we present the empirical log-likelihood ratio for the parametric components in model (1.2). First, we give estimators of the nonparametric components $\alpha_j(\cdot),\ j=1,\ldots,q$. Our basic idea is as follows: suppose β is known, then the model (1.2) can be reduced to a varying-coefficient regression model which can be written as

$$Y_i - \sum_{j=1}^{p} X_{ij}\beta_j = \sum_{j=1}^{q} Z_{ij}\alpha_j(t_i) + \varepsilon_i, \quad 1 \le i \le n. \tag{2.1}$$

Note that $E(Z_i^T)\alpha(t_i) = E(Y_i - X_i^T\beta) = E(Y_i - \xi_i^T\beta)$. Hence, the varying coefficient functions $\{\alpha_j(\cdot), j=1,\ldots,q\}$ in (2.1) can be estimated by a local linear regression technique. That is, for t in a small neighborhood of t_0, one can approximated $\alpha_j(t)$ locally by a linear function

$$\alpha_j(t) \approx \alpha_j(t_0) + \alpha_j'(t_0)(t-t_0) \equiv a_j + b_j(t-t_0), \quad j=1,\ldots,q,$$

where $\alpha_j'(t) = d\alpha_j(t)/dt$. This leads to the following weighted local least-squares problem: find $\{(a_j,b_j), j=1,\ldots,q\}$ to minimize

$$\sum_{i=1}^{n}\left\{\left(Y_i - \sum_{j=1}^{p}\xi_{ij}\beta_j\right) - \sum_{j=1}^{q}[a_j + b_j(t_i-t_0)]Z_{ij}\right\}^2 K_h(t_i-t_0), \tag{2.2}$$

where $K_h(\cdot) = K(\cdot/h)/h$, $K(\cdot)$ is a kernel function and $h := h_n$ is a sequence of positive numbers tending to zero, called bandwidth.

Set $Y = (Y_1,\ldots,Y_n)^T$, $X = (X_1,\ldots,X_n)^T$, $\xi = (\xi_1,\ldots,\xi_n)^T$, $W(t) = \mathrm{diag}(K_h(t_1-t),\ldots,K_h(t_n-t))$ and

$$D(t) = \begin{pmatrix} Z_1^T & \frac{t_1-t}{h}Z_1^\tau \\ \vdots & \vdots \\ Z_n^T & \frac{t_n-t}{h}Z_n^\tau \end{pmatrix}.$$

Then the solution to problem (2.2) is given by

$$(\hat\alpha_1(t),\ldots,\hat\alpha_q(t), h\hat b_1(t),\ldots,h\hat b_q(t))^T = [D^T(t)W(t)D(t)]^{-1}D^T(t)W(t)(Y-\xi\beta).$$

Hence we can estimate the coefficient functions $\{\alpha_j(\cdot), j=1,\ldots,q\}$ by

$$\hat\alpha(t) = (\hat\alpha_1(t),\ldots,\hat\alpha_q(t))^T = (I_q\ \mathbf{0}_q)[D^T(t)W(t)D(t)]^{-1}D^T(t)W(t)(Y-\xi\beta),$$

where I_q is the $q\times q$ identity matrix and $\mathbf{0}_q$ is the $q\times q$ null matrix.

Let $\widetilde X = (I_n - S)X$, $\widetilde\xi = (I_n-S)\xi$, $\widetilde Y = (I_n-S)Y$ and

$$H_i(\beta) = \big[\xi_i - E(\xi_i|Z_i)\big]\big[Y_i - E(Y_i|Z_i) - [\xi_i - E(\xi_i|Z_i)]^T\beta\big] + \Sigma_\eta\beta,\ i=1,\ldots,n,$$

where I_n is the $n \times n$ identity matrix,

$$S = \begin{pmatrix} (Z_1^T \quad \mathbf{0}^T)[D^T(t_1)W(t_1)D(t_1)]^{-1}D^T(t_1)W(t_1) \\ \vdots \\ (Z_n^T \quad \mathbf{0}^T)[D^T(t_n)W(t_n)D(t_n)]^{-1}D^T(t_n)W(t_n) \end{pmatrix}$$

and $\mathbf{0}$ is the $q \times 1$ null vector.

It is easy to verify that $E(H_i(\beta)) = 0$ for $i = 1, \ldots, n$, when β is the true parameter. Hence, the problem of testing whether β is the true parameter is equivalent to testing whether $E(H_i(\beta)) = 0$ for $i = 1, \ldots, n$. By [17], this can be done using the EL. However $E(Y_i|Z_i)$ and $E(\xi_i|Z_i)$ are usually unknown. Thus we should replace $E(Y_i|Z_i)$ and $E(\xi_i|Z_i)$ by their estimators $S_i^T Y$ and $S_i^T \xi$, respectively, where S_i is the $i-th$ row of the matrix S. We now introduce the auxiliary random vector

$$\eta_i(\beta) = \widetilde{\xi}_i[Y_i - \xi_i^T\beta - Z_i^T\hat{\alpha}(t_i)] + \Sigma_\eta\beta = \widetilde{\xi}_i(\widetilde{Y}_i - \widetilde{\xi}_i^T\beta) + \Sigma_\eta\beta. \tag{2.3}$$

Therefore, an empirical log-likelihood ratio function for β is defined as

$$\mathcal{L}_n(\beta) = -2\max\Big\{\sum_{i=1}^n \log(np_i) : \sum_{i=1}^n p_i\eta_i(\beta) = 0, p_i \geq 0, \sum_{i=1}^n p_i = 1\Big\}.$$

By the Lagrange multiplier method, one can obtain that $p_i = \frac{1}{n[1+\lambda^T\eta_i(\beta)]}$, and $\mathcal{L}_n(\beta)$ can be represented as

$$\mathcal{L}_n(\beta) = 2\sum_{i=1}^n \log\{1 + \lambda^T\eta_i(\beta)\}, \tag{2.4}$$

where $\lambda(\beta)$ is determined by

$$\frac{1}{n}\sum_{i=1}^n \frac{\eta_i(\beta)}{1 + \lambda^T\eta_i(\beta)} = 0. \tag{2.5}$$

3. Main results

In order to formulate the main results, we begin this section with making the following assumptions. These assumptions are quite mild and can be easily satisfied. They are also assumed in [11]. For convenience, let $V_Z = E[Z_1Z_1^T]$, $V_X = E[X_1X_1^T]$, $V_{ZX} = E[Z_1X_1^T]$ and $A^{\otimes 2} = AA^T$.

(A1) The kernel $K(\cdot)$ is a symmetric and Lipschitz continuous function with a compact support $[-1, 1]$.

(A2) (i) The matrices V_Z and V_X are no-singular, $E[\|Z_1\|^{2s}] < \infty$, $E[\|X_1\|^{2s}] < \infty$ and $E[\|\eta_1\|^{2s}] < \infty$ for some $s > 4$, where $\|\cdot\|$ is the L_2 norm.

(ii) $E[\varepsilon_1|X_1, Z_1] = 0$, $E[|\varepsilon_1|^{2s}|X_1, Z_1] < \infty$ $a.s.$ and $E\varepsilon_1^2 = \sigma^2$.

(A3) $\{\alpha_j(\cdot), j = 1, \ldots, q\}$ have continuous second derivatives.

(A4) The α-mixing coefficient $\alpha(k)$ and the bandwidth h satisfy that

$$\alpha(k) = O(k^{-\gamma}),\ \gamma > \frac{19s+12}{2(s-4)},\ s > 4 \ \text{ and } \ \frac{nh^2}{\log^2 n} \to \infty, \quad nh^8 \to 0.$$

With the assumptions above, we are ready to establish the main results.

Theorem 3.1. *Suppose that (A1)-(A4) hold. For model (1.2), let β_0 be the true value of the parameter β. Then we have*

$$\mathcal{L}_n(\beta_0) \xrightarrow{d} w_1\chi^2_{1,1} + \cdots + w_p\chi^2_{1,p} \quad \text{as} \quad n \to \infty, \tag{3.1}$$

where $\chi^2_{1,i}(1 \le i \le p)$ are p independent standard chi-square random variables with 1 degree of freedom, $w_i(1 \le i \le p)$ are eigenvalues of $D = \Sigma_1^{-1}\Sigma$ with

$$\begin{aligned}\Sigma = &\lim_{n\to\infty} \mathrm{Var}\Big[\frac{1}{\sqrt{n}}\sum_{i=1}^n (X_i - V_{ZX}^T V_Z^{-1} Z_i)\varepsilon_i\Big] + (V_X - V_{ZX}^T V_Z^{-1} V_{ZX})\beta_0^T\Sigma_\eta\beta_0 \\ &+ \sigma^2\Sigma_\eta + E\{(\eta_1\eta_1^T - \Sigma_\eta)\beta_0\}^{\otimes 2} \quad \text{and} \\ \Sigma_1 = &E\big[(X_1 - V_{ZX}^T V_Z^{-1} Z_1)(X_1 - V_{ZX}^T V_Z^{-1} Z_1)^T\varepsilon_1^2\big] \\ &+ \big(V_X - V_{ZX}^T V_Z^{-1} V_{ZX}\big)\beta_0^T\Sigma_\eta\beta_0 + \sigma^2\Sigma_\eta + E\{(\eta_1\eta_1^T - \Sigma_\eta)\beta_0\}^{\otimes 2}.\end{aligned}$$

Further, D is a consistent estimator of $\widehat{\Sigma}_{1n}^{-1}(\beta)\widehat{\Sigma}_n(\beta)$, where

$$\widehat{\Sigma}_n(\beta) = \frac{1}{n}\Big\{\sum_{i=1}^n \big[\tilde{\xi}_i(\tilde{Y}_i - \tilde{\xi}_i^T\beta) + \Sigma_\eta\beta\big]\Big\}^{\otimes 2}$$

and

$$\widehat{\Sigma}_{1n}(\beta) = \frac{1}{n}\sum_{i=1}^n \Big[\tilde{\xi}_i(\tilde{Y}_i - \tilde{\xi}_i^T\beta) + \Sigma_\eta\beta\Big]^{\otimes 2}.$$

Remark 3.1. If $(X_i, Z_i, \varepsilon_i)$ is a independent sequence and $E[\varepsilon_1^2|X_1, Z_1] = \sigma^2$ *a.s.*, then

$$\Sigma = \big(V_X - V_{ZX}^T V_Z^{-1} V_{ZX}\big)\{\sigma^2 + \beta_0^T\Sigma_\eta\beta_0\} + \sigma^2\Sigma_\eta + E\{(\eta_1\eta_1^T - \Sigma_\eta)\beta_0\}^{\otimes 2} = \Sigma_1.$$

So the conclusion of Theorem 3.1 becomes $\mathcal{L}_n(\beta_0) \xrightarrow{d} \chi^2_p$.

Remark 3.2. Theorem 3.1 can be used to construct the confidence regions of β_0 if the unknown weights $w_i(1 \le i \le n)$ can be estimated (see [22], for example). For Σ and Σ_1, we can take their consistent estimators $\widehat{\Sigma}_n(\hat{\beta}_I)$ and $\widehat{\Sigma}_{1n}(\hat{\beta}_I)$ respectively, where $\hat{\beta}_I$ is the modified profile least square estimator (see [35]) defined by

$$\hat{\beta}_I = (\widetilde{\xi}^T\widetilde{\xi} - n\Sigma_\eta)^{-1}\widetilde{\xi}^T\widetilde{Y}.$$

However, Σ_η may be unknown in practice and must be estimated. The usual method of doing so is by partial replication as mentioned by [13]. That is, suppose there are k_i replicate measurements of ξ_i, and $\bar{\xi}_i$ is their mean, so that we observe $\xi_{ij} = X_i + \eta_{ij}, j = 1, 2 \ldots, k_i$. Then a consistent estimator for Σ_η is

$$\hat{\Sigma}_\eta = \frac{\sum_{i=1}^n \sum_{j=1}^{k_i} (\xi_{ij} - \bar{\xi}_i)(\xi_{ij} - \bar{\xi}_i)^T}{\sum_{i=1}^n (k_i - 1)}.$$

Then we get the consistent estimator of D, i.e. $\widehat{D}_n = \widehat{\Sigma}_{1n}^{-1}(\hat{\beta}_I)\widehat{\Sigma}_n(\hat{\beta}_I)$. When Σ_η is estimated by the partial replication, Σ_η should be substituted by $\hat{\Sigma}_\eta$ and the estimator of β, i.e. $\hat{\beta}_I$ has to be modified too to accommodate the replicates. See [13] for details. Therefore the consistent estimators $\hat{w}_{i,n}$ of w_i are the eigenvalues of $\widehat{D}_n$. Let $c_{\alpha,n}$ be the $1-\alpha$ quantile of the conditional distribution of $\hat{w}_{1,n}\chi^2_{1,1} + \cdots + \hat{w}_{p,n}\chi^2_{1,p}$ given the data (X_i, Z_i, t_i, Y_i) for $i = 1, \ldots, n$. In practice, one can get $c_{\alpha,n}$ through Monte Carlo simulation. Then the confidence region with asymptotically correct coverage probability $1-\alpha$ can be defined as $\mathscr{C}_\alpha(\beta) = \{\beta \in R^p : \mathcal{L}_n(\beta) \le c_{\alpha,n}\}$.

Although Theorem 3.1 gives the confidence regions for β_0, it increases the burden of computing, this can be seen clearly from Remark 3.2. To attack this difficulty, we present an adjusted EL function which has an asymptotic standard chi-squared distribution. Let

$$\mathbb{S}_n(\beta) = \Big(\sum_{i=1}^n \eta_i(\beta)\Big)\Big(\sum_{i=1}^n \eta_i(\beta)\Big)^T.$$

Based on the idea of [24], the adjusted EL function can be defined as $\widehat{\mathcal{L}}_{n,ad}(\beta) = R_n(\beta)\mathcal{L}_n(\beta)$, where $R_n(\beta) = \mathrm{tr}[\widehat{\Sigma}_n^{-1}(\beta)\mathbb{S}_n(\beta)]/\mathrm{tr}[\widehat{\Sigma}_{1n}^{-1}(\beta)\mathbb{S}_n(\beta)]$, $\widehat{\Sigma}_n(\beta)$ and $\widehat{\Sigma}_{1n}(\beta)$ are defined in Theorem 3.1. The adjusted EL function has the following Wilks' phenomenon.

Theorem 3.2. *Under the conditions of (A1)-(A4), for model (1.2), if β_0 is the true value of the parameter, then*

$$\widehat{\mathcal{L}}_{n,ad}(\beta_0) \xrightarrow{d} \chi^2_p \quad as \quad n \to \infty,$$

where χ^2_p is a chi-square distributed with p degrees of freedom.

As a conclusion of Theorem 3.2, the confidence region for the parameter vector β can be constructed. For any $0 < \alpha < 1$, let c_α be the $1-\alpha$ quantile of chi-square distribution such that $P(\chi^2_p > c_\alpha) = \alpha$, then

$$\widehat{\mathscr{C}}_{\alpha,ad}(\beta) = \{\beta \in R^p : \widehat{\mathcal{L}}_{n,ad}(\beta) \le c_\alpha\}$$

constitutes a confidence region for β_0 with asymptotically correct coverage probability $1-\alpha$. If Σ_η is unknown, we can use its estimator $\hat{\Sigma}_\eta$ to replace it.

Remark 3.3. When $(X_i, Z_i, \varepsilon_i)$ is an independent sequence, by the proof of Lemmas 5.4-5.5, one can easily get $\widehat{\Sigma}_n(\beta) - \widehat{\Sigma}_{1n}(\beta) \xrightarrow{p} 0$, which implies $R_n(\beta) \xrightarrow{p} 1$. Thus, the adjusted EL asymptotically reduces to the EL for the partially time-varying coefficient EV model (1.2) with the absence of α-mixing.

4. Simulation study

In this section, some simulated examples and a real data example are provided to demonstrate the performance of the EL method in the paper.

4.1. Simulated examples

We first carry out some simulations to show the finite sample performance of the proposed EL confidence region of β_0.

Firstly, consider the following partially time-varying coefficient EV model:

$$\begin{cases} Y_i = X_i^T\beta_0 + Z_i^T\alpha(t_i) + \varepsilon_i, \\ \xi_i = X_i + \eta_i, \end{cases} \qquad i = 1, \ldots, n, \tag{4.1}$$

where $\beta_0 = (\beta_1, \beta_2)^T = (\sqrt{2}, \sqrt{3})^T/2$, $t_i = i/n$, $\alpha_1(t) = \sin(2\pi t)$, $\alpha_2(t) = (1 - e^{-2t})^{2/3}$, the measurement error $\eta_i \sim N(0, \Sigma_\eta)$ with $\Sigma_\eta = 0.2^2 I_2$, $0.4^2 I_2$ and $0.6^2 I_2$ respectively, I_2 is 2×2 identity matrix, ε_i, X_i and Z_i are generated by the AR(1) model as follows,

$$\begin{aligned} &\varepsilon_i = 0.5\varepsilon_{i-1} + \varrho_i, && \varrho_i \overset{i.i.d}{\sim} N(0, 0.25^2), \\ &X_{i,1} = 0.6X_{i-1,1} + u_{i,1}, && X_{i,2} = 0.4X_{i-1,2} + u_{i,2}, \\ &Z_{i,1} = 0.8Z_{i-1,1} + e_{i,1}, && Z_{i,2} = 0.2Z_{i-1,2} + e_{i,2}, \end{aligned}$$

where $u_i = (u_{i,1}, u_{i,2})^T \overset{i.i.d}{\sim} N((0,0)^T, \mathrm{diag}(1,1))$ and $e_i = (e_{i,1}, e_{i,2})^T \overset{i.i.d}{\sim} N((0,0)^T, \mathrm{diag}(1,1))$. It is easy to verify that $\{X_i, Z_i, \varepsilon_i\}$ is stationary and α-mixing.

For the weight function, we use the Gaussian kernel $K(t) = \exp(-t^2/2)/\sqrt{2\pi}$. The "leave-one-subject-out" cross-validation bandwidth h_{CV} is obtained by minimizing

$$CV(h) = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \xi_i^T\hat{\beta}_{[i]} - Z_i^T\hat{\alpha}_{[i]}(t_i)\right)^2,$$

where $\hat{\beta}_{[i]}$ and $\hat{\alpha}_{[i]}(\cdot)$ are estimators of β and $\alpha(\cdot)$ respectively, which are computed with all of the measurements but not the ith subject. The sample sizes n are chosen to be 100 and 150 respectively. The coverage probabilities and the average lengths of the confidence intervals for individual $\beta_i (i = 1, 2)$ are calculated based on 1000 replications with the nominal level $1 - \alpha = 0.90$ and 0.95 for three cases $\Sigma_\eta = 0.2^2 I_2$, $0.4^2 I_2$ and $0.6^2 I_2$. Some representative coverage probabilities and average lengths of confidence intervals are reported in Table 1.

Secondly, we generate ε_i, X_i and Z_i by the MA(1) model,

$$\begin{aligned} &\varepsilon_i = \varrho_i - 0.5\varrho_{i-1}, && \varrho_i \overset{i.i.d}{\sim} N(0, 0.2^2), \\ &X_{i,1} = u_{i,1} - 0.2u_{i-1,1}, && X_{i,2} = u_{i,2} - 0.5u_{i-1,2}, \\ &Z_{i,j} = e_{i,j} - 0.5e_{i-1,j}, && j =, 1, 2. \end{aligned}$$

TABLE 1
Coverage probabilities (CP) and average lengths (AL) of the confidence intervals with $\beta_0 = (\sqrt{2}, \sqrt{3})^T/2$

Σ_η	Parameters	n	$1-\alpha = 0.90$		$1-\alpha = 0.95$	
			CP	AL	CP	AL
$0.2^2 I_2$	β_1	100	0.886	0.1268	0.934	0.1764
		150	0.888	0.1103	0.942	0.1471
	β_2	100	0.877	0.1795	0.932	0.2513
		150	0.886	0.1604	0.948	0.2109
$0.4^2 I_2$	β_1	100	0.881	0.1609	0.923	0.2061
		150	0.885	0.1411	0.932	0.1712
	β_2	100	0.874	0.2490	0.930	0.3208
		150	0.884	0.2191	0.946	0.2705
$0.6^2 I_2$	β_1	100	0.860	0.2088	0.910	0.2572
		150	0.883	0.1746	0.936	0.2153
	β_2	100	0.836	0.3921	0.918	0.4543
		150	0.874	0.3116	0.934	0.3547

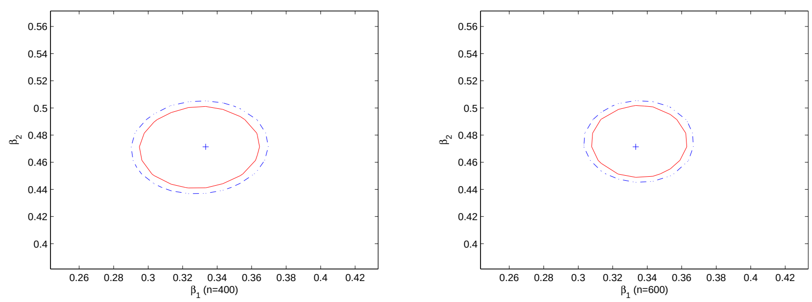

FIG 1. *The confidence regions of* $\beta_0 = (\beta_1, \beta_2)^T = (1, \sqrt{2})^T/3$ *with* $\Sigma_\eta = 0.2^2 I_2$ *for* $1-\alpha = 0.90$ *(solid curve) and* $1-\alpha = 0.95$ *(dashed curve).*

$\beta_0 = (1, \sqrt{2})^T/3$ and the other variables are taken as the same as that of model (4.1). In this example, we consider the confidence region of (β_1, β_2) with nominal confidence level $1-\alpha = 0.90$ and 0.95 for $n = 400$ and 600 respectively. The simulation results for the plots are presented in Figures 1-3 for $\Sigma_\eta = 0.2^2 I_2$, $0.4^2 I_2$ and $0.6^2 I_2$, respectively.

Table 1 and Figures 1-3 indicate the following four simulation results.

(1) From Table 1, in the three different choice of the measurement error, the coverage probabilities tend to nominal level $1-\alpha$ and the average lengths decrease as the sample size n increases both for $1-\alpha = 0.90$ and 0.95 cases.

(2) For the same sample size n and nominal level $1-\alpha$, Table 1 presents a comparison for the three cases of the measurement error covariance $\Sigma_\eta = 0.2^2 I_2$, $0.4^2 I_2$ and $0.6^2 I_2$. It is obvious to see that the first choice gives the best performance while the third choice offers the worst. Furthermore, the coverage probabilities of the confidence intervals tend to decrease and the average lengths increase as Σ_η gets larger.

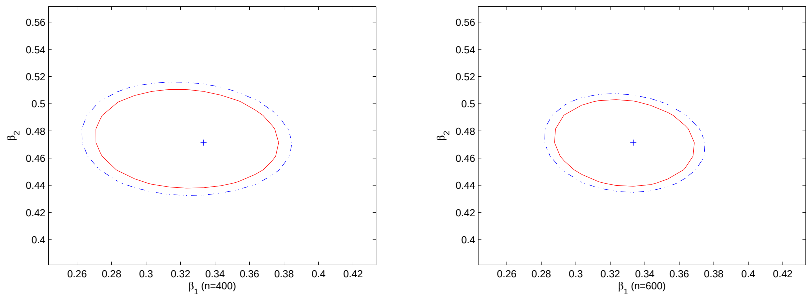

FIG 2. *The confidence regions of* $\beta_0 = (\beta_1, \beta_2)^T = (1, \sqrt{2})^T/3$ *with* $\Sigma_\eta = 0.4^2 I_2$ *for* $1 - \alpha = 0.90$ *(solid curve) and* $1 - \alpha = 0.95$ *(dashed curve).*

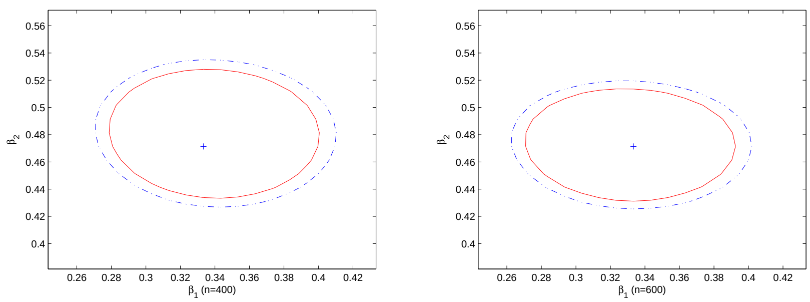

FIG 3. *The confidence regions of* $\beta_0 = (\beta_1, \beta_2)^T = (1, \sqrt{2})^T/3$ *with* $\Sigma_\eta = 0.6^2 I_2$ *for* $1 - \alpha = 0.90$ *(solid curve) and* $1 - \alpha = 0.95$ *(dashed curve).*

(3) In Figures 1-3, the confidence regions become narrower as n increases for three different choices of Σ_η. It is also interesting to note that the confidence regions of β_0 with $1 - \alpha = 0.95$ are wider than the case of $1 - \alpha = 0.90$ for three cases of Σ_η.

(4) In Figures 1-3, for the same sample size n and nominal level $1 - \alpha$, it is easy to see that the confidence regions in the case $\Sigma_\eta = 0.2^2 I_2$ are the narrowest and the confidence regions in the case $\Sigma_\eta = 0.6^2 I_2$ are the widest. That is, $\Sigma_\eta = 0.2^2 I_2$ gives the best performance while $\Sigma_\eta = 0.6^2 I_2$ offers the worst.

4.2. An application of the proposed method

We next study an Sydney CPI data set, which ia available from The Australian Bureau of Statistics (www.abs.gov.au). We use three quarterly CPI data series in Sydney: 'All groups', 'Food and non-alcoholic beverages' and 'Bread and cereal products' during the period 1981 to 2010. The data series are plotted in Figure 4.

Let y_k, Z_{k1} and Z_{k2} denote all-groups CPI variable, food and non-alcoholic beverages CPI variable and bread and cereal products CPI variable at time

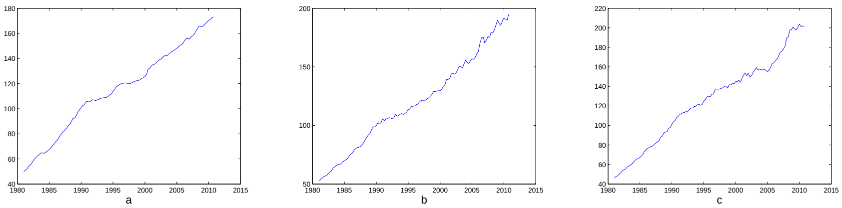

FIG 4. *(a) All groups CPI series; (b) Food and non-alcoholic beverages CPI series; (c) Bread and cereal products CPI series.*

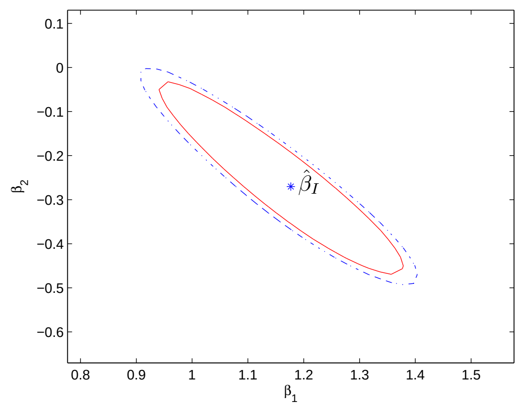

FIG 5. *The confidence regions of β with $1-\alpha=0.90$ (solid curve) and $1-\alpha=0.95$ (dashed curve).*

k, respectively. Denote $Z_k=(Z_{k1},Z_{k2})^T$ and $X_k=(y_{k-1},y_{k-2})^T$, where X_k stands for the lagged variables of the all-groups CPI variable y_k. In this example it is assumed that X_k is measured with additive error and Z_k is error free, i.e., we cannot observe X_k but we can observe ξ_k with $\xi_k=X_k+\eta_k$, where $\xi_k=(Y_{k-1},Y_{k-2})^T$ and Y_k is the observed value of all-groups index. We now use the following partially time-varying coefficient EV model to fit the data,

$$\begin{cases} Y_k = X_k^T\beta + Z_k^T\alpha(t_k)+\varepsilon_k, \\ \xi_k = X_k+\eta_k, \end{cases} \qquad k=1,\dots,120. \tag{4.2}$$

In order to identify the model, we assume $\Sigma_\eta=0.1^2 I_2$. The plots of confidence regions for β with $1-\alpha=0.90$ and 0.95 are put in Figure 5. From Remark 3.2, the modified profile least square estimator of β is $\hat{\beta}_I=(0.9208,-0.0769)^T$. It can be seen from Figure 5 that the EL confidence regions of β perform well and this may indicate the influence of government regulation. For example, if the CPI in the previous quarter increases too quickly, the government will usually take some measures to stabilize prices.

5. Proofs of main results

For the convenience, let $M = (M_1, \ldots, M_n)^T = \big(\alpha^T(t_1)Z_1, \ldots, \alpha^T(t_n)Z_n\big)^T$, $\widetilde{M} = (I-S)M$, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^T$, $\eta = (\eta_1, \ldots, \eta_n)^T$, $\widetilde{\varepsilon} = (I-S)\varepsilon$, $\widetilde{\eta} = (I-S)\eta$, $c_n = \{\log n/(nh)\}^{1/2} + h^2$ and C denote positive constant whose value may vary at each occurrence. Before proving the main theorems, we give a series of lemmas.

Lemma 5.1. **([11], Lemma A.2)** *Assume that conditions (A1)-(A4) are satisfied. Then*

$$\sup_{t\in(0,1)} \Big|\frac{1}{n}D^T(t)W(t)D(t) - V_Z \otimes \Lambda_1\Big| = O_p\bigg(\Big(\frac{\log n}{nh}\Big)^{1/2}\bigg)$$

and

$$\sup_{t\in(0,1)} \Big|\frac{1}{n}D^T(t)W(t)X - V_{ZX} \otimes \Omega_1\Big| = O_p\bigg(\Big(\frac{\log n}{nh}\Big)^{1/2}\bigg),$$

where $\Lambda_1 = \begin{pmatrix}\mu_0 & 0\\ 0 & \mu_2\end{pmatrix}, \Omega_1 = \begin{pmatrix}\mu_0\\ 0\end{pmatrix}$, $\mu_k = \int t^k K(t)dt$ *and* $\otimes$ *is the Kronecker product.*

Lemma 5.2. **([35], Lemma A.2)** *Let $D_1, \ldots, D_n$ be i.i.d. random variables. If $E|D_i|^s$ is bounded for $s > 1$, then $\max_{1\le i\le n}|D_i| = o(n^{1/s})$ a.s..*

Lemma 5.3. *Under the assumptions of (A1)-(A4), we have*

$$\max_{1\le l\le n}\|U_l\| = O_p(n^{1/2s}) \quad \textit{and} \quad \max_{1\le l\le n}\|\widetilde{U}_l\| = O_p(n^{1/2s}), \ \textit{where } U_l = X_l,\ Z_l.$$

Proof. Here we prove only the case $U_l = X_l$, the proof of the other case is analogous. According to Markov inequality and (A1), for all $n \ge 1$, we have

$$P\Big(\max_{1\le l\le n}\|X_l\| > c_0 n^{1/2s}\Big) \le \frac{1}{c_0^{2s} n}\sum_{l=1}^n E\|X_l\|^{2s} \le \frac{C}{c_0^{2s}} \to 0, \quad \text{as} \quad c_0 \to \infty,$$

which yields that $\max_{1\le l\le n}\|X_l\| = O_p(n^{1/2s})$.

As to $\max_{1\le l\le n}\|\tilde{X}_l\| = O_p(n^{1/2s})$, from Lemma 5.1, we have

$$\begin{aligned}\max_{1\le i\le n}\|\widetilde{X}_i\| &= \max_{1\le i\le n}\Big\|X_i - (Z_i^T\ \mathbf{0}^T)[D^T(t_i)W(t_i)D(t_i)]^{-1}D^T(t_i)W(t_i)X\Big\|\\ &= \max_{1\le i\le n}\Big\|X_i - V_{ZX}^T V_Z^{-1} Z_i\{1 + O_p(c_n)\}\Big\| = O_p(n^{1/2s}). \qquad (5.1)\end{aligned}$$

This completes the proof. □

Lemma 5.4. *Suppose that conditions (A1)-(A4) hold. Then we have*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n \eta_i(\beta_0) \xrightarrow{d} N(0, \Sigma),$$

where Σ is defined in Theorem 3.1.

Proof. From (2.3) and a simple calculation yields

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i(\beta_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big\{\widetilde{X}_i\tilde{\varepsilon}_i - \widetilde{X}_i\tilde{\eta}_i^T\beta_0 + \widetilde{X}_i\widetilde{M}_i + \tilde{\eta}_i\tilde{\varepsilon}_i - \tilde{\eta}_i\tilde{\eta}_i^T\beta_0 + \tilde{\eta}_i\widetilde{M}_i + \Sigma_\eta\beta_0\Big\}. \tag{5.2}$$

First, we establish that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widetilde{X}_i\tilde{\varepsilon}_i = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big[X_i - V_{ZX}^T V_Z^{-1}Z_i\Big]\varepsilon_i + o_p(1). \tag{5.3}$$

From the proof of Theorem 3.3 in [11], we have

$$\sup_{t\in(0,1)} e_{l,2q}^T[D^T(t)W(t)D(t)]^{-1}D^T(t)W(t)\varepsilon = O_p\Big(\big(\frac{\log n}{nh}\big)^{1/2}\Big),$$

where $e_{l,2q}$ is the $2q$-dimensional vector with 1 in the l-th position and 0 elsewhere. Then by Lemma 5.1 and equation (5.1), one can derive that

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widetilde{X}_i\tilde{\varepsilon}_i \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big[X_i - V_{ZX}^T V_Z^{-1}Z_i + O_p(c_n)\|Z_i\|\Big]\cdot\Big[\varepsilon_i + O_p\Big(\big(\frac{\log n}{nh}\big)^{1/2}\Big)Z_i^T\Big] \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big[X_i - V_{ZX}^T V_Z^{-1}Z_i\Big]\varepsilon_i + O_p\Big(\big(\frac{\log n}{nh}\big)^{1/2} + \sqrt{n}c_n\big(\frac{\log n}{nh}\big)^{1/2}\Big) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big[X_i - V_{ZX}^T V_Z^{-1}Z_i\Big]\varepsilon_i + o_p(1).
\end{aligned} \tag{5.4}$$

Similar to the proof of (5.4), from (A.21) in [11], Lemma 5.1 and equation (5.1), it is easily to prove that

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widetilde{X}_i\tilde{\eta}_i^T = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big[X_i - V_{ZX}^T V_Z^{-1}Z_i\Big]\eta_i^T + o_p(1), \\
&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\eta}_i\tilde{\varepsilon}_i = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i\varepsilon_i + o_p(1), \quad \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\eta}_i\tilde{\eta}_i^T = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i\eta_i^T + o_p(1), \\
&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widetilde{X}_i\widetilde{M}_i = o_p(1) \quad \text{and} \quad \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\eta}_i\widetilde{M}_i = o_p(1).
\end{aligned}$$

Therefore, we obtain that

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i(\beta_0) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big\{\big[X_i - V_{ZX}^T V_Z^{-1}Z_i\big](\varepsilon_i - \eta_i^T\beta_0) + \eta_i\varepsilon_i + (\Sigma_\eta - \eta_i\eta_i^T)\beta_0\Big\} + o_p(1)
\end{aligned}$$

$$:= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} H_i(\beta_0) + o_p(1). \tag{5.5}$$

Note that η_i is i.i.d. with mean zero and is independent of $(X_i, Z_i, \varepsilon_i)$, which together with (A2), yields that $\lim_{n\to\infty} \mathrm{Var}\big(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} H_i(\beta_0)\big) = \Sigma$. Then, applying Theorem 3.2.1 in [16] and noting that $\{H_i(\beta_0), i \geq 1\}$ is α-mixing with mixing coefficient $\alpha(k)$, we have

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} H_i(\beta_0) \xrightarrow{d} N(0, \Sigma),$$

which, combined with (5.5), completes the proof of Lemma 5.4. $\square$

Lemma 5.5. *Suppose that Assumptions (A1)-(A4) hold. Then we have*

$$\frac{1}{n} \sum_{i=1}^{n} \eta_i(\beta_0)\eta_i^T(\beta_0) \xrightarrow{p} \Sigma_1,$$

where Σ_1 is defined in Theorem 3.1.

Proof. From the definition of $\eta_i(\beta_0)$, according to (A.21) in [11], Lemmas 5.1-5.3 and (5.2), similar to the proof of (5.3), we can derive

$$\begin{aligned}
&\frac{1}{n} \sum_{i=1}^{n} \eta_i(\beta_0)\eta_i^T(\beta_0) \\
&= \frac{1}{n} \sum_{i=1}^{n} \Big\{ \widetilde{X}_i(\tilde{\varepsilon}_i - \tilde{\eta}_i^T\beta_0) + (\widetilde{X}_i + \tilde{\eta}_i)\widetilde{M}_i + \tilde{\eta}_i\tilde{\varepsilon}_i + (\Sigma_\eta - \tilde{\eta}_i\tilde{\eta}_i^T)\beta_0 \Big\}^{\otimes 2} \\
&= \frac{1}{n} \sum_{i=1}^{n} \widetilde{X}_i\widetilde{X}_i^T(\tilde{\varepsilon}_i - \tilde{\eta}_i^T\beta_0)^2 + \frac{1}{n} \sum_{i=1}^{n} \tilde{\eta}_i\tilde{\eta}_i^T\tilde{\varepsilon}_i^2 + \frac{1}{n} \sum_{i=1}^{n} \big\{(\Sigma_\eta - \tilde{\eta}_i\tilde{\eta}_i^T)\beta_0\big\}^{\otimes 2} + o_p(1) \\
&:= I_{1n} + I_{2n} + I_{3n} + o_p(1).
\end{aligned} \tag{5.6}$$

For I_{1n}, analogously to the proof of Lemma 5.4, from (5.3) and Lemma 5.3, it follows that

$$\begin{aligned}
I_{1n} &= \frac{1}{n} \sum_{i=1}^{n} \Big[X_i - V_{ZX}^T V_Z^{-1} Z_i + O_p(c_n)\|Z_i\|\Big] \\
&\quad \cdot \Big[X_i - V_{ZX}^T V_Z^{-1} Z_i + O_p(c_n)\|Z_i\|\Big]^T \cdot \Big[\varepsilon_i - \eta_i^T\beta_0 + O_p\Big(\big(\frac{\log n}{nh}\big)^{1/2}\Big)\|Z_i\|\Big]^2 \\
&= \frac{1}{n} \sum_{i=1}^{n} [X_i - V_{ZX}^T V_Z^{-1} Z_i][X_i - V_{ZX}^T V_Z^{-1} Z_i]^T \{\varepsilon_i^2 + (\eta_i^T\beta_0)^2\} + o_p(1).
\end{aligned} \tag{5.7}$$

Not that $\big\{[X_i - V_{ZX}^T V_Z^{-1} Z_i][X_i - V_{ZX}^T V_Z^{-1} Z_i]^T(\varepsilon_i^2 + (\eta_i^T\beta_0)^2), i \geq 1\big\}$ is α-mixing. Then according to the strong law of large numbers for α-mixing sequence (see

Remark 8.2.3. in [16]), one can verify easily that

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\big[X_i - V_{ZX}^T V_Z^{-1} Z_i\big]\big[X_i - V_{ZX}^T V_Z^{-1} Z_i\big]^T\{\varepsilon_i^2 + (\eta_i^T\beta_0)^2\} \\
&\xrightarrow{a.s.} \frac{1}{n}\sum_{i=1}^{n} E\Big\{\big[X_i - V_{ZX}^T V_Z^{-1} Z_i\big]\big[X_i - V_{ZX}^T V_Z^{-1} Z_i\big]^T\{\varepsilon_i^2 + (\eta_i^T\beta_0)^2\}\Big\} \\
&= E\big[(X_i - V_{ZX}^T V_Z^{-1} Z_i)(X_i - V_{ZX}^T V_Z^{-1} Z_i)^T\varepsilon_i^2\big] \\
&\quad + E\big[(X_i - V_{ZX}^T V_Z^{-1} Z_i)(X_i - V_{ZX}^T V_Z^{-1} Z_i)^T\big]\beta_0^T\Sigma_\eta\beta_0,
\end{aligned}$$

which, combining with (5.7), proves that $I_{1n} \xrightarrow{p} E\big[(X_i - V_{ZX}^T V_Z^{-1} Z_i)(X_i - V_{ZX}^T V_Z^{-1} Z_i)^T\varepsilon_i^2\big] + E\big[(X_i - V_{ZX}^T V_Z^{-1} Z_i)(X_i - V_{ZX}^T V_Z^{-1} Z_i)^T\big]\beta_0^T\Sigma_\eta\beta_0$.

Similarly, we can derive that

$$I_{2n} \xrightarrow{p} \sigma^2\Sigma_\eta \quad \text{and} \quad I_{3n} \xrightarrow{p} E[(\eta_1\eta_1^T - \Sigma_\eta)\beta_0]^{\otimes 2}.$$

Therefore, $\frac{1}{n}\sum_{i=1}^{n}\eta_i(\beta_0)\eta_i^T(\beta_0) \xrightarrow{p} \Sigma_1$. So the proof of Lemma 5.5 is completed. □

Lemma 5.6. *Suppose that Assumptions (A1)-(A5) hold. Then*

$$\max_{1\le i\le n}\|\eta_i(\beta_0)\| = o_p(n^{1/2}), \tag{5.8}$$

$$\lambda = O_p(n^{-1/2}). \tag{5.9}$$

Proof. From (5.2), we have

$$\begin{aligned}
\max_{1\le i\le n}\|\eta_i(\beta_0)\| \le &\max_{1\le i\le n}\|\widetilde{X}_i\|\cdot\max_{1\le i\le n}\|\tilde{\varepsilon}_i + \widetilde{M}_i - \tilde{\eta}_i^T\beta_0\| \\
&+ \max_{1\le i\le n}\|\tilde{\eta}_i\|\cdot\max_{1\le i\le n}\|\tilde{\varepsilon}_i + \widetilde{M}_i\| + C\max_{1\le i\le n}\|\tilde{\eta}_i\tilde{\eta}_i^T - \Sigma_\eta\|.
\end{aligned}$$

Similar to the proof of (5.1), from Lemmas 5.1-5.3 and (A2), one can derive that

$$\begin{aligned}
&\max_{1\le i\le n}\|\widetilde{X}_i\| = O_p(n^{1/(2s)}), \quad \max_{1\le i\le n}\|\tilde{\eta}_i\| = O_p(n^{1/(2s)}), \\
&\max_{1\le i\le n}\|\tilde{\varepsilon}_i + \widetilde{M}_i - \tilde{\eta}_i^T\beta_0\| = O_p(n^{1/(2s)}) \quad \text{and} \quad \max_{1\le i\le n}\|\tilde{\eta}_i\tilde{\eta}_i^T - \Sigma_\eta\| = o_p(n^{1/s}).
\end{aligned}$$

Therefore, (5.8) holds. Similarly to the arguments as in [17], from equation (2.5) and Lemmas 5.4-5.5, we can verify easily (5.9). □

Proof of Theorem 3.1. Applying the Taylor expansion to (2.4) and Lemmas 5.4-5.6, we obtain that

$$\mathcal{L}_n(\beta_0) = 2\sum_{i=1}^{n}\{\lambda^T\eta_i(\beta_0) - [\lambda^T\eta_i(\beta_0)]^2/2\} + o_p(1).$$

From (2.5), we have

$$0 = \frac{1}{n}\sum_{i=1}^{n}\frac{\eta_i(\beta_0)}{1+\lambda^T\eta_i(\beta_0)}$$
$$= \frac{1}{n}\sum_{i=1}^{n}\eta_i(\beta_0) - \frac{1}{n}\sum_{i=1}^{n}\eta_i(\beta_0)Z_{ni}(\beta_0)^T\lambda + \frac{1}{n}\sum_{i=1}^{n}\frac{\eta_i(\beta_0)[\lambda^T\eta_i(\beta_0)]^2}{1+\lambda^T\eta_i(\beta_0)}.$$

Using Lemmas 5.4-5.6, we find

$$\left\|\frac{1}{n}\sum_{i=1}^{n}\frac{\eta_i(\beta_0)[\lambda^T\eta_i(\beta_0)]^2}{1+\lambda^T\eta_i(\beta_0)}\right\| \le \frac{1}{n}\sum_{i=1}^{n}\frac{\|\eta_i(\beta_0)\|^3\|\lambda\|^2}{|1+\lambda^T\eta_i(\beta_0)|}$$
$$\le \|\lambda\|^2 \max_{1\le i\le n}\|\eta_i(\beta_0)\|\frac{1}{n}\sum_{i=1}^{n}\|\eta_i(\beta_0)\|^2 = O_p(n^{-1})o_p(n^{1/2})O_p(1) = o_p(n^{-1/2}).$$

Then $\sum_{i=1}^{n}[\lambda^T\eta_i(\beta_0)]^2 = \sum_{i=1}^{n}\lambda^T\eta_i(\beta_0) + o_p(1)$, and

$$\lambda = \Big[\sum_{i=1}^{n}\eta_i(\beta_0)\eta_i^T(\beta_0)\Big]^{-1}\sum_{i=1}^{n}\eta_i(\beta_0) + o_p(n^{-1/2}).$$

Thus

$$\mathcal{L}_n(\beta_0) = \Big(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i(\beta_0)\Big)^T\Big(\frac{1}{n}\sum_{i=1}^{n}\eta_i(\beta_0)\eta_i^T(\beta_0)\Big)^{-1}\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i(\beta_0)\Big) + o_p(1).$$

From Lemma 5.5, we obtain that

$$\mathcal{L}_n(\beta_0) = \Big(\Sigma^{-1/2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i(\beta_0)\Big)^T\Sigma^{1/2}\Sigma_1^{-1}\Sigma^{1/2}\Big(\Lambda^{-1/2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}Z_{ni}(\beta_0)\Big)+o_p(1).$$

Let $w_1,\ldots,w_p$ be the eigenvalues of $\Sigma_1^{-1}\Sigma$ and $\widetilde{\Sigma} = diag(w_1,\ldots,w_p)$. Note that $\Sigma_1^{-1}\Sigma$ and $\Sigma^{1/2}\Sigma_1^{-1}\Sigma^{1/2}$ have the same eigenvalues. Then there exists orthogonal matrix Q such that $Q^T\widetilde{\Sigma}Q = \Sigma^{1/2}\Sigma_1^{-1}\Sigma^{1/2}$. Hence

$$\mathcal{L}_n(\beta_0) = \Big(Q\Sigma^{-1/2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i(\beta_0)\Big)^T\widetilde{\Sigma}\Big(Q\Sigma^{-1/2}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i(\beta_0)\Big) + o_p(1).$$

Notice that Q is orthogonal matrix, which combing with Lemma 5.4, yields (3.1).

By Lemmas 5.1-5.3, the proof of Lemmas 5.4-5.5, it is easy to finish the proof of the consistency of $\widehat{\Sigma}_n(\beta)$ and $\widehat{\Sigma}_{1n}(\beta)$. □

Proof of Theorem 3.2. By [24], the distribution of $p(\sum_{i=1}^{p} w_i\chi^2_{1,i})/\sum_{i=1}^{p} w_i$ can be approximated by the standard χ^2_p distribution (also see [31], Page 474). Note that R_n is a consistent estimator of $p/\sum_{i=1}^{p} w_i$. This together with Theorem 3.1, implies that the asymptotic distribution of $p\cdot\mathcal{L}_n(\beta)/\mathrm{tr}[\Sigma_1^{-1}\widehat{\Sigma}_n(\beta)]$ can be approximated by χ^2_p. Therefore, from Theorem 3.1, we have

$$\widehat{\mathcal{L}}_{n,ad}(\beta_0) = R_n\mathcal{L}_n(\beta_0) \xrightarrow{d} \chi^2_p.$$

This completes the proof of Theorem 3.2. □

Acknowledgements

The authors are grateful to the Associate Editor and two anonymous referees for providing detailed lists of comments and suggestions which greatly improved the presentation of the paper. This research was supported by the National Natural Science Foundation of China (10871146, 71171003), Humanities and Social Science Planning Foundation of Ministry of Education of China (10YJA910005), Provincial Natural Science Research Project of Anhui Colleges (KJ2011A032), Anhui Provincial Natural Science Foundation (1208085QA04), Young Teachers Science Research Foundation of Anhui Polytechnic University (2009YQ035).

References

[1] Cai, Z. (2007). Trending time-varying coefficient timeseries models with serially correlated errors. *J. Econometrics* **136**, 163-188. MR2328589

[2] Chen, S. X. and Wong, C. M. (2009). Smoothed block empirical likelihood for quantiles of weakly dependent processes. *Statist. Sinica* **19**, 71-81. MR2487878

[3] Cui, H. J. and Li, R. C. (1998). On Parameter Estimation for Semi-linear Errors-in-Variables Models. *J. Multivariate Anal.* **64**, 1-24. MR1619970

[4] Cui, H. J. and Kong, E. F. (2006). Empirical likelihood confidence region for parameters in semi-linear errors-in-variables models. *Scand. J. Statist.* **33**, 153-168. MR2255115

[5] Doukhan, P. (1994). *Mixing: Properties and Examples.* Lecture Notes in Statistics, vol. 85. Springer, Berlin. MR1312160

[6] Fan, G. L. and Liang, H. Y. (2010). Empirical likelihood inference for semiparametric model with linear process errors. *J. Korean Statist. Soc.* **39**, 55-65. MR2752218

[7] Fan, G. L., Liang, H. Y. and Xu, H. X. (2011). Empirical likelihood for a heteroscedastic partial linear model. *Comm. Statist. Theory Methods* **8**, 1396-1417. MR2776239

[8] Hu, X., Wang Z. and Zhao, Z. (2009). Empirical likelihood for semiparametric varying-coefficient partially linear errors-in-variables models. *Statist. Probab. Lett.* **79**, 1044-1052. MR2510766

[9] Karlsen, H. A., Mykelbust, T. and Tjøsheim, D. (2007). Nonparametric estimation in a nonlinear cointegration typemodel. *Ann. Statist.* **35**, 252-299. MR2332276

[10] Kitamura, Y. (1997). Empirical likelihood methods with weakly dependent processes. *Ann. Statist.* **25**, 2084-2102. MR1474084

[11] Li, D., Chen, J. and Lin, Z. (2011). Statistical inference in partially time-varying coefficient models. *J. Statist. Plann. Inference* **141**, 995-1013. MR2732967

[12] Li, G. R. and Xue, L. G. (2008). Empirical likelihood confidence region for the parameter in a partially linear errors-in-variables model. *Comm. Statist. Theory Methods* **37**, 1552-1564. MR2440453

[13] Liang, H., Härdle, W., Carroll, R. J. (1999). Estimation in a semiparametric partially linear errors-in-variables model. *Ann. Statist.* **27**, 1519-1535. MR1742498

[14] Liang, H. Y. (2010). Empirical likelihood for conditional density under left truncation and α-mixing condition. *Comm. Statist. Theory Methods* **39**, 707-721. MR2745314

[15] Liang, H. Y. and de Uña-Álvarez, J. (2011). Empirical likelihood for conditional quantile with left-truncated and dependent data. *Ann. Inst. Statist. Math.* in Press. doi: 10.1007/s10463-011-0335-8

[16] Lin, Z. and Lu, C. (1996). *Limit Theory for Mixing Dependent Random Variables*. Science Press, Kluwer Academic Publisher, New York, Dordrecht. MR1486580

[17] Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75**, 237-249. MR0946049

[18] Owen, A. B. (2001). *Empirical likelihood*. Boca Raton, FL: Chapman & Hall/CRC.

[19] Park, J. and Phillips, P. C. B. (2001). Nonlinear regressions with integrated time series. *Econometrica* **69**, 117-162. MR1806536

[20] Phillips, P. C. B. (2001). Trending time series and macroeconomic activity: some present and future challengers. *J. Econometrics* **100**, 21-27. MR1807783

[21] Qin, G. S. and Jing, B. Y. (2001). Censored Partial Linear Models and Empirical Likelihood. *J. Multivariate Anal.* **78**, 37-61. MR1856265

[22] Qin, G. S. and Tsao, M. (2003). Empirical likelihood inference for median regression models for censored survival data. *J. Multivariate Anal.* **85**, 416-430. MR1983805

[23] Qin, J. and Lawless, J. F. (1994). Empirical likelihood and general estimating equations. *Ann. Statist.* **22**, 300-325. MR1272085

[24] Rao, J. N. K. and Scott, A. J. (1981). The analysis of categorical data from complex sample surveys: chi-squared tests for goodness of fit and independence in two-way tables. *J. Amer. Statist. Assoc.* **76**, 221-230. MR0624328

[25] Robinson, P. (1989). Nonparametric estimation of time-varying parameters. In: Hackl, P.(Ed.), *Statistical Analysis and Forecasting of Economic Structural Change*. Springer, Berlin, pp. 164-253.

[26] Shi, J. and Lau, T. S. (2000). Empirical likelihood for partially linear models. *J. Multivariate Anal.* **72**, 132-148. MR1747427

[27] Stute, W, Xue, L. G. and Zhu, L. X. (2007). Empirical likelihood inference in nonlinear errors-in-variables models with validation data. *J. Amer. Statist. Assoc.* **102**, 332-346. MR2345546

[28] Tsay, R. (2002). *Analysis of Financial Time Series.* Wiley, NewYork. MR2778591

[29] Wang, K. (2003). Asset pricing with conditioning information: a new test. *J. Finance* **58**, 161-196.

[30] Wang, Q. H. and Jing, B. Y. (2003). Empirical likelihood for partial linear models. *Ann. Inst. Statist. Math.* **55**, 585-595. MR2007800

[31] Wang, Q. H. and Li, G. (2002). Empirical likelihood semiparametric regression analysis under random censorship. *J. Multivariate Anal.* **83**, 469-486. MR1945964

[32] Wang, X., Li G. and Lin L. (2011). Empirical likelihood inference for semi-parametric varying coefficient partially linear EV models. *Metrika* **73**, 171-185. MR2769269

[33] Wang, S., Qian, L. and Carroll, R. J. (2010). Generalized empirical likelihood methods for analyzing longitudinal data. *Biometrika* **97**, 79-93. MR2594418

[34] Withers, C. S. (1981). Conditions for linear processes to be strong mixing. *Z. Wahrrcheinlichkcitstheos. Verwandte. Geb.* **57**, 477-480. MR0631371

[35] You, J. H. and Chen, G. M. (2006). Estimation of a semiparametric varying-coefficient partially linear errors-in-variables model. *J. Multivariate Anal.* **97**, 324-341. MR2234026

[36] You, J. H. and Zhou, Y. (2006). Empirical likeliood for semiparametric varying-coefficient partially linear regression models. *Statist. Probab. Lett.* **76**, 412-422. MR2232483

[37] Zhou, Y. and Liang, H. (2009). Statistical inference for semiparametric varying-coefficient partially linear models with error-prone linear covariates. *Ann. Statist.* **37**, 427-458. 166-188. MR2488358

[38] Zhu, L. X. and Xue, L. G. (2006). Empirical likelihood confidence regions in a partially linear single-index model. *J. Roy. Statist. Soc. Ser. B* **68**, 549-570. MR2278341