Journal of Mathematical Research with Applications
Jan., 2012, Vol. 32, No. 1, pp. 95–107
DOI:10.3770/j.issn:2095-2651.2012.01.011
Http://jmre.dlut.edu.cn

Empirical Likelihood for Semiparametric Varying-Coefficient Heteroscedastic Partially Linear Models

Guo Liang FAN[1,2], **Hong Xia XU**[1,*]
1. *School of Mathematics & Physics, Anhui Polytechnic University, Anhui* 241000*, P. R. China;*
2. *Department of Mathematics, Tongji University, Shanghai* 200092*, P. R. China*

Abstract Consider the semiparametric varying-coefficient heteroscedastic partially linear model $Y_i = X_i^\tau\beta + Z_i^\tau\alpha(T_i) + \sigma_i e_i$, $1 \le i \le n$, where $\sigma_i^2 = f(U_i)$, β is a $p \times 1$ column vector of unknown parameter, (X_i, Z_i, T_i, U_i) are random design points, Y_i are the response variables, $\alpha(\cdot)$ is a q-dimensional vector of unknown functions, e_i are random errors. For both cases that $f(\cdot)$ is known and unknown, we propose the empirical log-likelihood ratio statistics for the parameter β. For each case, a nonparametric version of Wilks' theorem is derived. The results are then used to construct confidence regions of the parameter. Simulation studies are carried out to assess the performance of the empirical likelihood method.

Keywords Empirical likelihood; heteroscedastic partially linear model; varying-coefficient model; local linear method; confidence region.

MR(2010) Subject Classification 62G15; 62E20

1. Introduction

Regression analysis is one of the most mature and widely applied branches of statistics. Various regression models (such as parametric regressions, nonparametric regressions and semiparametric regressions) have been extensively studied by many researchers. Recently, there has been increasing interest and activity in the general area of varying-coefficient partially linear model in statistics which has the following form (see Fan and Huang (2005)):

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

where Y is the response, $(X, Z) \in R^p \times R^q$ and $T \in R$ are regressors, $\beta = (\beta_1, \ldots, \beta_p)^\tau$ is a vector of p-dimensional unknown parameters, $\alpha(T) = (\alpha_1(T), \ldots, \alpha_q(T))^\tau$ is a q-dimensional vector of unknown functions and ε is the random error.

Obviously, model (1.1) includes many usual parametric, semiparametric, nonparametric and varying-coefficient regression models. For example, when $\alpha(\cdot) \equiv \alpha$ where α is a constant vector,

Received May 28, 2010; Accepted November 20, 2010
Supported by the National Natural Science Foundation of China (Grant No. 71171003), Natural Science Research Project of Anhui Provincial Colleges (Grant No. KJ2011A032), Anhui Polytechnic University Foundation for Recruiting Talent (Grant Nos. 2011YQQ004; 2009YQQ005), Young Teachers Science Research Foundation of Anhui Polytechnic University (Grant No. 2009YQ035) and Anhui Provincial Natural Science Foundation.
* Corresponding author
E-mail address: guoliangfan@yahoo.com (H. X. XU)

model (1.1) reduces to the usual linear regression model. When $q = 1$ and $Z = 1$, model (1.1) becomes the famous partially linear regression model, which was proposed by Engle et al. (1986) to study the effect of weather on electricity demand. The partially linear regression model has attracted considerable attention in the last decade and various methods and techniques have been proposed and studied (for example, kernel method; spline method; local polynomial type estimator; series estimator; two-stage estimator; and others). A survey of the estimation and application of partially linear regression model can be found in the monograph of Härdle et al. (2000). When $\beta = 0$, i.e., the parametric component is removed, model (1.1) reduces to the varying-coefficient regression model, which has been widely studied in the literature as well; see, for example, the work of Hastie and Tibshirani (1993), Fan and Zhang (1999) and Huang et al. (2002) among others.

It is well known that there are some striking advantages with the empirical likelihood (EL) method proposed by Owen (1988,1990) in the construction of confidence regions and intervals for unknown parameter. For example, the EL based inference does not involve covariance estimation and the EL method uses only the data to determine the shape and orientation of confidence regions. The EL method has been applied to partially linear model when the errors are assumed to be independent and identically distributed (i.i.d.), see Shi and Lau (2000), and Wang and Jing (1999, 2003). Meanwhile, when the errors are not i.i.d., the EL method has been successfully applied to partially linear EV model, e.g. Fan and Liang (2010) applied the EL method to partially linear model with linear process errors; Fan et al. (2010) proposed the empirical log-likelihood ratio for the parametric parts and the nonparametric version of the Wilks' theorems was derived in heteroscedastic partially linear model with martingale difference errors. Lu (2009) made empirical-likelihood-based inference for the parameters in heteroscedastic partially linear models.

Model (1.1) has been extensively studied by many authors. For example, Zhang et al. (2002) developed the procedures for estimation of the unknown parameters β and the unknown functions $\alpha(\cdot)$. Li et al. (2002) proposed a local least-squares method with a kernel weight function to estimate the model (1.1). They also used this model to estimate the production function of the nonmetal mineral industry in China. Fan and Huang (2005) proposed a profile least-squares technique for estimating parametric component and the asymptotic normality of the profile least-squares estimator was derived. Recently, You and Zhou (2006) used the EL method to study the confidence regions construction for β in model (1.1) when the errors are i.i.d. However, in many cases, the homoscedastic assumption for the errors is strict and has limited applications and heteroscedasticity is often found in residuals from both cross-sectional and time series modeling in applications. It is well known that if the errors are heteroscedastic, the least-squares estimator of β is inefficient and the conventional estimator of the covariance matrix is usually inconsistent. Heteroscedasticity has been applied to the partially linear model by many researchers, see Baltagi (1995), Carroll (1982), You and Chen (2005), You et al. (2007) and Fan et al. (2010) among others. Motivated by the above, in this paper, we assume the error in model (1.1) is heteroscedastic and a function (known or unknown) of random variables. To the best of our knowledge, this point has not been discussed in the literature.

Suppose that $\{Y_i, X_i, Z_i, T_i, 1 \le i \le n\}$ is a sample from (Y, X, Z, T) in the model (1.1) with $\varepsilon = \sqrt{f(U)}e$, where $f(\cdot)$ is a function defined on a compact subset Ω of the real line R, i.e.,

$$Y_i = X_i^\tau \beta + Z_i^\tau \alpha(T_i) + \varepsilon_i, \quad 1 \le i \le n, \tag{1.2}$$

where $\varepsilon_i = \sqrt{f(U_i)}e_i := \sigma_i e_i$, $X_i = (X_{i1}, \ldots, X_{ip})^\tau$. The model (1.2) belongs to semiparametric varying-coefficient heteroscedastic partially linear regression model. Let $\{X_i, Z_i, T_i, U_i, e_i, 1 \le i \le n\}$ be i.i.d. random variables with $E(e_i|X_i, Z_i, T_i, U_i) = 0$ a.s. and $\mathrm{Var}(e_i|X_i, Z_i, T_i, U_i) = 1$ a.s.

The rest of this paper is organized as follows. The empirical log-likelihood ratio for β is proposed in Section 2. Assumption conditions and main results are given in Section 3. Simulation studies are conducted in Section 4. The proofs of the main results are relegated to Section 5.

2. Empirical likelihood of the parametric components

In this section we construct the empirical likelihood confidence region of the parameter vector β when f is known and unknown, respectively. To apply the empirical likelihood method to the semiparametric varying-coefficient heteroscedastic partially linear model, we have to give estimators of $\alpha_j(\cdot)$ $(j = 1, \ldots, q)$ and $f(\cdot)$.

Our basic idea is as follows: without considering heteroscedasticity, suppose β is known, then the model (1.2) can be reduced to a varying-coefficient regression model which can be written as

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

where X_{ij} and Z_{ij} are the jth elements of X_i and Z_i, respectively.

Now, apply a local linear regression technique to estimate the varying coefficient functions $\{\alpha_j(\cdot), j = 1, \ldots, q\}$ in (2.1). For t in a small neighborhood of t_0, one can approximate $\alpha_j(t)$ 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 $\alpha_j'(t) = \partial\alpha_j(t)/\partial t$. This leads to the following weighted local least-squares problem: find $\{(a_j, b_j), j = 1, \ldots, q\}$ to minimize

$$\sum_{i=1}^{n} \left\{ \Big(Y_i - \sum_{j=1}^{p} X_{ij}\beta_j\Big) - \sum_{j=1}^{q} [a_j + b_j(T_i - t)]Z_{ij} \right\}^2 K_{h_1}(T_i - t), \tag{2.2}$$

where $K_{h_1}(\cdot) = K(\cdot/h_1)/h_1$, $K(\cdot)$ is a kernel function and $h_1 = h_{1n}$ is a sequence of positive numbers tending to zero, called bandwidth. Simple calculation yields

$$(\hat\alpha_1(t), \ldots, \hat\alpha_q(t), h_1\hat b_1(t), \ldots, h_1\hat b_q(t))^\tau = (D_t^\tau w_t D_t)^{-1} D_t^\tau w_t (Y - X\beta),$$

where

$$X=\begin{pmatrix} X_1^\tau \\ \vdots \\ X_n^\tau \end{pmatrix}=\begin{pmatrix} X_{11} & \cdots & X_{1p} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{np} \end{pmatrix},\quad D_t=\begin{pmatrix} Z_1^\tau & \frac{T_1-t}{h_1}Z_1^\tau \\ \vdots & \vdots \\ Z_n^\tau & \frac{T_n-t}{h_1}Z_n^\tau \end{pmatrix},$$

$Y=(Y_1,\dots,Y_n)^\tau$ and $w_t=\mathrm{diag}(K_{h_1}(T_1-t),\dots,K_{h_1}(T_n-t))$.

When the function $f(U_k)=\sigma_k^2$ is unknown, the local linear estimator of $f(\cdot)$ is defined by $\hat f_n(\cdot)=\hat\mu$, where

$$(\hat\mu,\hat\nu)=\arg\min_{\mu,\nu}\sum_{i=1}^n\Big\{\big[Y_i-\sum_{j=1}^p X_{ij}\beta_j-\sum_{j=1}^q\hat\alpha_j(T_i)Z_{ij}\big]^2-\mu-\nu(U_i-u)\Big\}^2K_{h_2}(U_i-u).$$

Then we get the estimator of σ_k^2,

$$\hat\sigma_{nk}^2=\hat f_n(U_k)=\sum_{i=1}^n W_{h_2i}(U_k)\Big(Y_i-\sum_{j=1}^p X_{ij}\beta_j-\sum_{j=1}^q\hat\alpha_j(T_i)Z_{ij}\Big)^2, \tag{2.3}$$

where $h_2=h_{2n}$ is a bandwidth and $W_{h_2i}(u)=\mathbf{U}_{h_2i}(u)/\sum_{j=1}^n\mathbf{U}_{h_2j}(u)$ with $\mathbf{U}_{h_2i}(u)=K_{h_2}(U_i-u)\{S_{h_2,2}(u)-(U_i-u)S_{h_2,1}(u)\}$, $S_{h_2,l}(u)=\frac{1}{nh_2}\sum_{i=1}^n K\Big(\frac{U_i-u}{h_2}\Big)(U_i-u)^l$, $l=0,1,2$.

Let

$$\tilde Y_i=Y_i-\sum_{k=1}^n S_{ik}Y_k,\quad \tilde X_{ij}=X_{ij}-\sum_{k=1}^n S_{ik}X_{kj},$$

where S_{ik} is the (i,k)th component of matrix S with

$$S=\begin{pmatrix} (Z_1^\tau\ \ 0_q^\tau)(D_{T_1}^\tau w_{T_1}D_{T_1})^{-1}D_{T_1}^\tau w_{T_1} \\ \vdots \\ (Z_n^\tau\ \ 0_q^\tau)(D_{T_n}^\tau w_{T_n}D_{T_n})^{-1}D_{T_n}^\tau w_{T_n} \end{pmatrix}.$$

In order to introduce two auxiliary random variables which will be used to define the empirical log-likelihood ratio functions, as to model (1.2), let $d_n^2(\beta)=\frac1n\sum_{i=1}^n\sigma_i^{-2}\big[(Y_i-X_i^\tau\beta-Z_i^\tau\hat\alpha(T_i)\big]^2$. An estimator of the true parameter β_0 should make $d_n^2(\beta)$ attain minimum, hence we have

$$\sum_{i=1}^n\sigma_i^{-2}\big[\tilde X_i(Y_i-X_i^\tau\beta-Z_i^\tau\hat\alpha(T_i))\big]=0.$$

We now introduce two auxiliary random variables $\Lambda_{ki}(\beta)$ $(k=1,2)$ under the conditions that f is known and unknown, respectively,

$$\Lambda_{1i}(\beta)=\sigma_i^{-2}\tilde X_i(Y_i-X_i^\tau\beta-Z_i^\tau\hat\alpha(T_i))=\sigma_i^{-2}\tilde X_i(\tilde Y_i-\tilde X_i^\tau\beta), \tag{2.4}$$
$$\Lambda_{2i}(\beta)=\hat\sigma_{ni}^{-2}\tilde X_i(Y_i-X_i^\tau\beta-Z_i^\tau\hat\alpha(T_i))=\hat\sigma_{ni}^{-2}\tilde X_i(\tilde Y_i-\tilde X_i^\tau\beta). \tag{2.5}$$

Similarly to Owen (1990), we define an empirical log-likelihood ratio function under the conditions that f is known and unknown, respectively, as follows.

$$\mathcal{L}_{kn}(\beta)=-2\max\Big\{\sum_{i=1}^n\log(np_{ki}):\sum_{i=1}^n p_{ki}\Lambda_{ki}(\beta)=0,p_{ki}\ge0,\sum_{i=1}^n p_{ki}=1\Big\}.$$

By the Lagrange multiplier method, one can obtain that $p_{ki} = \frac{1}{n[1+\lambda_k^\tau \Lambda_{ki}(\beta)]}$, and $\mathcal{L}_{kn}(\beta)$ can be represented as

$$\mathcal{L}_{kn}(\beta) = 2\sum_{i=1}^{n} \log\{1 + \lambda_k^\tau \Lambda_{ki}(\beta)\}, \quad k = 1, 2, \tag{2.6}$$

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

$$\frac{1}{n}\sum_{i=1}^{n} \frac{\Lambda_{ki}(\beta)}{1 + \lambda_k^\tau \Lambda_{ki}(\beta)} = 0, \quad k = 1, 2. \tag{2.7}$$

We will show in the next section that if β is the true parameter vector, $\Lambda_{kn}(\beta)$ $(k = 1, 2)$ is asymptotically χ^2-distributed when f is known and unknown, respectively.

3. The asymptotic results

In this section, we establish the nonparametric Wilks' theorem for $\Lambda_{kn}(\beta)$ $(k = 1, 2)$ when f is known and unknown, respectively. Let C denote positive constants whose values may vary at each occurrence. Before formulating the main results, we first give the following assumptions.

(A1) The random variables T and U both have a bounded support Ω; The density functions $p_T(t)$ of T and $p_U(u)$ of U both are Lipschitz continuous and bounded away from 0 on Ω.

(A2) The $q \times q$ matrix $\Gamma(T) \widehat{=} E(ZZ^\tau|T)$ is nonsingular for each $T \in \Omega$. $E(XX^\tau|T)$, $\Gamma^{-1}(T)$ and $\Phi(T) \widehat{=} E(ZX^\tau|T)$ are all Lipschitz continuous.

(A3) There is an $s > 2$ such that $E\|X\|^{2s} < \infty$, $E\|Z\|^{2s} < \infty$, $E\|U\|^{2s} < \infty$, $E|e|^{2s} < \infty$, and for some $\delta < 2 - s^{-1}$ there is $n^{2\delta-1}h_1 \to \infty$.

(A4) $\{\alpha_j(\cdot), j = 1, \ldots, q\}$ and $f(\cdot)$ have the continuous second derivative in Ω and $0 < m_0 \leq \min_{1\leq i\leq n} f(U_i) \leq \max_{1\leq i\leq n} f(U_i) \leq M_0 < \infty$ a.s.

(A5) The kernel $K(v)$ is a symmetric probability density function with a continuous derivative on its support $[-1, 1]$.

(A6) The bandwidth h_1 satisfies $nh_1^2/\log^2 n \to \infty$ and $nh_1^8 \to 0$.

(A7) The bandwidths h_1 and h_2 satisfy that $n^{1-\frac{3}{2s}}h_1/\log^2 n \to \infty$, $n^{3/s}h_1^8 \log^2 n = O(1)$, $nh_2^5 = O(1)$, $nh_2^{s/(s-2)} \to \infty$, $n^{1-1/s}h_1^{1/2}h_2/\log^{3/2} n \to \infty$ and $n^{-(1/2-1/s)}h_1^2 h_2^{-1} \log n = O(1)$.

Remark 2.1 *Assumptions (A1)–(A7), while looking a bit lengthy, are actually quite mild and can be easily satisfied. (A1)–(A5) can be found in Fan and Huang (2005) and Fan et al. (2010). The technical conditions of (A6) and (A7) are easily satisfied. For example, when $s = 3$, we can choose $h_1 = h_2 = Cn^{-1/4}$.*

Theorem 3.1 *Suppose that (A1)–(A6) hold. For model (1.2), if β_0 is the true value of the parameter β, then $\mathcal{L}_{1n}(\beta_0) \xrightarrow{d} \chi_p^2$, where χ_p^2 is a standard chi-square random variable with p degrees of freedom and $\xrightarrow{d}$ stands for convergence in distribution.*

Theorem 3.2 *Suppose that (A1)-(A7) hold. If β_0 is the true value of the parameter β in model (1.2), then $\mathcal{L}_{2n}(\beta_0) \xrightarrow{d} \chi_p^2$.*

Remark 2.2 *By using Theorems 3.1 and 3.2, an approximate $1 - \alpha$ level confidence region for*

β_0 can be taken as $I_{k\alpha} = \{\beta : \mathcal{L}_{kn}(\beta) \le c_\alpha\},\ k = 1, 2$ under the conditions that $f(\cdot)$ is known and unknown, respectively, where c_α is chosen to satisfy $P(\chi_p^2 \le c_\alpha) = 1 - \alpha$.

4. Simulation studies

In this section, we report some Monte Carlo experiments for the proposed EL method of varying-coefficient heteroscedastic partially linear model (1.2) in the cases $p = 1$ and $p = 2$.

First, we conduct a simulation study with $p = 1$. For simplicity we consider the following semiparametric varying-coefficient heteroscedastic partially linear model:

$$Y_i = X_i\beta_0 + Z_i^\tau \alpha(T_i) + \sqrt{f(U_i)}e_i,$$

where $\beta_0 = 1$. Take $f(u) = 1-0.5\cdot\sin(4\pi u)$ when $f(\cdot)$ is known. When $f(\cdot)$ is unknown, we utilize (2.3) to estimate $f(\cdot)$. The design points $X_i \sim N(1, 1)$, the covariate T_i is uniformly distributed on $[0, 1]$, $U_i \sim N(0, 1)$, $e_i \sim N(0, 1)$, the nonparametric component $\alpha(t) = (\alpha_1(t), \alpha_2(t))^\tau$ with $q = 2$ in which Z_{i1} and Z_{i2} are normal random variables with mean 1 and variance 1, and the coefficient functions are given as $\alpha_1(t) = \sin(2\pi t)$ and $\alpha_2(t) = t/(1 + t)$.

In the simulation below, let $K(t)$ be bi-weight kernel function $K(t) = \frac{15}{16}(1 - t^2)^2 I\{|t| \le 1\}$ and $h_1 = h_2 = n^{-1/5}$. The sample sizes n are chosen to be 20, 50 and 100, respectively. The coverage probabilities and the average lengths of the confidence intervals are calculated based on 1000 samples of simulated data. The nominal levels are taken to be $\alpha = 0.10,\ 0.05$. Some representative coverage probabilities and average lengths of confidence intervals are reported in Tables 1 and 2 when f is known and unknown, respectively.

n	$\alpha = 0.10$		$\alpha = 0.05$	
	CP	AL	CP	AL
20	0.780	0.5782	0.853	0.6254
50	0.863	0.4145	0.920	0.5087
100	0.881	0.2631	0.928	0.3127

Table 1 Coverage probabilities (CP) and average lengths (AL): when $f(\cdot)$ is known and $\beta_0 = 1$

n	$\alpha = 0.10$		$\alpha = 0.05$	
	CP	AL	CP	AL
20	0.778	0.5788	0.833	0.6282
50	0.852	0.4158	0.913	0.5136
100	0.865	0.2846	0.924	0.3502

Table 2 Coverage probabilities (CP) and average lengths (AL): when $f(\cdot)$ is unknown and $\beta_0 = 1$

Next, we conduct the simulation for the model (1.2) in the case $p = 2$. Consider the following semiparametric varying-coefficient heteroscedastic partially linear model:

$$Y_i = X_i^\tau\beta_0 + Z_i^\tau \alpha(T_i) + \sqrt{f(U_i)}e_i,$$

where $\beta_0 = (1,1)^\tau$, X_i i.i.d. $\sim N(1, I_2)$, where I_2 is the 2×2 unit matrix. The other settings are the same as the above $p = 1$ case.

When the sample sizes n are chosen to be 20, 50 and 100, the coverage probabilities of the confidence regions of β_0 are calculated from 500 runs and the nominal levels are taken to be $\alpha = 0.10$ and 0.05, respectively. The representative coverage probabilities of confidence regions are reported in Table 3. On the other hand, when the sample size $n = 50$, we plot the confidence region for β_0 which satisfies $\mathcal{L}_{kn}(\beta) \le c_\alpha (k = 1,2)$. Here c_α is the $(1-\alpha)$-quantile of standard chi-square distribution with 2 degrees of freedom with $\alpha = 0.10$ and 0.05. The confidence regions of β_0 are presented in Figures 1 and 2 when f is known and unknown, respectively.

From Tables 1–3, we see the EL method performs well in size for known and unknown f. It can be seen that the coverage probabilities of the confidence intervals (regions) tend to increase and the average lengths decrease as the sample size n becomes larger. At the same time, it is clear that the confidence intervals (regions) in the case for known f have bigger coverage probabilities and shorter average lengths than those in the case for unknown f. Figures 1 and 2 show that the confidence regions of β_0 with $\alpha = 0.05$ are wider than those with $\alpha = 0.10$ for known and unknown f.

n	$\alpha = 0.10$		$\alpha = 0.05$	
	f known	f unknown	f known	f unknown
20	0.778	0.776	0.851	0.831
50	0.856	0.845	0.917	0.912
100	0.878	0.862	0.925	0.920

Table 3 Coverage probabilities: when $\beta_0 = (1,1)^\tau$ and $f(\cdot)$ is known and unknown respectively

5. Proofs of the main results

For convenience and simplicity, let $c_n = \{(nh_1)^{-1}\log n\}^{1/2} + h_1^2$, $\mu_k = \int t^k K(t)dt$, and $v_k = \int t^k K^2(t)\mathrm{d}\,t$, $k = 0,1,2,4$. $\{j_1,\dots,j_n\}$ stands for any permutation of $\{1,\dots,n\}$.

Lemma 5.1 ([16, Lemma 1]) *Let $\{\zeta_i,\dots,\zeta_n\}$ be i.i.d. random variables with $E\zeta_1 = 0$, and $E|\zeta_1|^r \le C < \infty$ for some $r > 1$. Suppose that $\{a_{ij}, 1 \le i,j \le n\}$ is a series of real numbers such that $\max_{1\le j\le n}\sum_{i=1}^n |a_{ij}| \le C < \infty$. Set $d_n = \max_{1\le i,j\le n}|a_{ij}|$. Then $\max_{1\le j\le n}|\sum_{i=1}^n a_{ij}\zeta_i| = O((n^{1/r}d_n \vee d_n^{1/2})\log n)$ a.s.*

Lemma 5.2 *Let $\{\zeta_i,\dots,\zeta_n\}$ be i.i.d. random variables.*

(i) If $E\zeta_1 = 0$ and $E|\zeta_1|^3 < \infty$, then $\max_{1\le m\le n}|\sum_{i=1}^m \xi_{j_i}| = O(n^{1/2}\log n)$ a.s.

(ii) If $E|\zeta_1|^s$ is bounded for $s > 1$, then $\max_{1\le i\le n}|\zeta_i| = o(n^{1/s})$ a.s.

(i) in Lemma 5.2 is a particular situation of Lemma 4 in Sun et al. (2002) and the proof of (ii) in Lemma 5.2 can be found in Shi and Lau (2000).

Lemma 5.3 *Suppose that (A1)–(A6) hold. Then as $n\to\infty$ there hold uniformly in $T\in\Omega$*

$$D_T^\tau w_T D_T = np_T(T)\Gamma(T)\otimes \mathrm{diag}(1,\mu_2)\{1+O_p(c_n)\},$$

$$D_T^\tau w_T X = np_T(T)\Phi(T) \otimes (1,0)^\tau \{1 + O_p(c_n)\},$$

$$D_T^\tau w_T M = np_T(T)\Gamma(T) \otimes (1,0)^\tau \alpha(T)\{1 + O_p(c_n)\},$$

where $\otimes$ is the Cronecker product. See the proof of Lemma 2 in Fan and Huang (2005).

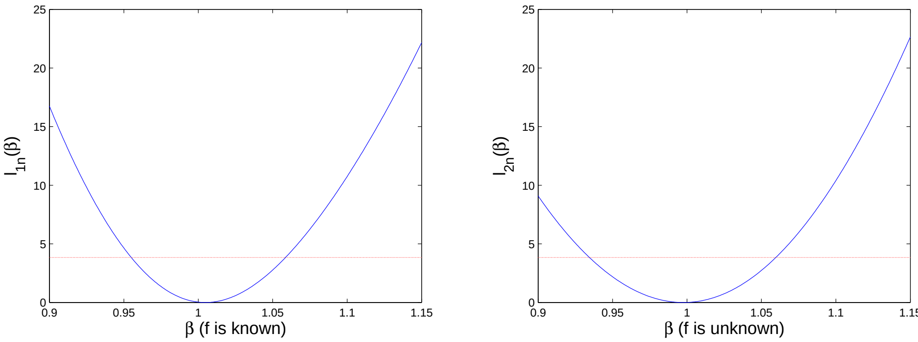

Figure 1 The plots of confidence regions for $\beta_0 = (\beta_1, \beta_2)^\tau = (1,1)^\tau$ when f is known

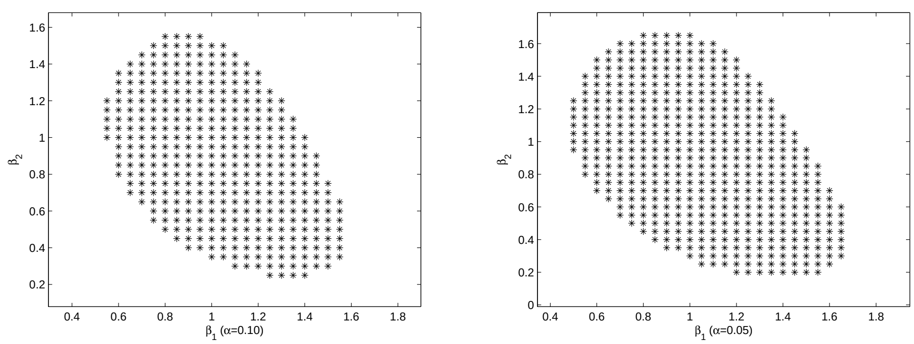

Figure 2 The plots of confidence regions for $\beta_0 = (\beta_1, \beta_2)^\tau = (1,1)^\tau$ when f is unknown

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

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \Lambda_{1i}(\beta_0) \xrightarrow{d} N(0, \Sigma), \tag{5.1}$$

$$\frac{1}{n} \sum_{i=1}^n \Lambda_{1i}(\beta_0)\Lambda_{1i}^\tau(\beta_0) \xrightarrow{p} \Sigma, \tag{5.2}$$

$$\max_{1 \le i \le n} \|\Lambda_{1i}(\beta_0)\| = o_p(n^{1/2}), \tag{5.3}$$

$$\lambda_1 = O_p(n^{-1/2}), \tag{5.4}$$

where $\Sigma = E\{f^{-1/2}(U)(X - \Phi^\tau(T)\Gamma^{-1}(T)Z)\}^{\otimes 2}$. Here $A^{\otimes 2} = AA^\tau$. Note that

$$\Sigma = E(f^{-1}(U))\big\{E(XX^\tau) - E\big[\Phi(T)^\tau\Gamma(T)^{-1}\Phi(T)\big]\big\}$$

if U is independent of (X, Z, T).

Proof Put $M = \big(\alpha^\tau(T_1)Z_1, \ldots, \alpha^\tau(T_n)Z_n\big)^\tau$ and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^\tau$. By the definition of $\Lambda_{1i}(\beta_0)$ and Lemma A.4 in Fang and Huang (2005), we can derive that

$$\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{i=1}^n \Lambda_{1i}(\beta_0) &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \sigma_i^{-2}\big(X_i - X^\tau S_i\big)\big(\varepsilon_i - \varepsilon^\tau S_i + Z_i^\tau \alpha(T_i) - M^\tau S_i\big) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n \sigma_i^{-2}\big(X_i - \Phi^\tau(T_i)\Gamma^{-1}(T_i)Z_i\big)\varepsilon_i + o_p(1) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n f^{-1/2}(U_i)\big(X_i - \Phi^\tau(T_i)\Gamma^{-1}(T_i)Z_i\big)e_i + o_p(1) \\
&:= \frac{1}{\sqrt{n}}\sum_{i=1}^n \eta_{in} + o_p(1).
\end{aligned} \tag{5.5}$$

Obviously, η_{in} are i.i.d. random vectors with $E\eta_{in} = 0$ and $\mathrm{Cov}(\eta_{in}) = \Sigma$. Then, by the multivariate central limit theorem, we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n \eta_{in} \xrightarrow{d} N(0, \Sigma). \tag{5.6}$$

Therefore, by (5.5) and (5.6), we obtain (5.1).

By (5.5) and $\mathrm{Cov}(\eta_{in}) = \Sigma$, we can easily prove (5.2). From (A4) and Lemma 5.3 in You and Zhou (2006), we can derive (5.3). Using (5.2) and (5.3), we can easily get (5.4). $\square$

Proof of Theorem 3.1 Applying the Taylor expansion, from (2.6) and Lemma 5.4, we obtain that

$$\mathcal{L}_{1n}(\beta_0) = 2\sum_{i=1}^n \{\lambda_1^\tau \Lambda_{1i}(\beta_0) - [\lambda_1^\tau \Lambda_{1i}(\beta_0)]^2/2\} + o_p(1). \tag{5.7}$$

By (2.7), it follows that

$$\begin{aligned}
0 &= \frac{1}{n}\sum_{i=1}^n \frac{\Lambda_{1i}(\beta_0)}{1 + \lambda_1^\tau \Lambda_{1i}(\beta_0)} \\
&= \frac{1}{n}\sum_{i=1}^n \Lambda_{1i}(\beta_0) - \frac{1}{n}\sum_{i=1}^n \Lambda_{1i}(\beta_0)\Lambda_{1i}^\tau(\beta_0)\lambda_1 + \frac{1}{n}\sum_{i=1}^n \frac{\Lambda_{1i}(\beta_0)[\lambda_1^\tau \Lambda_{1i}(\beta_0)]^2}{1 + \lambda_1^\tau \Lambda_{1i}(\beta_0)}.
\end{aligned} \tag{5.8}$$

In view of Lemma 5.4, we have

$$\begin{aligned}
&\Big\|\frac{1}{n}\sum_{i=1}^n \frac{\Lambda_{1i}(\beta_0)[\lambda_1^\tau \Lambda_{1i}(\beta_0)]^2}{1 + \lambda_1^\tau \Lambda_{1i}(\beta_0)}\Big\| \le \frac{1}{n}\sum_{i=1}^n \frac{\|\Lambda_{1i}(\beta_0)\|^3\|\lambda_1\|^2}{|1 + \lambda_1^\tau \Lambda_{1i}(\beta_0)|} \\
&\le \|\lambda_1\|^2 \max_{1\le i\le n}\|\Lambda_{1i}(\beta_0)\|\frac{1}{n}\sum_{i=1}^n \|\Lambda_{1i}(\beta_0)\|^2 = O_p(n^{-1})o_p(n^{1/2})O_p(1) = o_p(n^{-1/2}),
\end{aligned}$$

which, together with (5.8), yields that $\sum_{i=1}^n [\lambda_1^\tau \Lambda_{1i}(\beta_0)]^2 = \sum_{i=1}^n \lambda_1^\tau \Lambda_{1i}(\beta_0) + o_p(1)$ and

$$\lambda_1 = \Big[\sum_{i=1}^n \Lambda_{1i}(\beta_0)\Lambda_{1i}^\tau(\beta_0)\Big]^{-1}\sum_{i=1}^n \Lambda_{1i}(\beta_0) + o_p(n^{-1/2}).$$

Then, by (5.7), we have

$$\mathcal{L}_{1n}(\beta_0) = \Big(\frac{1}{\sqrt{n}}\sum_{i=1}^n \Lambda_{1i}(\beta_0)\Big)^\tau \Big(\frac{1}{n}\sum_{i=1}^n \Lambda_{1i}(\beta_0)\Lambda_{1i}^\tau(\beta_0)\Big)^{-1}\Big(\frac{1}{\sqrt{n}}\sum_{i=1}^n \Lambda_{1i}(\beta_0)\Big) + o_p(1),$$

which, combining with Lemma 5.4, yields that $\mathcal{L}_{1n}(\beta_0) \xrightarrow{d} \chi_p$. □

Lemma 5.5 *Assume that (A1), (A4), (A5) and (A7) hold, we have*

$$\max_{1\le k\le n}\Big|f(U_k) - \sum_{i=1}^n W_{h_2 i}(U_k) f(U_i)\Big| = O_p(h_2^2).$$

Proof Note that the kernel $K(\cdot)$ in (A5), U_i in (A1) and h_2 in (A7) satisfy Conditions 1, 2 and 7 in Xiao et al. (2003), hence Lemma 5.5 holds by (A.14) in Xiao et al. (2003). □

Lemma 5.6 *Under the conditions of Theorem 3.2, we have*

$$\max_{1\le k\le n}|\hat f_n(U_k) - f(U_k)| = O_p(a_n),$$

where $a_n = (nh_2)^{-1/2}\log n + n^{1/s}c_n^2 + n^{-(\frac{1}{2}-\frac{1}{2s})}h_2^{-1}c_n\log n.$

Proof Write

$$\begin{aligned}
\hat f_n(U_k) - f(U_k) =& \sum_{i=1}^n W_{h_2 i}(U_k)(Y_i - X_i^\tau\beta - \hat\alpha^\tau(T_i)Z_i)^2 - f(U_k)\\
=&\Big[\sum_{i=1}^n W_{h_2 i}(U_k)\sigma_i^2 e_i^2 - f(U_k)\Big] + \sum_{i=1}^n W_{h_2 i}(U_k)\Big[\alpha^\tau(T_i)Z_i - \hat\alpha^\tau(T_i)Z_i\Big]^2 +\\
&2\sum_{i=1}^n W_{h_2 i}(U_k)\Big[\alpha^\tau(T_i)Z_i - \hat\alpha^\tau(T_i)Z_i\Big]\sigma_i e_i\\
:=& A_{1n}(U_k) + A_{2n}(U_k) + A_{3n}(U_k).
\end{aligned}$$

Obviously,

$$\begin{aligned}
\max_{1\le k\le n}|A_{1n}(U_k)| \le& \max_{1\le k\le n}\Big|\sum_{i=1}^n W_{h_2 i}(U_k) f(U_i)(e_i^2 - 1)\Big| + \max_{1\le k\le n}\Big|f(U_k) - \sum_{i=1}^n W_{h_2 i}(U_k) f(U_i)\Big|\\
:=& A_{11n} + A_{12n}.
\end{aligned}$$

By Lemma 8.3 in Chiou and Müller (1999), we have

$$W_{h_2 j}(u) = \frac{1}{nh_2}K\Big(\frac{U_j - u}{h_2}\Big)\Big/ p_U(u) + O\Big(\frac{1}{n}\Big) \quad \text{a.s.},$$

which implies that

$$\max_{1\le i,j\le n}|W_{h_2 j}(U_i)| = O((nh_2)^{-1}) \quad \text{a.s.} \tag{5.9}$$

by (A1) and (A5), and

$$\begin{aligned}
\sup_{u\in\Omega}\sum_{j=1}^n |W_{h_2 j}(u)| \le& \sup_{u\in\Omega}\frac{C}{nh_2}\Big|\sum_{j=1}^n K\Big(\frac{U_j - u}{h_2}\Big) - E\sum_{j=1}^n K\Big(\frac{U_j - u}{h_2}\Big)\Big| +\\
&\sup_{u\in\Omega}\frac{C}{nh_2}\sum_{j=1}^n E\Big[K\Big(\frac{U_j - u}{h_2}\Big)\Big] + O(1)
\end{aligned}$$

$$:=\mathcal{K}_{1n}+\mathcal{K}_{2n}+O(1) \quad \text{a.s.}$$

From (A1) and (A5), it is easy to verify that $\mathcal{K}_{2n}=O(1)$. In view of the finite cover theorem and Bernstein's Inequality (cf. Härdle et al. (2000), Page 183), from (A1), (A5) and (A7) it follows that $\mathcal{K}_{1n}=o(1)$ a.s. Therefore

$$\sup_{0\le u\le 1}\sum_{j=1}^{n}\left|W_{h_2 j}(u)\right|=O(1) \quad \text{a.s.} \tag{5.10}$$

Thus, from (5.9), (5.10), Lemma 5.1, (A3) and (A7), we can obtain

$$A_{11n}=O\Big(\big(n^{-(1-1/s)}h_2^{-1}\vee(nh_2)^{-1/2}\big)\log n\Big)=O\big((nh_2)^{-1/2}\log n\big) \quad \text{a.s.},$$

which, together with (A7) and Lemma 5.5, yields that

$$\max_{1\le k\le n}|A_{1n}(U_k)|=O_p\big((nh_2)^{-1/2}\log n\big).$$

According to Theorem 3.5 in You and Chen (2006), we have

$$\max_{1\le j\le q}\max_{1\le i\le n}\left|\hat{\alpha}_j(T_i)-\alpha_j(T_i)\right|=O(c_n) \quad \text{a.s.} \tag{5.11}$$

Note that $\max_{1\le i\le n}\|Z_i\|=o(n^{1/2s})$ a.s. by (A3) and Lemma 5.2. Then, from (A3) and (5.9)–(5.11) by using the Abel inequality (see Härdle, Liang and Gao (2000), page 183), we find

$$\begin{aligned}
\max_{1\le k\le n}|A_{2n}(U_k)|\le&\max_{1\le j\le q}\max_{1\le i\le n}\left|\hat{\alpha}_j(T_i)-\alpha_j(T_i)\right|^2\cdot\max_{1\le i\le n}\|Z_i\|^2\cdot\max_{1\le k\le n}\sum_{i=1}^{n}|W_{h_2 i}(U_k)|\\
=&o(n^{1/s}c_n^2) \quad \text{a.s.}\\
\max_{1\le k\le n}|A_{3n}(U_k)|\le&2\cdot\max_{1\le j\le q}\max_{1\le i\le n}\left|\hat{\alpha}_j(T_i)-\alpha_j(T_i)\right|\cdot\max_{1\le i\le n}\|Z_i\|\cdot\max_{1\le i\le n}|\sigma_i|\max_{1\le i,k\le n}|W_{h_2 i}(U_k)|\cdot\\
&\max_{1\le m\le n}\left|\sum_{i=1}^{m}e_{j_i}\right|=o\big(n^{-(\frac{1}{2}-\frac{1}{2s})}h_2^{-1}c_n\log n\big) \quad \text{a.s.}
\end{aligned}$$

Thus, we obtain that $\max_{1\le k\le n}|\hat{f}_n(U_k)-f(U_k)|=O_p(a_n)$. □

Lemma 5.7 *Under the conditions of Theorem 3.2, we have*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Lambda_{2i}(\beta_0)\xrightarrow{d}N(0,\Sigma), \tag{5.12}$$

$$\frac{1}{n}\sum_{i=1}^{n}\Lambda_{2i}(\beta_0)\Lambda_{2i}^{\tau}(\beta_0)\xrightarrow{p}\Sigma, \tag{5.13}$$

$$\max_{1\le i\le n}\|\Lambda_{2i}(\beta_0)\|=o_p(n^{1/2}), \tag{5.14}$$

$$\lambda_2=O_p(n^{-1/2}). \tag{5.15}$$

Proof We only prove (5.12) and (5.13) here, since the proofs of (5.14) and (5.15) are similar to those of (5.3) and (5.4). We first establish (5.12). It is easy to see that $n^{-1/2}\sum_{i=1}^{n}\Lambda_{1i}(\beta_0)\xrightarrow{d}N(0,\Sigma)$ under the conditions of Theorem 3.2. Thus, it suffices to show that $n^{-1/2}\sum_{i=1}^{n}(\Lambda_{1i}(\beta_0)-\Lambda_{2i}(\beta_0))\xrightarrow{p}0$. Let $\ddot{\sigma}_{ni}=\sigma_i^{-2}-\hat{\sigma}_{ni}^{-2}$. By (A4) and Lemma 5.6 we have

$$\max_{1\le i\le n}|\ddot{\sigma}_{ni}|=O_p(a_n). \tag{5.16}$$

Observe that

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\Lambda_{1i}(\beta_0)-\Lambda_{2i}(\beta_0)) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\ddot{\sigma}_{ni}\tilde{X}_i(\tilde{Y}_i-\tilde{X}_i^\tau\beta_0)\\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\ddot{\sigma}_{ni}\tilde{X}_i\sigma_i e_i+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\ddot{\sigma}_{ni}\tilde{X}_i Z_i^\tau(\alpha(T_i)-\hat{\alpha}(T_i))\\ &:= B_{1n}+B_{2n}.\end{aligned}$$

Note that $\max_{1\le i\le n}\|\tilde{X}_i\|=o(n^{1/2s})$ a.s. Then, from (A3), (A6), (A7), (5.11), (5.16) and Lemma 5.2, by using the Abel inequality, we can derive that

$$\|B_{1n}\|\le Cn^{-1/2}\max_{1\le i\le n}|\ddot{\sigma}_{ni}|\cdot\max_{1\le i\le n}\|\tilde{X}_i\|\cdot\max_{1\le m\le n}\Big|\sum_{i=1}^{m}e_{j_i}\Big|=o_p(n^{\frac{1}{2s}}a_n\log n)=o_p(1),$$

$$\begin{aligned}\|B_{2n}\|&\le Cn^{-1/2}\max_{1\le i\le n}|\ddot{\sigma}_{ni}|\cdot\max_{1\le i\le n}\|\tilde{X}_i\|\cdot\max_{1\le i\le n}\big\|\alpha(T_i)-\hat{\alpha}(T_i)\big\|\cdot\max_{1\le m\le n}\Big\|\sum_{i=1}^{m}Z_{j_i}\Big\|\\ &=o_p\big(n^{\frac{1}{2s}}a_n c_n\log n\big)=o_p(1).\end{aligned}$$

Hence, (5.12) is verified.

As to (5.13), in view of (5.2), it suffices to show that

$$n^{-1}\sum_{i=1}^{n}\big(\Lambda_{2i}(\beta_0)\Lambda_{2i}^\tau(\beta_0)-\Lambda_{1i}(\beta_0)\Lambda_{1i}^\tau(\beta_0)\big)\xrightarrow{p}0. \tag{5.17}$$

Denote $\breve{\sigma}_{ni}=\sigma_i^{-4}-\hat{\sigma}_{ni}^{-4}$. (A4) and Lemma 5.6 imply that

$$\max_{1\le i\le n}|\breve{\sigma}_{ni}|=O_p(a_n). \tag{5.18}$$

We write

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n}\big(\Lambda_{2i}(\beta_0)\Lambda_{2i}^\tau(\beta_0)-\Lambda_{1i}(\beta_0)\Lambda_{1i}^\tau(\beta_0)\big)=\frac{1}{n}\sum_{i=1}^{n}\breve{\sigma}_{ni}\tilde{X}_i\tilde{X}_i^\tau(\tilde{Y}_i-\tilde{X}_i^\tau\beta_0)^2\\ &=\frac{1}{n}\sum_{i=1}^{n}\breve{\sigma}_{ni}\tilde{X}_i\tilde{X}_i^\tau\sigma_i^2e_i^2+\frac{1}{n}\sum_{i=1}^{n}\breve{\sigma}_{ni}\tilde{X}_i\tilde{X}_i^\tau\big[Z_i^\tau(\alpha(T_i)-\hat{\alpha}(T_i))\big]^2+\\ &\quad\frac{2}{n}\sum_{i=1}^{n}\breve{\sigma}_{ni}\tilde{X}_i\tilde{X}_i^\tau Z_i^\tau[\alpha(T_i)-\hat{\alpha}(T_i)]\sigma_i e_i\\ &:=R_{1n}+R_{2n}+R_{3n}.\end{aligned}$$

From Lemma 5.3 and simple calculation, it is easy to see that

$$(Z_i^\tau,0_q^\tau)(D_{T_i}^\tau w_{T_i}D_{T_i})^{-1}D_{T_i}^\tau w_{T_i}X=Z_i^\tau\Gamma^{-1}(T_i)\Phi(T_i)\{1+O_p(c_n)\}, \tag{5.19}$$

which together with (A3), (A6), (A7), Lemma 5.3 and the Abel inequality yields that, for any $p\times 1$ vector $\mathbf{a}$,

$$\begin{aligned}|\mathbf{a}^\tau R_{1n}\mathbf{a}| =&n^{-1}\Big|\mathbf{a}^\tau\sum_{i=1}^{n}\breve{\sigma}_{ni}\big[X_i-\Phi^\tau(T_i)\Gamma^{-1}(T_i)Z_i+O_p(c_n)\|Z_i\|\big]\cdot\\ &\big[X_i-\Phi^\tau(T_i)\Gamma^{-1}(T_i)Z_i+O_p(c_n)\|Z_i\|\big]^\tau\sigma_i^2e_i^2\mathbf{a}\Big|\end{aligned}$$

$$\le Cn^{-1} \max_{1\le i\le n} |\breve{\sigma}_{ni}| \cdot \max_{1\le i\le n} \|X_i\| \cdot \max_{1\le i\le n} e_i^2 \cdot \max_{1\le m\le n} \Big\| \sum_{i=1}^{m} X_{j_i} \Big\|$$
$$= o_p\big(n^{-(\frac{1}{2}-\frac{3}{2s})} a_n \log n\big) = o_p(1).$$

Thus, we obtain that $R_{1n} = o_p(1)$. Similarly, by (5.11), we can easily get $R_{2n} = o_p(1)$ and $R_{3n} = o_p(1)$. This completes the proof of Lemma 5.7. □

Proof of Theorem 3.2 According to Lemma 5.7 and following the arguments in the proof of Theorem 3.1, one can easily verify Theorem 3.2. □

References

[1] B. H. BALTAGI. *Econometric Analysis of Panel Data*. Wiley, New York, 1995.

[2] J. M. CHIOU, H. G. MÜLLER. *Nonparametric quasi-likelihood*. Ann. Statist, 1999, **27**(1): 36–64.

[3] R. ENGLE, C. GRANGER, J. RICE, et al. *Nonparametric estimates of the relation between weather and electricity sales*. J. Amer. Statist. Assoc., 1986, **81**: 310–320.

[4] Guoliang FAN, Hanying LIANG. *Empirical likelihood inference for semiparametric model with linear process errors*. J. Korean Statist. Soc., 2010, **39**: 55–65.

[5] Guoliang FAN, Hanying LIANG, Hongxia XU. *Empirical likelihood for heteroscedastic partial linear model*. Comm. Statist. Theory Methods, 2011, **40**(8): 1396–1417.

[6] Jianqing FAN, Tao HUANG. *Profile likelihood inferences on semiparametric varying-coefficient partially linear models*. Bernoulli, 2005, **11**(6): 1031–1057.

[7] Jianqing FAN, Chunming ZHANG, Jian ZHANG. *Generalized likelihood ratio statistics and Wilks phenomenon*. Ann. Statist, 2001, **29**(1): 153–193.

[8] Jianqing FAN, Wenyang ZHANG. *Statistical estimation in varying coefficient models*. Ann. Statist, 1999, **27**(5): 1491–1518.

[9] W. HÄRDLE, Hua LIANG, Jiti GAO. *Partially Linear Models*. Physica-Verlag, Heidelberg, 2000.

[10] T. HASTIE, R. TIBSHIRANI. *Varying-coefficient models*. Roy. Statist. Soc., 1993, **55**(4): 757–796.

[11] Jianhua Z. HUANG, C. O. WU, Lan ZHOU. *Varying-coefficient models and basis function approximations for the analysis of repeated measurements*. Biometrika, 2002, **89**(1): 111–128.

[12] Qi LI, C. J. HUANG, Dong LI, et al. *Semiparametric smooth coefficient models*. J. Bus. Econom. Statist, 2002, **20**(3): 412–422.

[13] Xuewen LU. *Empirical likelihood for heteroscedastic partially linear models*. J. Multivariate Anal., 2009, **100**(3): 387–396.

[14] A. B. OWEN. *Empirical likelihood ratio confidence intervals for a single functional*. Biometrika, 1988, **75**(2): 237–249.

[15] A. OWEN. *Empirical likelihood ratio confidence regions*. Ann. Statist., 1990, **18**(1): 90–120.

[16] Jian SHI, T. S. LAU. *Empirical likelihood for partially linear models*. J. Multivariate Anal., 2000, **72**(1): 132–148.

[17] Xiaoqian SUN, Jinhong YOU, Gemai CHEN, et al. *Convergence rates of estimators in partial linear regression models with MA(∞) error process*. Comm. Statist. Theory Methods, 2002, **31**(12): 2251–2273.

[18] Qihua WANG, Bingyi JING. *Empirical likelihood for partial linear models with fixed designs*. Statist. Probab. Lett., 1999, **41**(4): 425–433.

[19] Qihua WANG, Bingyi JING. *Empirical likelihood for partial linear models*. Ann. Inst. Statist. Math., 2003, **55**(3): 585–595.

[20] Zhijie XIAO, O. B. LINTON, R. J. CARROLL, et al. *More efficient local polynomial estimation in nonparametric regression with autocorrelated errors*. J. Amer. Statist. Assoc., 2003, **98**(464): 980–992.

[21] Jinhong YOU, Yong ZHOU. *Empirical likelihood for semiparametric varying-coefficient partially linear regression models*. Statist. Probab. Lett., 2006, **76**(4): 412–422.

[22] Wenyang ZHANG, S. Y. LEE, Xinyuan SONG. *Local polynomial fitting in semivarying coefficient model*. J. Multivariate Anal., 2002, **82**(1): 166–188.