Lecture 7: Spatial Panel Data Models

Zhenlin Yang

School of Economics, Singapore Management University
zlyang@smu.edu.sg
http://www.mysmu.edu.sg/faculty/zlyang/

ECON747 Spatial Econometric Models and Methods
Term I 2024-25

7.1. Introduction

Consider the following spatial panel data (SPD) model,

$$Y_{nt} = \lambda_0 W_{1n} Y_{nt} + X_{nt}\beta_0 + \boldsymbol{\mu}_{n0} + \alpha_{t0} 1_n + U_{nt}, \quad U_{nt} = \rho_0 W_{2n} U_{nt} + V_{nt}, \quad (7.1)$$

for $t = 1, 2, \ldots, T$, where for a given t,

- Y_{nt}: $n \times 1$ vector of observations on the response variable,
- X_{nt}: $n \times k$ matrix containing the values of k nonstochastic, individually and time (IT) varying regressors,
- V_{nt}: $n \times 1$ vector of errors; $\{v_{it}\}$ iid(0, σ_0^2) for all i and t,
- $\boldsymbol{\mu}_{n0}$: $n \times 1$ vector of unobserved individual-specific effects, and
- $\boldsymbol{\alpha}_{T0}$: $T \times 1$ vector of unobserved time-specific effects $(\alpha_{10}, \ldots, \alpha_{T0})'$.
- W_{1n} and W_{2n}: $n \times n$ spatial weights matrices, capturing the spatial interaction effects on the response, and the cross-sectional dependence among the disturbances.
- In practice, W_{1n} and W_{2n} may be the same.

The unobserved effects μ_{n0} and α_{T0} may be considered as,

- `fixed effects` (FE), in the sense that μ_{n0} and α_{T0} may be correlated with the IT-varying regressors in an arbitrary manner,
- `random effects` (RE), in the sense that μ_{n0} and α_{T0} are uncorrelated with the IT-varying regressors,
- `correlated random effects` (CRE), in the sense that μ_{n0} and α_{T0} are correlated with the IT-varying regressors linearly. It can be referred to as `linear fixed effects` (LFE).

The overall size of the data is $n \times T$, which may increase due to the

- increase of n but not T, giving the most interesting case for spatial econometrics, or micro-econometrics in general,
- increase of T but not n, giving a case of time-series VAR model,
- increase of both n and T (but n is faster than T), giving another interesting case for spatial econometrics.

It is clear that panel data allows researchers to control the unobserved heterogeneity (in intercepts) across individuals and time:

- FE specification gives a full control. However, it creates incidental parameters problem as the effects are treated as free parameters, and it does not allow separate estimation of time-invariant effects, such as `gender` and `race`, and individual-invariant effects, such as policy change;
- RE specification gives a minimum control. It does not suffer from these problems as the effects are uncorrelated with the IT-varying regressors and hence can be treated as iid random variables;
- The CRE or LFE specification goes in-between FE and RE. It does not suffer from these problems and at the same time gives 'enough' control for the unobserved heterogeneity.

Some tests are of interest: FE vs RE, Hausman test,
LFE vs RE, LM test.

The SPD model given above can be further extended to allow the individual- and time-specific effects to be interactive, giving

- the SPD model with interactive fixed effects (IFE): $\boldsymbol{\mu}_{n0}\alpha_{t0}$, where α_{t0} now can be a vector, referred to as common factors, and the corresponding $\boldsymbol{\mu}_{n0}$ becomes a matrix whose rows are referred to as factor loadings. The early FE-SDP model is then referred to as SDP model with additive fixed effects.
- the SPD model with interactive random effects (IRE). The early SDP model is then referred to as SDP model with additive random effects,
- the SPD model with correlated IRE.

These specifications have not been fully studied, in particular the RE and CRE specifications.

Different estimation and inference methods have not been fully explored.

Overview

In this lecture, we

1. introduce fully QML estimation and inference methods for FE-SPD model, based on the work of Lee and Yu (2010a, JOE);
2. introduce briefly QML estimation and inference methods for RE-SPD, by extending the work of Lee and Yu (2012), Baltagi et al. (2013);
3. introduce briefly QML estimation and inference methods CRE-SPD models, by further extending the work above;
4. discuss extended FE-SPD model: the SPD model with interactive FEs;
5. introduce various tests for spatial effects in panel data models;
6. present Monte Carlo results for the finite sample performance of the QMLEs and the tests.
7. present empirical applications to illustrate the methods introduced.

Under the fixed effects specifications, the **effects** $\{c_{i0}\}$ and $\{\alpha_{t0}\}$ are treated as free parameters, and hence must be eliminated in certain way to avoid the incidental parameters problem. Otherwise, joint estimation of these effects with model's common parameters will lead to inconsistency.

We follow the transformation approach of Lee and Yu (2010a) to introduce the QML estimation of the FE-SPD model given by (7.1).

To eliminate the individual and time effects, define the projection matrices:

- $J_T = (I_T - \frac{1}{T}1_T1_T')$, with orthonormal eigenvectors $[F_{T,T-1}, \frac{1}{\sqrt{T}}1_T]$;
- $J_n = (I_n - \frac{1}{n}1_n1_n')$, with orthonormal eigenvectors $[F_{n,n-1}, \frac{1}{\sqrt{n}}1_n]$.

Note: The eigenvalues of a projection matrix are either 1 or 0, with number of ones being the rank of the matrix. $F_{T,T-1}$ and $F_{n,n-1}$ are the submatrices of eigenvector matrices, corresponding to eigenvalues of one.

Now, for any $n \times T$ matrix, such as $[Y_{n1}, \cdots, Y_{nT}]$, define the $(n-1) \times (T-1)$ transformed matrix as

$$[Y_{n1}^*, \ldots, Y_{n,T-1}^*] = F_{n,n-1}'[Y_{n1}, \ldots, Y_{nT}]F_{T,T-1}. \tag{7.2}$$

This leads to the transformed vector Y_{nt}^*, for $t = 1, \ldots, T-1$. Similarly U_{nt}^*, V_{nt}^*, and $X_{nt,j}^*$ (for the jth regressor $X_{nt,j}$) are obtained, $j = 1, \cdots, k$.

- Let $X_{nt}^* = [X_{nt,1}^*, X_{nt,2}^*, \ldots, X_{nt,k}^*]$,
- Define $W_{rn}^* = F_{n,n-1}' W_{rn} F_{n,n-1}$, $r = 1, 2$,
- Assume W_{rn} is row-normalized. Then, $J_n W_{rn} = J_n W_{rn} J_n$.
- By *Spectral Theorem*, $J_n = F_{n,n-1} F_{n,n-1}'$, it follows that
- $F_{n,n-1}' W_{rn} = F_{n,n-1}' W_{rn} F_{n,n-1} F_{n,n-1}'$.

The transformed model we will work on thus takes the form:

$$Y_{nt}^* = \lambda_0 W_{1n}^* Y_{nt}^* + X_{nt}^* \beta_0 + U_{nt}^*, \quad U_{nt}^* = \rho_0 W_{2n}^* U_{nt}^* + V_{nt}^*, \tag{7.3}$$

for $t = 1, \ldots, T-1$.

After transformations, the effective sample size is $N = (n-1)(T-1)$.

When W_{jn} are not row normalized, the linear SARAR representation of (7.3) for the spatial panel model will no longer hold. In that case, a likelihood formulation would not be feasible.

Stacking the vectors and matrices, i.e., letting

- $\mathbf{Y}_N = (Y^{*\prime}_{n1}, \ldots, Y^{*\prime}_{n,T-1})'$,
- $\mathbf{U}_N = (U^{*\prime}_{n1}, \ldots, U^{*\prime}_{n,T-1})'$,
- $\mathbf{V}_N = (V^{*\prime}_{n1}, \ldots, V^{*\prime}_{n,T-1})'$,
- $\mathbf{X}_N = (X^{*\prime}_{n1}, \ldots, X^{*\prime}_{n,T-1})'$,

and denoting $\mathbf{W}_{hN} = I_{T-1} \otimes W^*_{hn}$, $h = 1, 2$, we have the following compact expression for the transformed model:

$$\mathbf{Y}_N = \lambda_0 \mathbf{W}_{1N} \mathbf{Y}_N + \mathbf{X}_N \beta_0 + \mathbf{U}_N, \quad \mathbf{U}_N = \rho_0 \mathbf{W}_{2N} \mathbf{U}_N + \mathbf{V}_N, \tag{7.4}$$

which is in form identical to the spatial autoregressive model with autoregressive errors (SARAR) or SLE model.

This shows that the QML estimation of the two-way fixed effects panel SARAR model is similar to that of the linear SARAR model.

The key difference is that the elements of $\mathbf{V}_N$ may not be independent although they are uncorrelated and homoskedastic as seen below:

- Letting $\otimes$ be the Kronecker product, i.e., $A \otimes B = \{a_{ij}B\}$, we have
$$(V_{n1}^{*\prime}, \ldots, V_{n,T-1}^{*\prime})' = (F_{T,T-1}' \otimes F_{n,n-1}')(V_{n1}', \ldots, V_{nT}')'.$$
- Then,
$$E(V_{n1}^{*\prime}, \ldots, V_{n,T-1}^{*\prime})'(V_{n1}^{*\prime}, \cdots, V_{n,T-1}^{*\prime})$$
$$= \sigma_0^2(F_{T,T-1}' \otimes F_{n,n-1}')(F_{T,T-1} \otimes F_{n,n-1}) = \sigma_0^2 I_N.$$
- Hence, $\{v_{it}^*\}$ are iid $N(0, \sigma_0^2)$ if the original errors $\{v_{it}\}$ are iid $N(0, \sigma_0^2)$.
- If the original errors $\{v_{it}\}$ are iid $(0, \sigma_0^2)$ (non-normal), then $\{v_{it}^*\}$ are in general only uncorrelated with mean 0 and constant variance σ_0^2.
- In the former, we have exact Gaussian likelihood, and in the latter it becomes Quasi-Gaussian likelihood.

The (quasi) Gaussian log likelihood function for (7.4) is,

$$\ell_N(\theta) = -\frac{N}{2}\ln(2\pi\sigma^2) + \ln|\mathbf{A}_N(\lambda)| + \ln|\mathbf{B}_N(\rho)| - \frac{1}{2\sigma^2}\mathbf{V}_N'(\beta,\delta)\mathbf{V}_N(\beta,\delta), \quad (7.5)$$

where $\theta = (\beta', \sigma^2, \lambda, \rho)'$, and $\delta = (\beta', \delta)'$;

$\mathbf{A}_N(\lambda) = I_N - \lambda\mathbf{W}_{1N}$, and $\mathbf{B}_N(\rho) = I_N - \rho\mathbf{W}_{2N}$;

$\mathbf{V}_N(\beta,\delta) = \mathbf{B}_N(\rho)[\mathbf{A}_N(\lambda)\mathbf{Y}_N - \mathbf{X}_N\beta]$.

1. Maximizing $\ell_N(\theta)$ gives the QMLE $\hat{\theta}_N$ of θ.
2. First, maximize w.r.t. β and σ^2 to give the constrained QMLEs:

$$\tilde{\beta}_N(\delta) = [\mathbb{X}_N'(\rho)\mathbb{X}_N(\rho)]^{-1}\mathbb{X}_N'(\rho)\mathbb{Y}_N(\delta), \quad (7.6)$$

$$\tilde{\sigma}_N^2(\delta) = \frac{1}{N}\mathbb{Y}_N'(\delta)\mathbb{M}_N(\rho)\mathbb{Y}_N(\delta), \quad (7.7)$$

where $\mathbb{Y}_N(\delta) = \mathbf{B}_N(\rho)\mathbf{A}_N(\lambda)\mathbf{Y}_N$, $\mathbb{X}_N(\rho) = \mathbf{B}_N(\rho)\mathbf{X}_N$,

$\mathbb{M}_N(\rho) = I_N - \mathbb{X}_N(\rho)[\mathbb{X}_N'(\rho)\mathbb{X}_N(\rho)]^{-1}\mathbb{X}_N'(\rho)$;

3. Then, maximize the concentrated loglikelihood:

$$\ell_N^c(\delta) = \ell_N(\tilde{\beta}_N(\delta), \tilde{\sigma}_N^2(\delta), \delta)$$
$$= -\frac{N}{2}(\ln(2\pi)+1) + \ln|\mathbf{A}_N(\lambda)| + \ln|\mathbf{B}_N(\rho)| - \frac{N}{2}\ln\tilde{\sigma}_N^2(\delta). \qquad (7.8)$$

to give the unconstrained QMLE $\hat{\delta}_N$ of δ.

4. The unconstrained QMLEs of β and σ^2 are thus,

$$\hat{\beta}_N = \tilde{\beta}_N(\hat{\delta}_N) \text{ and } \hat{\sigma}_N^2 = \tilde{\sigma}_N^2(\hat{\delta}_N).$$

5. Maximization of $\ell_N^c(\delta)$ can be computationally demanding if N is large due to the need for repeated calculation of the two determinants. Following simplifications help alleviate the computational burden:

$$\begin{aligned} |\mathbf{A}_N(\lambda)| &= |I_{n-1} - \lambda W_{1n}^*|^{T-1} \\ &= \left(\tfrac{1}{1-\lambda}|I_n - \lambda W_{1n}|\right)^{T-1}, \quad \text{(Lee and Yu, 2010b)} \\ &= \left(\tfrac{1}{1-\lambda}\prod_{i=1}^n(1-\lambda\omega_{1i})\right)^{T-1}, \quad \text{(Griffith, 1988)} \end{aligned}$$

where ω_{1i} are the eigenvalues of W_{1n}. Similarly, $|\mathbf{B}_N(\rho)|$ is calculated.

6. The **linear SARAR representation** (7.4) has greatly facilitated the QML estimation of the general FE-SPD model.
7. It is also very helpful for the subsequent developments in bias and variance corrections, in light of methods presented in Lecture 4.
8. Obviously, it contains the spatial regression models as special cases.
9. Based on this representation, the results developed for this general model can easily be reduced to suit simpler models:
   - setting ρ or λ to zero in (7.4) gives an FE-SPD model with only SLD, or an FE-SPD model with only SED;
   - dropping either α_{t0} or $\boldsymbol{\mu}_{n0}$ in (7.1) leads to a submodel with only individual-specific effects, or a submodel with only time-specific effects. The former is more interesting.
10. On the other hand, the spatial panel model considered in this chapter can also be extended to include more spatial lag terms in both the response and the disturbance, in particular the former.

Write $\hat{\theta}_N = (\hat{\beta}'_N, \hat{\sigma}^2_N, \hat{\delta}'_N)'$. Let θ_0 be the true value of θ, and δ_0 be the true value of δ. Lee and Yu (2010a) show that under some mild conditions $\hat{\theta}_N$ is $\sqrt{N}$-consistent and asymptotically normal.

Assumption A1. *W_{1n} and W_{2n} are row-normalized nonstochastic spatial weights matrices with zero diagonals.*

Assumption A2. *The disturbances $\{v_{it}\}$, $i = 1, 2, \ldots, n$ and $t = 1, 2, \ldots, T$, are iid across i and t with zero mean, variance σ_0^2 and $\mathrm{E}|v_{it}|^{4+\eta} < \infty$ for some $\eta > 0$.*

Assumption A3. *$A_n(\lambda)$ and $B_n(\rho)$ are invertible for all $\lambda \in \Lambda$ and $\rho \in \mathbb{P}$, where Λ and $\mathbb{P}$ are compact intervals. Furthermore, λ_0 is in the interior of Λ, and ρ_0 is in the interior of $\mathbb{P}$.*

Note: Due to the nonlinearity of λ and ρ in the model, compactness of Λ and $\mathbb{P}$ is needed. However, the compactness of the space of β and σ^2 is not necessary because the β and σ^2 estimates given λ and ρ are least squares type estimates.

Assumption A4. *The elements of X_{nt} are nonstochastic, and are bounded uniformly in n and t. Under the setting in Assumption A6, the limit of $\frac{1}{N}\mathbf{X}_N'\mathbf{X}_N$ exists and is nonsingular.*

Assumption A5. *W_{1n} and W_{2n} are uniformly bounded in both row and column sums in absolute value (for short, UB). Also $A_n^{-1}(\lambda)$ and $B_n^{-1}(\rho)$ are UB, in $\lambda \in \Lambda$ and $\rho \in \mathbb{P}$.*

Assumption A6. *n is large, where T can be finite or large.*

Assumption A7. *Either (a): $\lim_{n\to\infty} \mathcal{H}_N(\rho)$ is nonsingular $\forall \rho \in \mathbb{P}$ and $\lim_{n\to\infty} \mathcal{Q}_{1n}(\rho) \neq 0$ for $\rho \neq \rho_0$; or (b): $\lim_{n\to\infty} \mathcal{Q}_{2n}(\delta) \neq 0$ for $\delta \neq \delta_0$, where*

$$\mathcal{H}_N(\rho) = \tfrac{1}{N}(\mathbf{X}_N, \mathbf{W}_{1N}\mathbf{A}_N^{-1}\mathbf{X}_N\beta_0)'\mathbf{B}_N'(\rho)\mathbf{B}_N(\rho)(\mathbf{X}_N, \mathbf{W}_{1N}\mathbf{A}_N^{-1}\mathbf{X}_N\beta_0),$$
$$\mathcal{Q}_{1n}(\rho) = \tfrac{1}{n-1}\left(\ln\left|\sigma_0^2 B_n^{-1\prime} J_n B_n^{-1}\right| - \ln\left|\sigma_n^2(\rho) B_n^{-1}(\rho)' J_n B_n^{-1}(\rho)\right|\right),$$
$$\mathcal{Q}_{2n}(\delta) = \tfrac{1}{n-1}\left(\ln\left|\sigma_0^2 B_n^{-1\prime} A_n^{-1\prime} J_n A_n^{-1} B_n^{-1}\right| - \ln\left|\sigma_n^2(\delta) B_n^{-1}(\rho)' A_n^{-1}(\lambda)' J_n A_n^{-1}(\lambda) B_n^{-1}(\rho)\right|\right),$$
$$\sigma_n^2(\delta) = \tfrac{\sigma_0^2}{n-1}\mathrm{tr}[(B_n(\rho)A_n(\lambda)A_n^{-1}B_n^{-1})' J_n (B_n(\rho)A_n(\lambda)A_n^{-1}B_n^{-1})], \text{ and}$$
$$\sigma_n^2(\rho) = \sigma_n^2(\delta)|_{\lambda=\lambda_0}.$$

Assumption A8. *The limit of* $\frac{1}{(n-1)^2}\left[\mathrm{tr}(C_n^s C_n^s)\mathrm{tr}(D_n^s D_n^s) - \mathrm{tr}^2(C_n^s D_n^s)\right]$ *is strictly positive, where* $C_n = J_n\bar{G}_{1n} - \frac{\mathrm{tr}(J_n\bar{G}_{1n})}{n-1}J_n$, $D_n = J_n G_{2n} - \frac{\mathrm{tr}(J_n G_{2n})}{n-1}J_n$, $\bar{G}_{1n} = B_n G_{1n} B_n^{-1}$, $G_{1n} = W_{1n}A_n^{-1}$, *and* $G_{2n} = W_{2n}B_n^{-1}$.

Theorem

(Lee and Yu, 2010) Under Assumptions A1-A8, we have $\hat{\theta}_N \xrightarrow{p} \theta_0$, *and*

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \xrightarrow{D} N\left[0,\ \lim_{N\to\infty}\Sigma_N^{-1}(\theta_0)\Gamma_N(\theta_0)\Sigma_N^{-1}(\theta_0)\right], \tag{7.9}$$

where $\Sigma_N(\theta_0) = \frac{1}{N}\mathrm{E}[\frac{\partial^2}{\partial\theta_0\partial\theta_0'}\ell_N(\theta_0)]$ *assumed to be positive definite for large enough* N, *and* $\Gamma_N(\theta_0) = \frac{1}{N}\mathrm{E}[(\frac{\partial}{\partial\theta_0}\ell_N(\theta_0))(\frac{\partial}{\partial\theta_0}\ell_N(\theta_0))']$ *assumed to exist.*

The results of the theorem serve two purposes:

- they provide theory for asymptotic inferences;
- they provide crucial results for higher-order bias and variance corrections, and thus for refined inferences.

Statistical inferences for the parameters θ in the FE-SPD model require consistent estimators of $\Sigma_N(\theta_0)$ and $\Gamma_N(\theta_0)$.

We propose plug-in estimators and for this the analytical expressions of $\Sigma_N(\theta_0)$ and $\Gamma_N(\theta_0)$ are required.

The quasi score vector has a similar form as (2.27) for the SLE model, which is under the new notations designed for the FE-SPD model:

$$\frac{\partial \ell_N(\theta)}{\partial \theta} = \begin{cases} \frac{1}{\sigma^2}\mathbf{X}_N'\mathbf{B}_N(\rho)\mathbf{V}_N(\beta,\delta), \\ -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4}\mathbf{V}_N'(\beta,\delta)\mathbf{V}_N(\beta,\delta), \\ \frac{1}{\sigma^2}\mathbf{V}_N'(\beta,\delta)\mathbf{B}_N(\rho)\mathbf{W}_{1N}\mathbf{Y}_N - \mathrm{tr}[\mathbf{G}_{1N}(\lambda)], \\ \frac{1}{\sigma^2}\mathbf{V}_N'(\beta,\delta)\mathbf{G}_{2N}(\rho)\mathbf{V}_N(\beta,\delta) - \mathrm{tr}[\mathbf{G}_{2N}(\rho)]. \end{cases} \tag{7.10}$$

With this, $\Sigma_N(\theta_0) = \frac{1}{N}\mathrm{E}[\frac{\partial^2}{\partial\theta_0\partial\theta_0'}\ell_N(\theta_0)]$ can be derived. At the true θ_0, the score vector can be simplified so that $\Gamma_N(\theta_0) = \mathrm{E}[\frac{\partial \ell_N(\theta_0)}{\partial\theta}\frac{\partial \ell_N(\theta_0)}{\partial\theta'}]$, the VC matrix of the quasi score, can be derived.

First, with the notations introduced earlier, letting $\mathbf{G}_{1N}(\lambda) = \mathbf{W}_{1N}\mathbf{A}_N^{-1}(\lambda)$ and $\mathbf{G}_{2N}(\rho) = \mathbf{W}_{2N}\mathbf{B}_N^{-1}(\rho)$, and denoting $\mathbf{A}_N = \mathbf{A}_N(\lambda_0)$, $\mathbf{B}_N = \mathbf{B}_N(\rho_0)$, $\mathbf{G}_{1N} = \mathbf{G}_{1N}(\lambda_0)$, $\mathbf{G}_{2N} = \mathbf{G}_{2N}(\rho_0)$, $\mathbb{X}_N = \mathbb{X}_N(\rho_0)$, etc., we have:

$$\Sigma_N(\theta_0) = \begin{pmatrix} \frac{1}{N\sigma_0^2}\mathbb{X}_N'\mathbb{X}_N, & 0, & \frac{1}{N\sigma_0^2}\mathbb{X}_N'\boldsymbol{\eta}_N, & 0 \\ \sim, & \frac{1}{2\sigma_0^4}, & \frac{1}{N\sigma_0^2}\mathrm{tr}(\mathbf{G}_{1N}), & \frac{1}{N\sigma_0^2}\mathrm{tr}(\mathbf{G}_{2N}) \\ \sim, & \sim, & T_{1N} + \frac{1}{N\sigma_0^2}\boldsymbol{\eta}_N'\boldsymbol{\eta}_N, & T_{3N} \\ \sim, & \sim, & \sim, & T_{2N} \end{pmatrix}, \quad (7.11)$$

where $\boldsymbol{\eta}_N = \mathbf{B}_N\mathbf{G}_{1N}\mathbf{X}_N\beta_0$,

$$T_{1N} = \frac{1}{N}\mathrm{tr}(\bar{\mathbf{G}}_{1N}^s\bar{\mathbf{G}}_{1N}), \qquad \bar{\mathbf{G}}_{1N} = \mathbf{B}_N\mathbf{G}_{1N}\mathbf{B}_N^{-1},$$

$$T_{2N} = \frac{1}{N}\mathrm{tr}(\mathbf{G}_{2N}^s\mathbf{G}_{2N}), \qquad G^s = G + G', \text{ for a matrix } G,$$

$$T_{3N} = \frac{1}{N}\mathrm{tr}(\mathbf{G}_{2N}^s\bar{\mathbf{G}}_{1N}).$$

With the analytical expression, $\Sigma_N(\theta_0)$ can be consistently estimated by the plug-in estimator $\Sigma_N(\hat{\theta}_N)$

To obtain the other component $\Gamma_N(\theta_0)$ of the VC matrix of $\hat{\theta}_N$, Yang et al. (2016) give an expression of the score vector in terms of $\mathbb{V}_{nT} = (V'_{n1}, \ldots, V'_{nT})'$, the vector of original errors using the relation,

$$\mathbf{V}_N = \mathbb{F}'_{nT,N}\mathbb{V}_{nT}, \quad \text{where } \mathbb{F}_{nT,N} = F_{T,T-1} \otimes F_{n,n-1}:$$

$$\frac{\partial \ell_N(\theta_0)}{\partial \theta_0} = \begin{cases} \frac{1}{\sigma_0^2}\mathbb{A}'_{1nT}\mathbb{V}_{nT}, \\ \frac{1}{2\sigma_0^4}\mathbb{V}'_{nT}\mathbb{A}'_{2nT}\mathbb{V}_{nT} - \frac{N}{2\sigma_0^2}, \\ \frac{1}{\sigma_0^2}\mathbb{V}'_{nT}\mathbb{A}'_{3nT}\mathbb{V}_{nT} + \frac{1}{\sigma_0^2}\mathbf{b}'_{nT}\mathbb{V}_{nT} - \mathrm{tr}[\mathbf{G}_{1N}(\lambda)], \\ \frac{1}{\sigma_0^2}\mathbb{V}'_{nT}\mathbb{A}'_{4nT}\mathbb{V}_{nT} - \mathrm{tr}[\mathbf{G}_{2N}(\rho)], \end{cases} \tag{7.12}$$

where $\mathbf{b}_{nT} = \mathbb{F}_{nT,N}\boldsymbol{\eta}_N$,

$\mathbb{A}_{1nT} = \mathbb{F}_{nT,N}\mathbf{B}_N\mathbf{X}_N$,

$\mathbb{A}_{2nT} = \mathbb{F}_{nT,N}\mathbb{F}'_{nT,N}$,

$\mathbb{A}_{3nT} = \mathbb{F}_{nT,N}\mathbf{B}_N \cdot\mathbf{G}_N\mathbf{B}_N^{-1}\mathbb{F}'_{nT,N}$, and

$\mathbb{A}_{4nT} = \mathbb{F}_{nT,N}\mathbf{W}_{2N}\mathbf{B}_N^{-1}\mathbb{F}'_{nT,N}$.

Letting $\mathbf{a}_{inT}$ be the diagonal vector of $\mathbb{A}_{inT}$, and denoting

$$\Pi_{ij} = \frac{1}{N}\mathrm{tr}[\mathbb{A}_{inT}(\mathbb{A}_{jnT} + \mathbb{A}'_{jnT})] + \frac{1}{N}k_4\mathbf{a}'_{inT}\mathbf{a}_{jnT}, \tag{7.13}$$

where k_4 is the 4th cumulant of the original errors, Yang et al. (2016) give,

$$\Gamma_N(\theta_0) = \begin{pmatrix} \frac{1}{N\sigma_0^2}\mathbb{X}'_N\mathbb{X}_N, & 0, & \frac{1}{N\sigma_0^2}\mathbb{A}'_{1nT}\mathbf{b}_{nT}, & 0 \\ \sim, & \frac{1}{4\sigma_0^4}\Pi_{22}, & \frac{1}{2\sigma_0^2}\Pi_{23}, & \frac{1}{2\sigma_0^2}\Pi_{24} \\ \sim, & \sim, & \Pi_{33} + \frac{1}{N\sigma_0^2}\mathbf{b}'_{nT}\mathbf{b}_{nT}, & \Pi_{34} \\ \sim, & \sim, & \sim, & \Pi_{44} \end{pmatrix}. \tag{7.14}$$

- For practical applications, $\Sigma_N(\theta_0)$ is estimated by $\Sigma_N(\hat{\theta}_N)$;
- and $\Gamma_N(\theta_0)$ by plugging-in $\hat{\theta}_N$ for θ and $\hat{k}_4$ for k_4, where

$$\hat{k}_4 = \bar{a}_4^{-1}\kappa_4(\hat{\mathbf{V}}_N),$$

where $\kappa_4(\hat{\mathbf{V}}_N)$ is the fourth sample cumulant of the QML residuals $\hat{\mathbf{V}}_N$, and $\bar{a}_4 = \frac{1}{N}\sum_{i=1}^{N} a_{4,i}$ with $a_{4,i} = l'_{nT}\mathbf{f}_i^4$ and $\mathbf{f}_i$ is the ith column of $\mathbb{F}_{nT,N}$.

Lee and Yu (2010a) provide a useful identity:

$$(I_{n-1} - \lambda W_{hn}^*)^{-1} = F'_{n,n-1}(I_{n-1} - \lambda W_{hn})^{-1} F_{n,n-1}.$$

Based on this, the inverses of $\mathbf{A}_N(\lambda)$ and $\mathbf{B}_N(\lambda)$ can easily be calculated as they are block-diagonal.

As discussed in Lee and Yu (2010a, Footnote 12),

- the first difference and Helmert transformation have often been used to eliminate the individual effects.
- A special selection of $F_{T,T-1}$ gives rise to the Helmert transformation where $\{V_{nt}\}$ are transformed to

$$(\tfrac{T-t}{T-t+1})^{1/2}[V_{nt} - \tfrac{1}{T-t}(V_{n,t+1} + \cdots + V_{nT}),$$

  which is of particular interest for dynamic panel data models.

Under normality and homoskedasticity, Debarsy and Ertur (2010) developed LM tests for Model (7.4) of the following three hypotheses,

$$H_0^{\mathtt{SL}}: \lambda = 0 | \rho = 0; \;\; H_0^{\mathtt{SE}}: \rho = 0 | \lambda = 0; \;\; H_0^{\mathtt{SLE}}: \lambda = \rho = 0,$$

$$\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}} = \frac{N}{\sqrt{\mathbb{S}_1 + \tilde{\mathbb{D}}}} \frac{\tilde{\mathbf{V}}_N' \mathbf{W}_{1N} \mathbf{Y}_N}{\tilde{\mathbf{V}}_N' \tilde{\mathbf{V}}_N}, \tag{7.15}$$

$$\mathrm{LM}_{\mathtt{SE}}^{\mathtt{FI}} = \frac{N}{\sqrt{\mathbb{S}_2}} \frac{\tilde{\mathbf{V}}_N' \mathbf{W}_{2N} \tilde{\mathbf{V}}_N}{\tilde{\mathbf{V}}_N^{*\prime} \tilde{\mathbf{V}}_N}, \tag{7.16}$$

$$\mathrm{LM}_{\mathtt{SLE}}^{\mathtt{FI}} = \frac{1}{\tilde{\sigma}_N^4} \begin{pmatrix} \tilde{\mathbf{V}}_N^{*\prime} \mathbf{W}_{1N} \mathbf{Y}_N \\ \tilde{\mathbf{V}}_N^{*\prime} \mathbf{W}_{2N} \tilde{\mathbf{V}}_N \end{pmatrix}' \begin{pmatrix} \mathbb{S}_1 + \tilde{\mathbb{D}} & \mathbb{S}_3 \\ \mathbb{S}_3 & \mathbb{S}_2 \end{pmatrix}^{-1} \begin{pmatrix} \tilde{\mathbf{V}}_N' \mathbf{W}_{1N} \mathbf{Y}_N^* \\ \tilde{\mathbf{V}}_N' \mathbf{W}_{2N} \tilde{\mathbf{V}}_N \end{pmatrix}, \tag{7.17}$$

- $\tilde{\mathbf{V}}_N$ denotes the OLS residuals from regressing $\mathbf{Y}_N$ on $\mathbf{X}_N$,
- $\mathbb{S}_1 = \mathrm{tr}[(\mathbf{W}_{1N} + \mathbf{W}_{1N}')\mathbf{W}_{1N}]$, $\tilde{\mathbb{D}} = \tilde{\sigma}_N^{-2} \tilde{\eta}_N' \mathbf{M}_N \tilde{\eta}_N$, $\tilde{\eta}_N = \mathbf{W}_{1N} \mathbf{X}_N \tilde{\beta}_N$,
- $\mathbb{S}_2 = \mathrm{tr}[(\mathbf{W}_{2N} + \mathbf{W}_{2N}')\mathbf{W}_{2N}]$, $\mathbf{M}_N = I_N - \mathbf{X}_N(\mathbf{X}_N'\mathbf{X}_N)^{-1}\mathbf{X}_N'$,
- $\mathbb{S}_3 = \mathrm{tr}[(\mathbf{W}_{2N} + \mathbf{W}_{2N}')\mathbf{W}_{1N}]$,
- $\tilde{\beta}_N$ and $\tilde{\sigma}_N^2$ are the OLS estimators of β and σ^2, respectively.

Baltagi and Yang (2013a) present a standardized version of $\mathtt{LM}^{\mathtt{FI}}_{\mathtt{SE}}$, with a much improved finite sample performance, but not the other two tests.

Baltagi and Yang (2013b) give *outer-product-of-martingale-differences* (OPMD) variants of the three LM tests given in (7.15)-(7.17), robust to nonnormality and unknown heteroskedasticity:

$$\mathrm{LM}^{\mathtt{MD}}_{\mathtt{SAR}} = \frac{\tilde{\mathbf{V}}_N' \mathbf{W}_{1N} \mathbf{Y}_N}{(\tilde{\mathbf{V}}_N^{2\,\prime}\, \tilde{\xi}_{1N}^2)^{\frac{1}{2}}}, \tag{7.18}$$

$$\mathrm{LM}^{\mathtt{MD}}_{\mathtt{SED}} = \frac{\tilde{\mathbf{V}}_N' \mathbf{W}_{2N} \tilde{\mathbf{V}}_N}{(\tilde{\mathbf{V}}_N^{2\,\prime}\, \tilde{\xi}_{2N}^2)^{\frac{1}{2}}}, \text{ and} \tag{7.19}$$

$$\mathrm{LM}^{\mathtt{MD}}_{\mathtt{SLE}} = \begin{pmatrix} \tilde{\mathbf{V}}_N' \mathbf{W}_{1N} Y_N \\ \tilde{\mathbf{V}}_N' \mathbf{W}_{2N} \tilde{\mathbf{V}}_N \end{pmatrix}' \begin{pmatrix} \tilde{\mathbf{V}}_N^{2\,\prime} \tilde{\xi}_{1N}^{*2} & \tilde{\mathbf{V}}_N^{2\,\prime}(\tilde{\xi}_{1N} \otimes \tilde{\xi}_{2N}) \\ \sim & \tilde{\mathbf{V}}_N^{2\,\prime} \tilde{\xi}_{2N}^2 \end{pmatrix}^{-1} \begin{pmatrix} \tilde{\mathbf{V}}_N' \mathbf{W}_{1N} Y_N \\ \tilde{\mathbf{V}}_N' \mathbf{W}_{2N} \tilde{\mathbf{V}}_N \end{pmatrix}, \tag{7.20}$$

where $\tilde{\xi}_{1N} = (\mathbf{W}_{1N}^{l} + \mathbf{W}_{1N}^{u\prime})\tilde{\mathbf{V}}_N + \mathbf{M}_N \tilde{\eta}_N, \quad \tilde{\xi}_{2N} = (\mathbf{W}_{2N}^{l} + \mathbf{W}_{2N}^{u\prime})\tilde{\mathbf{V}}_N,$

$\mathbf{W}_{rN}^{l}$, $\mathbf{W}_{rN}^{u}$: lower- and upper-triangular matrices of $\mathbf{W}_{rN}$, $r = 1, 2$.

These tests may perform quite poorly in finite sample, but the OPMD ideas behind them are very important.

Batagi and Yang (3013b) went on to give OPMD-variants of the LM tests with **finite sample corrections**.

Define $\mathbb{A}_1 = \mathbf{M}_N\mathbf{W}_{1N}$ and $\mathbb{A}_2 = \mathbf{M}_N\mathbf{W}_{2N}\mathbb{M}$. Let $\mathbb{H}_r = \mathrm{diag}(\mathbb{A}_r)\mathrm{diag}(\mathbf{M}_N)^{-2}$ and $\mathbb{A}_r^\circ = \mathbb{A}_r - \mathbf{M}_N\mathbb{H}_r\mathbf{M}_N$, decomposed as $\mathbb{A}_r^\circ = \mathbb{A}_r^{\circ u} + \mathbb{A}_r^{\circ l} + \mathbb{A}_r^{\circ d}$, $r = 1, 2$.

- Let $\tilde{\xi}_{1N}^\circ = (\mathbb{A}_{1N}^{\circ u\prime} + \mathbb{A}_{1N}^{\circ l})\tilde{\mathbf{V}}_N + \mathbb{A}_{1N}^{\circ d}\tilde{\mathbf{V}}_N + M\tilde{\eta}_N$,
- and $\tilde{\xi}_{2N}^\circ = (\mathbb{A}_{2N}^{\circ u\prime} + \mathbb{A}_{2N}^{\circ l})\tilde{\mathbf{V}}_N + \mathbb{A}_{2N}^{\circ d}\tilde{\mathbf{V}}_N$.

The three OPMD-based tests are:

$$\mathrm{SLM}_{\mathrm{SL}}^{\mathrm{MD}} = \frac{\tilde{\mathbf{V}}_N'\mathbf{W}_{1N}\mathbf{Y}_N - \tilde{\mathbf{V}}_N'\mathbb{H}_1\tilde{\mathbf{V}}_N}{(\tilde{\mathbf{V}}_N^{2\,\prime}\ \tilde{\xi}_{1N}^{\circ 2})^{\frac{1}{2}}}, \tag{7.21}$$

$$\mathrm{SLM}_{\mathrm{SE}}^{\mathrm{MD}} = \frac{\tilde{\mathbf{V}}_N'(\mathbf{W}_{2N} - \mathbb{H}_2)\tilde{\mathbf{V}}_N}{(\tilde{\mathbf{V}}_N^{2\,\prime}\ \tilde{\xi}_{2N}^{\circ 2})^{\frac{1}{2}}}, \text{ and} \tag{7.22}$$

$$\mathrm{SLM}_{\mathrm{SLE}}^{\mathrm{MD}} = \mathbb{S}_N' \begin{pmatrix} \tilde{\mathbf{V}}_N^{2\,\prime}\tilde{\xi}_{1N}^{2\circ} & \tilde{\mathbf{V}}_N^{2\,\prime}(\tilde{\xi}_{1N}^\circ \odot \tilde{\xi}_{2N}^\circ) \\ \tilde{\mathbf{V}}_N^{2\,\prime}(\tilde{\xi}_{1N}^\circ \odot \tilde{\xi}_{2N}^\circ) & \tilde{\mathbf{V}}_N^{2\,\prime}\tilde{\xi}_{2N}^{\circ 2} \end{pmatrix}^{-1} \mathbb{S}_N, \tag{7.23}$$

where $\mathbb{S}_N = (\tilde{\mathbf{V}}_N'\mathbf{W}_{1N}\mathbf{Y}_N - \tilde{\mathbf{V}}_N'\mathbb{H}_1\tilde{\mathbf{V}}_N,\ \tilde{\mathbf{V}}_N'(\mathbf{W}_{2N} - \mathbb{H}_2)\tilde{\mathbf{V}}_N)'$. Under H_0, $\mathrm{SLM}_{\mathrm{SLD}}^{\mathrm{OPMD}} \xrightarrow{D} N(0, 1)$, $\mathrm{SLM}_{\mathrm{SED}}^{\mathrm{OPMD}} \xrightarrow{D} N(0, 1)$, and $\mathrm{SLM}_{\mathrm{SLE}}^{\mathrm{OPGMD}} \xrightarrow{D} \chi_2^2$.

Some Monte Carlo results are taken from Yang et al. (2016) to show

(*i*) finite sample performance of the QML and the bias-corrected QML estimators of the FE-SPD models,

(*ii*) finite sample performance of the LM tests for spatial effects.

For (i), the following model is used:

$$\begin{aligned} Y_{nt} &= \lambda_0 W_{1n} Y_{nt} + X_{1nt}\beta_{10} + X_{2nt}\beta_{20} + \boldsymbol{\mu}_{n0} + \alpha_{t0} l_n + U_{nt}, \\ U_{nt} &= \rho_0 W_{2n} U_{nt} + V_{nt},\ t = 1, \ldots, T. \end{aligned}$$

For (ii), the following model is used:

$$\begin{aligned} Y_{nt} &= \lambda_1 W_{1n} Y_{nt} + X_{1n}\beta_1 + X_{2n}\beta_2 + X_{3n}\beta_3 + \mu_n + u_{nt}, \\ u_{nt} &= \lambda_2 W_{2n} u_{nt} + \varepsilon_{nt}, \quad t = 1, \ldots, T. \end{aligned}$$

See Yang et al. (2016) for the detailed set-up of the experiments.

Table 7.1. Empirical Mean[rmse](sd) of Estimators of λ, 2FE-SPD Model with SLD

	(a) Queen Contiguity, REG1		(b) Group Interaction, REG2	
λ	$\hat{\lambda}_N$	$\hat{\lambda}_N^{bc2}$	$\hat{\lambda}_N$	$\hat{\lambda}_N^{bc2}$
	Normal Error, n=50, T=3			
.50	.484[.120](.119)	.502[.120](.120)	.469[.095](.089)	.497[.088](.088)
.25	.234[.142](.141)	.248[.143](.143)	.210[.130](.124)	.250[.123](.123)
.00	-.010[.158](.158)	.001[.161](.161)	-.049[.167](.159)	-.001[.160](.160)
-.25	-.258[.161](.161)	-.251[.164](.164)	-.303[.189](.182)	-.250[.184](.184)
-.50	-.504[.163](.163)	-.503[.166](.166)	-.565[.214](.204)	-.509[.208](.208)
	Normal Mixture, n=50, T=3			
.50	.483[.119](.117)	.500[.118](.118)	.470[.091](.086)	.498[.084](.084)
.25	.238[.139](.139)	.253[.141](.141)	.209[.128](.121)	.248[.120](.120)
.00	-.013[.155](.154)	-.002[.157](.157)	-.048[.160](.152)	-.001[.153](.153)
-.25	-.257[.158](.158)	-.251[.161](.161)	-.301[.188](.181)	-.248[.182](.182)
-.50	-.504[.163](.163)	-.503[.166](.166)	-.556[.206](.199)	-.500[.203](.203)
	Lognormal Error, n=50, T=3			
.50	.485[.111](.110)	.501[.111](.111)	.470[.090](.085)	.497[.083](.083)
.25	.239[.133](.133)	.253[.134](.134)	.212[.122](.116)	.249[.115](.115)
.00	-.010[.146](.146)	.001[.149](.149)	-.045[.154](.147)	.000[.147](.147)
-.25	-.255[.151](.151)	-.249[.154](.154)	-.302[.178](.171)	-.251[.173](.173)
-.50	-.498[.152](.152)	-.499[.155](.155)	-.556[.204](.196)	-.503[.200](.200)

Table 7.2. Empirical Mean[rmse](sd) of Estimators of ρ, 2FE-SPD Model with SED

	(a) Queen Contiguity, REG1		(b) Group Interaction, REG2	
ρ	$\hat{\rho}_N$	$\hat{\rho}_N^{bc2}$	$\hat{\rho}_N$	$\hat{\rho}_N^{bc2}$
	Normal Error, n=50, t=3			
.50	.481[.144](.142)	.500[.143](.143)	.457[.139](.132)	.503[.116](.116)
.25	.233[.171](.170)	.252[.171](.171)	.177[.202](.188)	.258[.167](.167)
.00	-.018[.190](.189)	-.001[.190](.190)	-.115[.266](.240)	-.004[.221](.221)
-.25	-.271[.202](.201)	-.255[.203](.203)	-.382[.299](.268)	-.250[.256](.256)
-.50	-.516[.203](.202)	-.503[.205](.205)	-.637[.321](.290)	-.496[.287](.287)
	Normal Mixture, n=50, T=3			
.50	.480[.139](.138)	.500[.138](.138)	.458[.137](.130)	.504[.114](.114)
.25	.233[.166](.165)	.252[.166](.166)	.168[.210](.194)	.251[.172](.172)
.00	-.016[.186](.185)	.002[.186](.186)	-.108[.258](.234)	.004[.214](.214)
-.25	-.267[.195](.194)	-.252[.196](.196)	-.381[.293](.262)	-.248[.251](.251)
-.50	-.511[.198](.197)	-.498[.200](.200)	-.636[.313](.282)	-.493[.280](.280)
	Lognormal Error, n=50, t=3			
.50	.483[.135](.133)	.504[.134](.134)	.454[.136](.128)	.502[.112](.112)
.25	.237[.160](.159)	.256[.161](.160)	.174[.196](.181)	.257[.160](.160)
.00	-.012[.179](.179)	.006[.180](.180)	-.105[.242](.218)	.009[.199](.199)
-.25	-.264[.186](.186)	-.248[.188](.188)	-.368[.273](.247)	-.233[.235](.235)
-.50	-.512[.191](.191)	-.499[.194](.194)	-.632[.305](.275)	-.489[.272](.272)

Table 7.3a. Empirical Sizes of LM Tests of $H_0 : \lambda = 0|\rho = 0$, FE-SPD with SLD

Normal Errors, W_{1n} = Group, $g = n^{0.5}$; XVal-B; $T = 3$.

	Heteroskedasticity $\propto$ group size					Heteroskedasticity = 1				
n	mean	sd	10%	5%	1%	mean	sd	10%	5%	1%
50	-0.3253	0.8610	.0607	.0209	.0032	-0.1970	0.9908	.1017	.0498	.0098
	-0.4452	0.9846	.1298	.0646	.0100	-0.2340	1.0235	.1186	.0591	.0112
	-0.0699	1.0249	.1096	.0534	.0075	-0.0453	1.0327	.1134	.0557	.0099
100	-0.2568	0.9231	.0817	.0372	.0056	-0.1633	0.9840	.0989	.0485	.0075
	-0.3465	0.9965	.1202	.0629	.0127	-0.1995	0.9999	.1102	.0547	.0107
	-0.0558	1.0059	.1038	.0536	.0098	-0.0311	1.0091	.1045	.0518	.0096
200	-0.2194	0.9466	.0851	.0364	.0063	-0.1599	0.9943	.1021	.0507	.0097
	-0.2834	1.0015	.1109	.0580	.0121	-0.1765	1.0046	.1052	.0547	.0113
	-0.0416	1.0123	.1027	.0504	.0105	-0.0180	1.0100	.1039	.0512	.0101
500	-0.1587	0.9665	.0904	.0434	.0081	-0.0901	0.9856	.0963	.0461	.0089
	-0.2023	1.0026	.1060	.0543	.0120	-0.0979	0.9891	.0987	.0467	.0091
	-0.0442	1.0086	.1025	.0515	.0107	0.0015	0.9913	.0966	.0458	.0089
1000	-0.1141	0.9576	.0869	.0426	.0088	-0.0705	0.9959	.0998	.0505	.0085
	-0.1472	1.0008	.1047	.0548	.0126	-0.0782	1.0006	.1030	.0512	.0092
	-0.0290	1.0043	.1023	.0525	.0124	-0.0120	1.0015	.1018	.0515	.0083

Note: Three rows under each n: $\text{LM}_{\text{SL}}^{\text{FI}}$, $\text{LM}_{\text{SL}}^{\text{MD}}$ and $\text{SLM}_{\text{SL}}^{\text{MD}}$.

Table 7.3b. Empirical Sizes of LM Tests of $H_0: \lambda = 0 | \rho = 0$, FE-SPD with SLD

Normal Mixture, W_{1n} = Group, $g = n^{0.5}$; XVal-B; $T = 3$.

	Heteroskedasticity $\propto$ group size					Heteroskedasticity = 1				
n	mean	sd	10%	5%	1%	mean	sd	10%	5%	1%
50	-0.3299	0.8416	.0577	.0190	.0024	-0.1623	0.9962	.1036	.0509	.0095
	-0.4366	0.9748	.1225	.0570	.0091	-0.1902	1.0287	.1179	.0588	.0089
	-0.0614	1.0211	.1062	.0497	.0062	-0.0066	1.0389	.1158	.0547	.0078
100	-0.2562	0.9227	.0785	.0350	.0067	-0.1706	0.9784	.0942	.0441	.0080
	-0.3378	1.0000	.1202	.0597	.0111	-0.2062	0.9999	.1062	.0524	.0083
	-0.0449	1.0136	.1040	.0507	.0092	-0.0380	1.0086	.1015	.0486	.0078
200	-0.2249	0.9235	.0770	.0349	.0063	-0.1542	0.9793	.0976	.0492	.0094
	-0.2839	0.9804	.1058	.0521	.0108	-0.1694	0.9928	.1047	.0518	.0102
	-0.0428	0.9920	.0950	.0484	.0103	-0.0112	0.9980	.0973	.0475	.0090
500	-0.1411	0.9710	.0948	.0452	.0079	-0.1102	1.0016	.1014	.0527	.0101
	-0.1835	1.0080	.1100	.0561	.0111	-0.1186	1.0039	.1037	.0521	.0106
	-0.0250	1.0146	.1047	.0546	.0097	-0.0192	1.0061	.1023	.0517	.0103
1000	-0.1230	0.9531	.0873	.0419	.0066	-0.0688	1.0029	.1009	.0529	.0095
	-0.1550	0.9993	.1011	.0542	.0108	-0.0764	1.0049	.1016	.0517	.0097
	-0.0366	1.0028	.0994	.0505	.0089	-0.0102	1.0061	.1019	.0515	.0104

Note: Three rows under each n: $\text{LM}_{\text{SL}}^{\text{FI}}$, $\text{LM}_{\text{SL}}^{\text{MD}}$ and $\text{SLM}_{\text{SL}}^{\text{MD}}$.

Table 7.3c. Empirical Sizes of LM Tests of $H_0 : \lambda = 0|\rho = 0$, FE-SPD with SLD

Lognormal Errors, W_{1n} = Group, $g = n^{0.5}$; XVal-B; $T = 3$.

	Heteroskedasticity ∝ group size					Heteroskedasticity = 1				
n	mean	sd	10%	5%	1%	mean	sd	10%	5%	1%
50	-0.3234	0.8164	.0554	.0186	.0030	-0.1856	0.9603	.0929	.0442	.0066
	-0.4302	0.9518	.1121	.0516	.0053	-0.2086	1.0116	.1107	.0503	.0069
	-0.0469	1.0001	.0988	.0447	.0057	-0.0293	1.0256	.1064	.0490	.0072
100	-0.2630	0.8978	.0716	.0324	.0055	-0.1404	0.9737	.0925	.0442	.0077
	-0.3345	0.9764	.1069	.0519	.0081	-0.1694	0.9988	.1022	.0488	.0079
	-0.0424	0.9938	.0966	.0432	.0068	-0.0039	1.0052	.0978	.0462	.0077
200	-0.2446	0.9243	.0814	.0375	.0063	-0.1699	0.9667	.0930	.0432	.0075
	-0.3003	0.9917	.1081	.0561	.0100	-0.1768	0.9834	.0964	.0466	.0068
	-0.0606	1.0058	.1000	.0480	.0088	-0.0216	0.9914	.0952	.0445	.0073
500	-0.1225	0.9450	.0836	.0393	.0069	-0.0776	0.9941	.0972	.0475	.0092
	-0.1650	0.9853	.0982	.0457	.0083	-0.0721	0.9968	.0993	.0474	.0082
	-0.0066	0.9921	.0968	.0465	.0075	0.0268	1.0020	.1003	.0464	.0079
1000	-0.0902	0.9596	.0868	.0398	.0080	-0.0622	0.9901	.0955	.0487	.0091
	-0.1186	1.0044	.1015	.0520	.0103	-0.0650	0.9938	.0974	.0468	.0079
	-0.0003	1.0079	.1011	.0496	.0105	0.0008	0.9955	.0986	.0482	.0080

Note: Three rows under each n: $\text{LM}_{\text{SL}}^{\text{FI}}$, $\text{LM}_{\text{SL}}^{\text{MD}}$ and $\text{SLM}_{\text{SL}}^{\text{MD}}$.

Table 7.4a. Empirical Sizes of LM Tests for $H_0 : \rho = 0 | \lambda = 0$, FE-SPD with SED

Normal Errors, W_{1n} = Group, $g = n^{0.5}$; XVal-B; $T = 3$.

	Heteroskedasticity $\propto$ group size					Heteroskedasticity = 1				
n	mean	sd	10%	5%	1%	mean	sd	10%	5%	1%
50	-0.3231	0.8613	.0524	.0173	.0043	-0.4076	0.9258	.0926	.0345	.0041
	-0.4803	1.0012	.1406	.0717	.0136	-0.5256	1.0103	.1499	.0816	.0170
	-0.1354	1.0579	.1224	.0632	.0126	-0.0887	1.0539	.1225	.0597	.0110
100	-0.2876	0.9175	.0773	.0295	.0048	-0.3306	0.9483	.0937	.0380	.0046
	-0.4094	1.0044	.1318	.0705	.0132	-0.4327	1.0129	.1378	.0732	.0146
	-0.1034	1.0239	.1120	.0580	.0100	-0.0874	1.0379	.1154	.0572	.0102
200	-0.2709	0.9169	.0739	.0285	.0052	-0.2827	0.9548	.0927	.0390	.0066
	-0.3835	0.9935	.1229	.0629	.0137	-0.3716	1.0051	.1273	.0676	.0147
	-0.0987	1.0152	.1073	.0548	.0100	-0.0668	1.0194	.1073	.0542	.0106
500	-0.2300	0.9333	.0790	.0334	.0063	-0.2451	0.9818	.1022	.0471	.0089
	-0.3352	1.0073	.1213	.0606	.0142	-0.3171	1.0163	.1229	.0654	.0156
	-0.0984	1.0155	.1067	.0542	.0105	-0.0773	1.0243	.1101	.0559	.0126
1000	-0.2367	0.9328	.0823	.0349	.0062	-0.1864	0.9716	.0978	.0447	.0077
	-0.3250	1.0003	.1168	.0608	.0145	-0.2437	0.9941	.1092	.0585	.0114
	-0.0891	1.0078	.1024	.0524	.0119	-0.0627	0.9999	.1015	.0517	.0096

Note: Three rows under each n: LM_{SE}^{FI}, LM_{SE}^{MD} and SLM_{SE}^{MD}.

Table 7.4b. Empirical Sizes of LM Tests for $H_0 : \rho = 0|\lambda = 0$, FE-SPD with SED

Normal Mixture, W_{1n} = Group, $g = n^{0.5}$; XVal-B; $T = 3$.

	Heteroskedasticity $\propto$ group size					Heteroskedasticity = 1				
n	mean	sd	10%	5%	1%	mean	sd	10%	5%	1%
50	-0.3185	0.8280	.0442	.0146	.0034	-0.4324	0.8930	.0913	.0321	.0024
	-0.4623	0.9745	.1248	.0608	.0086	-0.5414	0.9841	.1448	.0752	.0143
	-0.1098	1.0331	.1134	.0558	.0076	-0.0952	1.0373	.1141	.0547	.0088
100	-0.2767	0.9077	.0705	.0261	.0053	-0.3434	0.9299	.0878	.0369	.0060
	-0.3910	0.9956	.1272	.0624	.0113	-0.4399	0.9979	.1316	.0685	.0137
	-0.0790	1.0188	.1051	.0523	.0082	-0.0888	1.0267	.1081	.0548	.0105
200	-0.2822	0.9041	.0740	.0265	.0047	-0.2984	0.9253	.0813	.0345	.0054
	-0.3898	0.9922	.1211	.0613	.0129	-0.3816	0.9788	.1189	.0596	.0103
	-0.1015	1.0180	.1092	.0540	.0099	-0.0733	0.9943	.0975	.0457	.0090
500	-0.2451	0.9134	.0743	.0275	.0049	-0.2293	0.9686	.0942	.0431	.0064
	-0.3471	0.9970	.1170	.0597	.0133	-0.2980	1.0024	.1161	.0580	.0112
	-0.1091	1.0068	.1033	.0503	.0097	-0.0574	1.0110	.1052	.0492	.0094
1000	-0.2318	0.9306	.0814	.0319	.0057	-0.1797	0.9751	.0953	.0434	.0083
	-0.3199	1.0017	.1189	.0615	.0140	-0.2360	0.9952	.1100	.0551	.0100
	-0.0838	1.0094	.1050	.0528	.0109	-0.0546	1.0010	.1029	.0493	.0096

Note: Three rows under each n: $\text{LM}_{\text{SE}}^{\text{FI}}$, $\text{LM}_{\text{SE}}^{\text{MD}}$ and $\text{SLM}_{\text{SE}}^{\text{MD}}$.

Table 7.4c. Empirical Sizes of LM Tests for $H_0 : \rho = 0 | \lambda = 0$, FE-SPD with SED

Lognormal Errors, W_{1n} = Group, $g = n^{0.5}$; XVal-B; $T = 3$.

	Heteroskedasticity ∝ group size					Heteroskedasticity = 1				
n	mean	sd	10%	5%	1%	mean	sd	10%	5%	1%
50	-0.3231	0.8057	.0382	.0131	.0039	-0.3989	0.8701	.0706	.0242	.0032
	-0.4800	0.9669	.1253	.0583	.0073	-0.5309	0.9812	.1410	.0666	.0105
	-0.1099	1.0207	.1058	.0474	.0071	-0.0607	1.0242	.1035	.0470	.0066
100	-0.2792	0.8806	.0607	.0245	.0055	-0.3250	0.9069	.0763	.0333	.0068
	-0.4103	0.9920	.1252	.0614	.0091	-0.4399	0.9887	.1281	.0590	.0115
	-0.0788	1.0141	.1031	.0490	.0070	-0.0709	1.0129	.1017	.0462	.0070
200	-0.2910	0.8975	.0653	.0259	.0063	-0.2939	0.9230	.0801	.0339	.0068
	-0.4155	0.9985	.1305	.0641	.0113	-0.3968	0.9944	.1215	.0618	.0126
	-0.1160	1.0176	.1072	.0501	.0082	-0.0785	1.0078	.1024	.0470	.0083
500	-0.2188	0.9046	.0684	.0286	.0052	-0.2354	0.9472	.0879	.0370	.0060
	-0.3245	0.9938	.1153	.0571	.0119	-0.3145	0.9945	.1128	.0565	.0109
	-0.0810	1.0048	.1030	.0512	.0089	-0.0627	1.0004	.0990	.0462	.0081
1000	-0.2181	0.9361	.0806	.0316	.0055	-0.2000	0.9766	.0960	.0457	.0076
	-0.3127	1.0109	.1184	.0585	.0129	-0.2668	1.0107	.1173	.0595	.0119
	-0.0739	1.0188	.1072	.0524	.0100	-0.0642	1.0158	.1071	.0534	.0098

Note: Three rows under each n: $\text{LM}_{\text{SE}}^{\text{FI}}$, $\text{LM}_{\text{SE}}^{\text{MD}}$ and $\text{SLM}_{\text{SE}}^{\text{MD}}$.

Table 7.5a. Empirical Sizes of LM Tests of $H_0 : \lambda = \rho = 0$, FE-SPD with SLE

Normal Errors; W_{1n} = Queen, $r = 5$; W_{2n} = Group, $g = n^{0.5}$; XVal-B; $T = 3$.

	Heteroskedasticity ∝ group size					Heteroskedasticity = 1				
n	mean	sd	10%	5%	1%	mean	sd	10%	5%	1%
50	1.8475	1.7698	.0702	.0306	.0061	1.9882	1.8950	.0880	.0431	.0084
	2.2877	2.0462	.1310	.0620	.0100	2.2342	2.0594	.1236	.0624	.0104
	2.1617	1.9856	.1175	.0539	.0077	2.1093	1.9760	.1092	.0523	.0084
100	1.8967	1.8868	.0731	.0348	.0086	1.9887	1.8975	.0850	.0397	.0082
	2.2495	2.1328	.1310	.0661	.0124	2.2646	2.1610	.1286	.0654	.0136
	2.0986	2.0037	.1101	.0528	.0095	2.1072	2.0321	.1107	.0560	.0106
200	1.8844	1.8150	.0794	.0345	.0062	1.9774	1.9044	.0896	.0435	.0084
	2.2110	2.1588	.1236	.0628	.0130	2.1567	2.0882	.1170	.0620	.0117
	2.0704	2.0488	.1099	.0534	.0111	2.0467	1.9972	.1059	.0526	.0097
500	1.9370	1.9192	.0848	.0390	.0087	2.0093	2.0198	.0982	.0463	.0094
	2.1424	2.1107	.1222	.0613	.0114	2.1147	2.1101	.1144	.0576	.0126
	2.0377	2.0138	.1027	.0512	.0105	2.0492	2.0549	.1046	.0538	.0118
1000	1.9527	1.9511	.0907	.0444	.0090	1.9837	1.9384	.0952	.0434	.0086
	2.0930	2.0803	.1141	.0591	.0112	2.0706	2.0503	.1041	.0529	.0115
	2.0383	2.0335	.1065	.0532	.0107	2.0098	1.9949	.0999	.0491	.0108

Note: Three rows under each n: $\text{LM}^{\text{FI}}_{\text{SLE}}$, $\text{LM}^{\text{MD}}_{\text{SLE}}$ and $\text{SLM}^{\text{MD}}_{\text{SLE}}$.

Table 7.5b. Empirical Sizes of LM Tests of $H_0: \lambda = \rho = 0$, FE-SPD with SLE

Normal Mixture; W_{1n} = Queen, $r = 5$; W_{2n} = Group, $g = n^{0.5}$; XVal-B; $T = 3$.

	Heteroskedasticity $\propto$ group size					Heteroskedasticity = 1				
n	mean	sd	10%	5%	1%	mean	sd	10%	5%	1%
50	1.7835	1.7222	.0626	.0268	.0059	1.9417	1.9156	.0851	.0398	.0094
	2.2488	1.9511	.1190	.0563	.0068	2.2105	1.9364	.1172	.0554	.0066
	2.1386	1.9122	.1071	.0511	.0067	2.0945	1.8742	.1034	.0475	.0052
100	1.8511	1.7889	.0697	.0341	.0069	1.9745	1.8478	.0859	.0374	.0071
	2.2567	2.0837	.1243	.0618	.0112	2.2528	2.0556	.1230	.0592	.0109
	2.0949	1.9784	.1095	.0486	.0089	2.0979	1.9381	.1061	.0492	.0074
200	1.8491	1.8272	.0767	.0348	.0070	1.9458	1.8929	.0867	.0386	.0082
	2.1792	2.1047	.1181	.0621	.0128	2.1271	2.0206	.1137	.0542	.0085
	2.0437	1.9938	.1048	.0530	.0086	2.0275	1.9425	.1012	.0458	.0081
500	1.8883	1.8336	.0791	.0362	.0073	1.9872	1.9464	.0945	.0453	.0083
	2.1018	2.0185	.1092	.0561	.0101	2.0992	2.0569	.1114	.0565	.0104
	2.0081	1.9430	.0998	.0492	.0076	2.0345	2.0052	.1029	.0532	.0090
1000	1.9304	1.9345	.0864	.0417	.0091	2.0028	2.0047	.0985	.0512	.0101
	2.0690	2.0586	.1039	.0540	.0125	2.0891	2.1085	.1114	.0575	.0122
	2.0211	2.0064	.1008	.0491	.0105	2.0373	2.0604	.1070	.0549	.0103

Note: Three rows under each n: $\text{LM}^{\text{FI}}_{\text{SLE}}$, $\text{LM}^{\text{MD}}_{\text{SLE}}$ and $\text{SLM}^{\text{MD}}_{\text{SLE}}$.

Table 7.5c. Empirical Sizes of LM Tests of $H_0 : \lambda = \rho = 0$, FE-SPD with SLE

Lognormal Errors; W_{1n} = Queen, $r = 5$; W_{2n} = Group, $g = n^{0.5}$; XVal-B; $T = 3$.

	Heteroskedasticity ∝ group size					Heteroskedasticity = 1				
n	mean	sd	10%	5%	1%	mean	sd	10%	5%	1%
50	1.6484	1.6401	.0499	.0246	.0054	1.8401	1.9910	.0724	.0346	.0089
	2.2424	1.9181	.1149	.0534	.0060	2.2157	1.8932	.1122	.0486	.0065
	2.0917	1.8562	.0996	.0447	.0053	2.0671	1.8043	.0956	.0398	.0052
100	1.7922	1.8153	.0688	.0321	.0074	1.8906	1.8987	.0797	.0385	.0081
	2.2755	2.0395	.1235	.0591	.0105	2.2403	2.0305	.1188	.0579	.0099
	2.0908	1.9104	.1002	.0467	.0076	2.0575	1.8992	.0988	.0484	.0076
200	1.7899	1.7512	.0690	.0307	.0061	1.9355	1.9223	.0874	.0407	.0092
	2.1999	2.0088	.1174	.0571	.0094	2.1670	1.9633	.1133	.0531	.0075
	2.0485	1.9124	.1017	.0489	.0069	2.0503	1.8708	.1017	.0446	.0048
500	1.8536	1.9127	.0785	.0357	.0092	1.9202	1.8952	.0838	.0384	.0082
	2.1259	2.0422	.1127	.0553	.0108	2.0790	1.9645	.1002	.0508	.0084
	2.0156	1.9389	.0998	.0473	.0086	2.0117	1.9100	.0939	.0462	.0080
1000	1.9047	1.9584	.0856	.0436	.0089	1.9925	2.0059	.0999	.0480	.0093
	2.0683	1.9870	.1072	.0489	.0096	2.1012	2.0611	.1115	.0559	.0118
	2.0159	1.9403	.1010	.0465	.0079	2.0424	2.0051	.1036	.0512	.0096

Note: Three rows under each n: $\text{LM}_{\text{SLE}}^{\text{FI}}$, $\text{LM}_{\text{SLE}}^{\text{MD}}$ and $\text{SLM}_{\text{SLE}}^{\text{MD}}$.

7.5. An Empirical Application of FE-SPD Model

To facilitate the practical applications of the proposed methods, we provide an empirical illustration using the well known data set on public capital productivity of Munnell (1990).

- The dataset gives indicators related to public capital productivity for 48 US states observed over 17 years (1970-1986).
- The dataset can be downloaded from http://pages.stern.nyu.edu/∼wgreene/Text/Edition6/tablelist6.htm
- This dataset has been extensively used for illustrating the applications of the regular panel data models (see, e.g., Baltagi, 2013).
- In the spatial framework, it was used by Millo and Piras (2012) for illustrating the QML and GMM estimation of fixed effects and random effects spatial panel data models,
- and by Yang et al. (2016, Supplementary material) for illustrating the bias-correction and refined inferences for the FE-SPD models.

In Munnell (1990), the empirical model specified is a Cobb-Douglas production function of the form:

$$\ln Y = \beta_0 + \beta_1 \ln K_1 + \beta_2 \ln K_2 + \beta_3 \ln L + \beta_4 \text{Unemp} + \epsilon,$$

with state specific fixed effects, where

- Y is the gross social product of a given state,
- K_1 is public capital,
- K_2 is private capital,
- L is labour input and
- Unemp is the state unemployment rate.

This model is now extended by adding spatial effects and/or spatial Durbin effects. The spatial weights matrix W takes a contiguity form with its (i, j)th element being 1 if states i and j share a common border, otherwise 0. The final W is row normalized. For models with more than one spatial term, the corresponding $W's$ are taken to be the same.

- Table 7.6 gives the QMLEs and second-order bias-corrected QMLEs of model parameters for the full dataset spanning over the 17 years,
- fitted using the 2FE-SPD model with five different types of spatial specifications: SL, SE, SLE, Durbin-SL and Durbin-SE, see Lecture 4 for the bias-correction methodology.
- Spatial effect (SLD or SED) is highly significant when appeared in the model alone; But in the the model with both SLD and SED, the SLD effect is not significant.
- It is interesting to note that the spatial Durbin term $W \log(emp)$ is highly significant in the Durdin-SLD model but not significant in the Durbin-SED model.
- When the full dataset is considered, $N = (n-1)(T-1) = 752$ is relatively large, the difference between the original QMLE-based results and the bias-corrected results is not so much. This is in line with the theoretical results on the consistency of the QMLEs.

Table 7.6: QMLEs and BC-QMLEs for the SPD-FE Models Based on the Full Data (Years 1970-86).

	SLE		SLD		SED		Durbin-SLD		Durbin-SED	
	$\hat{\theta}_{\mathrm{ML}}$	$\hat{\theta}_{\mathrm{ML}}^{\mathrm{bc2}}$	$\hat{\theta}_{\mathrm{ML}}$	$\hat{\theta}_{\mathrm{ML}}^{\mathrm{bc2}}$	$\hat{\theta}_{\mathrm{ML}}$	$\hat{\theta}_{\mathrm{ML}}^{\mathrm{bc2}}$	$\hat{\theta}_{\mathrm{ML}}$	$\hat{\theta}_{\mathrm{ML}}^{\mathrm{bc2}}$	$\hat{\theta}_{\mathrm{ML}}$	$\hat{\theta}_{\mathrm{ML}}^{\mathrm{bc2}}$
SLD (λ)	0.0270	0.0294	0.2100	0.2129			0.4124	0.4205		
t-ratio	0.7037	0.7563	7.3923	7.4118			9.5186	9.5559		
SED (ρ)	0.4068	0.4095			0.4374	0.4403			0.4101	0.4168
t-ratio	7.5937	7.4208			10.2813	10.2547			9.4120	9.4037
$\ln(K_1)$	-0.0145	-0.0145	-0.0352	-0.0352	-0.0122	-0.0121	-0.0090	-0.0088	-0.0184	-0.0183
t-ratio	-0.5599	-0.5697	-1.3637	-1.3771	-0.4749	-0.4817	-0.3420	-0.3430	-0.6867	-0.6960
$\ln(K_2)$	0.1553	0.1553	0.1585	0.1583	0.1548	0.1547	0.1591	0.1591	0.1662	0.1662
t-ratio	5.8638	5.8822	5.9803	6.1402	5.8581	5.8662	5.9888	6.0288	6.1140	6.1957
$\ln(L)$	0.7555	0.7553	0.6824	0.6812	0.7584	0.7583	0.7514	0.7515	0.7539	0.7539
t-ratio	25.7262	25.3540	22.8939	22.5454	26.1169	25.7695	25.1208	24.9504	25.6309	25.4225
Unemp	-0.0012	-0.0012	-0.0015	-0.0015	-0.0012	-0.0012	-0.0006	-0.0006	-0.0009	-0.0009
t-ratio	-2.3652	-2.3935	-3.1327	-3.1614	-2.3511	-2.3807	-1.1295	-1.1574	-1.7158	-1.7599
$W\ln(K_1)$							-0.0567	-0.0558	-0.0750	-0.0745
t-ratio							-1.1809	-1.1734	-1.3044	-1.3110
$W\ln(K_2)$							0.0066	0.0046	0.0901	0.0897
t-ratio							0.1391	0.0950	1.5161	1.4757
$W\ln(L)$							-0.3159	-0.3228	-0.0130	-0.0141
t-ratio							-5.8105	-5.7999	-0.2559	-0.2758
W Unemp							-0.0013	-0.0013	-0.0017	-0.0017
t-ratio							-1.5365	-1.5334	-1.7525	-1.7440

- Table 7.7 gives the same results for a shorter time interval concentrating on the years 1982-84, allowing us to see the necessity of bias-correction and the effectiveness of the bias-correction methods, when the sample size is not so large (here, $N = 94$).
- As can be seen from the estimation results of Table 7.7, there is a clear difference between the original QMLE-based results and the bias-corrected results.
- Point estimates of the bias-corrected QMLEs of the spatial parameters can be significantly larger than the corresponding QMLEs, in line with the theoretical results that the QMLEs are downward biased.
- The bias-corrected t-ratios for the spatial effects and the covariate effects can be noticeably smaller compared to the original t-ratios, showing that the original QMLE-based inferences can be conservative (or over rejection) when sample size is not large, in line with the theoretical results reported in Lecture 4.

Table 7.7: QMLEs and BC-QMLEs for the SPD-FE Models Based on a Subset Data (Years 1982-84).

	SLE		SLD		SED		Durbin-SAR		Durbin-SED	
	$\hat{\theta}_{\mathrm{ML}}$	$\hat{\theta}^{\mathrm{bc2}}_{\mathrm{ML}}$	$\hat{\theta}_{\mathrm{ML}}$	$\hat{\theta}^{\mathrm{bc2}}_{\mathrm{ML}}$	$\hat{\theta}_{\mathrm{ML}}$	$\hat{\theta}^{\mathrm{bc2}}_{\mathrm{ML}}$	$\hat{\theta}_{\mathrm{ML}}$	$\hat{\theta}^{\mathrm{bc2}}_{\mathrm{ML}}$	$\hat{\theta}_{\mathrm{ML}}$	$\hat{\theta}^{\mathrm{bc2}}_{\mathrm{ML}}$
SLD (λ)	0.0552	0.0524	0.3074	0.3231			0.4963	0.5635		
t-ratio	0.4529	0.3948	4.0296	4.0207			4.4443	4.6018		
SED (ρ)	0.5516	0.5940			0.6160	0.6371			0.5230	0.5788
t-ratio	4.0558	3.8449			6.2920	6.1987			4.7379	4.5433
$\ln(K_1)$	-0.2469	-0.2361	-0.2839	-0.2799	-0.2322	-0.2262	-0.1069	-0.1108	-0.1168	-0.1161
t-ratio	-2.3605	-2.0841	-3.3297	-3.1705	-2.1801	-1.9933	-0.9088	-0.9007	-1.0261	-0.9876
$\ln(K_2)$	0.5663	0.5368	0.5132	0.4902	0.5522	0.5344	0.3309	0.3325	0.4619	0.4768
t-ratio	2.4170	2.1668	2.4694	2.2659	2.4118	2.2126	1.3570	1.3184	1.9837	1.9932
$\ln(L)$	1.1873	1.1880	1.1149	1.1113	1.1796	1.1795	1.1393	1.1385	1.1046	1.0963
t-ratio	13.9952	13.7878	12.7139	12.4853	14.2798	14.2521	13.1989	12.8462	12.1188	11.6417
Unemp	-0.0009	-0.0008	-0.0014	-0.0014	-0.0008	-0.0008	-0.0010	-0.0010	-0.0015	-0.0015
t-ratio	-1.0818	-1.0203	-1.7243	-1.6917	-1.0505	-1.0297	-1.3149	-1.2655	-1.7725	-1.8093
$W\ln(K_1)$							-0.0698	-0.0302	-0.1609	-0.1305
t-ratio							-0.3984	-0.1663	-0.7779	-0.6117
$W\ln(K2)$							0.3929	0.3195	0.9698	0.9718
t-ratio							1.0732	0.8280	2.3128	2.2267
$W\lg(L)$							-0.6881	-0.7761	-0.2377	-0.2664
t-ratio							-3.5131	-3.7384	-1.2768	-1.3755
W *Unemp*							-0.0023	-0.0022	-0.0034	-0.0033
t-ratio							-1.5803	-1.4514	-1.9087	-1.8428

When the unobserved individual and time specific effects $\boldsymbol{\mu}_n$ and $\boldsymbol{\alpha}_T$ are uncorrelated with IT-varying regressors, they can be treated as random vectors of iid elements with means 0 and variances $\sigma^2_{\mu 0}$ and $\sigma^2_{\alpha 0}$.

In this case, a set of time-invariant regressors including the constant term, Z_n, can be added to the Model (7.1) to give a RE-SPD model:

$$\begin{aligned} Y_{nt} &= \lambda_0 W_{1n} Y_{nt} + X_{nt}\beta_0 + Z_n\gamma_0 + \boldsymbol{\mu}_n + \alpha_t l_n + U_{nt}, \\ & U_{nt} = \rho_0 W_{2n} U_{nt} + V_{nt},\ \ t = 1, 2, \ldots, T. \end{aligned} \tag{7.24}$$

Various random effects specifications have been considered in the literature, but none contain the time effects α_t in the model.

- Anselin (1988), Baltagi et al. (2003): Model (7.24) without $W_{1n}Y_{nt}$ and α_t;
- Kapoor et al. (2007): $Y_{nt} = X_{nt}\beta_0 + Z_n\gamma_0 + U_{nt}$, $U_{nt} = \rho_0 W_{2n} U_{nt} + \boldsymbol{\mu}_n + V_{nt}$.
- Baltagi et al. (2013): $Y_{nt} = X_{nt}\beta_0 + Z_n\gamma_0 + U_{nt}$,
  $U_{nt} = \rho_{10} W_{2n} U_{nt} + V_{nt},\ \ \boldsymbol{\mu}_n = \rho_{20}\boldsymbol{\mu}_n W_{3n} + e_n.$

- Fingleton (2008): $Y_{nt} = \lambda_0 W_{1n} Y_{nt} + X_{nt}\beta_0 + Z_n\gamma_0 + U_{nt}$,
$$U_{nt} = (I_n + \rho_0 W_{2n})\xi_n, \;\; \xi_n = \boldsymbol{\mu}_n + V_{nt}.$$
- Lee and Yu (2012) provide a general model that embeds all the models above except (7.24) which contains random time effects $\{\alpha_t\}$:

$$\begin{aligned} Y_{nt} &= \lambda_{10} W_{1n} Y_{nt} + X_{nt}\beta_0 + Z_n\gamma_0 + \boldsymbol{\mu}_n + U_{nt}, \\ & U_{nt} = \lambda_{20} W_{2n} U_{nt} + (I_n + \delta_{10} M_{1n}) V_{nt}, \\ & V_{nt} = \rho_0 V_{n,t-1} + e_{nt}, t = 1, \ldots, T \\ & \boldsymbol{\mu}_n = \lambda_{30} W_{3n} \boldsymbol{\mu} + (I_n + \delta_{20} M_{2n})\epsilon_n. \end{aligned} \quad (7.25)$$

1. See Lee and Yu (2012) for the QML estimation of Model (7.25).
2. However, method for estimating the robust VC matrix of QMLEs of Model (7.25) is not given. Difficulty lies in the estimation of 3rd and 4th moments of two error components.
3. QML estimation of Model (7.24) has not been formally considered.
4. See Lee and Yu (2010b, 2015) for surveys on SPD models.

The Model (7.24) can be extended by allowing $\boldsymbol{\mu}_n$ to correlate with X_{nt} linearly. In particular,

$$\boldsymbol{\mu}_n = \delta_0 1_n + \bar{X}_n \beta_1 + \epsilon, \tag{7.26}$$

where $\bar{X}_n$ is the time mean of X_{nt}, as in Mundlak (1978).

1. Combining (7.26) with (7.24), the CRE-SPD model has the same form as the RE-SPD model (7.24), and hence can be estimated in the same manner.
2. Difficulty lies again in the estimation of the robust VC matrix of the QMLEs of the model.
3. Under nonnormality, the variance of the score function at the true parameters values involve 3rd and 4th moments of all three error components, and methods for consistent estimation of these higher moments are is not available.

References

- Anselin, L., 1988. *Spatial econometrics: Methods and models*. Dordrecht: Kluwer Academic.
- Baltagi, B. H., Egger, P., Pfaffermayr, M., 2013. A generalized spatial panel model with random effects. *Econometric Reviews* 32, 650-685. Publishers.
- Baltagi, B., Yang, Z. L., 2013a. Standardized LM tests for spatial error dependence in linear or panel regressions. *The Econometrics Journal* 16 103-134.
- Baltagi, B., Yang, Z. L., 2013b. Heteroskedasticity and non-normality robust LM tests of spatial dependence. *Regional Science and Urban Economics* 43, 725-739.
- Debarsy, N., C. Ertur, 2010. Testing for spatial autocorrelation in a fixed effects panel data model. *Regional Science and Urban Economics*, 40, 453-470.
- Lee, L. F., Yu, J. 2010a. Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics* 154, 165-185.

- Lee, L. F., Yu, J., 2010b. Some recent developments in spatial panel data models. *Regional Science and Urban Economics* 40, 255-271.
- Lee, L. F., Yu, J., 2012. Spatial panels: random components vs. fixed effects. *International Economic Review* 53, 1369-1412.
- Lee, L. F., Yu, J., 2015. Spatial panel data models. In: Baltagi, B. H. (Eds.), *The Oxford Handbook of Panel Data*. Oxford University Press, Oxford, pp. 363-401.
- Fingleton, B., 2008b. A generalized method of moments estimator for a spatial panel model with an endogenous spatial lag and spatial moving average errors. *Spatial Economic Analysis* 3, 27-44.
- Kapoor, M., Kelejian, H. H., Prucha, I. R., 2007. Panel data models with spatially correlated error components. *Journal of Econometrics* 140, 97-130.
- Mundlak, Y., 1978. On the pooling of time series and cross section data. *Econometrica* 46, 69-85.
- Yang, Z. L., Yu, J. H, Liu, S. F. (2016). Bias correction and refined inferences for fixed effects spatial panel data models. *Regional Science and Urban Economics* 61, 52-72.