Economics Bulletin

Volume 34, Issue 3

Nonparametric estimation of functional-coefficient partially linear dynamic panel data model with fixed effects

Kien C Tran
University of Lethrbidge

Abstract

In this paper we propose a stationary nonlinear dynamic functional coefficient panel data models with fixed effects and develops semiparametric estimation procedure using series approximation. Convergence rate and asymptotic distribution of the proposed series estimators are derived in which asymptotic biases are present. Bias corrections are developed using a heteroskedasticity and autocorrelation consistent (HAC) type estimator.

This paper was written while the author was a visiting scholar at the Department of Economics, Athens University of Economics and Business and The University of Macedonia, Greece. I would like to thank Bruce Hansen and Mike Tsionas for useful comments and suggestions on the earlier draft on this paper. The usual caveats apply.
Citation: Kien C Tran, (2014) "Nonparametric estimation of functional-coefficient partially linear dynamic panel data model with fixed effects", *Economics Bulletin*, Vol. 34 No. 3 pp. 1751-1761.
Contact: Kien C Tran - kien.tran@uleth.ca.
Submitted: April 28, 2014. **Published:** August 06, 2014.

1. Introduction

There is a large and growing literature on the fixed effect panel data analysis in econometrics, especially in semiparametric and nonparametric modeling, but very little attention has been paid to the nonparametric estimation in dynamic panel models with fixed effects. As Lee (2007) points out, one possible explanation for this is the difficult of treating autoregressive structure and the individual specific fixed effects simultaneously in the context of semiparametric and nonparametric estimation. In this paper, we propose a partially linear varying coefficient dynamic panel data model with fixed effect which is an extension and a generalization of Lee's (2007) nonparametric and partially linear fixed effects dynamic panel models. To estimate the functional coefficients of the model, we consider a within group type of series estimator and derive its convergence rate and its asymptotic normality. Our asymptotic results indicated that the within group series estimator is asymptotically biased and we also suggest a bias-corrected within group estimator.

Section 2 of this paper gives the specification of the model and derives the within group series estimator. Convergence rate and asymptotic normality of the proposed estimator are examined in Section 3. Section 4 concludes the paper. The mathematical proofs are gathered in the Appendix.

2. The Model and Within Group Series Estimation

We consider the following functional partially linear varying coefficient dynamic panel data model:

$$y_{it} = \theta(y_{it-1}) + x_{it}^{'}\beta(z_{it}) + \mu_i + e_{it},\ \ i = 1,\ldots,n;\ t = 1,\ldots,T \tag{1}$$

where $y_{it} \in \mathcal{Y} \subset \mathbb{R}$, $\theta(.) \in \Theta$ is unknown and Θ is a specified class function from Υ to $\mathbb{R}$; $x_{it} \in \mathbb{R}^d$ (excluding a constant), and $z_{it} \in \mathbb{R}^m$ is a vector of exogenous regressors, μ_i is individual specific effects and e_{it} is the usual random errors. We assume that $E(e_{it} \mid y_{it-1},\ldots,y_{i0},x_{it},z_{it},\mu_i) = 0$ and the realization of the initial values y_{i0} are observed. The individual specific effects μ_i is assumed to have finite variance and to be independent of u_{it} for all i and t. However, we allow μ_i to be correlated with y_{it-1}, x_{it} and z_{it} with unknown correlation structure.

Following Lee (2007, 2013), we assume that e_{it} is independently identically distributed with zero mean, finite variance and possess higher finite moments. It is further assumed that for each i, e_{it} is independent of $\{y_{it-s}\}_{s\leq t-1}$ and y_{it} is independent across i. Finally, conditional on μ_i, for each i, the process $\{y_{it}\}$ in (1) is geometric, stationary and β-mixing with exponential decay.

Model (1) includes the fixed effect functional coefficient panel data model of Sun, Carroll and Li (2009) when $\theta(.) = 0$, nonparametric fixed effect of Lee (2013) when $\beta(.) = 0$ and in a special case where $d = 1$ and $x_{it} = 1$ for all i and t, it reduces to the nonparametric additive model of Baglan (2010). The model is also a generalization of the fixed effects partially linear model of Lee (2007) obtain by replacing the parameters in the fixed effect partially linear model with some functions of covariates. Let $\overline{s}_{i.} = T^{-1}\sum_{t=1}^{T} s_{it}$ denotes the within group sample average of a random variable s_{it}, and let $\tilde{s}_{it} = s_{it} - \overline{s}_{i.}$. Also, let $g(x_{it}, z_{it}) = x_{it}^{'}\beta(z_{it})$, then by employing standard within transformation of (1) to eliminate the fixed effect, we obtain:

$$\begin{aligned} \tilde{y}_{it} &= \tilde{\theta}(y_{it-1}) + x_{it}^{'}\beta(z_{it}) - T^{-1}\sum_{s=1}^{T} x_{is}^{'}\beta(z_{is}) + \tilde{e}_{it} \\ &= \tilde{\theta}(y_{it-1}) + \tilde{g}(x_{it}, z_{it}) + \tilde{e}_{it} \end{aligned} \tag{2}$$

where $\tilde{\theta}(y_{it-1}) = \theta(y_{it-1}) - T^{-1}\sum_{s=1}^{T}\theta(y_{is-1})$ and $\tilde{g}(x_{it}, z_{it}) = x_{it}^{'}\beta(z_{it}) - T^{-1}\sum_{s=1}^{T} x_{is}^{'}\beta(z_{is})$. We define the general function $\xi(w_{it})$ to be the additive within transformation class of functions $\mathcal{G}$ if $\xi(w_{it}) = f(w_{it}) - T^{-1}\sum_{s=1}^{T} f(w_{is})$, $f(.)$ is twice differentiable in the interior of its support $\mathcal{W}$ which is a compact subset of $\mathbb{R}^{\nu}$ and $E\{f^2(w_{it})\} < \infty$. We will use the series $p^K(y)$ of $(K \times 1)$ dimension to approximate $\theta(y_{it-1})$ by $\theta(y_{it-1}) \approx p^K(y_{it-1})^{'}\alpha$, a linear combination of K known base functions, where $p^K(y_{it-1}) = [p_1(y_{it-1}), \ldots, p_K(y_{it-1})]^{'}$ is a $(K \times 1)$ vector of base function, and $\alpha = (\alpha_1, \ldots, \alpha_K)^{'}$ is a $(K \times 1)$ vector of unknown parameters. Similarly, for $l = 1, \ldots, d$, we approximate the varying coefficient $\beta_l(z_{it})$ by $\beta_l(z_{it}) \approx q_l^{r_l}(z_{it})^{'}\gamma_l^{r_l}$, a linear combination of r_l known base functions, where $q_l^{r_l}(z_{it}) = [q_{l1}(z_{it}), \ldots, q_{lr_l}(z_{it})]^{'}$ is a vector of $(r_l \times 1)$ vector of base function and $\gamma_l^{r_l} = (\gamma_{l1}, \ldots, \gamma_{lr_l})^{'}$ is a vector of $(r_l \times 1)$ vector of unknown parameters. Define the $(R \times 1)$ matrices

$$q^R(x_{it}, z_{it}) = [x_{it,1}q_1^{r_1}(z_{it})^{'}, \ldots, x_{it,d}q_d^{r_d}(z_{it})^{'}]^{'}$$

and $\gamma = (\gamma_1^{r_1^{'}}, \ldots, \gamma_d^{r_d^{'}})^{'}$, where $R = \sum_{l=1}^{d} r_l$. Thus, we use a linear combination of R functions, $q^R(x_{it}, z_{it})^{'}\gamma$, to approximate $g(x_{it}, z_{it}) = x_{it}^{'}\beta(z_{it})$.

Note that the approximating functions $p^K(y)$ and $q^R(x,z)$ have the following properties: (a) $\tilde{p}^K(y_{it-1}) \equiv p^K(y_{it-1}) - T^{-1}\sum_{s=1}^{T} p^K(y_{is-1}) \in \mathcal{G}$ and $\tilde{q}^R(x_{it}, z_{it}) = q^R(x_{it}, z_{it}) - T^{-1}\sum_{s=1}^{T} q^R(x_{is}, z_{is}) \in \mathcal{G}$; (b) as each of K and R grow, respectively, there is a linear combination of $\tilde{p}_k^K(y_{it-1})$ and a linear combination of $\tilde{q}_r^R(x_{it}, z_{it})$ that can approximate any function in $\mathcal{G}$ arbitrarily well in the sense of mean square error. Consequently, we can approximate $\tilde{\theta}(y_{it-1})$ by $\tilde{\theta}(y_{it-1}) \approx \tilde{p}^K(y_{it-1})^{'}\alpha$ and $\tilde{g}(x_{it}, z_{it})$ by $\tilde{g}(x_{it}, z_{it}) \approx \tilde{q}^R(x_{it}, z_{it})^{'}\gamma$ where $\tilde{p}^K(y_{it-1})$ and $\tilde{q}^R(x_{it}, z_{it})$ are the transformed basis functions. Hence we can rewrite (2) as

$$\tilde{y}_{it} = \tilde{p}^K(y_{it-1})^{'}\alpha + \tilde{q}^R(x_{it}, z_{it})^{'}\gamma + \tilde{\varepsilon}_{it} \tag{3}$$

where $\tilde{\varepsilon}_{it} = \tilde{e}_{it} + \{\tilde{\theta}(y_{it-1}) - \tilde{p}^K(y_{it-1})'\alpha\} + \{\tilde{g}^R(x_{it}, z_{it}) - \tilde{q}^R(x_{it}, z_{it})'\gamma\}$. Using matrix notation, we define $(nT \times 1)$ vector $\tilde{Y} = (\tilde{y}_{11}, \ldots, \tilde{y}_{nT})'$, an $(nT \times 1)$ vector $\tilde{\varepsilon} = (\tilde{\varepsilon}_{11}, \ldots, \tilde{\varepsilon}_{nT})'$, an $(nT \times K)$ matrix $\tilde{P}_K = (\tilde{p}^K(y_{10}), \ldots, \tilde{p}^K(y_{nT-1}))'$ and an $(nT \times R)$ matrix $\tilde{Q}_R = (\tilde{q}^R(x_{11}, z_{11}), \ldots, \tilde{q}^R(x_{nT}, z_{nT}))'$. We also define $(nT \times nT)$ matrices $M_P = I_{nT} - \tilde{P}_K(\tilde{P}_K'\tilde{P}_K)^{-1}\tilde{P}_K'$ and $M_Q = I_{nT} - \tilde{Q}_R(\tilde{Q}_R'\tilde{Q}_R)^{-1}\tilde{Q}_R'$ assuming that both $(\tilde{P}_K'\tilde{P}_K)$ and $(\tilde{Q}_R'\tilde{Q}_R)$ are nonsingular matrices for large n and T. Thus, (3) can be written using vector-matrix notation as

$$\tilde{Y} = \tilde{P}_K\alpha + \tilde{Q}_R\gamma + \tilde{\varepsilon} \tag{4}$$

Let $\hat{\alpha}$ and $\hat{\gamma}$ denote the least squares estimators of α and γ respectively, obtaining by regressing $\tilde{Y}$ on $(\tilde{P}_K, \tilde{Q}_R)$ from (4). Using partition inverse results, the estimators $\hat{\alpha}$ and $\hat{\gamma}$ are given by

$$\hat{\alpha}_K = (\tilde{P}_K' M_Q \tilde{P}_K)^{-1}\tilde{P}_K' M_Q \tilde{Y} \tag{5}$$

$$\hat{\gamma}_R = (\tilde{Q}_R' M_P \tilde{Q}_R)^{-1}\tilde{Q}_R' M_P \tilde{Y} \tag{6}$$

where we use the subscripts K and R to denote that these estimators are dependent of the number of approximating functions. We then estimate $\theta(y)$ by $\hat{\theta}(y) = p^K(y)'\hat{\alpha}_K$ and $\beta_l(z)$ by $\hat{\beta}_l(z) = q_l^{r_l}(z)'\hat{\gamma}_R$ for $l = 1, \ldots, d$.

3. Asymptotic Theory

For convenient, let Σ denote the $(K+R)\times(K+R)$ variance-covariance matrix of $\left(p^K(y_{it-1})', q^R(x_{it}, z_{it})'\right)'$ whose smallest eigenvalue is bounded above zero and the largest eigenvalue is bounded for every K and R. Following Lee (2007), we decompose Σ as follows

$$\begin{pmatrix} \Sigma_{pp} & \Sigma_{pq} \\ \Sigma_{qp} & \Sigma_{qq} \end{pmatrix}$$

conformably as $\left(p^K(y_{it-1})', q^R(x_{it}, z_{it})'\right)'$ and $\Sigma_{pq} = \Sigma_{qp}'$. Thus, the conditional variance of $p^K(y_{it-1})$ given $q^R(x_{it}, z_{it})$ can be defined as $\Sigma_{pp|q} = \Sigma_{pp} - \Sigma_{pq}\Sigma_{qq}^{-1}\Sigma_{qp}$, and likewise, the conditional variance of $q^R(x_{it}, z_{it})$ given $p^K(y_{it-1})$ is $\Sigma_{qq|p} = \Sigma_{qq} - \Sigma_{qp}\Sigma_{pp}^{-1}\Sigma_{pq}$. Let $p_*^K(y_{it}) = p^K(y_{it}) - E\{p^K(y_{it}) \mid \mu_i\}$ and $q_*^R(x_{it}, z_{it}) = q^R(x_{it}, z_{it}) - E\{q^R(x_{it}, z_{it}) \mid \mu_i\}$ be the demeaned processes of $p^K(y_{it})$ and $q^R(x_{it}, z_{it})$ respectively, such that $E[p_*^K(y_{it})] = 0$ and $E[q_*^R(x_{it}, z_{it})] = 0$, for all i and t. Also, we use "$\xrightarrow{\mathcal{D}}$" to denote convergence in distribution. We make the following assumptions:

Assumption 1: $\{y_{it}, x_{it}, z_{it}, i = 1,\ldots,n;\ t = 1,\ldots,T\}$ are independent across i index; (x_i, z_i) are $i.i.d.$ where $x_i = (x_{i1},\ldots,x_{iT})'$ and $z_i = (z_{i1},\ldots,z_{iT})'$, and the support of (x_i, z_i) is compact subset of $\mathbb{R}^{d+m}$.

***Assumption 2*:** The error $\{e_{it}\}$ is $i.i.d.$ with zero mean, variance σ^2 and $E(|e_{it}|^{\nu}) < \infty$ for some $\nu > 4$, and e_{it} is independent of μ_i as well as x_{it} and z_{it}, for all i and t. Furthermore, $\{e_{it}\}$ has positive density almost everywhere and an absolutely continuous marginal distribution with respect to the Lebesgue measure on $\mathbb{R}$.

***Assumption 3*:** *(i)* For each i, the Markov process $\{y_{it}\}$ has a homogeneous transition probability $\mathcal{F}_i$ and the initial value y_{i0} is drawn from the invariant distribution π_i. *(ii)* Conditional on μ_i, for each i, the process $\{y_{it}\}$ is geometrically ergodic over t and hence β-mixing with exponentially decay mixing coefficients $\beta_i(\tau)$ such that $n^{-1}\sum_{i=1}^{n}\sum_{\tau=1}^{\infty}\beta_i(\tau)^{1-4/\nu} < \infty$ a.s., for some $\nu > 4$.

Assumption 4: *(i)* $\theta(0) = 0$ and *(ii)* $\lim_{n,T\to\infty}\frac{n}{T} = \kappa$ where $0 < \kappa < \infty$.

Assumption 5: *(i)* For every K, there exists positive integers n^* and T^* such that for all $n \geq n^*$ and $T \geq T^*$, the $(nT \times K)$ matrix $\tilde{P}_K$ is of full column rank. *(ii)* For each K, there exists a $(K \times K)$ matrix $\Gamma_K = E[p_*^K(y)p_*^K(y)']$ such that Γ_K has the smallest eigenvalue bounded away from zero and the bounded largest eigenvalue; and in addition, $\{p^K(y)\}$ is fourth order stationary with $\int p^K(y)p^K(y)'\pi_i 9(dy) < \infty$ for each i. *(iii)* $K = K(n,T)$ is nonrandom satisfying $K \to \infty$ and $K^2 / nT \to 0$ as $n,T \to \infty$. *(iv)* There exists a constant ζ satisfying $\sup_{1\leq i\leq n}\sup_{1\leq k\leq K} E\{|p_*^{Kk}(y_{it})|^{\nu}|\mu_i\} \leq \zeta < \infty$ for some $\nu > 4$.

Assumption 6: *(i)* For every R, there exists positive integers n^* and T^* such that for all $n \geq n^*$ and $T \geq T^*$, the $(nT \times R)$ matrix $\tilde{Q}_R$ is of full column rank. *(ii)* For each R, there exists a $(R \times R)$ matrix $\Gamma_R = E[q_*^R(x_{it}, z_{it})q_*^R(x_{it}, z_{it})']$ such that Γ_R has the smallest eigenvalue bounded away from zero and the bounded largest eigenvalue. *(iii)* $r_l = r_l(n,T)$ is nonrandom satisfying $r_l \to \infty$ as $n,T \to \infty$. *(iv)* There exists a sequence of nondecreasing constant $\zeta_0(R): \mathcal{S} \to \mathbb{R}_+$ satisfying $\sup_{(x,z)}\left\|Q_R(x,z)\right\| < \zeta_0(R)$ for every R such that $\zeta_0^4(R)R^2 / nT \to 0$ as $n,T \to \infty$ where $\mathcal{S}$ is the support of (x,z).

Assumption 7: *(i)* There exists a parameter vector $\alpha_K \in \mathbb{R}^K$ and a constant $\rho > 0$ such that $\sup_{y\in\mathcal{Y}_c}\left|\theta(y) - P_K(y)\alpha_K\right| = O(K^{-\rho})$ for every K; *(ii)* For $g(x,z) = \sum_{l=1}^{d} x_l\beta_l(z)$ there exists a parameter vector $\gamma_g = \gamma_{gR} = (\gamma_1^{r_1'},\ldots,\gamma_d^{r_d'})'$ and some constant $\delta_l > 0$ such that $\sup_{(x,z)\in\mathcal{S}} \left| g(x,z) - Q_R(x,z)\gamma_g \right| = O(\sum_{l=1}^{d} r_l^{-\delta_l})$.

Most of the above assumptions are very similar to the ones that are used in Lee (2007) and Ahmas et al. (2005), we modify a few of them for the purpose of our analysis. Assumption 1 is standard for panel data series estimation. Assumption 2 implies that e_{it} is independent of $\{y_{is}\}_{s\le t-1}$ as well as μ_i and (x_{it}, z_{it}). Assumption 3 is a standard condition for a stationary Markov chain to be geometrically ergodic. In time series context, a wide class of nonlinear autoregressive functions satisfied this assumption. See for example, Chen and Shen (1998), Fan and Yao (2003). Assumption 4(i) gives the identification condition for $\theta(.)$, whilst assumption 4(ii) says that the time series T should not be too small compared with n. Assumption 5 and 6 usually imply that the density of y and the density of (x,z) are each bounded from below by a positive constant. Assumption 7 states that there exist some positive constants such that the uniform approximation errors to the functions shrink at particular rates. It is known that many series function satisfy assumptions 5-7, for example, power series, orthogonal polynomial, trigonometric series and splines. Under the above assumptions, we can now state our main asymptotic results.

Theorem 3.1: *Under Assumptions 1-7, as* $n,T\to\infty$ *jointly we have*

$$E\int_{y\in\mathcal{Y}_c}\{\hat\theta(y)-\theta(y)\}^2\pi_i(dy) = O\left\{\left(\frac{K}{nT}+\frac{R}{nT}\right)+(K^{-2\rho}+\sum_{l=1}^{d} r_l^{-2\delta_l})+\left(\frac{\zeta_0^2(K)K}{nT}+\frac{\zeta_0^2(R)R}{nT}\right)\right\}$$

and for $l=1,\ldots,d$, *let* $\mathcal{S}_z$ *denotes the support of* z_{it} *,then*

$$E\int_{z\in\mathcal{S}_z}\{\hat\beta_l(z)-\beta(z)\}^2\pi_i(dz) = O\left\{\left(\frac{K}{nT}+\frac{R}{nT}\right)+(K^{-2\rho}+\sum_{l=1}^{d} r_l^{-2\delta_l})+\left(\frac{\zeta_0^2(K)K}{nT}+\frac{\zeta_0^2(R)R}{nT}\right)\right\}$$

where π_i *is the marginal distribution of* y_{it} *and* z_{it}.

The proof of Theorem 3.1 is a straightforward extension of Theorem 2 of Lee (2007), and we outline the key steps of the proof in Appendix A. Theorem 3.1 implies that the convergence rate of $\hat\theta(y)$ and $\hat\beta_l(z)$ $(l=1,\ldots,d)$ depends on both K and R, and it consists of three terms. The first term $((K/nT)+(R/nT))$ is essentially due to the convergence rate of the variance whereas the remaining terms $(K^{-2\rho}+\sum_{l=1}^{d} r_l^{-2\delta_l})$ and $((\zeta_0^2(K)K/nT)+(\zeta_0^2(R)R/nT))$ corresponds to the convergence rate of the squared bias. As discussed in Lee (2007), the last term is new and it does not appear in the standard series estimators for the cross-section case. Apparently, it reflects the asymptotic order of squared bias where it corresponds to the demeaning component of the within transformation.

Theorem 3.2: *Let* $\Phi_K = \sum_{j=0}^{\infty} \mathrm{cov}(p_K(y_{it+j}), e_{it+j})$ *where the* k^{th} *component* Φ_{Kk} *satisfies* $\left\|\Phi_{Kk}\right\| < \infty \ (k = 1, \ldots, K)$ *for every* K. *Under Assumptions 1-7, and in addition,* $\sqrt{nT}K^{-\rho} \to 0$, $\sqrt{nT} \ \sum_{l=1}^{d} r_l^{-2\delta_l} \to 0$ *as* $n, T \to \infty$ *jointly, we have*

(i) $\quad v(y, K, n, T)^{-1/2} \ \ \hat{\theta}(y) - \theta(y) + T^{-1} b_K(y) \ \xrightarrow{\ \mathcal{D}\ } N(0,1)$ *for* $y \in \mathcal{Y}_c$ *and*

(ii) $\quad V(y, K, n, T)^{-1/2} \ \ \hat{\beta}(z) - \beta(z) + T^{-1} b_R(z) \ \xrightarrow{\ \mathcal{D}\ } N(0, I_d)$ *for* $z \in \mathcal{S}_z$

where $v(y, K, n, T) = \sigma^2 P_K(y)' \Sigma_{pp.q}^{-1} P_K(y) / nT$,

$$V(y, x, z, R) = \sigma^2 Q_R(x, z)' \Sigma_{qq.p}^{-1} \Sigma_{qp} \Sigma_{pp}^{-1} Q_R(x, z) / nT$$

$$b_K(y) = P_K(y)' \Sigma_{pp.q}^{-1} \Phi_K$$

$$b_R = Q_R(x, z)' \Sigma_{qq.p}^{-1} \Sigma_{qp} \Sigma_{pp}^{-1} \Phi_K$$

The proof of Theorem 3.2 is a straightforward extension of Theorem 3 of Lee (2007) and the key steps of the proof are given in Appendix A. As in Lee (2007), Theorem 3.2 implies that the WG series estimators are asymptotically biased just as in the case of fixed effects linear dynamic panel models when the within transformation of the data is used, see for example Alvarez and Arellano (2003). However, since the forms of the biases are known, this suggests the following bias corrected estimators for $\theta(y)$ and $\beta(z)$.

Let $\hat{\theta}^*(y) = \hat{\theta}(y) + T^{-1}\hat{b}_K(y)$ and $\hat{\beta}^*(z) = \hat{\beta}(z) + T^{-1}\hat{b}_R(z)$ where $\hat{b}_K(y) = P_K(y)' \hat{\Sigma}_{pp.q}^{-1} \hat{\Phi}_K$, $\hat{b}_R = Q_R(x, z)' \hat{\Sigma}_{qq.p}^{-1} \hat{\Sigma}_{qp} \hat{\Sigma}_{pp}^{-1} \hat{\Phi}_K$, $\hat{\Phi}_K = (nT)^{-1} \sum_{i=1}^{n} \sum_{j=1}^{J} w(i, J) \sum_{t=1}^{T} \tilde{P}_K(y_{it+j}) \tilde{\varepsilon}_{it+j}$ for some proper weighting function $w(i, J)$, and $\hat{\Sigma}_{qq.p}^{-1}$, $\hat{\Sigma}_{pp}^{-1}$ and $\hat{\Sigma}_{qp}$ can be obtained from the estimated partitioned matrices of $\hat{\Sigma} = (nT)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[p^K(y_{it-1})', q^R(x_{it}, z_{it})' \right]' \left[p^K(y_{it-1})', q^R(x_{it}, z_{it})' \right]$.
Here $\hat{\Phi}_K$ is simply a heteroskedasticity and autocorrelation consistent (HAC) estimator of the long-run variance Φ_K, and a simple Barlett weighting function $w(i, J) = 1 - (j / (J + 1))$ can be used. Under the assumption 1-7, it can be shown that the results in Theorem 3.2 remain valid when replacing $\hat{\theta}(y)$ and $\hat{\beta}(z)$ by $\hat{\theta}^*(y)$ and $\hat{\beta}^*(z)$, respectively.

4. Conclusion

This paper considers a within-group type series estimation of a functional-coefficient partially linear dynamic panel data model with fixed effects. The convergence rates and asymptotic normality of the proposed estimator is developed. It is found that the within-group series estimator is asymptotically biased, as in the case of fixed effects linear dynamic panel model when the within-group transformation of data is used. A bias corrected type of series estimator is proposed.

References

Ahmad, I., S. Leelahanon and Q. Li (2005) "Efficient estimation of a semiparametric partially linear varying coefficient model" *Annals of Statistics* **33**, 258-283.

Alvares, J. and M. Arelleno (2003) "The time series and cross-section asymptotics of dynamic panel data estimator" *Econometrica* **71**, 1795-1843.

Baglan, D. (2010) "Efficient estimation of a partially linear dynamic panel data model with fixed effects: an application to unemployment dynamics in the U.S." *Working Paper*, Department of Economics, Howard University, Washington DC 20059.

Chen, X. and X. Shen (1998) "Sieve extremum estimators for weakly dependent data" *Econometrica* **66**, 289-314.

Fan, J. and Q. Yao (2003) *Nonlinear Time Series: Nonparametric and Parametric Methods*, New York: Springer-Verlag.

Lee, Y. (2007) "Nonparametric estimation of dynamic panel models with fixed effects" *Working Paper*, Department of Economics, University of Michigan, Ann Arbor MI 48109-1220.

Lee, Y. (2013) "Nonparametric estimation of dynamic panel models with fixed effects" *Econometric Theory*, Forthcoming.

Sun, Y., R. Carroll, and D. Li (2009) "Semiparametric estimation of fixed effects panel data with smooth coefficient models" *Advances in Econometrics* **25**, 101-130.

Appendix: Mathematical Proofs

The proofs of Theorem 1 and 2 are similar to that of Lee (2007), therefore we only discuss the heuristic ideas of the proofs here.

Proof of Theorem 1:

(i) By (5) and (6), we can write

$$\hat{\alpha} = (P'P)^{-1}P'(Y - Q\hat{\gamma})$$

$$= (P'P)^{-1}P'(P\alpha + Q\gamma + e + (\Theta - P\alpha) + (G - Q\gamma) - Q\hat{\gamma})$$

$$\hat{\alpha} - \alpha = (P'P/nT)^{-1}(P'e/nT) + (P'P/nT)^{-1}(P'(\Theta - P\alpha)/nT) +$$

$$(P'P/nT)^{-1}(P'(G - Q\gamma)/nT) - (P'P/nT)^{-1}(P'Q(\hat{\gamma} - \gamma)/nT)$$

Hence,

$$1_n \left\|\hat{\alpha} - \alpha\right\|^2 \le 1_n \left\|(P'P/nT)^{-1}(P'e/nT)\right\|^2 + 1_n \left\|(P'P/nT)^{-1}(P'(\Theta - P\alpha)/nT)\right\|^2 +$$

$$1_n \left\|(P'P/nT)^{-1}(P'(G - Q\gamma)/nT)\right\|^2 - 1_n \left\|(P'P/nT)^{-1}(P'Q(\hat{\gamma} - \gamma)/nT)\right\|^2$$

Following the proof of Theorem 3.1 in Lee (2007), it can be shown that the first term and the second term respectively, are $1_n \left\|(P'P/nT)^{-1}(P'e/nT)\right\|^2 \le O_p((K/nT) + \zeta_0^2(K)K/nT)$ and $1_n \left\|(P'P/nT)^{-1}(P'(\Theta - P\alpha)/nT)\right\|^2 \le O_p(K^{-2\rho} + \zeta_0^2(K)K^{1-2\rho}/nT)$. The convergence rate of the third term follows similarly but its convergence rate depends on $R = \sum_{j=1}^{d} r_j$ because the function G is approximated by $Q\gamma$, thus $1_n \left\|(P'P/nT)^{-1}(P'(G - Q\gamma)/nT)\right\|^2 = O_p(\sum_{j=1}^{d} r_j^{-2\delta_j} + \zeta_0^2(R)\sum_{j=1}^{d} r_j^{1-2\delta_j}/nT)$. The last term looks more complicated because it involves the convergence rate of $(\hat{\gamma} - \gamma)$. However from (4), $(\hat{\gamma} - \gamma)$ can be expressed explicitly as a function of $(\hat{\alpha} - \alpha)$ and by substituting this expression into the last term and solve for $(\hat{\alpha} - \alpha)$, it can be shown that

$$1_n \left\|(P'P/nT)^{-1}(P'Q(\hat{\gamma} - \gamma)/nT)\right\|^2 = O_p((R/nT) + (\zeta_0^2(R)R/nT)).$$

Thus, by combining the above results, also by noting that $1_n \to 1$ almost surely, we have

$$\left\|\hat{\alpha} - \alpha\right\|^2 \le O_p((K/nT) + \zeta_0^2(K)K/nT) + O_p(K^{-2\rho} + \zeta_0^2(K)K^{1-2\rho}/nT) +$$

$$O_p(\sum_{j=1}^{d} r_j^{-2\delta_j} + \zeta_0^2(R)\sum_{j=1}^{d} r_j^{1-2\delta_j}/nT) + O_p((R/nT) + (\zeta_0^2(R)R/nT))$$

$$= O_p((K/nT) + K^{-2\rho} + \zeta_0^2(K)K/nT) + O_p((R/nT) + \sum_{j=1}^{d} r_j^{-2\delta_j} + \zeta_0^2(R)R/nT)$$

Next, by triangle inequality, we have

$$
\begin{aligned}
E\int[\hat{\theta}(u)-\theta(u)]^2 dF_u(u) &= E\int[p^L(u)'(\hat{\alpha}-\alpha)+(p^L(u)'\alpha-\theta(u))]^2\, dF_u(u) \\
&\le 2CE\left\|\hat{\alpha}-\alpha\right\|^2 + 2E\int[p^L(u)'\alpha-\theta(u)]^2\, dF_u(u) \\
&= O_p((K/nT)+K^{-2\rho}+\zeta_0^2(K)K/nT)+O_p((R/nT)+ \\
&\quad \sum\nolimits_{j=1}^{d} r_j^{-2\delta_j}+\zeta_0^2(R)R/nT)+O(K^{-2\rho}) \\
&= O_p((K/nT)+(R/nT)+K^{-2\rho}+\sum\nolimits_{j=1}^{d} r_j^{-2\delta_j}+ \\
&\quad \zeta_0^2(K)K/nT+\zeta_0^2(R)R/nT)
\end{aligned}
$$

for some positive constant $C<\infty$ and Assumption 5. Thus, we have proved Theorem 1(i). The proof of Theorem 1(ii) follows the same arguments as above and hence omitted here. □

Proof of Theorem 2: First observe that

$$
\begin{aligned}
1_n\sqrt{nT}(\hat{\theta}(u)-\theta(u)) &= 1_n\sqrt{nT}P(\hat{\alpha}-\alpha) \\
&= 1_n\sqrt{nT}P\{(P'M_QP)^{-1}P'M_Q(\Theta-P\alpha)\}+ \\
&\quad 1_n\sqrt{nT}P\{(P'M_QP)^{-1}P'M_Q(G-Q\gamma)\}+ \\
&\quad 1_n\sqrt{nT}P\{(P'M_QP)^{-1}P'M_Qe\}
\end{aligned} \tag{A.1}
$$

and

$$
\begin{aligned}
1_n\sqrt{nT}(\hat{\beta}(z)-\beta(z)) &= 1_n\sqrt{nT}Q(\hat{\gamma}-\gamma) \\
&= 1_n\sqrt{nT}Q\{(Q'M_PQ)^{-1}Q'M_P(G-Q\gamma)\}+ \\
&\quad 1_n\sqrt{nT}Q\{(Q'M_PQ)^{-1}Q'M_P(\Theta-P\alpha)\}+ \\
&\quad 1_n\sqrt{nT}Q\{(Q'M_PQ)^{-1}Q'M_Pe\}
\end{aligned} \tag{A.2}
$$

By Lemma A1.2 of Lee (2007), we have $\left\|\hat{\Sigma}-\Sigma\right\|\to_p 0$ as $n,T\to\infty$, thus, the first two terms in $(A.1)$ and $(A.2)$ can be shown to be asymptotically negligible by Assumption 7. By combining the last term in $(A.1)$ and $(A.2)$ along with the result of partitioned regression, yields

$$
\begin{pmatrix} 1_n\sqrt{nT}P\{(P'M_QP)^{-1}P'M_Qe\} \\ 1_n\sqrt{nT}Q\{(Q'M_PQ)^{-1}Q'M_Pe\} \end{pmatrix} = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}\hat{\Sigma}^{-1}\begin{pmatrix} 1_nP'e/\sqrt{nT} \\ 1_nQ'e/\sqrt{nT} \end{pmatrix}
$$

First note that $\left\|\hat{\Sigma}^{-1}-\Sigma^{-1}\right\|\to_p 0$ as $n,T\to\infty$ and z_{it} is strictly exogenous for all i and t. Second, following the proof of Theorem 3.2 and Corollary 4.1 of Lee (2007), and the fact that $1_n\to 1$ almost surely, the limit distribution of the quantity $\begin{pmatrix} P'e+(1/T)\Phi_K/\sqrt{nT} \\ Q'e/\sqrt{nT} \end{pmatrix}$ is approximately normal with mean zero and variance $\sigma^2\Sigma$. Therefore, by using the inverse matrix formula of the partitioned matrix, we have

$$\Sigma^{-1} = \begin{pmatrix} \Sigma_{pp} & \Sigma_{pq} \\ \Sigma_{qp} & \Sigma_{qq} \end{pmatrix}^{-1} = \begin{pmatrix} \Sigma_{pp|q}^{-1} & -\Sigma_{pp|q}^{-1}\Sigma_{pq}\Sigma_{qq}^{-1} \\ -\Sigma_{qq}^{-1}\Sigma_{qp}\Sigma_{pp|q}^{-1} & \Sigma_{qq}^{-1} + \Sigma_{qq}^{-1}\Sigma_{qp}\Sigma_{pp|q}^{-1}\Sigma_{pq}\Sigma_{qq}^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} \Sigma_{pp}^{-1} + \Sigma_{pp}^{-1}\Sigma_{pq}\Sigma_{qq|p}^{-1}\Sigma_{qp}\Sigma_{pp}^{-1} & -\Sigma_{pp}^{-1}\Sigma_{pq}\Sigma_{qq|p}^{-1} \\ -\Sigma_{qq|p}^{-1}\Sigma_{qp}\Sigma_{pp}^{-1} & \Sigma_{qq|p}^{-1} \end{pmatrix}$$

and the desired result is obtained using the above expression. $\square$