应用概率统计 第 34 卷
第 2 期 2018 年 4 月

Chinese Journal of Applied Probability and Statistics
Apr., 2018, Vol. 34, No. 2, pp. 111-134
doi: 10.3969/j.issn.1001-4268.2018.02.0001

Weighted Profile LSDV Estimation of Fixed Effects Panel Data Partially Linear Regression Models *

ZHU NengHui⋆

(*School of Applied Mathematics, Xiamen University of Technology, Xiamen, 361024, China*)

LI Xiao

(*School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai, 200433, China*)

SHI YaFeng

(*School of Science, Ningbo University of Technology, Ningbo, 315211, China*)

Abstract: This paper concerns with the estimation of a fixed effects panel data partially linear regression model with the idiosyncratic errors being an autoregressive process. For fixed effects short time series panel data, the commonly used autoregressive error structure fitting method will not result in a consistent estimator of the autoregressive coefficients. Here we propose an alternative estimation and show that the resulting estimator of the autoregressive coefficients is consistent and this method is workable for any order autoregressive error structure. Moreover, combining the B-spline approximation, profile least squares dummy variable (PLSDV) technique and consistently estimated the autoregressive error structure, we develop a weighted PLSDV estimator for the parametric component and a weighted B-spline series (BS) estimator for the nonparametric component. The weighted PLSDV estimator is shown to be asymptotically normal and more asymptotically efficient than the one which ignores the error autoregressive structure. In addition, this paper derives the asymptotic bias of the weighted BS estimator and establish its asymptotic normality as well. Simulation studies and an example of application are conducted to illustrate the finite sample performance of the proposed procedures.

Keywords: panel data partially linear varying-coefficient model; fixed effect; profile least squares dummy variable technique; semiparametric; auto-regressive process

2010 Mathematics Subject Classification: Primary 62H12; Secondary 62G05

Citation: ZHU N H, LI X, SHI Y F. Weighted profile LSDV estimation of fixed effects panel data partially linear regression models [J]. Chinese J Appl Probab Statist, 2018, 34(2): 111–134.

---

*The project was supported by the High-Level Personnel Fund of Xiamen University of Technology (Grant No. YKJ15031R) and the Graduate Innovation Fund of Shanghai University of Finance and Economics (Grant No. CXJJ-2013-459).

⋆Corresponding author, E-mail: zhunenghui163@163.com.

Received January 7, 2015. Revised June 29, 2015.

§1. Introduction

Various parametric regression models and statistical tools have been developed for panel data analysis; see [1–4] and the references therein. It is often desirable to relax the assumption of a parametric function and estimate the conditional nonparametrically or semiparametrically. Especially, semiparametric regressions have gained increasing popularity, which can increase the flexibility of modeling and avoid the "curse of dimensionality". One of the most important panel data semiparametric regression models is the panel data partially linear varying-coefficient regression model, which has the following form:

$$Y_{it} = \boldsymbol{X}_{it}^{\mathsf{T}}\boldsymbol{\beta} + \boldsymbol{Z}_{it}^{\mathsf{T}}\boldsymbol{\alpha}(U_{it}) + \mu_i + \nu_{it}, \qquad i = 1, 2, \ldots, N \ \text{ and } \ t = 1, 2, \ldots, T, \tag{1}$$

where Y_{it} is the response, $\boldsymbol{X}_{it} = (X_{it1}, X_{it2}, \ldots, X_{itp})^{\mathsf{T}}$ and U_{it} are explanatory variables, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)^{\mathsf{T}}$ is an unknown p-vector parameter, $\boldsymbol{\alpha}(\cdot) = (\alpha_1(\cdot), \alpha_2(\cdot), \ldots, \alpha_q(\cdot))^{\mathsf{T}}$ is an unknown q-vector function, $\boldsymbol{Z}_{it} = (Z_{it1}, Z_{it2}, \ldots, Z_{itq})^{\mathsf{T}}$, μ_i represents the individual effect. For the purpose of identification, we assume $\sum\limits_{i=1}^{n} \mu_i = 0$. ν_{it} is the idiosyncratic error and "T" denotes a transpose of a matrix or vector.

For panel data analysis, there are usually two approaches. One is the "fixed-effects" approach in which the individual effects are taken as individual-specific constants and can be estimated by entering individual-specific dummies as explanatory variables. An alternative approach, the "random-effects" technique in which the individual effects are treated as random components of the error terms and generalized least squares method is often used. Fan et al.[5] applied the "random-effects" technique to investigate the efficient estimation for model (1). Different from [5], in this paper we assume that the individual effects μ_i in model (1) are individual-specific constants and will implement the "fixed-effects" approach to investigate the estimating problem of model (1).

Fixed effects panel data partially linear modeling has attracted great attention in last several years. Baltagi and Li[6] proposed difference-based series estimators for the parametric component and the nonparametric component of the fixed effects panel data partially linear model. By applying the back-fitting method, Fan et al.[7] proposed a profile least squares estimator for the parametric component and a local linear estimator for the nonparametric component. Su and Ullah[8] proposed the profile likelihood estimation for the fixed effects panel data partially linear model. Henderson et al.[9] and You et al.[10] also considered the estimation problem of the fixed effects panel data partially linear model. Zhang et al.[11] proposed an empirical likelihood method for application to a partially linear panel data model with fixed effects. In addition, Sun et al. [12] considered the estimating problem of the fixed effects panel data varying-coefficient model. All the work mentioned above assumes that the idiosyncratic errors are not serially uncorrelated. As argued in [4], this assumption is quite likely to be violated as the dynamic effect of shocks to the dependent variable is often distributed over several time periods. Same as [13], we assume that for fixed i the idiosyncratic errors ν_{it} in (1) follow an auto-regressive process:

$$
\begin{gathered}
\nu_{it} = \rho_1 \nu_{i,t-1} + \rho_2 \nu_{i,t-2} + \cdots + \rho_m \nu_{i,t-m} + e_{it}, \\
1 - \rho_1 z - \rho_2 z^2 - \cdots - \rho_m z^m \neq 0 \ \text{ for } \ |z| \leqslant 1, \ t = 1, 2, \ldots, T,
\end{gathered} \tag{2}
$$

where e_{it} is i.i.d. random variable over (i,t) with mean zero and variance σ_e^2. $(\rho_1, \rho_2, \ldots, \rho_m)^\top$ is an unknown autoregressive vector coefficient.

Nickell [14] derived an exact expression for the inconsistency of the first-order autotocorrelation estimator in the particular case of idiosyncratic error follows a first-order autoregression and this expression can be used to correct the inconsistency of the estimator. His correcting technique is not easy to be extended to a high-order autoregression error model, especially when the order is greater than 2. Becker [13] proposed a differencing procedure to estimate the autocorrelation. However, his method will result in a more complicated autocorrelation structure and may obscure the original error structure.

In this paper, by extending the method of [15] we propose a new estimation for the autoregressive coefficients of the error structure and show that the resulting estimator of the autoregressive coefficients is consistent and our method is workable for any order autoregressive error structure. Based on consistently estimating the error autoregressive structure, by combining the B-spline series approximation and profile least squares dummy variable (PLSDV) technique we further propose a weighted PLSDV estimator for the parametric component and a weighted B-spline series (BS) estimator for the nonparametric component. The weighted PLSDV estimator is shown to be asymptotically normal and more asymptotically efficient than the one which ignores the error auto-regressive structure. In addition, we derive the asymptotic bias of the weighted BS estimator and establish its asymptotic normality as well.

The rest of this paper is organized as follows. In Section 2 we describe a B-spline based PLSDV estimation. In Section 3, the fitting of the error structure is investigated. In Section 4 we propose a weighted B-spline based PLSDV estimation. Section 5 presents results from simulation studies. An example of application is illustrated in Section 6. All proofs of main results are relegated to Appendix.

§2. B-Spline Based Profile LSDV Estimation

Without lose of generality, we consider the estimation of $\alpha(\cdot)$ on a compact set, say $\mathscr{U}=[0,1]$. It is well known that the polynomial spline is a good way to approximate the unknown functions. Let $0=\varsigma_0<\varsigma_1<\cdots<\varsigma_{M_n}=1$ be a sequence of quasi-uniform spaced points on $[0,1]$. Denote $\mathscr{U}_i=[\varsigma_{i-1},\varsigma_i)$, $i=1,2,\ldots,M_n-1$, and $\mathscr{U}_{M_n}=[\varsigma_{M_n-1},1]$. A polynomial spline of degree s is a polynomial function on each interval $\mathscr{U}_i$ for $i=1,2,\ldots,M_N$ and globally has continuous derivatives of order $s-1$ for $s\geqslant 1$. For given degree and knots sequence, the collection of polynomial spline form a linear function space. For this space, two famous basis function series are the truncated power basis functions and B-spline basis functions. Usually, B-spline basis owns better numerical properties. The books by [16] and [17] are good references for spline functions.

For given positive integer s, M_N and the knots sequence $\{\varsigma_i\}_{i=0}^{M_N}$, $\mathscr{S}_{s,M_N}$ denotes the space of polynomial spline on $\mathscr{U}$ and the dimension is $\kappa_n=M_n+s$ $(s\geqslant 2)$. Let $\{\zeta_l(u)$, $l=1,2,\ldots,\kappa_n\}$ be a B-spline basis of $\mathscr{S}_{s,M_N}$. Then we can approximate each $\alpha_j(u)$ by

$$\alpha_j(u)\approx\sum_{l=1}^{\kappa_n}\theta_{jl}\zeta_l(u)=(\boldsymbol{\zeta}(u))^{\top}\boldsymbol{\theta}_j,$$

where $\boldsymbol{\zeta}(\cdot)=(\zeta_1(\cdot),\zeta_2(\cdot),\ldots,\zeta_{\kappa_n}(\cdot))^{\top}$ and $\boldsymbol{\theta}_j=(\theta_{j1},\theta_{j2},\ldots,\theta_{j\kappa_n})^{\top}$. Thus, model (1) can be approximated by

$$Y_{it}\approx\boldsymbol{X}_{it}^{\top}\boldsymbol{\beta}+\sum_{j=1}^{q}Z_{itj}\Big\{\sum_{l=1}^{\kappa_n}\theta_{jl}\zeta_l(U_{it})\Big\}+\mu_i+\nu_{it},\qquad i=1,2,\ldots,n\ \text{ and }\ t=1,2,\ldots,T. \tag{3}$$

To present a matrix form of (3), let

$$\begin{aligned}
\boldsymbol{Z}_{it}^{*}&=(Z_{it1}\zeta_1(U_{it}),\ldots,Z_{it1}\zeta_{\kappa_n}(U_{it}),\ldots,Z_{itq}\zeta_1(U_{it}),\ldots,Z_{itq}\zeta_{\kappa_n}(U_{it}))^{\top},\\
\boldsymbol{Z}_{i}^{*}&=(\boldsymbol{Z}_{i1}^{*},\boldsymbol{Z}_{i2}^{*},\ldots,\boldsymbol{Z}_{iT}^{*})^{\top},\qquad \boldsymbol{Z}^{*}=(\boldsymbol{Z}_{1}^{*},\boldsymbol{Z}_{2}^{*},\ldots,\boldsymbol{Z}_{n}^{*})^{\top},\\
\boldsymbol{X}_{i}&=(\boldsymbol{X}_{i1},\boldsymbol{X}_{i2},\ldots,\boldsymbol{X}_{iT})^{\top},\qquad \boldsymbol{X}=(\boldsymbol{X}_{1},\boldsymbol{X}_{2},\ldots,\boldsymbol{X}_{n})^{\top},\\
\boldsymbol{Y}_{i}&=(Y_{i1},Y_{i2},\ldots,Y_{iT})^{\top},\qquad \boldsymbol{Y}=(\boldsymbol{Y}_{1}^{\top},\boldsymbol{Y}_{2}^{\top},\ldots,\boldsymbol{Y}_{n}^{\top})^{\top},\\
\boldsymbol{\theta}&=(\boldsymbol{\theta}_{1}^{\top},\boldsymbol{\theta}_{2}^{\top},\ldots,\boldsymbol{\theta}_{q}^{\top})^{\top},\qquad \boldsymbol{\mu}=(\mu_1,\mu_2,\ldots,\mu_n)^{\top},\\
\boldsymbol{\nu}_{i}&=(\nu_{i1},\nu_{i2},\ldots,\nu_{iT})^{\top},\qquad \boldsymbol{\nu}=(\boldsymbol{\nu}_{1}^{\top},\boldsymbol{\nu}_{2}^{\top},\ldots,\boldsymbol{\nu}_{n}^{\top})^{\top}.
\end{aligned}$$

Then (3) can be written as

$$\boldsymbol{Y}\approx\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{Z}^{*}\boldsymbol{\theta}+\boldsymbol{B}\boldsymbol{\mu}+\boldsymbol{\nu}, \tag{4}$$

with $\boldsymbol{B}=\boldsymbol{I}_n\otimes\boldsymbol{1}_T$ $(\boldsymbol{1}_T=(1,1,\ldots,1)^{\top})$. Define $\boldsymbol{M_B}=\boldsymbol{I}_{nT}-\boldsymbol{B}(\boldsymbol{B}^{\top}\boldsymbol{B})^{-1}\boldsymbol{B}^{\top}$. Then $\boldsymbol{M_B}\boldsymbol{B}\boldsymbol{\mu}=\boldsymbol{0}$. Thus (4) leads to

$$\boldsymbol{M_B}\boldsymbol{Y}\approx\boldsymbol{M_B}\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{M_B}\boldsymbol{Z}^{*}\boldsymbol{\theta}+\boldsymbol{M_B}\boldsymbol{\nu}. \tag{5}$$

If we take $\boldsymbol{M_B\nu}$ as the residuals, then model (5) is a version of the usual linear regression. An estimator of $(\boldsymbol{\beta}^\top, \boldsymbol{\theta}^\top)^\top$ is

$$\begin{pmatrix} \widehat{\boldsymbol{\beta}}_n \\ \widehat{\boldsymbol{\theta}}_n \end{pmatrix} = \{(\boldsymbol{X}, \boldsymbol{Z}^*)^\top \boldsymbol{M_B}(\boldsymbol{X}, \boldsymbol{Z}^*)\}^{-1}(\boldsymbol{X}, \boldsymbol{Z}^*)^\top \boldsymbol{M_B Y}.$$

Therefore,

$$\widehat{\boldsymbol{\beta}}_n = \left(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M_B}} \boldsymbol{X}\right)^{-1} \boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M_B}} \boldsymbol{Y} \quad \text{and} \quad \widehat{\boldsymbol{\theta}}_n = (\boldsymbol{Z}^{*\top} \boldsymbol{M_B Z}^*)^{-1} \boldsymbol{Z}^{*\top} \boldsymbol{M_B}(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}_n).$$

with $\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M_B}} = \boldsymbol{M_B} - \boldsymbol{M_B Z}^*(\boldsymbol{Z}^{*\top} \boldsymbol{M_B Z}^*)^{-1} \boldsymbol{Z}^{*\top} \boldsymbol{M_B}$. Then $\widehat{\alpha}_{jn}(u) = (\boldsymbol{\zeta}(u))^\top \widehat{\boldsymbol{\theta}}_{jn}$.

In the following, we introduce some notations and technical assumptions. c denotes a generic constants in different cases. Let $\|\boldsymbol{a}\|$ be the Euclidean norm of a real valued vector $\boldsymbol{a}$. Denotes "$\xrightarrow{\mathrm{d}}$" as the convergence in distribution and "$\xrightarrow{\mathsf{P}}$" as the convergence in probability. Define $\boldsymbol{\Pi}_{it} = \boldsymbol{X}_{it} - \mathsf{E}(\boldsymbol{X}_{it} \,|\, U_{it})$, $h_j(U_{it}) = \mathsf{E}(X_{itj} \,|\, U_{it})$, $j = 1, 2, \ldots, p$.

Assumption 1 $(\boldsymbol{X}_{i1}^\top, \boldsymbol{X}_{i2}^\top, \ldots, \boldsymbol{X}_{iT}^\top, \boldsymbol{Z}_{i1}^\top, \boldsymbol{Z}_{i2}^\top, \ldots, \boldsymbol{Z}_{iT}^\top, U_{i1}, U_{i2}, \ldots, U_{iT})^\top$ are independent and identically distributed (i.i.d.) over $i = 1, 2, \ldots, n$. In addition, $\sum\limits_{t=1}^{T} \mathsf{E}(\|\boldsymbol{X}_{it}\|^{2+\delta}) \leqslant c < \infty$, $\sum\limits_{t=1}^{T} \mathsf{E}(\|\boldsymbol{Z}_{it}\|^{2+\delta}) \leqslant c < \infty$ for some δ, $c > 0$.

Assumption 2 $\alpha(\cdot)$ and $h_j(\cdot)$ are r-times continuously differentiable for $j = 1, 2, \ldots, p$ and $r \geqslant 2$.

Assumption 3 Let $\iota_j = \varsigma_j - \varsigma_{j-1}$, then $\iota = \max\limits_{1 \leqslant j \leqslant M_n} \iota_j = O(\kappa_n^{-1})$ and $\max(\iota_{j+1}, j = 0, 1, \ldots, M_n - 1)/\min(\iota_{j+1}, j = 0, 1, \ldots, M_n - 1) \leqslant c$.

Assumption 4 (i) The individual effects $\{\mu_i, 1 \leqslant i \leqslant n\}$ are i.i.d. with mean and variance σ_μ^2. Further, $\mathsf{E}(|\mu_i|^{4+\delta_\mu}) \leqslant c < \infty$ for some δ_μ, $c > 0$.

(ii) The random errors $\{e_{it}, 1 \leqslant i \leqslant n, 1 \leqslant t \leqslant T\}$ are i.i.d. with mean and variance σ_e^2. Further, $\mathsf{E}(|e_{it}|^{4+\delta_e}) \leqslant c < \infty$ for some δ_e, $c > 0$.

(iii) $\{\mu_i\}$ and $\{e_{it}\}$ are independent.

(iv) $1 - \rho_1 z - \rho_2 z^2 - \cdots - \rho_m z^m \neq 0$, for $|z| \leqslant 1$.

Assumption 5 U_{it}s' are generated from a distribution which has bounded support $\mathscr{U}$ and a Lipschitz continuous density function $p(\cdot)$ such that $0 < \inf\limits_{\mathscr{U}} p(\cdot) \leqslant \sup\limits_{\mathscr{U}} p(\cdot) < \infty$.

Assumption 6 $\kappa_n = o(n^{1/2})$ and $n^{1/2}\kappa_n^{-4} = o(1)$.

For $\widehat{\boldsymbol{\beta}}_n$, $\widehat{\boldsymbol{\theta}}_{jn}$ and $\widehat{\alpha}_{jn}(u)$, we have the following results.

Theorem 7 Suppose that Assumptions 1 to 6 hold, then

$$\sqrt{nT}(\widehat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) \xrightarrow{\mathrm{d}} N(0, \boldsymbol{\Omega}_1^{-1}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Omega}_1^{-1}),$$

where

$$\boldsymbol{\Omega}_1=\lim_{n\to\infty}\frac{1}{nT}\boldsymbol{X}^{\top}\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}}\mathbf{X},\qquad \boldsymbol{\Omega}_2=\lim_{n\to\infty}\frac{1}{nT}\boldsymbol{X}^{\top}\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma})\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}}\mathbf{X},$$

with $\boldsymbol{\Sigma}=(\sigma_{t_1t_2}^2)_{t_1,t_2=1}^T=\mathsf{Cov}\,(\boldsymbol{\nu}_i)$ and $\boldsymbol{\nu}_i=(\nu_{i1},\nu_{i2},\ldots,\nu_{iT})^{\top}$.

Theorem 8 Suppose that Assumptions 1 to 6 hold, then

(i) $\|\widehat{\boldsymbol{\theta}}_{jn}-\boldsymbol{\theta}_j\|=O_p(\sqrt{\kappa_n/n}+\kappa_n^{-2})$ for $j=1,2,\ldots,q$.

(ii) $\int_{u\in\mathscr{U}}[\widehat{\alpha}_{jn}(u)-\alpha(u)]p(u)\mathrm{d}u=O_p(\kappa_n/n+\kappa_n^{-4})$ for $j=1,2,\ldots,q$.

Remark 9 $\widehat{\boldsymbol{\beta}}_n$ and $(\widehat{\alpha}_{1n}(\cdot),\widehat{\alpha}_{2n}(\cdot),\ldots,\widehat{\alpha}_{qn}(\cdot))^{\top}$ do not take the contemporaneous correlation into account, hence may not be asymptotically efficient. We will construct more efficient estimators in the following sections.

§3. Estimation of Autoregressive Error Structure

Based on $\widehat{\boldsymbol{\beta}}_n$ and $(\widehat{\alpha}_{1n}(\cdot),\widehat{\alpha}_{2n}(\cdot),\ldots,\widehat{\alpha}_{qn}(\cdot))^{\top}$, we can obtain the estimated residuals as

$$\widehat{\varepsilon}_{it}=Y_{it}-\boldsymbol{X}_{it}^{\top}\widehat{\boldsymbol{\beta}}_n-Z_{it1}\widehat{\alpha}_{1n}(U_{it})-\cdots-Z_{itq}\widehat{\alpha}_{qn}(U_{it}),\qquad i=1,2,\ldots,n,\ t=1,2,\ldots,T.$$

Define

$$\widehat{\boldsymbol{D}}_0=\frac{1}{n\{T-(m+1)\}}\sum_{i=1}^{n}\sum_{t=1}^{T-(m+1)}(\widehat{\varepsilon}_{i,t+(m-1)},\ldots,\widehat{\varepsilon}_{it})^{\top}(\widehat{\varepsilon}_{i,t+(m-1)},\ldots,\widehat{\varepsilon}_{it}),$$

$$\widehat{\boldsymbol{D}}_1=\frac{1}{n(T-(m+1))}\sum_{i=1}^{n}\sum_{t=1}^{T-(m+1)}(\widehat{\varepsilon}_{i,t+(m-1)},\ldots,\widehat{\varepsilon}_{it})^{\top}(\widehat{\varepsilon}_{i,t+m},\ldots,\widehat{\varepsilon}_{i,t+1}),$$

$$\widehat{\boldsymbol{D}}_3=\frac{1}{n(T-(m+1))}\sum_{i=1}^{n}\sum_{t=1}^{T-(m+1)}(\widehat{\varepsilon}_{i,t+(m-1)},\ldots,\widehat{\varepsilon}_{it})^{\top}\widehat{\varepsilon}_{i,t+m},$$

and

$$\widehat{\boldsymbol{D}}_4=\frac{1}{n(T-(m+1))}\sum_{i=1}^{n}\sum_{t=1}^{T-(m+1)}(\widehat{\varepsilon}_{i,t+(m-1)},\ldots,\widehat{\varepsilon}_{it})^{\top}\widehat{\varepsilon}_{i,t+m+1}.$$

Then, we can estimate $(\rho_1,\rho_2,\ldots,\rho_m)^{\top}$ by extensting of the method of [15] $(\widehat{\rho}_{1n},\widehat{\rho}_{2n},\ldots,\widehat{\rho}_{mn})^{\top}=(\widehat{\boldsymbol{D}}_0-\widehat{\boldsymbol{D}}_1)^{-1}(\widehat{\boldsymbol{D}}_3-\widehat{\boldsymbol{D}}_4)$. In addition, we notice that $\varepsilon_{i,t+m}=\mu_i+\nu_{i,t+m}=\mu_i+\rho_1\nu_{i,t+(m-1)}+\cdots+\rho_m\nu_{it}+e_{it}$. Therefore,

$$\varepsilon_{i,t+m}-\rho_1\varepsilon_{i,t+(m-1)}-\cdots-\rho_m\varepsilon_{it}=(\mu_i-\rho_1\mu_i-\cdots-\rho_m\mu_i)+e_{it}.$$

Denote $\ell_{it}=\varepsilon_{i,t+m}-\rho_1\varepsilon_{i,t+(m-1)}-\cdots-\rho_m\varepsilon_{it}$ for $t=1,2,\ldots,T-m$. Based on the fact that $\sigma_e^2=\mathsf{E}(\ell_{it}^2)-\mathsf{E}(\ell_{it}\ell_{i,t+1})$, we can define an estimator of σ_e^2 by

$$\widehat{\sigma}_{en}^2=\frac{1}{n(T-m)}\sum_{i=1}^{n}\sum_{t=1}^{T-m}\widehat{\ell}_{it}^2-\frac{1}{n(T-m-1)}\sum_{i=1}^{n}\sum_{t=1}^{T-m-1}\widehat{\ell}_{it}\widehat{\ell}_{i,t+1},$$

where $\widehat{\ell}_{it} = \widehat{\varepsilon}_{i,t+m} - \widehat{\rho}_{1n}\widehat{\varepsilon}_{i,t+(m-1)} - \cdots - \widehat{\rho}_{mn}\widehat{\varepsilon}_{it}$ for $t = 1, 2, \ldots, T-m$.

The following theorem shows that $(\widehat{\rho}_{1n}, \widehat{\rho}_{2n}, \ldots, \widehat{\rho}_{mn})^\top$ and $\widehat{\sigma}^2_{en}$ are the consistent estimator of $(\rho_1, \rho_2, \ldots, \rho_m)^\top$ and σ_e^2, respectively.

Theorem 10 Under Assumptions 1 to 6, we have $(\widehat{\rho}_{1n}, \widehat{\rho}_{2n}, \ldots, \widehat{\rho}_{mn})^\top \xrightarrow{\mathsf{P}} (\rho_1, \rho_2, \ldots, \rho_m)^\top$ and $\widehat{\sigma}^2_{en} \xrightarrow{\mathsf{P}} \sigma_e^2$.

In order to obtain $\widehat{\boldsymbol{\beta}}_n$, $(\widehat{\alpha}_{1n}(\cdot), \widehat{\alpha}_{2n}(\cdot), \ldots, \widehat{\alpha}_{qn}(\cdot))^\top$, $(\widehat{\rho}_{1n}, \widehat{\rho}_{2n}, \ldots, \widehat{\rho}_{qn})^\top$ and $\widehat{\sigma}^2_{en}$, we need to select the degrees of splines and the numbers and locations of knots. Similar to [18], we use splines with equally spaced knots and fixed degrees and select κ_n, the numbers of knots, by a data-driven cross-validation method. Here $\kappa_n = M_n + 1 + s$.

§4. Weighted Profile LSDV Estimation

Define

$$\widehat{\boldsymbol{\Sigma}}^{-1} = \begin{pmatrix} \widehat{\gamma}_\nu(0) & \widehat{\gamma}_\nu(1) & \ldots & \widehat{\gamma}_\nu(T-1) \\ \widehat{\gamma}_\nu(1) & \widehat{\gamma}_\nu(0) & \ldots & \widehat{\gamma}_\nu(T-2) \\ \vdots & \vdots & \vdots & \vdots \\ \widehat{\gamma}_\nu(T-1) & \widehat{\gamma}_\nu(T-2) & \ldots & \widehat{\gamma}_\nu(0) \end{pmatrix}^{-1}.$$

Pre-multiplying (4) by $\widehat{\boldsymbol{\Sigma}}^{-1/2}$ leads to

$$\begin{aligned}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})\boldsymbol{Y} \approx{}& (\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})\boldsymbol{X}\boldsymbol{\beta} + (\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})\boldsymbol{Z}^*\boldsymbol{\theta} + (\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})\boldsymbol{B}\boldsymbol{\mu} \\ &+ (\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})\boldsymbol{\nu}. \end{aligned} \tag{6}$$

Denote $\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}^{-1}} = \boldsymbol{I}_{nT} - (\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})\boldsymbol{B}\{\boldsymbol{B}^\top(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1})\boldsymbol{B}\}^{-1}\boldsymbol{B}^\top(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})$. Then (6) becomes

$$\begin{aligned}\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})\boldsymbol{Y} \approx{}& \boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})\boldsymbol{Z}^*\boldsymbol{\theta} \\ &+ \boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})\boldsymbol{\nu}. \end{aligned} \tag{7}$$

Denote

$$\begin{aligned} &\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})} \\ ={}& \boldsymbol{I}_{nT} - \boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})\boldsymbol{Z}^*\big\{\boldsymbol{Z}^{*\top}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})\boldsymbol{Z}^*\big\}^{-1} \\ &\cdot \boldsymbol{Z}^{*\top}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}^{-1}}. \end{aligned}$$

Then, we can obtain the weighted semiparametric least squares estimator of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ by

$$\widehat{\boldsymbol{\beta}}_n^w = \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})}\boldsymbol{X}\Big)^{-1}\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})}\boldsymbol{Y},$$

$$\widehat{\boldsymbol{\theta}}_n^w = \big\{\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})\boldsymbol{Z}^*\big\}^{-1}\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}_n^w),$$

where $\boldsymbol{M}_{\boldsymbol{Z}^*}^{\widehat{\boldsymbol{\Sigma}}^{-1}} = \widehat{\boldsymbol{\Sigma}}^{-1} - \widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{Z}^*(\boldsymbol{Z}^{*\mathsf{T}}\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{Z}^*)^{-1}\boldsymbol{Z}^{*\mathsf{T}}\widehat{\boldsymbol{\Sigma}}^{-1}$. This gives the weighted BS estimator of $\alpha_j(\cdot)$ as $\widehat{\alpha}_{jn}^w(u) = (\boldsymbol{\zeta}(u))^{\mathsf{T}}\widehat{\boldsymbol{\theta}}_{jn}^w$. For $\widehat{\boldsymbol{\beta}}_n^w$ and $(\widehat{\alpha}_{1n}^w(\cdot), \widehat{\alpha}_{2n}^w(\cdot), \ldots, \widehat{\alpha}_{qn}^w(\cdot))^{\mathsf{T}}$ we have the following asymptotic results.

Theorem 11 Suppose that Assumptions 1 to 6 hold, then

$$\sqrt{nT}(\widehat{\boldsymbol{\beta}}_n^w - \boldsymbol{\beta}) \xrightarrow{\mathrm{d}} N(0, \boldsymbol{\Omega}_3^{-1}),$$

where

$$\boldsymbol{\Omega}_3 = \lim_{n\to\infty}\frac{1}{nT}\boldsymbol{X}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})}\boldsymbol{X}.$$

Theorem 12 Suppose that Assumptions 1 to 6 hold, then for $j = 1, 2, \ldots, q$,

(i) $\|\widehat{\boldsymbol{\theta}}_{jn}^w - \boldsymbol{\theta}_j\| = O_p(\sqrt{\kappa_n/n} + \kappa_n^{-2})$;

(ii) $\int_{u\in\mathscr{U}}[\widehat{\alpha}_{jn}^w(u) - \alpha(u)]p(u)\mathrm{d}u = O_p(\kappa_n/n + \kappa_n^{-4})$.

In order to apply Theorem 12 to make statistical inferences, a consistent estimator of $\boldsymbol{\Omega}_3$ is needed. This is given by $\widehat{\boldsymbol{\Omega}}_3 = (nT)^{-1}\boldsymbol{X}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}^{-1/2})}\boldsymbol{X}$ via the following theorem.

Theorem 13 Suppose that Assumptions 1 to 6 hold, then $\widehat{\boldsymbol{\Omega}}_3 \xrightarrow{\mathsf{P}} \boldsymbol{\Omega}_3$.

§5. Simulation Studies

We will investigate the finite sample performance of the proposed estimators by conducting some simulation studies. The data are generated from the following:

$$Y_{it} = X_{1it}\beta_1 + X_{2it}\beta_2 + X_{3it}\beta_3 + Z_{1it}\alpha_1(U_{it}) + Z_{2it}\alpha_2(U_{it}) + \mu_i + \nu_{it},$$
$$i = 1, 2, \ldots, n,\ t = 1, 2, \ldots, T, \tag{8}$$

where $X_{1it} = \eta_{1it} + \eta_i + \eta_{it}$ with $\eta_{1it} \overset{\text{i.i.d.}}{\sim} \mathrm{U}(0, 2)$, $\eta_i \overset{\text{i.i.d.}}{\sim} \mathrm{U}(0, 1)$, $\eta_{it} \overset{\text{i.i.d.}}{\sim} \mathrm{U}(0, 0.5)$, $X_{2it} = \eta_{2it} + (1.1\eta_i + \eta_{it})/3.1$ with $\eta_{2it} \overset{\text{i.i.d.}}{\sim} \mathrm{U}(0, 2)$, $X_{3it} = \eta_{3it} + (1.2\eta_i + \eta_{it})/3.2$ with $\eta_{3it} \overset{\text{i.i.d.}}{\sim} \mathrm{U}(0, 2)$, $Z_{1it} = \eta_{4it} + (1.3\eta_i + \eta_{it})/3.3$ with $\eta_{4it} \overset{\text{i.i.d.}}{\sim} \mathrm{U}(0, 2)$, $Z_{2it} = \eta_{5it} + (1.4\eta_i + \eta_{it})/3.4$ with $\eta_{5it} \overset{\text{i.i.d.}}{\sim} \mathrm{U}(0, 2)$, $U_{it} \overset{\text{i.i.d.}}{\sim} \mathrm{U}(0, 1)$, $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 0.75$, $\alpha_1(U_{it}) = 3\{0.1\sin(2\pi U_{it}) + 0.2\cos(2\pi U_{it}) + 0.3(\sin(2\pi U_{it}))^2 + 0.4(\cos(2\pi U_{it}))^4\}$, $\alpha_2(U_{it}) = U_{it}^2 + 0.5(\sin(2\pi U_{it}))^3$, μ_i are generated by

$$\mu_i = \frac{1}{T}\sum_{t=1}^{T}(X_{1it} + X_{2it} + X_{3it}) - \frac{1}{T}\sum_{t=1}^{T}\mathsf{E}(X_{1it} + X_{2it} + X_{3it}),$$

which are correlated with $(X_{1it}, X_{2it}, X_{3it})^{\mathsf{T}}$, and $\nu_{it} = \rho\nu_{i,t-1} + e_{it}$ with $e_{it} \overset{\text{i.i.d.}}{\sim} \mathrm{N}(0, 1-\rho^2)$. We take $\rho = -0.5, 0, 0.5$, $n = 100, 200$ and 300, and $T = 3, 4, 5$ and 10. We use the cubic polynomial splines with uniform knots as our basis, and the number κ is chosen by the deleting group cross-validation criterion ([19]). Since κ equals the number of knots + 3, to choose κ is to choose the number of knots. The rang of the number of knots is $[\,[0.1(nT)^{1/4}\ln(nT)], [(nT)^{1/4}\ln(nT)]\,]$, where $[a]$ denotes the integer part of a.

For the estimators $\widehat{\rho}_n$, $\widehat{\sigma}^2_{n\nu}$, $\widehat{\sigma}^2_{ne}$ of the parameters ρ, σ^2_ν, σ^2_e, given a sample size, the sample mean (sm), standard deviation (std) which is calculated based on the estimators of the error structure parameters are summarized in Table 1. From Table 1, the estimators $\widehat{\rho}_n$, $\widehat{\sigma}^2_{n\nu}$, $\widehat{\sigma}^2_{ne}$ of the error structure parameters ρ, σ^2_ν, σ^2_e work well and their performance is improved with the increasing of sample size.

Table 1 Finite sample performance of the estimators $\widehat{\rho}_n$, $\widehat{\sigma}^2_{n\nu}$, $\widehat{\sigma}^2_{ne}$

			$n=100$				$n=200$				$n=300$			
ρ			$T=3$	$T=4$	$T=5$	$T=10$	$T=3$	$T=4$	$T=5$	$T=10$	$T=3$	$T=4$	$T=5$	$T=10$
-0.5	$\widehat{\rho}_n$	sm	-0.4422	-0.4443	-0.4376	-0.4280	-0.4661	-0.4629	-0.4624	-0.4520	-0.4679	-0.4703	-0.4687	-0.4570
		std	0.1165	0.0657	0.0693	0.0491	0.0829	0.0475	0.0500	0.0349	0.0686	0.0396	0.0388	0.0283
1	$\widehat{\sigma}^2_{n\nu}$	sm	1.0280	1.0605	1.0867	1.1817	1.0398	1.0542	1.0679	1.1234	1.0390	1.0538	1.0612	1.1057
		std	0.1215	0.0981	0.0934	0.0925	0.0873	0.0695	0.0666	0.0587	0.0725	0.0559	0.0504	0.0465
0.75	$\widehat{\sigma}^2_{ne}$	sm	0.8169	0.8460	0.8735	0.8630	0.8087	0.8254	0.8368	0.8330	0.8083	0.8187	0.8265	0.8440
		std	0.1629	0.0948	0.0998	0.0964	0.1179	0.0644	0.0713	0.0645	0.1008	0.0530	0.0544	0.0495
0	$\widehat{\rho}_n$	sm	0.0168	0.0011	0.0035	0.0009	0.0063	0.0034	-0.0017	0.0017	0.0031	0.0019	0.0053	0.0005
		std	0.1710	0.0958	0.0910	0.0642	0.1202	0.0688	0.0618	0.0446	0.0982	0.0551	0.0512	0.0366
1	$\widehat{\sigma}^2_{n\nu}$	sm	1.0843	1.0827	1.1094	1.2105	1.0657	1.0655	1.0793	1.1428	1.0593	1.0642	1.0791	1.1198
		std	0.2152	0.1201	0.1200	0.1070	0.1403	0.0845	0.0801	0.0702	0.1120	0.0681	0.0664	0.0530
1	$\widehat{\sigma}^2_{ne}$	sm	1.0487	1.0726	1.0999	1.2055	1.0495	1.0604	1.0752	1.1405	1.0488	1.0610	1.0762	1.1183
		std	0.1858	0.1174	0.1156	0.1054	0.1310	0.0829	0.0798	0.0699	0.1083	0.0674	0.0657	0.0528
0.5	$\widehat{\rho}_n$	sm	0.4041	0.4177	0.4377	0.4453	0.4130	0.4247	0.4355	0.4377	0.4280	0.4392	0.4203	0.4363
		std	0.2344	0.1435	0.1175	0.0918	0.1689	0.1045	0.0894	0.0699	0.1361	0.0861	0.0734	0.0561
1	$\widehat{\sigma}^2_{n\nu}$	sm	1.0696	1.0580	1.0506	1.1125	1.1066	1.0330	1.0223	1.0509	1.0678	1.0144	1.0231	1.0424
		std	0.5163	0.4238	0.2357	0.1662	0.4894	0.2313	0.1624	0.1134	0.4158	0.1585	0.1384	0.0967
0.75	$\widehat{\sigma}^2_{ne}$	sm	0.8238	0.8356	0.8573	0.8725	0.8123	0.8249	0.8351	0.8617	0.8112	0.8179	0.8290	0.8419
		std	0.1551	0.1056	0.0952	0.1022	0.1075	0.0743	0.0664	0.0645	0.0892	0.0602	0.0543	0.0512

For the weighted PLSDV estimator $(\widehat{\beta}^w_{1n}, \widehat{\beta}^w_{2n}, \widehat{\beta}^w_{3n})^{\mathsf{T}}$, the unweighted PLSDV estimator $(\widehat{\beta}_{1n}, \widehat{\beta}_{2n}, \widehat{\beta}_{3n})^{\mathsf{T}}$ which neglects the serially correlated error component structure, and the benchmark estimator $(\breve{\beta}^w_{1n}, \breve{\beta}^w_{2n}, \breve{\beta}^w_{3n})^{\mathsf{T}}$ which has the same definition as the weighted PLSDV estimator except for the known the error structure parameters, the sample mean (sm), standard deviation (std) which is calculated based on the estimators of parametric components, mean of the estimate of the standard deviation (mstd) and coverage of the 95% nominal confidence intervals (cps) are summarized in Tables 2−4. In Tables 2−4, we also present the sm, std of the naive estimator $(\widetilde{\beta}_{1n}, \widetilde{\beta}_{2n}, \widetilde{\beta}_{3n})^{\tau}$ which ignores the fixed effects.

From Tables 2−4 we make the following observations:

1. The weighted PLSDV estimator, unweighted PLSDV estimator and benchmark estimator of the parametric components are unbiased. However, the naive estimator is unbiased.
2. The weighted PLSDV estimator has smaller std than the unweighted PLSDV estimator. The std of the weighted PLSDV estimator is very close to that of the benchmark estimator.
3. The weighted PLSDV estimator can substantially improve the estimation of the parametric components over the unweighted PLSDV estimator, especially when the contemporaneous correlation is strong or the number of the observed time points is large.
4. The coverage of the confidence interval is very close to the 95% nominal level.

For the estimators of the nonparametric components, we compute a measure of estimation accuracy referred to as the *root average squared error* (RASE), given by

$$\mathrm{RASE}_j = \Big[\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\{\overline{\alpha}_{jn}(U_{it}) - \alpha_j(U_{it})\}^2\Big]^{1/2}, \qquad j = 1, 2,$$

where $\overline{\alpha}_{jn}(u)$ is either the naive estimator $\widetilde{\alpha}_{jn}(u)$ which ignores the fixed effects, or $\widehat{\alpha}_{jn}(u)$, or $\widehat{\alpha}^{w}_{jn}(u)$ or the benchmark estimator $\widetilde{\alpha}^{w}_{jn}(u)$ which has the same definition as $\widehat{\alpha}^{w}_{jn}(u)$. The sample mean and standard deviation of the RASE are summarized in Table 5. In addition, we plot the estimator $\widetilde{\alpha}_{jn}(u)$, $\widehat{\alpha}_{jn}(u)$ and $\widehat{\alpha}^{w}_{jn}(u)$ for $n = 100, 200, 300$, $T = 4$ and $\rho = -0.5$ and 0.5, respectively.

From Table 5, the RASE values of $\widehat{\alpha}_{jn}(u)$, $\widehat{\alpha}^{w}_{jn}(u)$, $\breve{\alpha}^{w}_{jn}(u)$ always have smaller RASE values than $\widetilde{\alpha}_{jn}(u)$. This is corroborated by Figure 1 and 2. In addition, $\widehat{\alpha}^{w}_{jn}(u)$ and $\breve{\alpha}^{w}_{jn}(u)$ always have smaller RASE values than $\widehat{\alpha}_{jn}(u)$ as well, and most importantly $\widehat{\alpha}^{w}_{jn}(u)$ performs nearly as well as $\breve{\alpha}^{w}_{jn}(u)$. All of these are consistent with the theoretical results.

§6. Application

The data set was extracted from the STARS database of the World Bank, in which measures of GDP and the aggregate physical capital stock for 81 countries over the period

Table 2 Finite sample performance of the parametric component estimators under scenario of $\rho = -0.5$

		$n = 100$				$n = 200$				$n = 300$			
		$T = 3$	$T = 4$	$T = 5$	$T = 10$	$T = 3$	$T = 4$	$T = 5$	$T = 10$	$T = 3$	$T = 4$	$T = 5$	$T = 10$
$\widetilde{\beta}_{1n}$	sm	2.0211	2.0640	2.0960	2.2734	2.0217	2.0468	2.0830	2.2718	2.0241	2.0505	2.0849	2.2605
	std	0.4000	0.3903	0.4011	0.4612	0.2711	0.2735	0.2716	0.3352	0.2291	0.2218	0.2203	0.2698
$\widetilde{\beta}_{2n}$	sm	1.4833	1.5111	1.5124	1.6539	1.4694	1.4992	1.5287	1.6579	1.5040	1.5098	1.5428	1.6748
	std	0.4124	0.4110	0.4056	0.4905	0.2743	0.2833	0.2857	0.3343	0.2289	0.2364	0.2411	0.2766
$\widetilde{\beta}_{3n}$	sm	1.1974	1.2142	1.2507	1.3191	1.1992	1.2093	1.2422	1.3184	1.1926	1.2232	1.2286	1.3082
	std	0.4236	0.4142	0.4246	0.4948	0.2929	0.3006	0.2867	0.3679	0.2410	0.2422	0.2411	0.2877
$\widehat{\beta}_{1n}$	sm	1.5200	1.5032	1.4937	1.5049	1.4855	1.4893	1.5036	1.4977	1.5012	1.4886	1.4978	1.5001
	std	0.4237	0.3241	0.2801	0.1744	0.2885	0.2347	0.1907	0.1234	0.2387	0.1938	0.1537	0.1063
	mstd	0.4237	0.3370	0.2883	0.1915	0.2951	0.2347	0.2001	0.1316	0.2390	0.1907	0.1624	0.1065
	cp	0.9490	0.9680	0.9570	0.9720	0.9490	0.9530	0.9630	0.9610	0.9540	0.9490	0.9610	0.9530
$\widehat{\beta}_{2n}$	sm	0.9961	1.0032	0.9919	1.0034	0.9932	1.0034	1.0035	0.9965	1.0136	0.9927	1.0065	1.0015
	std	0.4216	0.3395	0.2901	0.1805	0.2950	0.2422	0.1997	0.1309	0.2564	0.1885	0.1687	0.1066
	mstd	0.4345	0.3476	0.2962	0.1970	0.3028	0.2414	0.2055	0.1353	0.2460	0.1964	0.1670	0.1095
	cp	0.9620	0.9530	0.9560	0.9640	0.9630	0.9560	0.9580	0.9600	0.9400	0.9540	0.9510	0.9460
$\widehat{\beta}_{3n}$	sm	0.7799	0.7514	0.7667	0.7544	0.7653	0.7457	0.7510	0.7519	0.7497	0.7528	0.7471	0.7522
	std	0.4473	0.3419	0.2989	0.1775	0.3149	0.2462	0.2033	0.1329	0.2426	0.1979	0.1694	0.1085
	mstd	0.4472	0.3559	0.3040	0.2025	0.3114	0.2475	0.2113	0.1391	0.2523	0.2016	0.1716	0.1125
	cp	0.9410	0.9650	0.9630	0.9700	0.9580	0.9430	0.9560	0.9650	0.9540	0.9500	0.9550	0.9530
$\widehat{\beta}^w_{1n}$	sm	1.5167	1.5021	1.4984	1.4997	1.4868	1.4919	1.5010	1.4992	1.5035	1.4985	1.4982	1.5018
	std	0.3921	0.2867	0.2412	0.1405	0.2674	0.2046	0.1606	0.1003	0.2231	0.1685	0.1296	0.0853
	mstd	0.3964	0.3021	0.2535	0.1635	0.2740	0.2077	0.1729	0.1102	0.2218	0.1679	0.1397	0.0887
	cp	0.9480	0.9710	0.9690	0.9680	0.9520	0.9510	0.9670	0.9570	0.9460	0.9480	0.9570	0.9590
$\widehat{\beta}^w_{2n}$	sm	0.9947	1.0078	0.9996	1.0021	0.9921	1.0049	1.0010	0.9999	1.0140	0.9927	1.0049	0.9994
	std	0.3902	0.2826	0.2418	0.1447	0.2704	0.2083	0.1689	0.1022	0.2369	0.1673	0.1385	0.0874
	mstd	0.4066	0.3115	0.2605	0.1682	0.2812	0.2136	0.1775	0.1132	0.2284	0.1729	0.1436	0.0912
	cp	0.9550	0.9690	0.9650	0.9730	0.9630	0.9540	0.9580	0.9670	0.9420	0.9580	0.9540	0.9450
$\widehat{\beta}^w_{3n}$	sm	0.7796	0.7424	0.7610	0.7550	0.7690	0.7506	0.7517	0.7511	0.7468	0.7499	0.7493	0.7515
	std	0.4127	0.2894	0.2515	0.1386	0.2905	0.2109	0.1728	0.1048	0.2245	0.1677	0.1424	0.0873
	mstd	0.4185	0.3189	0.2674	0.1728	0.2891	0.2191	0.1826	0.1163	0.2342	0.1775	0.1475	0.0937
	cp	0.9550	0.9690	0.9600	0.9830	0.9540	0.9540	0.9600	0.9730	0.9470	0.9600	0.9570	0.9640
$\check{\beta}^w_{1n}$	sm	1.5116	1.4993	1.4970	1.4981	1.4845	1.4905	1.4995	1.4987	1.5020	1.4979	1.4973	1.5016
	std	0.3904	0.2859	0.2404	0.1391	0.2669	0.2036	0.1596	0.1003	0.2236	0.1682	0.1298	0.0850
	mstd	0.3909	0.2868	0.2352	0.1421	0.2682	0.1989	0.1636	0.0999	0.2167	0.1613	0.1329	0.0813
	cp	0.9450	0.9620	0.9450	0.9500	0.9440	0.9450	0.9570	0.9440	0.9400	0.9400	0.9500	0.9340
$\check{\beta}^w_{2n}$	sm	0.9913	1.0064	0.9980	1.0006	0.9916	1.0033	1.0001	0.9998	1.0127	0.9919	1.0041	0.9991
	std	0.3895	0.2809	0.2393	0.1445	0.2699	0.2079	0.1691	0.1016	0.2362	0.1679	0.1379	0.0874
	mstd	0.4066	0.3115	0.2605	0.1682	0.2812	0.2136	0.1775	0.1132	0.2284	0.1729	0.1436	0.0912
	cp	0.9610	0.9640	0.9530	0.9470	0.9600	0.9490	0.9440	0.9500	0.9360	0.9510	0.9450	0.9200
$\check{\beta}^w_{3n}$	sm	0.7790	0.7379	0.7589	0.7551	0.7669	0.7492	0.7510	0.7508	0.7461	0.7486	0.7490	0.7511
	std	0.4098	0.2880	0.2503	0.1384	0.2912	0.2104	0.1725	0.1043	0.2234	0.1669	0.1420	0.0871
	mstd	0.4125	0.3029	0.2481	0.1502	0.2828	0.2097	0.1728	0.1055	0.2288	0.1705	0.1405	0.0859
	cp	0.9520	0.9580	0.9420	0.9680	0.9510	0.9480	0.9520	0.9560	0.9460	0.9540	0.9440	0.9470

Table 3 Finite sample performance of the parametric component estimators under scenario of $\rho = 0$

		$n = 100$				$n = 200$				$n = 300$			
		$T = 3$	$T = 4$	$T = 5$	$T = 10$	$T = 3$	$T = 4$	$T = 5$	$T = 10$	$T = 3$	$T = 4$	$T = 5$	$T = 10$
$\widetilde{\beta}_{1n}$	sm	2.0124	2.0586	2.0901	2.2504	2.0208	2.0498	2.0752	2.2655	2.0265	2.0610	2.1022	2.2526
	std	0.3991	0.3872	0.3872	0.4639	0.2697	0.2670	0.2685	0.3233	0.2319	0.2243	0.2259	0.2732
$\widetilde{\beta}_{2n}$	sm	1.4825	1.4797	1.5236	1.6798	1.4895	1.5071	1.5325	1.6630	1.4818	1.5172	1.5251	1.6688
	std	0.4150	0.4104	0.4119	0.4928	0.2865	0.2898	0.2866	0.3307	0.2368	0.2393	0.2326	0.2705
$\widetilde{\beta}_{3n}$	sm	1.1791	1.2235	1.2390	1.2969	1.2100	1.2208	1.2384	1.3220	1.2044	1.2116	1.2427	1.3203
	std	0.4364	0.4124	0.4114	0.4826	0.2813	0.2920	0.2876	0.3423	0.2354	0.2379	0.2470	0.2783
$\widehat{\beta}_{1n}$	sm	1.5043	1.5017	1.4969	1.5043	1.5006	1.5016	1.4996	1.4974	1.5074	1.5010	1.5014	1.5054
	std	0.3688	0.3137	0.2589	0.1743	0.2583	0.2090	0.1790	0.1208	0.2057	0.1701	0.1465	0.0999
	mstd	0.3845	0.3148	0.27307	0.1882	0.2664	0.2169	0.1887	0.1286	0.2158	0.1763	0.1532	0.1038
	cp	0.9550	0.9470	0.9620	0.9660	0.9560	0.9600	0.9590	0.9680	0.9560	0.9550	0.9600	0.9480
$\widehat{\beta}_{2n}$	sm	0.9890	1.0008	1.0130	1.0029	0.9970	1.0021	0.9954	0.9976	0.9910	0.9954	0.9934	0.9986
	std	0.3788	0.3262	0.2701	0.1812	0.2657	0.2215	0.1785	0.1271	0.2192	0.1832	0.1528	0.1007
	mstd	0.3956	0.3231	0.2808	0.1936	0.2738	0.2233	0.1942	0.1323	0.2221	0.1814	0.1577	0.1068
	cp	0.9530	0.9450	0.9540	0.9620	0.9560	0.9560	0.9720	0.9570	0.9510	0.9490	0.9530	0.9670
$\widehat{\beta}_{3n}$	sm	0.7100	0.7724	0.7513	0.7538	0.7484	0.7510	0.7522	0.7546	0.7558	0.7428	0.7524	0.7494
	std	0.4095	0.3215	0.2628	0.1898	0.2738	0.2295	0.1826	0.1283	0.2198	0.1815	0.1560	0.1000
	mstd	0.4067	0.3317	0.2889	0.1986	0.2815	0.2293	0.1992	0.1359	0.2278	0.1862	0.1620	0.1097
	cp	0.9490	0.9520	0.9650	0.9630	0.9470	0.9510	0.9690	0.9540	0.9560	0.9510	0.9550	0.9730
$\widehat{\beta}^w_{1n}$	sm	1.5087	1.5035	1.5007	1.5061	1.5027	1.5032	1.5010	1.4979	1.5084	1.5010	1.5022	1.5059
	std	0.3682	0.3149	0.2605	0.1754	0.2591	0.2090	0.1794	0.1211	0.2059	0.1702	0.1468	0.0999
	mstd	0.3823	0.3137	0.2719	0.1876	0.2656	0.2165	0.1883	0.1284	0.2154	0.1761	0.1530	0.1037
	cp	0.9510	0.9470	0.9630	0.9640	0.9560	0.9550	0.9570	0.9650	0.9560	0.9550	0.9570	0.9500
$\widehat{\beta}^w_{2n}$	sm	0.9935	1.0040	1.0150	1.0035	0.9994	1.0037	0.9965	0.9982	0.9924	0.9969	0.9941	0.9989
	std	0.3827	0.3278	0.2704	0.1822	0.2665	0.2218	0.1794	0.1274	0.2187	0.1831	0.1526	0.1009
	mstd	0.3933	0.3219	0.2796	0.1930	0.2729	0.2229	0.1938	0.1321	0.2217	0.1812	0.1575	0.1067
	cp	0.9520	0.9400	0.9510	0.9650	0.9530	0.9570	0.9700	0.9570	0.9520	0.9440	0.9550	0.9670
$\widehat{\beta}^w_{3n}$	sm	0.7127	0.7733	0.7523	0.7549	0.7515	0.7521	0.7530	0.7553	0.7567	0.7438	0.7531	0.7499
	std	0.4112	0.3204	0.2638	0.1902	0.2752	0.2294	0.1827	0.1289	0.2200	0.1817	0.1565	0.0998
	mstd	0.4044	0.3305	0.2877	0.1980	0.2807	0.2288	0.1988	0.1357	0.2273	0.1860	0.1617	0.1096
	cp	0.9460	0.9530	0.9620	0.9600	0.9460	0.9520	0.9690	0.9540	0.9520	0.9510	0.9540	0.9740
$\check{\beta}^w_{1n}$	sm	1.5043	1.5017	1.4969	1.5043	1.5006	1.5016	1.4996	1.4974	1.5074	1.5010	1.5014	1.5054
	std	0.3688	0.3137	0.2589	0.1743	0.2583	0.2090	0.1790	0.1208	0.2057	0.1701	0.1465	0.0999
	mstd	0.3775	0.3039	0.2603	0.1713	0.2606	0.2107	0.1819	0.1205	0.2109	0.1712	0.1478	0.0981
	cp	0.9520	0.9430	0.9520	0.9480	0.9530	0.9490	0.9500	0.9510	0.9510	0.9450	0.9500	0.9390
$\check{\beta}^w_{n2}$	sm	0.9890	1.0008	1.0130	1.0029	0.9970	1.0021	0.9954	0.9976	0.9910	0.9954	0.9934	0.9986
	std	0.3788	0.3262	0.2701	0.1812	0.2657	0.2215	0.1785	0.1271	0.2192	0.1832	0.1528	0.1007
	mstd	0.3933	0.3219	0.2796	0.1930	0.2729	0.2229	0.1938	0.1321	0.2217	0.1812	0.1575	0.1067
	cp	0.9550	0.9320	0.9420	0.9430	0.9530	0.9490	0.9650	0.9410	0.9460	0.9360	0.9480	0.9510
$\check{\beta}^w_{3n}$	sm	0.7100	0.7724	0.7513	0.7538	0.7484	0.7510	0.7522	0.7546	0.7558	0.7428	0.7524	0.7494
	std	0.4095	0.3215	0.2628	0.1898	0.2738	0.2295	0.1826	0.1283	0.2198	0.1815	0.1560	0.1000
	mstd	0.3994	0.3202	0.2755	0.1808	0.2754	0.2228	0.1920	0.1272	0.2227	0.1809	0.1563	0.1037
	cp	0.9450	0.9470	0.9590	0.9460	0.9480	0.9470	0.9650	0.9430	0.9520	0.9460	0.9470	0.9640

Table 4 Finite sample performance of the parametric component estimators under scenario of $\rho = 0.5$

		$n = 100$				$n = 200$				$n = 300$			
		$T = 3$	$T = 4$	$T = 5$	$T = 10$	$T = 3$	$T = 4$	$T = 5$	$T = 10$	$T = 3$	$T = 4$	$T = 5$	$T = 10$
$\widetilde{\beta}_{1n}$	sm	2.0088	2.0428	2.0810	2.2588	2.0267	2.0522	2.0797	2.2717	2.0263	2.0594	2.0933	2.2656
	std	0.4043	0.3845	0.4068	0.4694	0.2820	0.2736	0.2743	0.3239	0.2189	0.2206	0.2273	0.2663
$\widetilde{\beta}_{2n}$	sm	1.4779	1.4919	1.5291	1.6455	1.4829	1.4976	1.5249	1.6575	1.4900	1.5181	1.5419	1.6671
	std	0.4115	0.4081	0.3868	0.4903	0.2999	0.2908	0.2903	0.3549	0.2229	0.2252	0.2265	0.2729
$\widetilde{\beta}_{3n}$	sm	1.2217	1.2370	1.2130	1.3009	1.2025	1.2105	1.2446	1.3033	1.1997	1.2192	1.2263	1.3194
	std	0.4284	0.4158	0.4029	0.4926	0.2807	0.2903	0.2908	0.3753	0.2417	0.2420	0.2419	0.2941
$\widehat{\beta}_{1n}$	sm	1.4831	1.4866	1.4958	1.4984	1.5047	1.4998	1.5002	1.5025	1.5018	1.5007	1.5001	1.5009
	std	0.2881	0.2509	0.2229	0.1532	0.1949	0.1703	0.1489	0.1081	0.1603	0.1352	0.1241	0.0907
	mstd	0.3090	0.2613	0.2338	0.1703	0.2105	0.1793	0.1597	0.1158	0.1697	0.1448	0.1296	0.0937
	cp	0.9610	0.9580	0.9660	0.9710	0.9720	0.9610	0.9580	0.9690	0.9660	0.9710	0.9550	0.9480
$\widehat{\beta}_{2n}$	sm	1.0110	1.0068	0.9882	1.0018	0.9959	1.0023	1.0011	1.0020	1.0031	0.9974	0.9988	0.9999
	std	0.3076	0.2461	0.2210	0.1684	0.2076	0.1800	0.1570	0.1142	0.1687	0.1467	0.1273	0.0945
	mstd	0.3178	0.2692	0.2403	0.1752	0.2166	0.1841	0.1643	0.1191	0.1745	0.1488	0.1331	0.0964
	cp	0.9550	0.9640	0.9650	0.9580	0.9650	0.9590	0.9580	0.9590	0.9500	0.9520	0.9620	0.9520
$\widehat{\beta}_{3n}$	sm	0.7711	0.7468	0.7552	0.7527	0.7496	0.7513	0.7513	0.7451	0.7532	0.7491	0.7438	0.7553
	std	0.3045	0.2618	0.2261	0.1604	0.2175	0.1706	0.1593	0.1149	0.1710	0.1455	0.1261	0.0951
	mstd	0.3265	0.2759	0.2466	0.1798	0.2220	0.1889	0.1688	0.1224	0.1795	0.1530	0.1368	0.0989
	cp	0.9700	0.9650	0.9660	0.9730	0.9510	0.9730	0.9590	0.9590	0.9650	0.9610	0.9650	0.9570
$\widehat{\beta}^w_{1n}$	sm	1.4881	1.4918	1.5002	1.5001	1.5089	1.4989	1.5030	1.5000	1.5013	1.5010	1.4999	1.5007
	std	0.2792	0.2342	0.2002	0.1310	0.1878	0.1568	0.1352	0.0922	0.1545	0.1254	0.1120	0.0772
	mstd	0.2997	0.2486	0.2188	0.1556	0.2039	0.1693	0.1483	0.1043	0.1642	0.1365	0.1197	0.0836
	cp	0.9610	0.9660	0.9650	0.9800	0.9690	0.9690	0.9690	0.9730	0.9710	0.9700	0.9540	0.9700
$\widehat{\beta}^w_{2n}$	sm	1.0110	1.0085	0.9935	1.0006	0.9972	1.0045	1.0038	1.0000	1.0043	1.0009	0.9981	0.9994
	std	0.3009	0.2295	0.2018	0.1444	0.1998	0.1663	0.1412	0.0976	0.1613	0.1369	0.1159	0.0777
	mstd	0.3083	0.2562	0.2248	0.1601	0.2098	0.1738	0.1525	0.1073	0.1688	0.1404	0.1230	0.0860
	cp	0.9540	0.9650	0.9670	0.9720	0.9610	0.9690	0.9660	0.9700	0.9530	0.9580	0.9600	0.9690
$\widehat{\beta}^w_{2n}$	sm	0.7701	0.7506	0.7550	0.7544	0.7480	0.7509	0.7520	0.7476	0.7516	0.7486	0.7452	0.7529
	std	0.2976	0.2445	0.2083	0.1411	0.2085	0.1589	0.1443	0.0975	0.1642	0.1335	0.1170	0.0798
	mstd	0.3167	0.2625	0.2308	0.1643	0.2150	0.1784	0.1567	0.1102	0.1736	0.1443	0.1264	0.0883
	cp	0.9620	0.9640	0.9710	0.9770	0.9540	0.9770	0.9620	0.9700	0.9700	0.9660	0.9680	0.9730
$\check{\beta}^w_{1n}$	sm	1.4887	1.4907	1.4995	1.4996	1.5084	1.4973	1.5012	1.4990	1.5005	1.5001	1.4993	1.5001
	std	0.2770	0.2323	0.1975	0.1287	0.1868	0.1572	0.1339	0.0920	0.1539	0.1253	0.1115	0.0766
	mstd	0.2778	0.2258	0.1960	0.1306	0.1910	0.1571	0.1363	0.0917	0.1545	0.1274	0.1109	0.0748
	cp	0.9440	0.9510	0.9480	0.9600	0.9620	0.9500	0.9490	0.9490	0.9590	0.9550	0.9400	0.9400
$\check{\beta}^w_{2n}$	sm	1.0098	1.0075	0.9926	0.9980	0.9959	1.0037	1.0034	0.9989	1.0036	1.0003	0.9975	0.9987
	std	0.2964	0.2277	0.2000	0.1415	0.1991	0.1656	0.1406	0.0968	0.1609	0.1367	0.1155	0.0769
	mstd	0.3083	0.2562	0.2248	0.1601	0.2098	0.1738	0.1525	0.1073	0.1688	0.1404	0.1230	0.0860
	cp	0.9460	0.9600	0.9520	0.9380	0.9530	0.9520	0.9450	0.9380	0.9380	0.9390	0.9440	0.9490
$\check{\beta}^w_{3n}$	sm	0.7664	0.7483	0.7541	0.7528	0.7469	0.7498	0.7513	0.7473	0.7508	0.7475	0.7446	0.7518
	std	0.2976	0.2428	0.2071	0.1399	0.2075	0.1586	0.1438	0.0969	0.1642	0.1328	0.1167	0.0794
	mstd	0.2936	0.2383	0.2067	0.1380	0.2014	0.1656	0.1440	0.0969	0.1634	0.1346	0.1170	0.0789
	cp	0.9460	0.9500	0.9530	0.9470	0.9430	0.9630	0.9450	0.9520	0.9510	0.9550	0.9560	0.9560

Table 5 Finite sample performance of the nonparametric component estimators under the scenarios of $\rho = -0.5, 0$ and 0.5, respectively

			$n = 100$				$n = 200$				$n = 300$			
ρ			$T = 3$	$T = 4$	$T = 5$	$T = 10$	$T = 3$	$T = 4$	$T = 5$	$T = 10$	$T = 3$	$T = 4$	$T = 5$	$T = 10$
0.5	$\widetilde{\alpha}_{1n}(\cdot)$	sm(RASE)	1.1451	1.1281	1.1440	1.2894	0.8903	0.8863	0.8913	0.9985	0.8084	0.7960	0.8035	0.8686
		std(RASE)	0.3089	0.2925	0.3082	0.3493	0.2384	0.2314	0.2376	0.2651	0.2072	0.1944	0.2024	0.2256
	$\widetilde{\alpha}_{2n}(\cdot)$	sm(RASE)	1.1738	1.1678	1.1959	1.3569	0.9318	0.9272	0.9411	1.0573	0.8369	0.8336	0.8424	0.9491
		std(RASE)	0.3209	0.3194	0.3437	0.3680	0.2330	0.2409	0.2603	0.2909	0.2155	0.2161	0.2048	0.2461
	$\widehat{\alpha}_{1n}(\cdot)$	sm(RASE)	0.9963	0.7878	0.6678	0.4257	0.6967	0.5545	0.4688	0.3080	0.5651	0.4510	0.3810	0.2556
		std(RASE)	0.2551	0.1996	0.1740	0.1086	0.1813	0.1450	0.1190	0.0718	0.1463	0.1107	0.0927	0.0574
	$\widehat{\alpha}_{2n}(\cdot)$	sm(RASE)	1.0172	0.7928	0.6682	0.4241	0.6950	0.5568	0.4681	0.2963	0.5654	0.4523	0.3725	0.2437
		std(RASE)	0.2650	0.2076	0.1756	0.1126	0.1851	0.1481	0.1243	0.0773	0.1569	0.1147	0.0972	0.0648
	$\widehat{\alpha}^w_{1n}(\cdot)$	sm(RASE)	0.9267	0.6846	0.5741	0.3501	0.6461	0.4854	0.3962	0.2512	0.5251	0.3910	0.3271	0.2144
		std(RASE)	0.2372	0.1782	0.1490	0.0866	0.1683	0.1262	0.0985	0.0559	0.1360	0.0942	0.0774	0.0444
	$\widehat{\alpha}^w_{2n}(\cdot)$	sm(RASE)	0.9486	0.6907	0.5711	0.3437	0.6391	0.4840	0.3929	0.2356	0.5225	0.3889	0.3150	0.1962
		std(RASE)	0.2495	0.1872	0.1476	0.0902	0.1694	0.1311	0.1054	0.0615	0.1403	0.0989	0.0792	0.0511
	$\breve{\alpha}^w_{1n}(\cdot)$	sm(RASE)	0.9227	0.6821	0.5703	0.3482	0.6447	0.4850	0.3949	0.2504	0.5236	0.3905	0.3270	0.2143
		std(RASE)	0.2364	0.1793	0.1490	0.0860	0.1673	0.1258	0.0978	0.0557	0.1356	0.0940	0.0774	0.0444
	$\breve{\alpha}^w_{2n}(\cdot)$	sm(RASE)	0.9445	0.6886	0.5683	0.3417	0.6382	0.4837	0.3917	0.2351	0.5213	0.3882	0.3146	0.1959
		std(RASE)	0.2475	0.1877	0.1464	0.0895	0.1707	0.1310	0.1041	0.0614	0.1404	0.0989	0.0791	0.0512
0	$\widetilde{\alpha}_{1n}(\cdot)$	sm(RASE)	1.1403	1.1340	1.1496	1.2774	0.8942	0.8869	0.9060	0.9949	0.8122	0.8060	0.8042	0.8633
		std(RASE)	0.3174	0.3074	0.3140	0.3459	0.2314	0.2328	0.2299	0.2631	0.1970	0.2032	0.2026	0.2137
	$\widetilde{\alpha}_{2n}(\cdot)$	sm(RASE)	1.1738	1.1626	1.1892	1.3660	0.9419	0.9169	0.9389	1.0632	0.8409	0.8407	0.8622	0.9545
		std(RASE)	0.3146	0.3158	0.3281	0.3738	0.2473	0.2465	0.2491	0.2821	0.2129	0.2146	0.2105	0.2301
	$\widehat{\alpha}_{1n}(\cdot)$	sm(RASE)	0.8893	0.7127	0.6265	0.4078	0.6174	0.5024	0.4406	0.2975	0.4989	0.4122	0.3611	0.2507
		std(RASE)	0.2362	0.1828	0.1571	0.1026	0.1603	0.1264	0.1071	0.0714	0.1283	0.1019	0.0860	0.0575
	$\widehat{\alpha}_{2n}(\cdot)$	sm(RASE)	0.8964	0.7182	0.6287	0.4017	0.6198	0.5052	0.4371	0.2904	0.5011	0.4095	0.3504	0.2359
		std(RASE)	0.2445	0.1922	0.1714	0.1080	0.1559	0.1294	0.1112	0.0761	0.1315	0.1066	0.0897	0.0613
	$\widehat{\alpha}^w_{1n}(\cdot)$	sm(RASE)	0.8930	0.7134	0.6269	0.4095	0.6189	0.5032	0.4413	0.2977	0.4993	0.4128	0.3612	0.2510
		std(RASE)	0.2370	0.1835	0.1564	0.1035	0.1614	0.1269	0.1073	0.0716	0.1280	0.1021	0.0861	0.0575
	$\widehat{\alpha}^w_{2n}(\cdot)$	sm(RASE)	0.8994	0.7192	0.6299	0.4036	0.6206	0.5059	0.4376	0.2907	0.5022	0.4098	0.3503	0.2361
		std(RASE)	0.2441	0.1925	0.1721	0.1087	0.1561	0.1303	0.1115	0.0762	0.1314	0.1067	0.0896	0.0614
	$\breve{\alpha}^w_{1n}(\cdot)$	sm(RASE)	0.8893	0.7127	0.6265	0.4078	0.6174	0.5024	0.4406	0.2975	0.4989	0.4122	0.3611	0.2507
		std(RASE)	0.2362	0.1828	0.1571	0.1026	0.1603	0.1264	0.1071	0.0714	0.1283	0.1019	0.0860	0.0575
	$\breve{\alpha}^w_{2n}(\cdot)$	sm(RASE)	0.8964	0.7182	0.6287	0.4017	0.6198	0.5052	0.4371	0.2904	0.5011	0.4095	0.3504	0.2359
		std(RASE)	0.2445	0.1922	0.1714	0.1080	0.1559	0.1294	0.1112	0.0761	0.1315	0.1066	0.0897	0.0613
-0.5	$\widetilde{\alpha}_{1n}(\cdot)$	sm(RASE)	1.1732	1.1383	1.1427	1.2698	0.9018	0.8868	0.8888	0.9869	0.8014	0.8006	0.8002	0.8627
		std(RASE)	0.3227	0.3101	0.3218	0.3483	0.2353	0.2329	0.2351	0.2618	0.2043	0.2008	0.1968	0.2280
	$\widetilde{\alpha}_{2n}(\cdot)$	sm(RASE)	1.1825	1.1898	1.1870	1.3525	0.9108	0.9254	0.9427	1.0591	0.8425	0.8403	0.8560	0.9472
		std(RASE)	0.3094	0.3413	0.3287	0.3748	0.2517	0.2459	0.2577	0.2802	0.2088	0.2119	0.2106	0.2414
	$\widehat{\alpha}_{1n}(\cdot)$	sm(RASE)	0.7018	0.5858	0.5155	0.3783	0.4696	0.4083	0.3670	0.2715	0.3913	0.3362	0.3085	0.2335
		std(RASE)	0.1871	0.1492	0.1272	0.0936	0.1255	0.0981	0.0905	0.0619	0.0981	0.0815	0.0713	0.0526
	$\widehat{\alpha}_{2n}(\cdot)$	sm(RASE)	0.7038	0.5849	0.5151	0.3734	0.4702	0.4063	0.3623	0.2587	0.3824	0.3279	0.2983	0.2149
		std(RASE)	0.1951	0.1579	0.1409	0.0976	0.1248	0.1032	0.0960	0.0698	0.1007	0.0865	0.0806	0.0574
	$\widehat{\alpha}^w_{1n}(\cdot)$	sm(RASE)	0.6844	0.5521	0.4717	0.3260	0.4569	0.3815	0.3384	0.2363	0.3766	0.3148	0.2807	0.2037
		std(RASE)	0.1813	0.1434	0.1154	0.0762	0.1175	0.0900	0.0832	0.0518	0.0964	0.0734	0.0635	0.0420
	$\widehat{\alpha}^w_{2n}(\cdot)$	sm(RASE)	0.6835	0.5496	0.4692	0.3175	0.4565	0.3768	0.3310	0.2196	0.3684	0.3052	0.2693	0.1824
		std(RASE)	0.1913	0.1474	0.1272	0.0819	0.1205	0.0952	0.0884	0.0571	0.0998	0.0787	0.0733	0.0485
	$\breve{\alpha}^w_{1n}(\cdot)$	sm(RASE)	0.6783	0.5496	0.4670	0.3206	0.4555	0.3806	0.3373	0.2348	0.3755	0.3143	0.2796	0.2028
		std(RASE)	0.1798	0.1439	0.1152	0.0739	0.1166	0.0897	0.0824	0.0510	0.0961	0.0728	0.0628	0.0413
	$\breve{\alpha}^w_{2n}(\cdot)$	sm(RASE)	0.6783	0.5462	0.4655	0.3117	0.4547	0.3758	0.3298	0.2173	0.3675	0.3050	0.2687	0.1813
		std(RASE)	0.1879	0.1465	0.1250	0.0802	0.1202	0.0948	0.0874	0.0561	0.0992	0.0777	0.0726	0.0482

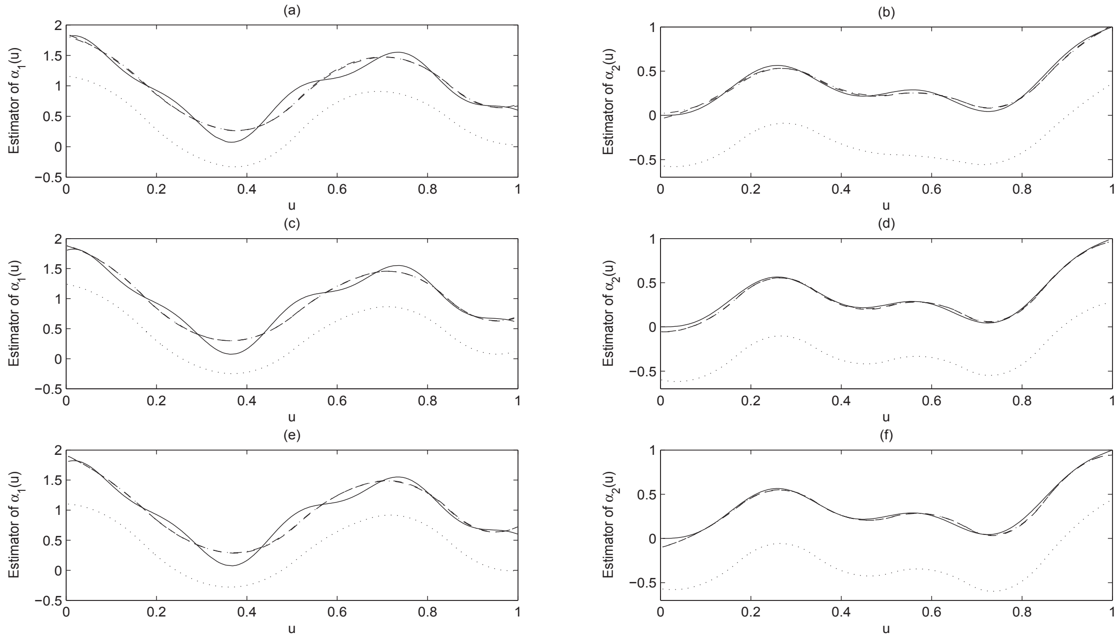

Figure 1 **The estimators $\widetilde{\alpha}_{jn}(u)$ (dotted curve), $\widehat{\alpha}_{jn}(u)$ (dash-dotted curve), $\widehat{\alpha}^{w}_{jn}(u)$ (dashed curve) and $\alpha_j(u)$ (solid curve) for $j = 1, 2$. (a) The estimators of $\alpha_1(u)$ with $n = 100$, $T = 4$, $\rho = -0.5$; (b) The estimators of $\alpha_2(u)$ with $n = 100$, $T = 4$, $\rho = -0.5$; (c) The estimators of $\alpha_1(u)$ with $n = 200$, $T = 4$, $\rho = -0.5$; (d) The estimators of $\alpha_2(u)$ with $n = 200$, $T = 4$, $\rho = -0.5$; (e) The estimators of $\alpha_1(u)$ with $n = 300$, $T = 4$, $\rho = -0.5$; (f) The estimators of $\alpha_2(u)$ with $n = 300$, $T = 4$, $\rho = -0.5$.**

from 1968 to 1987 were recorded. More details can be found in [20]. Same as [21] and [22], we will analyze this data by following an autoregressive structure:

$$Y_{it} = X_{1it}\beta_1 + X_{2it}\beta_2 + \alpha(U_{it}) + \mu_i + \nu_{it}, \quad \nu_{it} = \rho\nu_{i,t-1} + e_{it},\ i = 1, 2, \ldots, 8w,\ t = 1, 2, \ldots, 20,$$

where Y_{it} is the log real GDP of country i in year t (with $t = 1$ for year 1968, and so on), X_{1it} is the log real capital, X_{2it} is the log labor supply, U_{it} is the log mean years of schooling for the workforce, and μ_i is the individual effect of country i.

Fitting this model to the data set, we find $\widehat{\rho} = 0.8161$, $\widehat{\sigma}^2_{\nu} = 0.0072$ and $\widehat{\sigma}^2_e = 0.0024$. The unweighted PLSDV estimator $\widehat{\boldsymbol{\beta}}_n = (\widehat{\beta}_{1n}, \widehat{\beta}_{2n})^{\mathsf{T}} = (0.5433, 0.2585)^{\mathsf{T}}$ with s-tandard error $(0.0247, 0.0756)^{\mathsf{T}}$. The corresponding confidence intervals of β_1 and β_2 are $(0.4949, 0.5918)$ and $(0.1104, 0.4066)$. The weighted PLSDV estimator $\widehat{\boldsymbol{\beta}}^w_n = (\widehat{\beta}^w_{1n}, \widehat{\beta}^w_{2n})^{\mathsf{T}} = (0.5805, 0.1580)^{\mathsf{T}}$ with standard error $(0.0216, 0.0648)^{\mathsf{T}}$. The corresponding confidence intervals of β_1 and β_2 are $(0.5381, 0.6229)$ and $(0.0309, 0.2851)$. We see that $\widehat{\boldsymbol{\beta}}^w_n$ has smaller

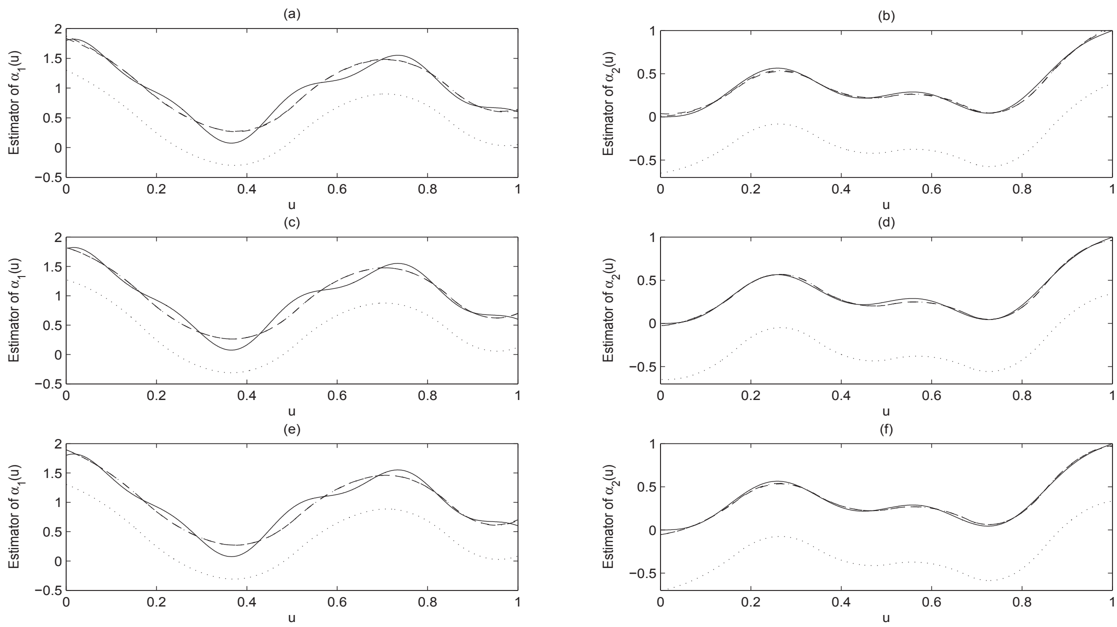

Figure 2 **The estimators $\widetilde{\alpha}_{jn}(u)$ (dotted curve), $\widehat{\alpha}_{jn}(u)$ (dash-dotted curve), $\widehat{\alpha}_{jn}^{w}(u)$ (dashed curve) and $\alpha_j(u)$ (solid curve) for $j = 1, 2$. (a) The estimators of $\alpha_1(u)$ with $n = 100$, $T = 4$, $\rho = 0.5$; (b) The estimators of $\alpha_2(u)$ with $n = 100$, $T = 4$, $\rho = 0.5$; (c) The estimators of $\alpha_1(u)$ with $n = 200$, $T = 4$, $\rho = 0.5$; (d) The estimators of $\alpha_2(u)$ with $n = 200$, $T = 4$, $\rho = 0.5$; (e) The estimators of $\alpha_1(u)$ with $n = 300$, $T = 4$, $\rho = 0.5$; (f) The estimators of $\alpha_2(u)$ with $n = 300$, $T = 4$, $\rho = 0.5$.**

standard error than $\widehat{\boldsymbol{\beta}}_n$. $\widehat{\boldsymbol{\beta}}_n$ and $\widehat{\boldsymbol{\beta}}_n^w$ show that both log real capital and log labor supply are significant. This implies that both log real capital and log labor supply positively affect log GDP.

The estimators of the effect of log mean years of schooling for the workforce on GDP are presented in Figure 3. From Figure 3 (b) we see that the weighted PLSDV estimator $\widehat{\alpha}_n^w(u)$ has smaller (point wise) standard errors than the unweighted PLSDV estimator $\widehat{\alpha}_n(u)$. Figure 3 (a), (c) and (d) show that when the workforce has log mean years of schooling between -2.5 and 1 (or between 0 and 2.7 years in real time) GDP changes not much, however, when the workforce has log mean years of schooling over 1 (or over 2.7 years), GDP increases very quickly and nonlinearly. This is consistent with former results.

Appendix

To present the proofs of the main results we first introduce two lemmas.

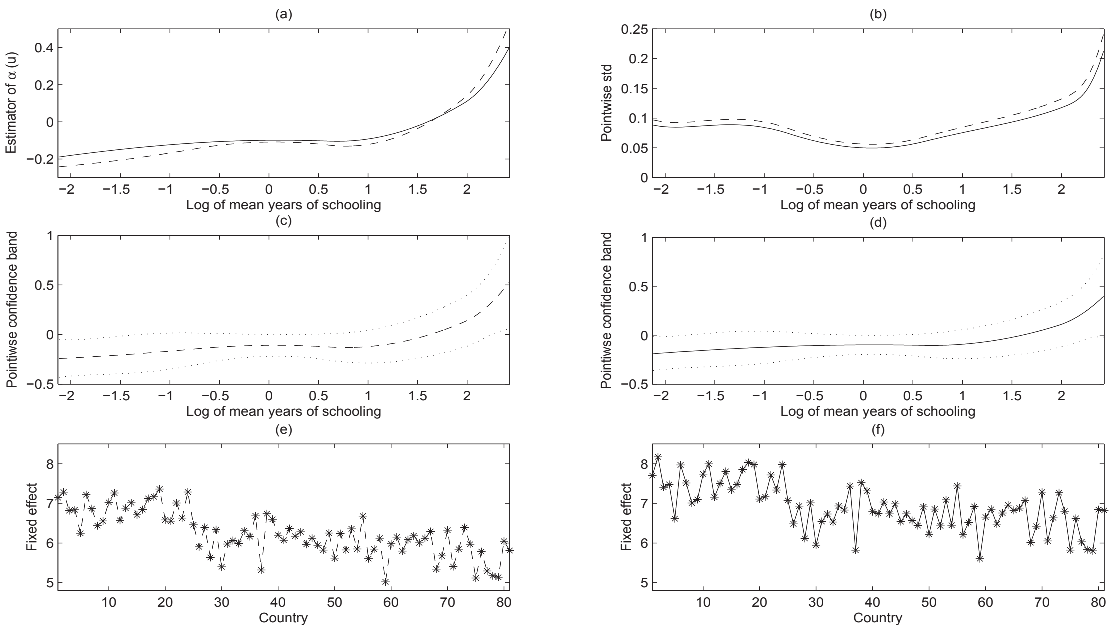

Figure 3 Estimators of the effect of log mean years of schooling for the workforce on log GDP and the fixed effects. (a) the unweighted profile LSDV estimator $\widehat{\alpha}_n(u)$ (dashed curve) and the weighted profile LSDV estimator $\widehat{\alpha}_n^w(u)$ (solid curve); (b) the point std of the unweighted profile LSDV estimator $\widehat{\alpha}_n(u)$ (dashed curve) and the point std of the weighted profile LSDV estimator $\widehat{\alpha}_n^w(u)$ (solid curve); (c) the point confidence band of $\alpha(u)$ based on the unweighted profile LSDV estimator $\widehat{\alpha}_n(u)$; (d) the point confidence band of $\alpha(u)$ based on the weighted profile LSDV estimator $\widehat{\alpha}_n^w(u)$; (e) the estimator of the fixed effects from unweighted profile LSDV fitting; (f) the estimator of the fixed effects from weighted profile LSDV fitting.

Lemma 14 There are positive constants c_1 and c_2 such that

$$c_1\|\theta\|^2 \leqslant \int_{\mathscr{U}} \Big\{ \sum_{l=1}^{\kappa_n} \theta_{jl}\zeta_l(u)\Big\}^2 \mathrm{d}u \leqslant c_2\|\theta\|^2.$$

Proof It is a basic property of B-splines ([16] and [17]). $\square$

Let

$$\widetilde{\boldsymbol{\beta}}_n^w = \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})}\boldsymbol{X}\Big)^{-1}\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})}\boldsymbol{Y}$$

and

$$\widetilde{\boldsymbol{\theta}}_n^w = \big\{\boldsymbol{Z}^{*\top}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{Z}^*\big\}^{-1}\boldsymbol{Z}^{*\top}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})(\boldsymbol{Y}-\boldsymbol{X}\widetilde{\boldsymbol{\beta}}_n^w).$$

In addition, let $\widetilde{\alpha}_{jn}^w(u) = (\boldsymbol{\zeta}(u))^\top\widetilde{\boldsymbol{\theta}}_{jn}^w$. For $\widetilde{\boldsymbol{\beta}}_n^w$, $\widetilde{\boldsymbol{\theta}}_{jn}^w$ and $\widetilde{\alpha}_{jn}^w(\cdot)$ we have the following lemma.

Lemma 15 Suppose that Assumptions 1 to 6 hold, then

(i) $\sqrt{nT}(\widetilde{\boldsymbol{\beta}}_n^w - \boldsymbol{\beta}) \xrightarrow{\mathrm{d}} N(0, \boldsymbol{\Omega}_3^{-1})$, where $\boldsymbol{\Omega}_3$ is defined in Theorem 12.

(ii) $\|\widetilde{\boldsymbol{\theta}}_{jn}^w - \boldsymbol{\theta}_j\| = O_p(\sqrt{\kappa_n/n} + \kappa_n^{-2})$ for $j = 1, 2, \ldots, n$.

(iii) $\int_{u\in\mathscr{U}}[\widetilde{\alpha}_{jn}^w(u) - \alpha_j(u)]p(u)\mathrm{d}u = O_p(\kappa_n/n + \kappa_n^{-4})$.

Proof

(i)

$$\begin{aligned}\sqrt{nT}(\widetilde{\boldsymbol{\beta}}_n^w - \boldsymbol{\beta}) &= \sqrt{nT}\Big(\boldsymbol{X}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})}\boldsymbol{X}\Big)^{-1}\boldsymbol{X}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})}\boldsymbol{\nu}\\ &\quad+\sqrt{nT}\Big(\boldsymbol{X}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})}\mathbf{X}\Big)^{-1}\boldsymbol{X}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})}\boldsymbol{M}.\end{aligned}$$

Since $\boldsymbol{X}_{it} = \boldsymbol{\Pi}_{it} + \boldsymbol{H}_{it}(U_{it})$ with $\boldsymbol{H}_{it}(U_{it}) = (h_1(U_{it}), h_2(U_{it}), \ldots, h_p(U_{it}))^{\mathsf{T}}$ and $\boldsymbol{X} = \boldsymbol{\Pi} + \boldsymbol{H}$ with $\boldsymbol{X} = (\boldsymbol{X}_{11}, \boldsymbol{X}_{12}, \ldots, \boldsymbol{X}_{1T}, \boldsymbol{X}_{21}, \boldsymbol{X}_{22}, \ldots, \boldsymbol{X}_{2T}, \ldots, \boldsymbol{X}_{n1}, \boldsymbol{X}_{n2}, \ldots, \boldsymbol{X}_{nT})^{\mathsf{T}}$, $\boldsymbol{\Pi} = (\boldsymbol{\Pi}_{11}, \boldsymbol{\Pi}_{12}, \ldots, \boldsymbol{\Pi}_{1T}, \boldsymbol{\Pi}_{21}, \boldsymbol{\Pi}_{22}, \ldots, \boldsymbol{\Pi}_{2T}, \ldots, \boldsymbol{\Pi}_{n1}, \boldsymbol{\Pi}_{n2}, \ldots, \boldsymbol{\Pi}_{nT})^{\mathsf{T}}$ and $\boldsymbol{H} = (\boldsymbol{H}_{11}, \boldsymbol{H}_{12}, \ldots, \boldsymbol{H}_{1T}, \boldsymbol{H}_{21}, \boldsymbol{H}_{22}, \ldots, \boldsymbol{H}_{2T}, \ldots, \boldsymbol{H}_{n1}, \boldsymbol{H}_{n2}, \ldots, \boldsymbol{H}_{nT})^{\mathsf{T}}$, it is easy to see that $\boldsymbol{X}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})}\boldsymbol{\nu}$ can be decomposed as

$$\begin{aligned}&\boldsymbol{\Pi}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{\nu} - \boldsymbol{\Pi}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{Z}^*\big\{\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{Z}^*\big\}^{-1}\\ &\cdot\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{\nu} + \boldsymbol{H}\big\{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\\ &-\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{Z}^*\big\{\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{Z}^*\big\}^{-1}\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\big\}\boldsymbol{\nu}\\ =\;& J_1 - J_2 + J_3, \qquad \text{say,}\end{aligned}$$

where $\boldsymbol{\nu} = (\nu_{11}, \nu_{12}, \ldots, \nu_{1T}, \nu_{21}, \nu_{22}, \ldots, \nu_{2T}, \ldots, \nu_{n1}, \nu_{n2}, \ldots, \nu_{nT})^{\mathsf{T}}$. Applying Lemma 14, it holds that

$$\begin{aligned}\frac{1}{n}\|J_2\|^2 \leqslant{}& O_p(n^{-3})\cdot\Big\{\lambda_{\min}\Big(\frac{1}{n}\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{Z}^*\Big)\Big\}^{-2}\\ &\cdot\big\{\boldsymbol{\nu}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{Z}^*\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{\nu}\big\}\\ &\cdot\big\|\boldsymbol{\Pi}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{Z}^*\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{\Pi}\big\|\\ &\cdot\big\|\boldsymbol{\Pi}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{Z}^*\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{\Pi}\big\|\\ &\cdot\mathsf{E}\big[\mathrm{tr}\big\{\boldsymbol{Z}^*\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\mathsf{E}_{\mathscr{D}}(\boldsymbol{\Pi}\boldsymbol{\Pi}^{\mathsf{T}})\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\big\}\big]\\ ={}& O_p(n^{-3})\cdot\{\mathsf{E}\|\boldsymbol{Z}^*\|^2\}^2 = o_p(1).\end{aligned}$$

This implies that $J_2 = o_p(n^{1/2})$. For J_3, also by Lemma 14 we have $J_3 = o_p(n^{1/2})$. Therefore,

$$\boldsymbol{X}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})}\boldsymbol{\nu} = \boldsymbol{\Pi}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{\nu} + o_p(n^{1/2}).$$

In addition,

$$\boldsymbol{X}^{\top}\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})}\boldsymbol{M} = \boldsymbol{\Pi}^{\top}\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})}\boldsymbol{M} + \boldsymbol{H}^{\top}\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})}\boldsymbol{M}$$
$$= J_4 + J_5.$$

By the same argument, $J_4 = o_p(n^{1/2})$ and $J_5 = o_p(n^{1/2})$. Therefore, the proof of (i) is complete.

(ii) From the definition of $\widetilde{\boldsymbol{\theta}}_n^w$ it holds that

$$\begin{aligned}
\widetilde{\boldsymbol{\theta}}_n^w - \boldsymbol{\theta} = &\left\{\boldsymbol{Z}^{*\top}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{Z}^*\right\}^{-1}\boldsymbol{Z}^{*\top}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{\nu}\\
&+\left\{\boldsymbol{Z}^{*\top}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{Z}^*\right\}^{-1}\boldsymbol{Z}^{*\top}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{M} - \boldsymbol{\theta}\\
&+\left\{\boldsymbol{Z}^{*\top}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{Z}^*\right\}^{-1}\boldsymbol{Z}^{*\top}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{X}(\boldsymbol{\beta} - \widetilde{\boldsymbol{\beta}}_n^w)\\
= &J_1 + J_2 + J_3,
\end{aligned}$$

where $\boldsymbol{M} = (\boldsymbol{Z}_{11}^{\top}\boldsymbol{\alpha}(U_{11}), \boldsymbol{Z}_{12}^{\top}\boldsymbol{\alpha}(U_{12}), \ldots, \boldsymbol{Z}_{1T}^{\top}\boldsymbol{\alpha}(U_{1T}), \boldsymbol{Z}_{21}^{\top}\boldsymbol{\alpha}(U_{21}), \boldsymbol{Z}_{22}^{\top}\boldsymbol{\alpha}(U_{22}), \ldots, \boldsymbol{Z}_{2T}^{\top}\boldsymbol{\alpha}(U_{2T}), \ldots, \boldsymbol{Z}_{n1}^{\top}\boldsymbol{\alpha}(U_{n1}), \boldsymbol{Z}_{n2}^{\top}\boldsymbol{\alpha}(U_{n2}), \ldots, \boldsymbol{Z}_{nT}^{\top}\boldsymbol{\alpha}(U_{nT}))^{\top}$. It is easy to show that $\|J_1\| = O_p(\sqrt{\kappa_n/n})$. In addition, $\|J_2\| = O_p(\kappa_n^{-2})$, $\|J_3\| = O_p(\sqrt{\kappa_n/n})$. Together, (ii) holds.

(iii) By the definition of $\widetilde{\alpha}_{jn}^w(u)$ it holds that

$$\begin{aligned}
&\int_{u\in\mathscr{U}}\{\widetilde{\alpha}_{jn}^w(u) - \alpha_j(u)\}^2 p(u)\mathrm{d}u\\
=&\int_{u\in\mathscr{U}}[(\widetilde{\alpha}_{jn}^w(u) - (\boldsymbol{\zeta}(u))^{\top}\boldsymbol{\theta}_j) - \{\alpha_j(u) - (\boldsymbol{\zeta}(u))^{\top}\boldsymbol{\theta}_j\}]^2 p(u)\mathrm{d}u\\
\leqslant& 2\lambda_{\max}\Big(\int_{u\in\mathscr{U}}\boldsymbol{\zeta}(u)(\boldsymbol{\zeta}(u))^{\top}p(u)\mathrm{d}u\Big)\cdot\|\widetilde{\boldsymbol{\theta}}_{jn} - \boldsymbol{\theta}_j\|^2 + 2\int_{u\in\mathscr{U}}(\alpha_j(u) - (\boldsymbol{\zeta}(u))^{\top}\boldsymbol{\theta}_j)^2 p(u)\mathrm{d}u\\
=&O_p(\kappa_n^{-4} + \kappa_n/n) + O_p(\kappa_n^{-4}),
\end{aligned}$$

so (iii) holds. □

Proof of Theorems 7 and 8 Theorems 7 and 8 can be proved in the same way we prove Lemma 15. We here omit the details. □

Proof of Theorem 10 Let

$$\boldsymbol{D}_0 = \frac{1}{n\{T-(s+1)\}}\sum_{i=1}^{n}\sum_{t=1}^{T-(m+1)}(\varepsilon_{i,t+(m-1)},\ldots,\varepsilon_{it})^{\top}(\varepsilon_{i,t+(m-1)},\ldots,\varepsilon_{it}),$$

and $\Delta_{it} = \boldsymbol{X}_{it}^{\top}(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}_n) - Z_{it1}(\alpha_1(U_{it}) - \widehat{\alpha}_{1n}(U_{it})) - \cdots - Z_{itq}(\alpha_q(U_{it}) - \widehat{\alpha}_{qn}(U_{it}))$. Since

$$\widehat{\varepsilon}_{it} = Y_{it} - \boldsymbol{X}_{it}^{\top}\widehat{\boldsymbol{\beta}}_n - Z_{it1}\widehat{\alpha}_{1n}(U_{it}) - \cdots - Z_{itq}\widehat{\alpha}_{qn}(U_{it}), \qquad i = 1,2,\ldots,n,\ t = 1,2,\ldots,T, \quad (9)$$

we have

$$\begin{aligned}\widehat{\boldsymbol{D}}_0 &= \boldsymbol{D}_0 + \frac{1}{n\{T-(m+1)\}}\sum_{i=1}^{n}\sum_{t=1}^{T-(m+1)}(\Delta_{i,t+(m-1)},\ldots,\Delta_{it})^{\mathsf{T}}(\Delta_{i,t+(m-1)},\ldots,\Delta_{it})\\ &\quad+\frac{1}{n\{T-(m+1)\}}\sum_{i=1}^{n}\sum_{t=1}^{T-(m+1)}(\varepsilon_{i,t+(m-1)},\ldots,\varepsilon_{it})^{\mathsf{T}}(\Delta_{i,t+(m-1)},\ldots,\Delta_{it})\\ &\quad+\frac{1}{n\{T-(m+1)\}}\sum_{i=1}^{n}\sum_{t=1}^{T-(m+1)}(\Delta_{i,t+(m-1)},\ldots,\Delta_{it})^{\mathsf{T}}(\varepsilon_{i,t+(m-1)},\ldots,\varepsilon_{it})\\ &= \boldsymbol{D}_0 + J_1 + J_2 + J_3.\end{aligned}$$

According to Theorems 7 and 8, $J_1 = o_p(n^{-1/2})$. In addition,

$$\begin{aligned}J_2 &= \frac{1}{n\{T-(m+1)\}}\sum_{i=1}^{n}\sum_{t=1}^{T-(s+1)}(\varepsilon_{i,t+(m-1)},\ldots,\varepsilon_{it})^{\mathsf{T}}\big(\boldsymbol{X}_{i,t+(m-1)}^{\mathsf{T}}(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}_n),\ldots,\boldsymbol{X}_{it}^{\mathsf{T}}(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}_n)\big)\\ &\quad+\frac{1}{n\{T-(m+1)\}}\sum_{i=1}^{n}\sum_{t=1}^{T-(m+1)}(\varepsilon_{i,t+(m-1)},\ldots,\varepsilon_{it})^{\mathsf{T}}\\ &\quad\cdot\Big(\sum_{j=1}^{q}Z_{i,t+(m-1),j}(\alpha_j(U_{i,t+(m-1)})-\widehat{\alpha}_{jn}(U_{i,t+(m-1)})),\ldots,\sum_{j=1}^{q}Z_{itj}(\alpha_j(U_{it})-\widehat{\alpha}_{jn}(U_{it}))\Big)\\ &= J_{21}+J_{22}.\end{aligned}$$

By Theorem 7, for $0 \leqslant d \leqslant m-1$,

$$\frac{1}{n\{T-(m+1)\}}\sum_{i=1}^{n}\sum_{t=1}^{T-(m+1)}\boldsymbol{X}_{i,t+d}(\varepsilon_{i,t+(m-1)},\ldots,\varepsilon_{it})^{\mathsf{T}} = O_p(n^{-1/2}).$$

Therefore, Theorem 7 leads to $J_{21} = o_p(n^{-1/2})$. Furthermore,

$$\begin{aligned}&\sum_{j=1}^{q}Z_{itj}(\alpha_j(U_{it})-\widehat{\alpha}_{jn}(U_{it}))\\ &= \sum_{j=1}^{q}Z_{itj}(\alpha_j(U_{it})-(\boldsymbol{\zeta}(U_{it}))^{\mathsf{T}}\widehat{\boldsymbol{\theta}}_{jn})\\ &= \sum_{j=1}^{q}Z_{itj}\{\alpha_j(U_{it})-(\boldsymbol{\zeta}(U_{it}))^{\mathsf{T}}(\mathbf{0}_{\kappa_n\times(j-1)},\boldsymbol{I}_{\kappa_n},\mathbf{0}_{\kappa_n\times(q-j)})(\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}\boldsymbol{Z}^*)^{-1}\boldsymbol{Z}^{\mathsf{T}^*}\boldsymbol{M}_{\boldsymbol{B}}\boldsymbol{M}\}\\ &\quad-\sum_{j=1}^{q}Z_{itj}(\boldsymbol{\zeta}(U_{it}))^{\mathsf{T}}(\mathbf{0}_{\kappa_n\times(j-1)},\boldsymbol{I}_{\kappa_n},\mathbf{0}_{\kappa_n\times(q-j)})(\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}\boldsymbol{Z}^*)^{-1}\boldsymbol{Z}^{\mathsf{T}^*}\boldsymbol{M}_{\boldsymbol{B}}\boldsymbol{\varepsilon}\\ &\quad-\sum_{j=1}^{q}Z_{itj}(\boldsymbol{\zeta}(U_{it}))^{\mathsf{T}}(\mathbf{0}_{\kappa_n\times(j-1)},\boldsymbol{I}_{\kappa_n},\mathbf{0}_{\kappa_n\times(q-j)})(\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}\boldsymbol{Z}^*)^{-1}\boldsymbol{Z}^{\mathsf{T}^*}\boldsymbol{M}_{\boldsymbol{B}}\boldsymbol{X}(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}_n).\end{aligned}$$

Hence, set

$$A = \frac{1}{n\{T-(m-1)\}}\sum_{i=1}^{n}\sum_{t=1}^{T-(m-1)}\varepsilon_{it},$$

then

$$\begin{aligned}
J_{22} = & A\sum_{j=1}^{q} Z_{i,t+m,j}\{\alpha_j(U_{i,t+m}) - (\boldsymbol{\zeta}(U_{i,t+m}))^{\mathsf{T}}(\mathbf{0}_{\kappa_n\times(j-1)}, \boldsymbol{I}_{\kappa_n}, \mathbf{0}_{\kappa_n\times(q-j)}) \\
& \cdot (\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}\boldsymbol{Z}^{*})^{-1}\boldsymbol{Z}^{\mathsf{T}^*}\boldsymbol{M}_{\boldsymbol{B}}\boldsymbol{M}\} \\
& - A\sum_{j=1}^{q} Z_{i,t+m,j}(\boldsymbol{\zeta}(U_{i,t+m}))^{\mathsf{T}}(\mathbf{0}_{\kappa_n\times(j-1)}, \boldsymbol{I}_{\kappa_n}, \mathbf{0}_{\kappa_n\times(q-j)})(\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}\boldsymbol{Z}^{*})^{-1}\boldsymbol{Z}^{\mathsf{T}^*}\boldsymbol{M}_{\boldsymbol{B}}\boldsymbol{\varepsilon} \\
& - A\sum_{j=1}^{q} Z_{i,t+m,j}(\boldsymbol{\zeta}(U_{i,t+m}))^{\mathsf{T}}(\mathbf{0}_{\kappa_n\times(j-1)}, \boldsymbol{I}_{\kappa_n}, \mathbf{0}_{\kappa_n\times(q-j)}) \\
& \cdot (\boldsymbol{Z}^{*\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}\boldsymbol{Z}^{*})^{-1}\boldsymbol{Z}^{\mathsf{T}^*}\boldsymbol{M}_{\boldsymbol{B}}\boldsymbol{X}(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}_n) \\
= & J_{221} + J_{222} + J_{223}.
\end{aligned}$$

Further, $J_{221} = o_p(n^{-1/2})$, $J_{222} = o_p(n^{-1/2})$, $J_{223} = O(n^{-1}\kappa_n)$. So $J_{22} = o_p(n^{-1/2})$. As a result, $J_2 = o_p(n^{-1/2})$. Following the same line, $J_3 = O_p(n^{-1/2})$. This implies that $\widehat{\boldsymbol{Q}}_0 = \boldsymbol{Q}_0 + o_p(n^{-1/2})$. By the same argument, $\widehat{\boldsymbol{Q}}_1 = \boldsymbol{Q}_1 + o_p(n^{-1/2})$, $\widehat{\boldsymbol{Q}}_2 = \boldsymbol{Q}_2 + o_p(n^{-1/2})$ and $\widehat{\boldsymbol{Q}}_3 = \boldsymbol{Q}_3 + o_p(n^{-1/2})$. As a result, for $\widehat{\boldsymbol{\rho}}_n = (\widehat{\rho}_{1n}, \widehat{\rho}_{2n}, \ldots, \widehat{\rho}_{mn})^{\mathsf{T}}$,

$$\begin{aligned}
\widehat{\boldsymbol{\rho}}_n - \boldsymbol{\rho} &= \{(\widehat{\boldsymbol{Q}}_2 - \widehat{\boldsymbol{Q}}_3)^{-1}(\widehat{\boldsymbol{Q}}_0 - \widehat{\boldsymbol{Q}}_1) - \boldsymbol{\rho}\} \\
&= \Big[\Big\{\sum_{i=1}^{n}\sum_{t=1}^{T-(m-1)}((\varepsilon_{i,t+(m-1)}, \ldots, \varepsilon_{it})^{\mathsf{T}}(\varepsilon_{i,t+(m-1)}, \ldots, \varepsilon_{it}) \\
&\quad - (\varepsilon_{i,t+(m-1)}, \ldots, \varepsilon_{it})^{\mathsf{T}}(\varepsilon_{i,t+m}, \ldots, \varepsilon_{i,t+1}))\Big\}^{-1} \\
&\quad \cdot \Big\{\sum_{i=1}^{n}\sum_{t=1}^{T-(m-1)}((\varepsilon_{i,t+(m-1)}, \ldots, \varepsilon_{it})^{\mathsf{T}}\varepsilon_{i,t+m} \\
&\quad - (\varepsilon_{i,t+(m-1)}, \ldots, \varepsilon_{it})^{\mathsf{T}}\varepsilon_{i,t+(m+1)})\Big\} - \rho\Big] + o_p(1).
\end{aligned}$$

According to (2) we have $\varepsilon_{i,t+m} = \rho_1\varepsilon_{i,t+(m-1)} + \cdots + \rho_m\varepsilon_{it} + (1 - \rho_1 - \cdots - \rho_m)\mu_i + e_{i,t+m}$, and $\varepsilon_{i,t+(m+1)} = \rho_1\varepsilon_{i,t+m} + \cdots + \rho_m\varepsilon_{i,t+1} + (1 - \rho_1 - \cdots - \rho_m)\mu_i + e_{i,t+(m+1)}$. Therefore,

$$\begin{aligned}
\widehat{\boldsymbol{\rho}}_n - \boldsymbol{\rho} = & \Big\{\sum_{i=1}^{n}\sum_{t=1}^{T-(m+1)}((\varepsilon_{i,t+(m-1)}, \ldots, \varepsilon_{it})^{\mathsf{T}}(\varepsilon_{i,t+(m-1)}, \ldots, \varepsilon_{it}) \\
& - (\varepsilon_{i,t+(m-1)}, \ldots, \varepsilon_{it})^{\mathsf{T}}(\varepsilon_{i,t+m}, \ldots, \varepsilon_{i,t+1}))\Big\}^{-1} \\
& \cdot \Big[\sum_{i=1}^{n}\sum_{t=1}^{T-(m+1)}\{(\varepsilon_{i,t+(m-1)}, \ldots, \varepsilon_{it})^{\mathsf{T}}(e_{i,t+m} + (1 - \rho_1 - \cdots - \rho_m)\mu_i) \\
& - (\varepsilon_{i,t+(m-1)}, \ldots, \varepsilon_{it})^{\mathsf{T}}(e_{i,t+(m+1)} + (1 - \rho_1 - \cdots - \rho_m)\mu_i)\}\Big] + o_p(1).
\end{aligned}$$

Therefore the result of $\widehat{\boldsymbol{\rho}}_n$ holds. By the same argument, we can show the result of $\widehat{\sigma}^2_{en}$ holds. $\square$

Proof of Theorems 11, 12 and 13 From to the definitions of $\widehat{\boldsymbol{\beta}}_n^w$ and $\widetilde{\boldsymbol{\beta}}_n^w$ and the fact that $a_1b_1 - a_2b_2 = (a_1 - b_1)(a_2 - b_2) + (a_1 - b_1)b_2 + b_1(a_2 - b_2)$ it holds that

$$
\begin{aligned}
&\widehat{\boldsymbol{\beta}}_n^w - \boldsymbol{\beta} \\
={}& \widetilde{\boldsymbol{\beta}}_n^w - \boldsymbol{\beta} + \Big\{ \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} \boldsymbol{X} \Big)^{-1} - \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{-1/2})} \boldsymbol{X} \Big)^{-1} \Big\} \\
&\cdot \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} - \boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{-1/2})} \Big) \boldsymbol{M} \\
&+ \Big\{ \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} \boldsymbol{X} \Big)^{-1} - \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{-1/2})} \boldsymbol{X} \Big)^{-1} \Big\} \\
&\cdot \boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} \boldsymbol{M} \\
&+ \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} \boldsymbol{X} \Big)^{-1} \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} - \boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{-1/2})} \Big) \boldsymbol{M} \\
&+ \Big\{ \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} \boldsymbol{X} \Big)^{-1} - \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{-1/2})} \boldsymbol{X} \Big)^{-1} \Big\} \\
&\cdot \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} - \boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{-1/2})} \Big) \boldsymbol{\nu} \\
&+ \Big\{ \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} \boldsymbol{X} \Big)^{-1} - \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{-1/2})} \boldsymbol{X} \Big)^{-1} \Big\} \\
&\cdot \boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} \boldsymbol{\nu} \\
&+ \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} \boldsymbol{X} \Big)^{-1} \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} - \boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{-1/2})} \Big) \boldsymbol{\nu}.
\end{aligned}
$$

By Lemma 15 and the fact that $(\boldsymbol{A} + a\boldsymbol{B})^{-1} = \boldsymbol{A}^{-1} - a\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1} + O(a^2)$ as $a \to 0$, to prove Theorems 11, 12 and 13, we only need to prove

$$
\frac{1}{nT} \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} \boldsymbol{X} - \boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{-1/2})} \boldsymbol{X} \Big) = O_p(n^{-1/2}), \tag{10}
$$

$$
\frac{1}{nT} \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} - \boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{-1/2})} \Big) \boldsymbol{M} = o_p(n^{-1/2}), \tag{11}
$$

$$
\frac{1}{nT} \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} - \boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{-1/2})} \Big) \boldsymbol{\nu} = o_p(n^{-1/2}), \tag{12}
$$

$$
\frac{1}{nT} \boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{-1/2})} \boldsymbol{X} = O_p(1), \tag{13}
$$

$$
\frac{1}{nT} \boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} \boldsymbol{M} = o_p(n^{-1/2}), \quad \frac{1}{nT} \boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} \boldsymbol{\nu} = O_p(n^{-1/2}). \tag{14}
$$

According to the proof of Lemma 15 and the root-n consistency of $\widehat{\boldsymbol{\Sigma}}$ it holds that

$$
\frac{1}{nT} \Big(\boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_n^{-1/2})} \boldsymbol{X} - \boldsymbol{X}^\top \boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{-1/2})} \boldsymbol{X} \Big)
$$

$$
\begin{aligned}
&= \frac{1}{nT}\big\{\boldsymbol{\Pi}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n\otimes\widehat{\boldsymbol{\Sigma}}_n^{-1/2})\boldsymbol{\Pi}-\boldsymbol{\Pi}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{\Pi}\big\}+o_p(n^{-1/2})\\
&= O_p(n^{-1/2}).
\end{aligned}
$$

This implies that (10) holds. By the same argument, we can show that (11) holds. Moreover, from the proof of Lemma 15 we know that (13) and (14) hold as well. Therefore, to complete the proof we just need to prove (12). By the proof of Lemma 14 again and the root-n consistency of $\widehat{\boldsymbol{\Sigma}}_n$, we have

$$
\begin{aligned}
&\frac{1}{nT}\Big(\boldsymbol{X}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n\otimes\widehat{\boldsymbol{\Sigma}}_n^{-1/2})}\boldsymbol{\nu}-\boldsymbol{X}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{Z}^*}^{\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})}\boldsymbol{\nu}\Big)\\
&= \frac{1}{nT}\big\{\boldsymbol{\Pi}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\widehat{\boldsymbol{\Sigma}}_n^{-1}}(\boldsymbol{I}_n\otimes\widehat{\boldsymbol{\Sigma}}_n^{-1/2})\boldsymbol{\nu}-\boldsymbol{\Pi}^{\mathsf{T}}\boldsymbol{M}_{\boldsymbol{B}}^{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{I}_n\otimes\boldsymbol{\Sigma}^{-1/2})\boldsymbol{\nu}\big\}+o_p(n^{-1/2})\\
&= O_p(n^{-1/2}).
\end{aligned}
$$

This shows that (12) holds. Thus, the proof is completed. □

References

[1] DIGGLE P J, LIANG K Y, ZEGER S L. *Analysis of Longitudinal Data* [M]. New York: Oxford University Press, 1994.

[2] AHN S C, LEE Y H, SCHMIDT P. GMM estimation of linear panel data models with time-varying individual effects [J]. *J Econometrics*, 2001, **101(2)**: 219–255.

[3] VERBEKE G, MOLENBERGHS G. *Linear Mixed Models for Longitudinal Data* [M]. 2000th ed. New York: Springer-Verlag, 2009.

[4] BALTAGI B H. *Econometric Analysis of Panel Data* [M]. 4th ed. New York: Wiley, 2008.

[5] FAN J Q, HUANG T, LI R Z. Analysis of longitudinal data with semiparametric estimation of covariance function [J]. *J Amer Statist Assoc*, 2007, **102(478)**: 632–641.

[6] BALTAGI B H, LI D. Series estimation of partially linear panel data models with fixed effects [J]. *Ann Econ Financ*, 2002, **3(1)**: 103–116.

[7] FAN J Q, PENG H, HUANG T. Semilinear high-dimensional model for normalization of microarray data: a theoretical analysis and partial consistency (with discussion) [J]. *J Amer Statist Assoc*, 2005, **100(471)**: 781–813.

[8] SU L J, ULLAH A. Profile likelihood estimation of partially linear panel data models with fixed effects [J]. *Econom Lett*, 2006, **92(1)**: 75–81.

[9] HENDERSON D J, CARROLL R J, LI Q. Nonparametric estimation and testing of fixed effects panel data models [J]. *J Econometrics*, 2008, **144(1)**: 257–275.

[10] YOU J H, ZHOU X, ZHOU Y. Statistical inference for panel data semiparametric partially linear regression models with heteroscedastic errors [J]. *J Multivariate Anal*, 2010, **101(5)**: 1079–1101.

[11] ZHANG J H, FENG S Y, LI G R, et al. Empirical likelihood inference for partially linear panel data models with fixed effects [J]. *Econom Lett*, 2011, **113(2)**: 165–167.

[12] SUN Y G, CARROLL R J, LI D D. Semiparametric estimation of fixed-effects panel data varying coefficient models [M] // LI Q, RACINE J S. *Nonparametric Econometric Methods* (*Advances in Econometrics, Volume 25*). United Kingdom: Emerald Group Publishing Limited, 2009: 101–129.

[13] BECKER G S. A theory of competition among pressure groups for political influence [J]. *Quart J Econom*, 1983, **98(3)**: 371–400.

[14] NICKELL S. Biases in dynamic models with fixed effects [J]. *Econometrica*, 1981, **49(6)**: 1417–1426.

[15] BALTAGI B H, LI Q. Monte Carlo results on pure and pretest estimators of an error component model with autocorrelated disturbances [J]. *Ann Économ Statist*, 1997, **48**: 69–82.

[16] DE BOOR C. *A Practical Guide to Splines* [M]. New York: Springer-Verlag, 1978.

[17] SCHUMAKER L L. *Spline Functions: Basic Theory* [M]. New York: Wiley, 1981.

[18] RICE J A, WU C O. Nonparametric mixed effects models for unequally sampled noisy curves [J]. *Biometrics*, 2001, **57(1)**: 253–259.

[19] HUANG J Z, WU C O, ZHOU L. Polynomial spline estimation and inference for varying coefficient models with longitudinal data [J]. *Statist Sinica*, 2004, **14(3)**: 763–788.

[20] DUFFY J, PAPAGEORGIOU C. A cross-country empirical investigation of the aggregate production function specification [J]. *J Economic Growth*, 2000, **5(1)**: 87–120.

[21] BHARGAVA A, FRANZINI L, NARENDRANATHAN W. Serial correlation and the fixed effects model [J]. *Rev Econom Stud*, 1982, **49(4)**: 533–549.

[22] SOLON G. Estimating autocorrelations in fixed-effects models (technical working paper no. 32) [R]. Cambridge, Mass.: National Bureau of Eonomic Research, 1984.

固定效应面板数据部分线性模型的加权截面 LSDV 估计

朱能辉 李 肖

(厦门理工学院应用数学学院, 厦门, 361024) (上海财经大学统计与管理学院, 上海, 200433)

施雅丰

(宁波工程学院理学院, 宁波, 315211)

摘 要: 本文考虑误差为自回归过程的固定效应面板数据部分线性回归模型的估计. 对于固定效应短时间序列面板数据, 通常使用的自回归误差结构拟合方法不能得到一个一致的自回归系数估计量. 因此本文提出一个替代估计并证明所提出的自回归系数估计是一致的, 且该方法在任何阶的自回归误差下都是可行的. 进一步, 通过结合 B 样条近似, 截面最小二乘虚拟变量 (LSDV) 技术和自回归误差结构的一致估计, 本文使用加权截面 LSDV 估计参数部分和加权 B 样条 (BS) 估计非参数部分, 所得到的加权截面 LSDV 估计量被证明是渐近正态的, 且比可忽略误差的自回归结构模型更渐近有效. 另外, 加权 BS 估计量被推导出具有渐近偏差和渐近正态性. 模拟研究和实际例子相应地说明了所估计程序的有限样本性.

关键词: 面板部分线性变系数模型; 固定效应; 截面最小二乘虚拟变量法; 半参数; 自回归过程

中图分类号: O212.4