ARTICLE IN PRESS

Journal of the Korean Statistical Society ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

journal homepage: www.elsevier.com/locate/jkss

Empirical likelihood for semi-varying coefficient models for panel data with fixed effects

Bang-Qiang He [a,b], Xing-Jian Hong [a,*], Guo-Liang Fan [b]

[a] *School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou 310018, China*

[b] *School of Mathematics & Physics, Anhui Polytechnic University, Wuhu 241000, China*

ARTICLE INFO

Article history:
Received 15 August 2015
Accepted 7 January 2016
Available online xxxx

AMS 2000 subject classifications:
primary 62G15
secondary 62G20

Keywords:
Empirical likelihood
Semi-varying coefficient model
α-mixing
Confidence region

ABSTRACT

The empirical likelihood inference for semi-varying coefficient models for panel data with fixed effects is investigated in this paper. We propose an empirical log-likelihood ratio function for the regression parameters in the model under α-mixing condition. The empirical log-likelihood ratio is proven to be asymptotically chi-squared. We also obtain the maximum empirical likelihood estimator of the parameters of interest, and prove that it is the asymptotically normal under some suitable conditions. A simulation study and a real data application are undertaken to assess the finite sample performance of our proposed method.

1. Introduction

Panel data is a dataset in which a given sample of individuals is observed across time. Thus it provides multiple observations on each individual in the sample. There are two kinds of information in panel data: the cross-sectional information reflected in the differences between subjects, and the time-series or within-subject information reflected in the changes within subjects over time. With the increasing availability of panel data, both theoretical and applied works in panel data analysis have become more popular in recent years. Baltagi (2005) and Hsiao (2003) provided an excellent overview of statistical inference and econometric interpretation of this widely used class of parametric panel data models. As in both the cross-section and time-series analysis, however, parametric panel data models may be misspecified and estimators obtained from misspecified models are often inconsistent. To deal with this problem, some nonparametric panel data models have been introduced and studied, such as the fixed effects nonparametric panel data models (Henderson, Carroll, & Li, 2008), the random effects nonparametric panel data models (Henderson & Ullah, 2005) and dynamic nonparametric panel data models (Cai & Li, 2008). While the nonparametric approach is useful in exploring hidden structures and reducing modeling biases, it can be too flexible to draw concise conclusions, and faces the *curse of dimensionality* due to a large number of covariates. How to circumvent the curse of dimensionality is an important topic in both nonlinear time series and panel data analysis. Many useful approaches are developed to avoid this problem (see, recent book by Fan & Yao, 2003 for example) and various semiparametric panel data models have been proposed. For example, Hu and Li (2011) studied semiparametric varying-coefficient partially linear model with longitudinal data. Hu (2014) considered semi-varying coefficient model for panel

* Corresponding author.
E-mail address: hbangqiang@126.com (X.-J. Hong).

http://dx.doi.org/10.1016/j.jkss.2016.01.001
1226-3192/

Please cite this article in press as: He, B.-Q., et al., Empirical likelihood for semi-varying coefficient models for panel data with fixed effects. Journal of the Korean Statistical Society (2016), http://dx.doi.org/10.1016/j.jkss.2016.01.001

data. Varying-coefficient models are well known in the statistic literatures and also have a lot of applications in economics and finance (cf. Cai & Li, 2008; Fan & Zhang, 1999; Hu & Li, 2011). One of the main advantages of the varying-coefficient models is that it allows the coefficients to depend on some informative variables and then balances the dimension reduction and model flexibility. The semi-varying coefficient model covers two important statistical models: the varying-coefficient model and the semiparametric partially linear model. Relevant studies can be found in Fan and Huang (2005), Hu (2014), You and Zhou (2006) and so on. Hu (2014) proposed the profile likelihood procedure to estimate semi-varying coefficient model for panel data with fixed effects.

Our aim in the paper is to apply the empirical likelihood (EL) method to construct confidence regions of β in the semi-varying coefficient model for panel data with fixed effects. Firstly, we shall propose an empirical log-likelihood ratio function for the regression parameter in the model under the α-mixing condition. Based on the this, one can immediately construct an approximate confidence region for the regression parameter. One motivation is that EL inference does not involve the asymptotic covariance of the estimators, which is rather complex structure for the semi-varying coefficient model for panel data with fixed effects under the α-mixing condition. Another motivation is that the confidence region based on EL approach does not impose prior constraints on the region shape, and the shape and orientation of confidence regions are determined completely by the data. The EL method has been used by many authors up to now, such as Fan and Liang (2010), Hu and Li (2011) Shi and Lau (2000), Wang and Jing (1999), You and Zhou (2006), among others. Most of the above papers related to the EL method have always assumed that the data are independent and identically distributed. However, the independence assumption for the data is not always appropriate in application, especially for sequentially collected economic data, which often exhibit evident dependence, and α-mixing condition has been assumed by many authors. For example, Fan, Liang, and Huang (2012) employed the EL method to construct confidence regions for partially time-varying coefficient models with dependent observations.

Throughout this paper, we assume that $\left\{(X_{it}, Z_{it}, Y_{it}, U_{it}, \upsilon_{it}),\ t \geq 1\right\}$ is stationary α-mixing for each i. Recall that a sequence $\{\xi_k, k \geq 1\}$ is said to be α-mixing if the α-mixing coefficient

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

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 literature, the α-mixing is reasonably weak and is known to be fulfilled by many stochastic processes including many time series models. For example, Auestad and Tjøstheim (1990) provided illuminating discussions on the role of α-mixing for model identification in nonlinear time series analysis. Further, Masry and Tjøstheim (1995) showed that under some mild conditions, both autoregressive conditional heteroskedastic (ARCH) process and additive autoregressive process with exogenous variables, which are particularly popular in finance, are stationary and α-mixing.

The rest of this paper is organized as follows. Section 2 introduces the methodology and empirical log-likelihood ratio function for β. Assumption conditions and the main result are given in Section 3. Some simulation studies and a real-data example are conducted in Section 4. The proofs of the main results are relegated to Section 5. In Section 6, we present a brief discussion of the results and methods. Some preliminary lemmas, which are used in the proofs of the main results, are collected in the Appendix.

2. The model and methodology

This paper considers the following semi-varying coefficient model for panel data with fixed effects:

$$Y_{it} = Z_{it}^{\tau}\beta + X_{it}^{\tau}\alpha(U_{it}) + \mu_i + \upsilon_{it}, \quad i = 1, \ldots, n,\ t = 1, \ldots, T, \tag{2.1}$$

where Y_{it} is the response, $(Z_{it}, X_{it}) \in R^p \times R^q$ and $U_{it} \in R$ are strictly exogenous regressors, $\beta = (\beta_1, \ldots, \beta_p)^{\tau}$ is a vector of p-dimensional unknown parameters, and the superscript τ denotes the transpose of a vector or matrix. $\alpha(U_{it}) = (\alpha_1(U_{it}), \ldots, \alpha_q(U_{it}))^{\tau}$ is a q-dimensional vector of unknown functions and μ_i is the unobserved individual effects, υ_{it} is the random model error. Here, we assume μ_i to be i.i.d. with zero mean and finite variance $\sigma^2 > 0$. We allow μ_i to be correlated with Z_{it}, X_{it}, and U_{it} with an unknown correlation structure.

To introduce our estimation, we assume that model holds with the restriction $\sum_{i=1}^n \mu_i = 0$. Let $\mu = (\mu_2, \ldots, \mu_n)^{\tau}$ and $\mu_0 = (-\sum_{i=2}^n \mu_i, \mu_2, \ldots, \mu_n)^{\tau}$. We rewrite model (2.1) in a matrix format yields

$$Y = Z\beta + X\alpha(U) + H\mu + V, \tag{2.2}$$

where $H = [-i_{n-1}\ I_{n-1}] \otimes i_T$ is an $nT \times (n-1)$ matrix. I_n denotes the $n \times n$ identity matrix, and i_n denotes the $n \times 1$ vector of ones. There are many approaches to estimating the parameters $\{\beta_j, j = 1, \ldots, p\}$ and the varying coefficient functions $\{\alpha_i(\cdot),\ i = 1, \ldots, q\}$. The main idea is from the profile least squares approach proposed by Fan and Huang (2005): suppose that we have a random sample $\{(U_{it}, Z_{it1}, \ldots, Z_{itp}, X_{it1}, \ldots, X_{itq}, Y_{it}),\ i = 1, \ldots, n,\ t = 1, \ldots, T\}$ from model (2.1). Let $\theta = (\mu^{\tau}, \beta^{\tau})^{\tau}$. Given θ, one can apply a local linear regression technique to estimate the varying coefficient functions $\{\alpha_j(\cdot), j = 1, \ldots, q\}$ in (2.1). For U_{it} in a small neighborhood of u_0, one can approximate $\alpha_j(U_{it})$ locally by a linear function

$$\alpha_j(U_{it}) \approx \alpha_j(u_0) + \alpha_j'(u_0)(U_{it} - u_0) \equiv a_j + b_j(U_{it} - u_0), \quad j = 1, \ldots, q,$$

where $\alpha_j'(u) = \partial\alpha_j(u)/\partial u$. This leads to the following weighted local least-squares problem: find $\{(a_j, b_j),\ j = 1, \ldots, q\}$ to minimize

$$\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\left(Y_{it} - Z_{it}^{\tau}\beta - \mu_i\right) - \sum_{j=1}^{q}\left[a_j + b_j(U_{it} - u_0)\right]X_{it}\right\}^2 K_h(U_{it} - u_0), \tag{2.3}$$

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

$$(\hat{\alpha}_1(u), \ldots, \hat{\alpha}_q(u), h\hat{a'}_1(u), \ldots, h\hat{a'}_q(u))^{\tau} = (D_u^{\tau}W_uD_u)^{-1}D_u^{\tau}W_u(Y - Z\beta - H\mu),$$

where

$$Z = \begin{pmatrix} Z_{11}^{\tau} \\ \vdots \\ Z_{1T}^{\tau} \\ \vdots \\ Z_{nT}^{\tau} \end{pmatrix}, \qquad X = \begin{pmatrix} X_{11}^{\tau} \\ \vdots \\ X_{1T}^{\tau} \\ \vdots \\ X_{nT}^{\tau} \end{pmatrix}, \qquad D_u = \begin{pmatrix} X_{11}^{\tau} & \frac{U_{11} - u}{h}X_{11}^{\tau} \\ \vdots & \vdots \\ X_{1T}^{\tau} & \frac{U_{1T} - u}{h}X_{1T}^{\tau} \\ \vdots & \vdots \\ X_{nT}^{\tau} & \frac{U_{nT} - u}{h}X_{nT}^{\tau} \end{pmatrix},$$

$Y = (Y_{11}, \ldots, Y_{1T}, \ldots, Y_{nT})^{\tau}$ and $W_u = \mathrm{diag}(K_h(U_{11} - u), \ldots, K_h(U_{1T} - u), \ldots, K_h(U_{nT} - u))$.

The profile likelihood estimator of parameter θ is given by

$$\hat{\theta} = \arg\min_{\theta}\left[Y - Z\beta - H\mu - S(Y - Z\beta - H\mu)\right]^{\tau}[Y - Z\beta - H\mu - S(Y - Z\beta - H\mu)], \tag{2.4}$$

where the smoothing matrix S is

$$S = \begin{pmatrix} (X_{11}^{\tau}\ 0_q^{\tau})(D_{U_{11}}^{\tau}W_{U_{11}}D_{U_{11}})^{-1}D_{U_{11}}^{\tau}W_{U_{11}} \\ \vdots \\ (X_{1T}^{\tau}\ 0_q^{\tau})(D_{U_{1T}}^{\tau}W_{U_{1T}}D_{U_{1T}})^{-1}D_{U_{1T}}^{\tau}W_{U_{1T}} \\ \vdots \\ (X_{nT}^{\tau}\ 0_q^{\tau})(D_{U_{nT}}^{\tau}W_{U_{nT}}D_{U_{nT}})^{-1}D_{U_{nT}}^{\tau}W_{U_{nT}} \end{pmatrix} = \begin{pmatrix} s_{11} \\ \vdots \\ s_{1T} \\ \vdots \\ s_{nT} \end{pmatrix}. \tag{2.5}$$

Let $M = (m_1^{\tau}, \ldots, m_n^{\tau})^{\tau}$, $m_i = (x_{i1}\alpha(U_{i1}), \ldots, x_{iT}\alpha(U_{iT}))$, $Z_i = (Z_{i1}^{\tau}, \ldots, Z_{iT}^{\tau})^{\tau}$, $Y_i = (Y_{i1}, \ldots, Y_{iT})^{\tau}$, $\widetilde{Z} = (I_{nT} - S)Z$, $\widetilde{H} = (I_{nT} - S)H$, $\widetilde{Y} = (I_{nT} - S)Y$, $\widetilde{\mu} = (\widetilde{H}^{\tau}\widetilde{H})^{-1}\widetilde{H}^{\tau}(\widetilde{Y} - \widetilde{Z}\beta)$ and $\Lambda = I_{nT} - \widetilde{H}(\widetilde{H}^{\tau}\widetilde{H})^{-1}\widetilde{H}^{\tau}$. We introduce the following auxiliary random vector:

$$\eta_i(\beta) = \widetilde{Z}_i\Lambda(\widetilde{Y}_i - \widetilde{Z}_i\beta), \quad i = 1, \ldots, n. \tag{2.6}$$

Note that $E(\eta_i(\beta)) = 0$ if β is the true parameter. Therefore, similar to Owen (1990), we define an empirical log-likelihood ratio as follows.

$$\log\mathcal{L}_n(\beta) = -2\max\left\{\sum_{i=1}^{n}\log(np_i) : \sum_{i=1}^{n}p_i\eta_i(\beta) = 0, p_i \geq 0, \sum_{i=1}^{n}p_i = 1\right\}. \tag{2.7}$$

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

$$\log\mathcal{L}_n(\beta) = 2\sum_{i=1}^{n}\log\{1 + \lambda^{\tau}\eta_i(\beta)\}, \tag{2.8}$$

where λ is determined by

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\eta_i(\beta)}{1 + \lambda^{\tau}\eta_i(\beta)} = 0. \tag{2.9}$$

Please cite this article in press as: He, B.-Q., et al., Empirical likelihood for semi-varying coefficient models for panel data with fixed effects. Journal of the Korean Statistical Society (2016), http://dx.doi.org/10.1016/j.jkss.2016.01.001

3. The asymptotic result

In this section, we will show that if β is the true parameter vector, $\log \mathcal{L}_n(\beta)$ is asymptotically χ^2-distributed. Before formulating the main results, we first give the following some assumptions.

(A1) The kernel $K(v)$ is a symmetric probability density function with a continuous derivative on its compact support $[-1, 1]$.
(A2) The random variable U has a bounded support Ω. The matrix $\Gamma(U)\widehat{=}E(ZZ^\tau|U)$ is nonsingular for each $U \in \Omega$. $G(U)\widehat{=}E(XX^\tau|U)$, $\Gamma^{-1}(U)$ and $\Phi(U)\widehat{=}E(ZX^\tau|U)$ are all Lipschitz continuous.
(A3) $\{\alpha_j(\cdot),\ j = 1, \dots, q\}$ have the continuous second derivative in $U \in \Omega$.
(A4) (i) $\{X_{it}, Y_{it}, Z_{it}, U_{it}, v_{it}\}$ are independent and identically distributed across the i index for each fixed t, and strictly stationary over t for each fixed i.
(ii) For each fixed i, the processes $\{X_{it}, Y_{it}, Z_{it}, U_{it}, v_{it}\}$ are a stationary sequence of α-mixing with the mixing coefficient satisfying the condition $\alpha(k) = O(k^{-\kappa})$, where $\kappa = \frac{(2+\delta)(1+\delta)}{\delta}$ and $\delta > 2$.
(A5) There exists some $\delta > 2$ such that $E\|X_i\|^{2+\delta} < \infty$, $E\|Z_i\|^{2+\delta} < \infty$, $E\|v_i\|^{2+\delta} < \infty$, where $\|\cdot\|$ is the L_2-distance.
(A6) Let $N = nT$, the bandwidth h satisfies $Nh \to \infty$, $Nh^6 \to 0$, $(\log N)^{\kappa+1/2}N^{-((\kappa-\frac{1}{2})-\frac{1}{2\delta})}T^{\frac{\kappa+1}{\delta}}h^{\kappa-3} \to 0$. where κ and δ are defined in A4(ii) above.
(A7) $E|\breve{Z}_{it}|^{2+\delta} < \infty$, $\Psi = \sum_{t=1}^T[\breve{Z}_{it}(\breve{Z}_{it} - \sum_{s=1}^T \frac{\breve{Z}_{is}}{T})^\tau]$ is positive definite, where $\breve{Z}_{it} = Z_{it} - G^{-1}(u)\Phi(u)X_{it}$.

Remark 3.1. Assumptions (A1)–(A7) which look a bit lengthy, are actually quite mild and can be easily satisfied. (A1)–(A3) can be founded in Fan and Huang (2005). (A4)–(A5) have been used by many authors (see Cai & Li, 2008; Chen, Gao, & Li, 2013 for example). (A7) has been used by Hu (2014). The technical conditions of (A6) are easily satisfied. For example, when $T \sim N^{1/5}$ and $h \sim N^{-\theta}$, it can be shown that $Nh \to \infty$, $Nh^6 \to 0$, and $(\log N)^{\kappa+1/2}N^{-((\kappa-\frac{1}{2})-\frac{1}{2\delta})}T^{\frac{\kappa+1}{\delta}}h^{\kappa-3} \to 0$ are all satisfied when $\frac{1}{6} < \theta < \frac{1}{5}$.

Theorem 3.1. *Suppose that* (A1)–(A7) *hold. For model* (2.1), *if* β *is the true value of the parameter, then* $\log \mathcal{L}_n(\beta) \xrightarrow{d} \chi_p^2$, *as* $n \to \infty$, *where* χ_p^2 *is a standard chi-square random variable with* p *degrees of freedom and* $\xrightarrow{d}$ *stands for convergence in distribution.*

As a consequence of the theorem, confidence regions for the parameter β can be constructed. More precisely, for any $0 \le \alpha < 1$, let c_α be such that $P(\chi_p^2 > c_\alpha) \le 1-\alpha$. Then $\ell_{EL}(\alpha) = \{\beta \in R^p : \log \mathcal{L}_n(\beta) \le c_\alpha\}$ constitutes a confidence region for β with asymptotic coverage α because the event that β belongs to $\ell_{EL}(\alpha)$ is equivalent to the event that $\log \mathcal{L}_n(\beta) \le c_\alpha$.

Maximizing $\{-\log \mathcal{L}_n(\beta)\}$ can obtain the maximum empirical likelihood estimator (MELE) $\widehat{\beta}$ of β. Write $\widehat{\Psi} = \frac{1}{n}\sum_{i=1}^n\sum_{t=1}^T \widetilde{Z}_{it}^\tau(\widetilde{Z}_{it} - \sum_{s=1}^T \frac{\widetilde{Z}_{is}}{T})$. If the matric $\widehat{\Psi}$ is invertible, with the similar proof theorem 1 in Qin and Lawless (1994), then MELE $\widehat{\beta}$ of β can be represented as

$$\widehat{\beta} = \widehat{\Psi}^{-1}\frac{1}{n}\sum_{i=1}^n \widetilde{Z}_i^\tau \Lambda \widetilde{Y}_i + o_p\left(n^{-\frac{1}{2}}\right).$$

Then the asymptotic normality of $\widehat{\beta}$ is stated in the following theorem.

Theorem 3.2. *Suppose that* (A1)–(A7) *hold. Then we have* $\sqrt{n}(\widehat{\beta} - \beta) \xrightarrow{d} N(0, V)$, *where* $V = \Psi^{-1}\Theta\Psi^{-1}$, $\Theta = \sum_{t=1}^T\sum_{s=1}^T E\left\{\left[\breve{Z}_{it}(\breve{Z}_{is} - \sum_{l=1}^T \frac{\breve{Z}_{il}}{T})^\tau\right](v_{it}v_{is})\right\}$.

4. Numerical studies

4.1. Simulation study

In this section, we carry out a simulation study to show the finite sample performance of the proposed EL confidence regions of β.

Firstly, we consider the following semi-varying coefficient model for panel data with fixed effects:

$$Y_{it} = Z_{it}^\tau\beta + X_{it,1}\alpha_1(U_{it}) + X_{it,2}\alpha_2(U_{it}) + \mu_i + v_{it}, \quad i = 1, \dots, n,\ t = 1, \dots, T, \tag{4.1}$$

where $\beta = (\beta_1, \beta_2)^\tau = (1, \sqrt{2})^\tau/\sqrt{3}$, $\alpha_1(U_{it}) = \sin(2\pi U_{it})$, $\alpha_2(U_{it}) = \cos(\pi U_{it})$, $U_{it} \sim U(0, 1)$, $\mu_i = \rho\bar{U}_i + w_i$ with $\rho = 0.5,\ 1$ and $w_i \sim N(0, 1)$ for $i = 2, 3, \dots, n$, and $\mu_i = -\sum_{i=2}^n \mu_i$, $\bar{U}_i = \frac{1}{T}\sum_{t=1}^T U_{it}$. We use ρ to control the correlation between μ_i and $\bar{U}_i$. X_{it} and Z_{it} are generated by the AR(1) model as follows,

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

Table 1
Coverage probabilities (CP) of confidence regions for (β_1, β_2) based on the EL and the NA with nominal confidence level $1-\alpha = 0.90$ and 0.95.

ρ	$1-\alpha$	(n, T)	CP_{EL}	CP_{NA}	ρ	$1-\alpha$	(n, T)	CP_{EL}	CP_{NA}
0.5	0.90	(50, 4)	0.881	0.683	1	0.90	(50, 4)	0.874	0.693
		(100, 2)	0.876	0.670			(100, 2)	0.866	0.706
		(100, 4)	0.892	0.716			(100, 4)	0.882	0.713
	0.95	(50, 4)	0.929	0.848		0.95	(50, 4)	0.928	0.835
		(100, 2)	0.926	0.832			(100, 2)	0.920	0.810
		(100, 4)	0.936	0.856			(100, 4)	0.934	0.854

Note: "CP_{EL} and CP_{NA}" denote the coverage probabilities based on the EL and the NA, respectively.

Table 2
Coverage probabilities (CP) and average lengths (AL) of the confidence intervals based on the EL and the NA with nominal confidence level $1-\alpha = 0.90$ and 0.95.

ρ	Para	(n, T)	$1-\alpha = 0.90$				$1-\alpha = 0.95$			
			CE	CN	AE	AN	CE	CN	AE	AN
0.5	β_1	(50, 4)	0.885	0.770	0.2890	0.5844	0.944	0.895	0.3804	0.7471
		(100, 2)	0.870	0.738	0.3786	0.5891	0.932	0.870	0.5427	0.7588
		(100, 4)	0.895	0.795	0.2309	0.5006	0.946	0.916	0.3199	0.7282
	β_2	(50, 4)	0.884	0.623	0.3713	0.8436	0.932	0.705	0.4943	1.0681
		(100, 2)	0.882	0.610	0.4586	0.7579	0.925	0.697	0.5971	0.9659
		(100, 4)	0.887	0.650	0.2993	0.7525	0.938	0.745	0.4102	0.9137
1	β_1	(50, 4)	0.850	0.778	0.3286	0.9286	0.934	0.868	0.4847	1.2076
		(100, 2)	0.836	0.766	0.4220	0.9361	0.926	0.906	0.6329	1.3677
		(100, 4)	0.872	0.815	0.2672	0.9057	0.945	0.910	0.3373	1.1380
	β_2	(50, 4)	0.864	0.732	0.4321	1.3571	0.930	0.868	0.6275	1.7960
		(100, 2)	0.855	0.742	0.4982	1.3260	0.923	0.860	0.7315	1.6739
		(100, 4)	0.882	0.764	0.3243	0.8871	0.936	0.870	0.4816	1.6095

Note: "CE and CN" denote the coverage probabilities based on the EL and the NA, respectively; "AE and AN" denote the average lengths based on the EL and the NA, respectively.

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

In the simulation below, we choose the kernel function $K(u) = \frac{15}{16}(1-u^2)^2 I\{|u| \le 1\}$. The "leave-one-subject-out" cross-validation bandwidth h_{CV} is obtained by minimizing

$$CV(h) = \frac{1}{N}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(Y_{it} - Z_{it}^\tau \hat{\beta}_i - X_{it}^\tau \hat{\alpha}_i(U_{it})\right)^2,$$

where $(\hat{\beta}_1, \hat{\beta}_2)^\tau$ and $\hat{\alpha}_j(U_{it})$ with $j = 1, 2$ are estimators of $(\beta_1, \beta_2)^\tau$ and $\alpha_j(U_{it})$ with $j = 1, 2$ respectively, which are computed with all of the measurements but not the ith subject.

We present a consistent estimator of the covariance of $\hat{\beta}$ in Theorem 3.2. A consistent estimator of Θ is given by replacing β with $\hat{\beta}$, namely,

$$\hat{\Theta} = \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\widetilde{Z}_{it}\left(\widetilde{Z}_{it} - \sum_{s=1}^{T}\frac{\widetilde{Z}_{is}}{T}\right)^\tau (\widetilde{Y}_{it} - \widetilde{Z}_{it}\hat{\beta})(\widetilde{Y}_{it} - \widetilde{Z}_{it}^\tau\hat{\beta})^\tau\right]$$

and a consistent estimator of $\widehat{\Psi} = \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\widetilde{Z}_{it}(\widetilde{Z}_{it} - \sum_{s=1}^{T}\frac{\widetilde{Z}_{is}}{T})^\tau$. Therefore, an $1-\alpha$ level confidence region based normal approximation is $\ell_{NA}(\alpha) = \left\{\hat{\beta} - z_{\frac{\alpha}{2}}\widehat{\Psi}\sqrt{n^{-1}\hat{\Theta}}, \hat{\beta} + z_{\frac{\alpha}{2}}\widehat{\Psi}\sqrt{n^{-1}\hat{\Theta}}\right\}$.

We consider two approaches to compute the coverage probabilities (CP) and the average lengths (AL) of the confidence intervals for individual β_i $(i = 1, 2)$: the empirical likelihood (EL) and the normal approximation (NA). The sample sizes (n, T) are chosen to be (50, 4), (100, 2) and (100, 4) respectively. The CP and AL of the confidence intervals for β_i $(i = 1, 2)$ are calculated based on 500 replications with the nominal level $1-\alpha = 0.90$ and 0.95, respectively. Some representative CP of confidence regions for (β_1, β_2) are reported in Table 1. Simultaneously, we give the CP and AL of confidence intervals for β_1 and β_2 in Table 2 with the nominal level $1-\alpha = 0.90$ and 0.95, respectively.

From Tables 1 and 2, we can be seen that the method based on the EL performs better than the NA method since the EL method gives higher coverage probabilities and shorter average lengths of confidence intervals than the NA method. Also, it is obvious to see that the CP of the confidence regions/intervals tend to increase and the AL decrease as the sample size (n, T), especially T, becomes larger. It is also interesting to note that the coverage probabilities of the confidence intervals tend to decrease and the AL increase as ρ gets larger.

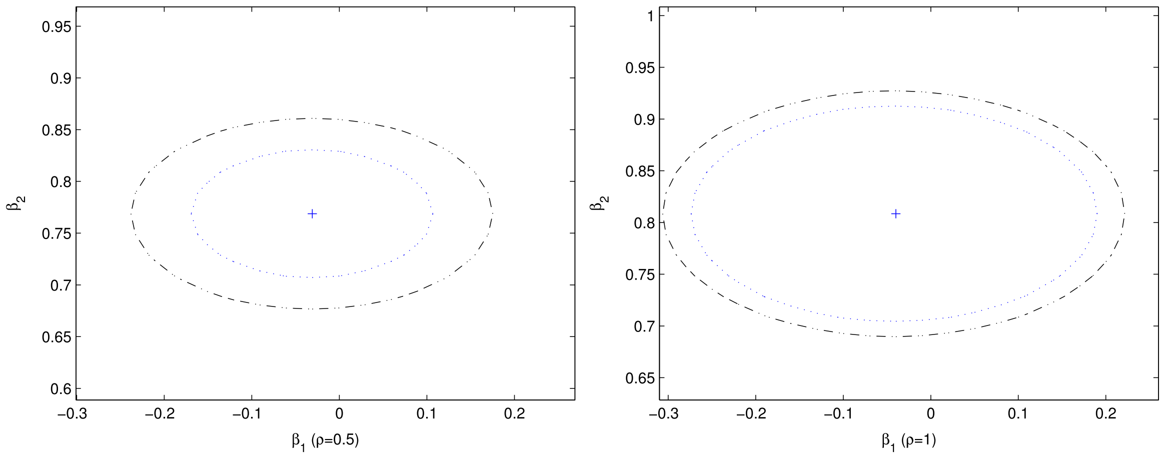

Fig. 1. The plots of confidence regions based on the NA (dot-dashed curve) and EL (dotted curve) for $\beta = (\beta_1, \beta_2)$.

4.2. Application to CD4 data

To examine the performance of the proposed method, we analyze a CD4 data from the Multi-Center AIDS Cohort study. The dataset contains cigarette smoking status, age at HIV infection, pre-infection CD4 percentage and repeatedly measured post-infection CD4 percentage of 283 homosexual men who were infected with HIV during the follow-up period between 1984 and 1991. Details about the related design, methods and medical implications of the Multi-Center AIDS Cohort study have been described by Kaslow, Ostrow, Detels, Polk, and Rinaldo (1987). This dataset has been studied by many authors, such as Huang, Wu and Zhou (2002), Xue and Zhu (2007) and Zhou and Lin (2014). Although all the individuals were scheduled to have clinical visits semi-annually for taking the measurement of CD4 percentage, due to the various reasons, some individuals missed scheduled visits, which resulted in unequal numbers of measurements and different measurement times across individuals. We select 158 individuals in which there have the first six measurements. Hence, we can obtain equal numbers of measurements for each individuals.

The objectives of the study are to describe the trend of the mean CD4 percentage depletion over time and to evaluate the effects of cigarette smoking, the pre-HIV infection CD4 percentage, and age at HIV infection on the mean CD4 percentage after infection. Let Y_{it} be the ith individual's CD4 percentage measured at time t, Z_{it1} be 1 or 0 if the ith individual ever or never smoked cigarettes, respectively, after the HIV infection. Z_{it2} be the ith individual's CD4 percentage measured at time t. We consider the following model:

$$Y_{it} = Z_{it1}\beta_1 + Z_{it2}\beta_2 + X_{it,1}\alpha_1(U_{it}) + X_{it,2}\alpha_2(U_{it}) + \mu_i + \upsilon_{it}, \quad i = 1, \ldots, n,\ t = 1, \ldots, T, \tag{4.2}$$

where $X_{it,1}$, the baseline CD4 percentage, represents the mean CD4 percentage after the infection, $X_{it,2}$ represents age at HIV infection on the mean CD4 percentage after infection, and β_1 and β_2 describe the effects for cigarette smoking and pre-infection CD4 percentage, respectively, on the post-infection CD4 percentage at time t. A profile least square estimator of (β_1, β_2) is $(-0.0372, 0.7795)$, which shows that cigarette smoking of HIV infection has no significant effect on the post-infection CD4 percentage, but preCD4 percentage is highly positively associated with post-infection CD4 percentage, which basically agree with that was discovered by Huang et al. (2002), Xue and Zhu (2007) and Zhou and Lin (2014). Using the EL and NA methods, we obtained $1 - \alpha = 0.95$ confidence regions for (β_1, β_2) that are presented in Fig. 1. These figures show that EL gives more smaller confidence regions than the NA does, and the confidence regions become wider as ρ gets larger.

5. Proofs of the main result

For the convenience and simplicity, let $\varepsilon_N = \{(Nh)^{-1}\log N\}^{1/2}$, $c_N = \varepsilon_N + h^2$, $\vartheta_k = \int t^k K(t)dt$, $\upsilon_k = \int t^k K^2(t)dt$ and $W = (I_{nT} - S)^\tau (I_{nT} - S)$. Note that $S = (s_{11}, \ldots, s_{1T}, \ldots, s_{nT})^\tau$. Denote a typical entry of $s(u)$ by $s_{it}(u) = (X_{it}\ 0)(D_u^\tau W_u D_u)^{-1} D_{it} K_h(U_{it} - u)$. Define $N_{it}(u, j) = X_{it}X_{it}^\tau (\frac{U_{it}-u}{h})^j K_h(U_{it} - u)$, $R_{it}(u, j) = X_{it}(\frac{U_{it}-u}{h})^j K_h(U_{it} - u)Y_{it}$.

Lemma 5.1. *Suppose that conditions* (A1)–(A6) *hold. Then we have*

$$\sup_{u\in[-1,1]} \left| \frac{1}{N}\sum_{i=1}^{n}\sum_{t=1}^{T} N_{it}(u, j) - G(u)f(u)\vartheta_j \right| = O_P(c_N), \tag{5.1}$$

$$\sup_{u\in[-1,1]} \left| \frac{1}{N}\sum_{i=1}^{n}\sum_{t=1}^{T} R_{it}(u, j) - \Phi(u)f(u)\vartheta_j \right| = O_P(c_N). \tag{5.2}$$

Proof. To prove (5.1), it suffices to show that

$$\sup_{u\in[-1,1]}\left|\frac{1}{N}\sum_{i=1}^{n}\sum_{t=1}^{T}\{N_{it}(u,j)-E[N_{it}(u,j)]\}\right| = O_P[(Nh)^{-1}\log N]^{\frac{1}{2}} \tag{5.3}$$

and

$$\sup_{u\in[-1,1]}|E[N_{it}(u,j)]-G(u)f(u)\vartheta_j| = O_P(h^2). \tag{5.4}$$

By conditions (A2), we have

$$\begin{aligned} &E\left[\frac{1}{N}\sum_{i=1}^{n}\sum_{t=1}^{T}X_{it}X_{it}^{\tau}\left(\frac{U_{it}-u}{h}\right)^{j}K_h(U_{it}-u)\right] = E\left[X_{it}X_{it}^{\tau}\left(\frac{U_{it}-u}{h}\right)^{j}K_h(U_{it}-u)\right] \\ &= \int G(U_{it})\left(\frac{U_{it}-u}{h}\right)^{j}K_h(U_{it}-u)f(U_{it})dU_{it} = [G(u)+O(h)]\int u^jK(u)du[f(u)+O(h)] \\ &= G(u)f(u)\vartheta_j + O(h^2), \end{aligned}$$

which implies that (5.4) holds.

Let us now turn to the proof of (5.3). We need only to show that

$$\sup_{u\in[-1,1]}\left|\frac{1}{N}\sum_{i=1}^{n}\sum_{t=1}^{T}\{N_{it}(u,j)-E[N_{it}(u,j)]\}\right| = o(l(N)\varepsilon_N), \tag{5.5}$$

for any $l(N)\to\infty$, as $N\to\infty$. The main idea in proving (5.5) is to consider covering of the compact interval $D=[-1,1]$ by a finite number of subinterval W_k, which are centered at w_k with length $\xi_n = o(h^2)$. Denote the total number these intervals by Γ_n. Then $\Gamma_n = O(\xi_n^{-1})$. It is easy to show that

$$\begin{aligned} &\sup_{u\in[-1,1]}\left|\frac{1}{N}\sum_{i=1}^{n}\sum_{t=1}^{T}\{N_{it}(u,j)-E[N_{it}(u,j)]\}\right| = \max_{1\le k\le\Gamma_n}\sup_{u\in W_k}\left|\frac{1}{N}\sum_{i=1}^{n}\sum_{t=1}^{T}\{N_{it}(u,j)-E[N_{it}(u,j)]\}\right| \\ &\le \max_{1\le k\le\Gamma_n}\sup_{u\in W_k}\left|\frac{1}{N}\sum_{i=1}^{n}\sum_{t=1}^{T}\{N_{it}(u,j)-N_{it}(w_k,j)\}\right| + \max_{1\le k\le\Gamma_n}\left|\frac{1}{N}\sum_{i=1}^{n}\sum_{t=1}^{T}\{N_{it}(w_k,j)-E(N_{it}(w_k,j))\}\right| \\ &\quad + \max_{1\le k\le\Gamma_n}\sup_{u\in W_k}\left|\frac{1}{N}\sum_{i=1}^{n}\sum_{t=1}^{T}\{E[N_{it}(w_k,j)]-E(N_{it}(u,j))\}\right| = \Pi_1+\Pi_2+\Pi_3. \end{aligned} \tag{5.6}$$

Noting that $k(\cdot)$ is Lipschitz continuous by assumption (A1) and taking $\xi_n = O(\varepsilon_N h^2)$, we have

$$E(\Pi_1) \le C\frac{\xi_n}{h^2}E[G(u)] = O(\varepsilon_N) = o(l(N)\varepsilon_N). \tag{5.7}$$

From (5.5) we find immediately that

$$E(\Pi_3) = o(l(N)\varepsilon_N). \tag{5.8}$$

For Π_2, we apply the truncation method, Define

$$\bar{G}_{it} = X_{it}X_{it}^{\tau}\left\{\|X_{it}X_{it}^{\tau}\| \le CN^{\frac{1}{\delta}}\right\} \quad \text{and} \quad \bar{G}_{it}^{c} = G_{it}-\bar{G}_{it}.$$

It is easy to check that

$$\begin{aligned} \Pi_2 \le &\max_{1\le k\le\Gamma_n}\left|\frac{1}{N}\sum_{i=1}^{n}\sum_{t=1}^{T}(\bar{G}_{it}-E[\bar{G}_{it}])K_h(U_{it}-w_k)\right| \\ &+ \max_{1\le k\le\Gamma_n}\left|\frac{1}{N}\sum_{i=1}^{n}\sum_{t=1}^{T}(\bar{G}_{it}^{c}-E[\bar{G}_{it}^{c}])K_h(U_{it}-w_k)\right| = \Pi_{21}+\Pi_{22}. \end{aligned} \tag{5.9}$$

By the

$$P(\Pi_{22} > l(N)\varepsilon_N) \le \sum_{i=1}^{n}\sum_{t=1}^{T} P\left(\|X_{it}X_{it}^{\tau}\| > l(N)N^{\frac{1}{\delta}}\right) = \sum_{i=1}^{n}\sum_{t=1}^{T}\left(\frac{E\|X_{it}X_{it}^{\tau}\|^{\delta}}{\left[l(N)N^{\frac{1}{\delta}}\right]^{\delta}}\right) = O\left(\frac{1}{l(N)^{\delta}}\right) = o(1).$$

Thus, we have

$$\Pi_{22} = o_p(l(N)\varepsilon_N). \tag{5.10}$$

Observe that

$$|(\bar{G}_{it} - E[\bar{G}_{it}])K_h(U_{it} - w_k)| \le \frac{2N^{\frac{1}{\delta}}l(N)}{h} \quad \text{for any } i,\ t.$$

By taking $q = [T^{1+1/\delta}\varepsilon_N]$ and Bernstein inequality for α-mixing processes (see Lemma A.4), we have, for any

$$\begin{aligned} &P\left(\left|\frac{1}{T}\sum_{t=1}^{T}(\bar{G}_{it} - E[\bar{G}_{it}])K_h(U_{it} - s_l)\right| > l(N)\varepsilon_N\right) \\ &\le 4\exp(-l(N)\varepsilon_N^2 Th) + C\left\{\frac{2l(N)N^{\frac{1}{\delta}}}{[l(N)\varepsilon_N h]}\right\}^{\frac{1}{2}}\left[T^{1+\frac{1}{\delta}}\varepsilon_N\right]\left[T^{-\frac{1}{\delta}}\varepsilon_N^{-1}\right]^{-\kappa} \\ &= 4\exp\left[-\frac{l(N)}{n}\log N\right] + C(\log N)^{\kappa+\frac{1}{2}}N^{-\left(\left(\kappa+\frac{1}{2}\right)-\frac{1}{2\delta}\right)}T^{1+\frac{\kappa+1}{\delta}}h^{\kappa-1}. \end{aligned}$$

By conditions (A6), which implies that

$$\begin{aligned} P(|\Pi_{21}| > l(N)\varepsilon_N) &\le P\Big(\max_{1\le k\le \Gamma_n}\max_{1\le i\le n}\left|\frac{1}{T}\sum_{t=1}^{T}(\bar{G}_{it} - E[\bar{G}_{it}])K_h(U_{it} - w_k)\right| > l(N)\varepsilon_N\Big) \\ &\le C\Gamma_n N^{-l(N)} + Cn\Gamma_n(\log N)^{\kappa+1/2}N^{-\left(\left(\kappa+\frac{1}{2}\right)-\frac{1}{2\delta}\right)}T^{1+\frac{\kappa+1}{\delta}}h^{\kappa-1} \\ &= o(1) + O\left[(\log N)^{\kappa+1/2}N^{-\left(\left(\kappa-\frac{1}{2}\right)-\frac{1}{2\delta}\right)}T^{\frac{\kappa+1}{\delta}}h^{\kappa-3}\right] = o(1). \end{aligned}$$

Thus, we have

$$\Pi_{21} = o_p(l(N)\varepsilon_N). \tag{5.11}$$

By Eqs. (5.5)–(5.11), we know that (5.3) holds. Combining (5.3) and (5.4), we obtain (5.1). The same argument can be used to prove (5.2). Hence, the proof of Lemma 5.1 is completed.

Lemma 5.2. *Suppose that conditions* (A1)–(A6) *hold. Then*

(a) $s_{it} = O\left(\frac{1}{Nh}\right)$.

(b) $(H^{\tau}WH)^{-1} = (H^{\tau}H)^{-1} + O(\zeta_n) = T^{-1}I_n + O(\zeta_n)$ *for sufficiently large n, where* $\zeta_n = \frac{\sqrt{lnn}}{nh}i_n i_n^{\tau}$.

Proof. (a) By the law of large numbers and Lemma 5.1, $N^{-1}D_u^{\tau}W_uD_u = f(u)G(u)\otimes\begin{pmatrix}1 & 0\\ 0 & \mu_2\end{pmatrix}\{1+O_P(c_n)\}$, consequently, $s_{it}(u) = N^{-1}(X_{it}\ 0)[N^{-1}D_u^{\tau}W_uD_u]^{-1}D_{it}K_h(U_{it}-u) = N^{-1}K_h(U_{it}-u)f^{-1}(u)\{1+O_P(c_n)\} = O(\frac{1}{Nh})$.

(b) $H^{\tau}WH = H^{\tau}(I_{nT}-S)^{\tau}(I_{nT}-S)H = H^{\tau}H - H^{\tau}SH - H^{\tau}S^{\tau}H + H^{\tau}S^{\tau}SH = \Omega_{11} - \Omega_{12} - \Omega_{13} + \Omega_{14}$. By (a), we can show that the (i,j)th element of Ω_{1l} is $[\Omega_{1l}]_{ij} = O(\frac{1}{Nh}),\ l = 2, 3, 4$. Similar to the proof of Lemma A.2 in Su and Ullah (2006). We can prove $(H^{\tau}WH)^{-1} = (H^{\tau}H)^{-1} + O(\zeta_n) = T^{-1}I_n + O(\zeta_n)$.

Lemma 5.3. *Suppose that conditions* (A1)–(A7) *hold. Let* e_{it} *is the* $nT \times 1$ *vector having* 1 *in the it entry and all other entries* 0 *. Then*

(a) $Z_{it}e_{it}^{\tau}Wv_{it} = (Z_{it} - G^{-1}(u)\Phi(u)X_{it})v_{it} + o_P(1)$.

(b) $Z_{it}e_{it}^{\tau}WH(H^{\tau}H)^{-1}H^{\tau}Wv_{it} = \frac{1}{T}\sum_{s=1}^{T}(Z_{it} - G^{-1}(u)\Phi(u)X_{it})v_{is} + o_P(1)$.

(c) *Let* $B_{it} = \widetilde{Z}_{it}e_{it}^{\tau}\Lambda(I_{nT} - S)v_{it}$, *Then* $\mathrm{Var}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{n}\sum_{s=1}^{T}B_{it}\right) \xrightarrow{d} \Delta$.

Proof. (a), (b) are the direct results of the Lemma 4.5 in Hu (2014).

(c) By (a) and (b), we have $Z_{it}e_{it}^{\tau}\Lambda(I_{nT} - S)v_{it} = \widetilde{Z}_{it}\widetilde{v}_{it} + o_p(1)$. Clearly, $E(B_{it}) = 0$ and

$$\mathrm{Var}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{n}\sum_{t=1}^{T}B_{it}\right) = \frac{1}{T}\sum_{t=1}^{T}\mathrm{Var}(B_{it}) + \frac{2}{T}\sum_{t=1}^{T-1}(T-t)\mathrm{cov}(B_{i1}, B_{i(t+1)}) = J_1 + J_2.$$

By Lemma 4.6 in Hu (2014), it is easy to obtain $J_1 = \mathrm{Var}(\widetilde{Z}_{it}e_{it}^{\tau}\Lambda(I_{nT} - S)v_{it}) \xrightarrow{d} \Delta$. Hence, in order to prove (c), it suffices to show $J_2 = o(1)$. Let d_n be a sequence of positive integers such that $d_n h \to 0$. Define

$$J_{21} = \frac{2}{T}\sum_{l=1}^{d_n-1}|\mathrm{Cov}(B_{i1}, B_{il})| \quad \text{and} \quad J_{22} = \frac{2}{T}\sum_{l=d_n}^{T}|\mathrm{Cov}(B_{i1}, B_{il})|.$$

From (A5) and (A7), by the choice of d_n, we have

$$J_{21} = \frac{2}{T}\sum_{l=1}^{d_n-1}|\mathrm{Cov}(B_{i1}, B_{il})| \le \frac{2d_n}{T}E|\widetilde{Z}_{i1}\widetilde{Z}_{il}\widetilde{v}_{i1}\widetilde{v}_{il}| \le C\frac{2d_n}{T} = o(1).$$

Next we consider the upper bound of J_{22}. To this end, by using Lemma A.2, we obtain

$$|\mathrm{Cov}(B_{i1}, B_{il})| \le C[E|B_{i1}|^{2+\delta}]^{\frac{1}{2+\delta}}[E|B_{il}|^{2+\delta}]^{\frac{1}{2+\delta}}[\alpha(l)]^{\frac{\delta}{2+\delta}}.$$

From (A4) and (A5), by the choice of d_n, we have

$$J_{22} = \frac{2}{T}\sum_{l=d_n}^{T}|\mathrm{Cov}(B_{i1}, B_{il})| \le \frac{C}{T}\sum_{l=d_n}^{\infty}[\alpha(l)]^{\frac{\delta}{2+\delta}} \le \frac{C}{T}\sum_{l=d_n}^{\infty}d_n^{-(1+\delta)} = o(1).$$

Therefore, $J_2 = o(1)$. This completes the proof of (c).

Lemma 5.4. *Under the conditions of Theorem 3.1, we have*

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

Proof. By the definition of $\eta_i(\beta)$, we have

$$\begin{aligned}\frac{1}{\sqrt{N}}\sum_{i=1}^{n}\eta_i(\beta) &= \frac{1}{\sqrt{N}}\sum_{i=1}^{n}\sum_{t=1}^{T}\widetilde{Z}_{it}e_{it}^{\tau}\Lambda(I_{nT} - S)v_{it} + \frac{1}{\sqrt{N}}\sum_{i=1}^{n}\sum_{t=1}^{T}\widetilde{Z}_{it}e_{it}^{\tau}\Lambda[X_{it}\alpha(u_{it}) - s_{it}^{\tau}M] \\ &= I_{1n} + I_{2n}.\end{aligned}$$

First, we verify that $I_{1n} \xrightarrow{d} N(0, \Delta)$. In view of the Cramér–Wold device, it is sufficient to show that for any $d \times 1$ vector $a = (a_1, a_2, \ldots, a_d)^{\tau} \neq 0$,

$$a^{\tau}I_{1n} = \frac{1}{\sqrt{N}}\sum_{i=1}^{n}\sum_{t=1}^{T}a^{\tau}\widetilde{Z}_{it}e_{it}^{\tau}\Lambda(I_{nT} - S)v_{it} = \frac{1}{\sqrt{N}}\sum_{i=1}^{n}\sum_{t=1}^{T}\widetilde{B}_{it} \to N(0, a^{\tau}\Delta a). \tag{5.12}$$

To prove (5.12), we employ the Doob's small-block and large-block technique. Partition the set $\{1, 2, \ldots, T\}$ into $2q_n + 1$ subsets with large-block of size $r = r_T$ and small-block of size $s = s_T$. Set $q = q_T = [\frac{T}{r_T + s_T}]$ and define the random variables, for $0 \le j \le q$,

$$\pi_{ij} = \sum_{t=k_m}^{k_m+r-1}\widetilde{B}_{it}, \qquad \xi_{ij} = \sum_{t=l_m}^{l_m+s-1}\widetilde{B}_{it}, \quad \text{and} \quad \zeta_{iq} = \sum_{t=q(r+s)+1}^{n}\widetilde{B}_{it}$$

Please cite this article in press as: He, B.-Q., et al., Empirical likelihood for semi-varying coefficient models for panel data with fixed effects. Journal of the Korean Statistical Society (2016), http://dx.doi.org/10.1016/j.jkss.2016.01.001

where $k_m = (m-1)(r+s)+1$, $l_m = (m-1)(r+s)+r+1$, $m = 1, \ldots, q$. Then

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{n}\sum_{t=1}^{T}\widetilde{B}_{it} = \frac{1}{\sqrt{N}}\left\{\sum_{i=1}^{n}\sum_{j=0}^{q}\pi_{ij} + \sum_{i=1}^{n}\sum_{j=0}^{q}\xi_{ij} + \sum_{i=1}^{n}\zeta_{iq}\right\} = \frac{1}{\sqrt{N}}\{S_{n,1}+S_{n,2}+S_{n,3}\}.$$

In order to prove (5.12), it suffices to show the following: as $N \to \infty$,

$$\frac{1}{N}E(S_{n,2})^2 \to 0, \qquad \frac{1}{N}E(S_{n,3})^2 \to 0, \tag{5.13}$$

$$\left|E\exp\left(itS_{n,1}\right) - \prod_{m=1}^{q}E\exp(it\pi_{ij})\right| \to 0, \tag{5.14}$$

$$\mathrm{Var}[(N)^{-1/2}S_{n,1}] \to a^{\tau}\Delta a, \tag{5.15}$$

$$\frac{1}{N}\sum_{i=1}^{n}\sum_{j=1}^{q}E[\pi_{ij}^2 I|\pi_{ij}| > \epsilon\sqrt{n}) \to 0, \quad \forall\epsilon > 0. \tag{5.16}$$

Relation (5.13) implies that $S_{n,2}/\sqrt{N}$ and $S_{n,3}/\sqrt{N}$ are asymptotically negligible in probability; (5.14) shows that the summands π_j in $S_{n,1}/\sqrt{N}$ are asymptotically independent, and (5.15) and (5.16) are the standard Lindeberg–Feller conditions for asymptotic normality of $S_{n,1}/\sqrt{N}$ for the independent setup.

First, we establish (5.13). For this purpose, we choose the large-block size r_T by $r_T = [T^{1/t}]$ and the small-block size by $s_T = [T^{1/(1+t)}]$, where t is given in Condition (A4) and $[x]$ denotes the integer part of x. Then, it can easily be shown from Condition (A4) that:

$$\frac{s_T}{r_T} \to 0, \qquad \frac{r_T}{n} \to 0, \quad \text{and} \quad q_T\alpha(s_T) \le CT^{-1/(t+1)t} \to 0. \tag{5.17}$$

Observe that

$$\frac{1}{N}E(S_{n,2})^2 = \sum_{j=0}^{q}\mathrm{Var}(\xi_{ij}) + 2\sum_{0\le j_1<j_2\le q}\mathrm{Cov}(\xi_{ij_1}, \xi_{ij_2}) \equiv \Pi_1 + \Pi_2. \tag{5.18}$$

It follows from stationarity and Lemma 5.3 that

$$\Pi_1 = q_n\mathrm{Var}(\xi_{ij}) = q_n\mathrm{Var}\left(\sum_{j=1}^{s_T}\widetilde{B}_{ij}\right) = q_T s_T[a^{\tau}\Delta a + o(1)]. \tag{5.19}$$

Next consider the term Π_2. Let $\widetilde{r}_j = j(r_n+s_n)$, then $\widetilde{r}_j - \widetilde{r}_i \ge r_T$ for all $j_2 > j_1$, we therefore have

$$|\Pi_2| \le \sum_{0\le j_1<j_2\le q}\sum_{j_1=1}^{s_T}\sum_{j_2=1}^{s_T}|\mathrm{Cov}(\widetilde{B}_{\widetilde{r}_j+r_T+j_1}, \widetilde{B}_{\widetilde{r}_j+r_T+j_2}) \le \sum_{j_1=1}^{T-r_n}\sum_{j_2=j_1+r_T}^{T}|\mathrm{Cov}(\widetilde{B}_{ij_1}, \widetilde{B}_{ij_2})|.$$

By stationarity and Lemma 5.3, we have

$$|\Pi_2| \le 2T\sum_{j=r_T+1}^{n}|\mathrm{Cov}(\widetilde{B}_{i1}, \widetilde{B}_{ij})| = o(T). \tag{5.20}$$

Hence, by (5.17)–(5.20), we have

$$\frac{1}{N}E(S_{n,2})^2 = O\left(\frac{q_T s_T}{T}\right) + o(1) = o(1). \tag{5.21}$$

As to $\frac{1}{N}E(S_{n,3})^2$, from (A4) and (A7). It follows that

$$\begin{aligned}\frac{1}{N}E(S_{n,3})^2 &= \frac{1}{N}\sum_{i=1}^{n}\sum_{i=q(r+s)+1}^{T}\mathrm{Var}(\widetilde{B}_i) + \frac{1}{N}\sum_{i=1}^{n}\sum_{q(r+s)+1\le j_1<j_2\le T}\mathrm{cov}(\widetilde{B}_{ij_1}, \widetilde{B}_{ij_2}) \\ &\le c_1\frac{T-q(r+s)}{T} + c_2\sum_{q(r+s)+1\le j_1<j_2\le T}\mathrm{cov}(\widetilde{B}_{ij_1}, \widetilde{B}_{ij_2}) = o(1).\end{aligned} \tag{5.22}$$

Combining (5.21) and (5.22), we establish (5.13).

Please cite this article in press as: He, B.-Q., et al., Empirical likelihood for semi-varying coefficient models for panel data with fixed effects. Journal of the Korean Statistical Society (2016), http://dx.doi.org/10.1016/j.jkss.2016.01.001

To establish (5.14), according to Lemma A.1, we have

$$\left|E\exp\Big(it\sum_{j=1}^{q_T}\pi_{ij}\Big)-\Pi_{j=1}^{q_T}E\big[\exp(it\pi_{ij})\big]\right|\le 16\frac{T}{r_T}\alpha(s_T),$$

which goes to zero as $T\to\infty$ by (5.17).

As for (5.15), by stationarity, (5.17) and Lemma 5.3, it is easily seen that

$$\sum_{j=0}^{q}E(\pi_{ij}^2)=q_TE(\pi_{i1}^2)=q_Tr_T\frac{1}{r_T}\mathrm{Var}\Big(\sum_{j=0}^{r_T}\widetilde{B}_{ij}\Big)\to a^\tau\Sigma a.$$

It remains to establish (5.16). For this purpose, by employing Lemma A.3, we have

$$\begin{aligned}E[\pi_{ij}^2I|\pi_{ij}|>\epsilon\sqrt{T})&\le cT^{-\frac{\delta}{2}}E\Big|\sum_{t=k_m}^{k_m+r-1}\widetilde{B}_{it}\Big|^{2+\delta}\\&\le cT^{-\frac{\delta}{2}}\Big\{r^{1+\frac{\delta}{2}}\max_{1\le t\le T}[E(\widetilde{B}_{it})^4]^{\frac{2+\delta}{4}}+r^{1+\epsilon_1}[E(\widetilde{B}_{it})^4]^{\frac{2+\delta}{4}}\Big\}\\&\le cT^{-\frac{\delta}{2}}\Big\{c_1r^{1+\frac{\delta}{2}}+c_2r^{1+\epsilon_1}\Big\}.\end{aligned}$$

Thus, by the definition of r, we obtain

$$\frac{1}{N}\sum_{i=1}^{n}\sum_{j=1}^{q}E[\pi_{ij}^2I|\pi_{ij}|>\epsilon\sqrt{T})\le cT^{-1-\frac{\delta}{2}}\Big\{c_1r^{1+\frac{\delta}{2}}+c_2r^{1+\epsilon_1}\Big\}=o_p(1).$$

Hence, $I_{1n}\xrightarrow{d}N(0,\Delta)$ holds. Similar to the proof of Lemma 4.4 in Hu (2014), we have $\frac{1}{\sqrt{N}}\sum_{i=1}^n\sum_{t=1}^T\widetilde{Z}_{it}e_{it}^\tau\Lambda[X_{it}\alpha(u_{it})-s_{it}^\tau M]=o_p(1)$. Then it is easy to get that $I_{2n}=o_p(1)$. Therefore, the proof of Lemma 5.4 is completed.

Lemma 5.5. *Under the conditions of Theorem 3.1, we have*

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{n}\eta_i(\beta)\eta_i^\tau(\beta)\xrightarrow{p}\Delta.$$

Proof. We also use the notations in the proof of Lemma 5.4, and denote $R_{it}=\widetilde{Z}_{it}e_{it}^\tau\Lambda[X_{it}\alpha(u_{it})-s_{it}^\tau M]$, then

$$\begin{aligned}\frac{1}{\sqrt{N}}\sum_{i=1}^{n}\eta_i(\beta)\eta_i^\tau(\beta)&=\frac{1}{\sqrt{N}}\sum_{i=1}^{n}\sum_{t=1}^{T}B_{it}B_{it}^\tau+\frac{1}{\sqrt{N}}\sum_{i=1}^{n}\sum_{t=1}^{T}R_{it}R_{it}^\tau+\frac{1}{\sqrt{N}}\sum_{i=1}^{n}\sum_{t=1}^{T}B_{it}R_{it}^\tau\\&\quad+\frac{1}{\sqrt{N}}\sum_{i=1}^{n}\sum_{t=1}^{T}R_{it}B_{it}^\tau=\Xi_1+\Xi_2+\Xi_3+\Xi_4.\end{aligned}$$

One can similarly obtain that $\Xi_1\xrightarrow{p}\Delta$ by the law of large numbers and $\Xi_v\xrightarrow{p}0$, $v=2,3,4$. Thus, the proof of Lemma 5.5 is completed.

Lemma 5.6. *Under the conditions of Theorem 3.1, we have*

$$\text{(i)}\ \max_{1\le i\le n}\|\eta_i(\beta)\|=o_p(N^{1/2}),\tag{5.23}$$

$$\text{(ii)}\ \lambda(\beta)=O_p(N^{-1/2}).\tag{5.24}$$

Proof. Similar to the arguments given Owen (1988), from Eq. (2.9) and Lemmas 5.4 and 5.5, we can easily verify Eq. (5.24). We now prove (5.23).

$$\begin{aligned}\max_{1\le i\le n}\|\eta_i(\beta)\|&\le\max_{1\le i\le n}\Big\|\sum_{t=1}^{T}\widetilde{Z}_{it}e_{it}^\tau\Lambda(I_{nT}-S)v_{it}\Big\|+\max_{1\le i\le n}\Big\|\sum_{t=1}^{T}\widetilde{Z}_{it}e_{it}^\tau\Lambda[X_{it}\alpha(u_{it})-s_{it}^\tau M]\Big\|\\&=M_1+M_2.\end{aligned}$$

For M_1, note that $E[\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{Z}_{it}e_{it}^{\tau}\Lambda(I_{nT}-S)v_{it}\widetilde{Z}_{is}e_{is}^{\tau}\Lambda(I_{nT}-S)v_{is}]=\Delta_{s,s}+o(1)<\infty$, where $\Delta_{s,s}$ denotes the sth diagonal element of Δ. B_{it} is independent and identically distributed across the i index for each fixed t. By Lemma 3 in Owen (1990), we obtain $M_1=o_P(N^{1/2})$. Now see the second term M_2, by the Markov inequality, we obtain

$$\begin{aligned}
P(M_2>N^{1/2}) &\le \frac{1}{N}\sum_{i=1}^{n}E\Big\|\sum_{t=1}^{T}\widetilde{Z}_{it}e_{it}^{\tau}\Lambda[X_{it}\alpha(u_{it})-s_{it}^{\tau}M]\Big\|^2\\
&=\frac{1}{N}\sum_{i=1}^{n}E\Big\|\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{Z}_{it}\widetilde{Z}_{is}e_{it}^{\tau}\Lambda[X_{it}\alpha(u_{it})-s_{it}^{\tau}M]e_{is}^{\tau}\Lambda[X_{is}\alpha(u_{is})-s_{is}^{\tau}M]\Big\|\\
&\le \frac{1}{N}\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{s=1}^{T}\Big\{E\big[\widetilde{Z}_{it}e_{it}^{\tau}\Lambda[X_{it}\alpha(u_{it})-s_{it}^{\tau}M]\big]^2\Big\}^{1/2}\Big\{E\big[\widetilde{Z}_{is}e_{is}^{\tau}\Lambda[X_{is}\alpha(u_{is})-s_{is}^{\tau}M]\big]^2\Big\}^{1/2}\\
&=o_p(1)\to 0.
\end{aligned}$$

Therefore, $M_2=o_p(N^{1/2})$ and $\max_{1\le i\le n}\|\eta_i(\beta)\|=o_p(N^{1/2})$.

Proof of Theorem 3.1. Applying the Taylor expansion, from (2.8), we obtain that

$$\log\mathcal{L}_n(\beta)=2\sum_{i=1}^{n}\left\{\lambda^{\tau}\eta_i(\beta)-\frac{[\lambda^{\tau}\eta_i(\beta)]^2}{2}\right\}+o_p(1). \tag{5.25}$$

By (2.9), it follows that

$$\begin{aligned}
0&=\frac{1}{n}\sum_{i=1}^{n}\frac{\eta_i(\beta)}{1+\lambda^{\tau}\eta_i(\beta)}\\
&=\frac{1}{n}\sum_{i=1}^{n}\eta_i(\beta_0)-\frac{1}{n}\sum_{i=1}^{n}\eta_i(\beta)\eta_i^{\tau}(\beta)\lambda+\frac{1}{n}\sum_{i=1}^{n}\frac{\eta_i(\beta)[\lambda^{\tau}\eta_i(\beta)]^2}{1+\lambda^{\tau}\eta_i(\beta)}.
\end{aligned}\tag{5.26}$$

In view of Lemmas 5.4–5.6, we have

$$\begin{aligned}
\left\|\frac{1}{n}\sum_{i=1}^{n}\frac{\eta_i(\beta)[\lambda^{\tau}\eta_i(\beta_0)]^2}{1+\lambda^{\tau}\Lambda_i(\beta)}\right\| &\le \frac{1}{n}\sum_{i=1}^{n}\frac{\|\eta_i(\beta)\|^3\|\lambda\|^2}{|1+\lambda^{\tau}\eta_i(\beta)|}\\
&\le \|\lambda\|^2\max_{1\le i\le n}\|\eta_i(\beta)\|\frac{1}{n}\sum_{i=1}^{n}\|\eta_i(\beta)\|^2=O_p(N^{-1})o_p(N^{1/2})O_p(1)=o_p(N^{-1/2}),
\end{aligned}$$

which, together with (5.26), yields that $\sum_{i=1}^{n}[\lambda^{\tau}\eta_i(\beta)]^2=\sum_{i=1}^{n}\lambda^{\tau}\eta_i(\beta)+o_p(1)$ and

$$\lambda=\left[\sum_{i=1}^{n}\eta_i(\beta)\eta_i^{\tau}(\beta)\right]^{-1}\sum_{i=1}^{n}\eta_i(\beta_0)+o_p(N^{-1/2}).$$

Then, by (5.25), we have

$$\log\mathcal{L}_n(\beta)=\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i(\beta)\right)^{\tau}\left(\frac{1}{n}\sum_{i=1}^{n}\eta_i(\beta)\eta_i^{\tau}(\beta)\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i(\beta)\right)+o_p(1),$$

which, combining with Lemmas 5.4 and 5.5, proves Theorem 3.1.

Proof of Theorem 3.2. Following the similar arguments as were used in the proof of theorem 2 in Hu and Li (2011), we have

$$\widehat{\beta}-\beta=\widehat{\Psi}^{-1}\frac{1}{n}\sum_{i=1}^{n}\eta_i(\beta)+o_p(n^{-1/2}).$$

By Lemma 5.3, we can prove $\widehat{\Psi}\xrightarrow{P}\Psi$ by the law of large numbers. Together with Lemma 5.4 and the Slutsky Theorem, this proves Theorem 2.

6. Conclusion and discussion

In the current paper, the EL method was applied to semi-varying coefficient models for panel data with fixed effects. By using a local linear regression approach and weighted local least-squares method, the fixed effects were removed and the EL statistic for the unknown parameter in the model was constructed. At the same time, we assumed that the observations were stationary α-mixing. The empirical log-likelihood ratio was proven to be asymptotically chi-squared. We also obtained the maximum EL estimator for the parameter of interest, and proved it followed asymptotically normal distribution. The simulation study indicated that, in terms of coverage probabilities and average lengths of the confidence regions, the proposed method performed better than the NA method. An application to a real dataset had been also included. In addition, an interesting topic of further research is that the case of the covariates has measurement errors.

Acknowledgments

The authors are grateful to the Editor and two anonymous referees for providing detailed lists of comments and suggestions which greatly improved the presentation of the paper. This research is supported by the National Natural Science Foundation of China (11401006); the Financial engineering research and development center of Anhui Polytechnic University (JRGCKF201504); Anhui Polytechnic University Youth Science Foundation (2009YQ036); Zhejiang Provincial Key Research Base for Humanities and Social Science Research (JYTtjrc20151202) and the National Statistical Science Research Program of China (2015LY55).

Appendix

This appendix states some lemmas, which are used in the proofs of the main results in Section 5. Let $\{\mathbb{Z}_i,\ i \geq 1\}$ be a stationary α-mixing sequence of random variables with mixing coefficients $\{\alpha(m)\}$.

Lemma A.1 (Volkonskii & Rozanov, 1959)**.** *Let $U_1, \dots, U_n$ be α-mixing random variables measurable with respect to the σ-algebra $\mathcal{F}_{i_1}^{j_1}, \dots, \mathcal{F}_{i_m}^{j_m}$, respectively, with $1 \leq i_1 < j_1 < \cdots < j_m \leq n,\ i_{l+1} - j_l \geq w \geq 1$ and $|U_j| \leq 1$ for $l, j = 1, 2, \dots, m$. Then*

$$\left| E\Big(\prod_{j=1}^{m} U_j\Big) - \prod_{j=1}^{m} EU_j \right| \leq 16(m-1)\alpha_w,$$

where $\mathcal{F}_a^b = \sigma\{U_i, a < i \leq b\}$ denotes σ-field generated by $U_{a+1}, U_{a+2}, \dots, U_b$, α_w is the mixing coefficient.

Lemma A.2 (Hall & Heyde, 1980, Corollary A.2, p. 278)**.** *Suppose that X and Y are random variables such that $E|X|^p < \infty$, $E|Y|^q < \infty$, where $p, q > 1$, $p^{-1} + q^{-1} < 1$. Then*

$$|EXY - EXEY| \leq 8\|X\|_p \|Y\|_q \left\{ \sup_{A\in\sigma(X), B\in\sigma(Y)} |P(A\cap B) - P(A)P(B)| \right\}^{1-p^{-1}-q^{-1}}.$$

Lemma A.3 (Yang, 2007, Theorem 2.2)**.** *Let $\lambda > 2$, $\mu > 0$, $E\mathbb{Z}_i = 0$ and $E|\mathbb{Z}_i|^{\lambda+\mu} < \infty$. Suppose that $\alpha(n) = O(n^{-r})$ for $r > \lambda(\lambda+\mu)/(2\mu)$. Then, for any given $\varepsilon > 0$, there exists constant $C = C(r, \mu, \varepsilon, \lambda)$ such that $E \max_{1\leq k\leq n} |\sum_{i=1}^{k} \mathbb{Z}_i|^{\lambda} \leq C\{n^{\varepsilon} \sum_{i=1}^{n} E|\mathbb{Z}_i|^{\lambda} + (\sum_{i=1}^{n} \|\mathbb{Z}_i\|_{\lambda+\mu}^2)^{\lambda/2}\}$.*

Lemma A.4 ((Fan & Yao, 2003), Theorem 2.18(ii), p. 73)**.** *Suppose that $P(|X_t| \leq b) = 1$. Then*
(ii) For each $q = 1, \dots, [n/2]$ and $\varepsilon > 0$,

$$P(|S_n| > n\varepsilon) \leq 4\exp\Big(-\frac{\varepsilon^2 q}{8v^2(q)}\Big) + 22\left(1 + \frac{4b}{\varepsilon}\right)^{\frac{1}{2}} q\alpha\left(\left[\frac{n}{2q}\right]\right),$$

where $v^2(q) = 2\sigma^2(q)/p^2 + b\varepsilon/2$, $p = n/(2q)$, and

$$\sigma^2(q) = \max_{0\leq j\leq 2q-1} E\,([jp] + 1 - jp)X_1 + X_2 + \cdots + X_{[(j+1)p_{[jp]}]} + (jp + p - [jp+p])X_{[(j+1)p]-[jp]+1}{}^2.$$

References

Auestad, B., & Tjøstheim, D. (1990). Identification of nonlinear time series: First order characterization and order determination. *Biometrika*, *77*, 669–687.
Baltagi, B. H. (2005). *Econometric analysis of panel data* (second ed.). New York: Wiley.
Cai, Z., & Li, Q. (2008). Nonparametric estimation and varying coefficient dynamic panel data models. *Econometric Theory*, *24*, 1321–1342.
Chen, J., Gao, J. T., & Li, D. G. (2013). Estimation in single-index panel data models with heterogeneous link functions. *Econometric Reviews*, *32*(8), 928–955.
Fan, J., & Huang, T. (2005). Profile likelihood inferences on semiparametric varying-coefficient partially linear models. *Bernoulli*, *11*(6), 1031–1057.
Fan, G. L., & Liang, H. Y. (2010). Empirical likelihood inference for semiparametric model with linear process errors. *Journal of the Korean Statistical Society*, *39*, 55–65.
Fan, G. L., Liang, H. Y., & Huang, Z. S. (2012). Empirical likelihood for partially time-varying coefficient models with dependent observations. *Journal of Nonparametric Statistics*, *24*, 71–84.
Fan, J., & Yao, Q. W. (2003). *Nonlinear times series: Nonparametric and parametric methods*. New York: Springer.
Fan, J., & Zhang, W. (1999). Statistical estimation in varying coefficient models. *Annals of Statistics*, *27*, 1491–1518.
Hall, P., & Heyde, C. C. (1980). *Martingale limit theory and its application*. New York: Academic Press.

Henderson, D. J., Carroll, R. J., & Li, Q. (2008). Nonparametric estimation and testing of fixed effects panel data models. *Journal of Econometrics*, *144*, 257–275.
Henderson, D. J., & Ullah, A. (2005). A nonparametric random effects estimator. *Economics Letters*, *88*, 403–407.
Hsiao, C. (2003). *Analysis of panel data* (second ed.). Cambridge University Press.
Hu, X. M. (2014). Estimation in a semi-varying coefficient model for panel data with fixed effects. *Journal of the System Science and Complexity*, *27*, 594–604.
Hu, Y., & Li, T. T. (2011). Empirical likelihood for semiparametric varying coefficient partially linear models with longitudinal data. *Statistics & Probability Letters*, *80*, 111–121.
Huang, J. Z., Wu, C. O., & Zhou, L. (2002). Varying-coefficient models and basis function approximations for the analysis of repeated measurements. *Biometrika*, *89*, 118–128.
Kaslow, R. A., Ostrow, D. G., Detels, J. P., Polk, B. F., & Rinaldo, C. R. (1987). The multicenter AIDS cohort study: rational,oranization and selected characteristics of the prticipants. *American Journal of Epidemiology*, *126*, 310–318.
Masry, E., & Tjøstheim, D. (1995). Nonparametric estimation and identification of nonlinear ARCH time series: Strong convergence and asymptotic normality. *Econometric Theory*, *11*, 258–289.
Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, *75*, 237–249.
Owen, A. B. (1990). Empirical likelihood ratio confidence regions. *Annals of Statistics*, *18*, 90–120.
Qin, J., & Lawless, B. (1994). Empirical likelihood and general estimating equations. *Annals of Statistics*, *22*, 300–325.
Shi, J., & Lau, T. S. (2000). Empirical likelihood for partially linear models. *Journal of Multivariate Analysis*, *72*, 132–148.
Su, L. J., & Ullah, A. (2006). Profile likelihood estimation of partially linear panel data models with models with fixed effects. *Econometrics Letters*, *92*, 75–81.
Volkonskii, V. A., & Rozanov, Y. A. (1959). Some limit theorems for random functions. *Theory of Probabiity and its Applications*, *4*, 178–197.
Wang, Q. H., & Jing, B. Y. (1999). Empirical likelihood for partial linear models with fixed designs. *Statistics & Probability Letters*, *41*, 425–433.
Xue, L. G., & Zhu, L. X. (2007). Empirical likelihood semiparametric regression analysis for longitudinal data. *Biometrika*, *94*, 921–937.
Yang, S. C. (2007). Maximal moment inequality for partial sums of strong mixing sequences and application. *Acta Mathematica Sinica (English Series)*, *23*, 1013–1024.
You, J. H., & Zhou, Y. (2006). Empirical likeliood for semiparametric varying-coefficient partially linear regression models. *Statistics & Probability Letters*, *76*, 412–422.
Zhou, X. C., & Lin, J. G. (2014). Empirical likeliood for varying-coefficient semiparametric mixed-effects errors-in-variables models with longitudinal data. *Statistical Methods & Applications*, *23*, 51–69.