Chapter 10: Spatial Dynamic Panel Data Models

Zhenlin Yang

School of Economics, Singapore Management University

ECON6002 Panel Data and Spatial Econometrics
Term II 2020-21
10.1 Introduction

The materials presented in this chapter are drawn from Yang (2018a, JOE) and its Supplement Yang (2018b). To simplify the notation, we suppress the subscript $n$ in a vector or a matrix. Consider the following spatial dynamic panel data (SDPD) model of the form:

$$
\begin{align*}
    y_t &= \rho y_{t-1} + \lambda_1 W_1 y_t + \lambda_2 W_2 y_{t-1} + X_t' \beta + Z \gamma + \mu + \alpha_t 1_n + u_t, \\
    u_t &= \lambda_3 W_3 u_t + v_t, \quad t = 1, \ldots, T, 
\end{align*}
$$

(1)

where for $r = 1, 2, 3$, $W_r$ are the given $n \times n$ spatial weight matrices, $\lambda_r$ are, respectively, the spatial lag (SL), space-time lag (STL), and spatial error (SE) parameter,

- $y_t$: $n \times 1$ vector of response
- $X_t$: $n \times p$ matrix of time-varying regressors,
- $\mu$: $n \times 1$ vector of individual-specific effects,
- $v_t$: $n \times 1$ vector of idiosyncratic errors, iid$(0, \sigma^2_v)$. 
We focus on the fixed-effects (FE) SDPD model, i.e., $\mu$ is allowed to be correlated with $X_t$ in an arbitrary manner, leading to an

- FE-SDPD model with $SL$, $STL$, and $SE$ dependence, or simply $STLE$;
- or FE-SDPD model with $SL$ and $SE$, or $SLE$, by setting $\lambda_2 = 0$;
- or FE-SDPD model with $SL$ and $STL$ by setting $\lambda_3 = 0$;
- or FE-SDPD model with only $SL$ by setting $\lambda_2 = \lambda_3 = 0$;
- or FE-SDPD model with only $SE$ by setting $\lambda_1 = \lambda_3 = 0$.

We consider the large-$n$ and small-$T$ setting (*short panels*), and assume:

1. data collection starts from time point $t = 0$,
2. **initial observations**, $\{y_0, X_0, Z\}$, are available,
3. process started $m$ periods before $t = 0$ from positions $y_{-m}$, treated as **exogenous**, with $m$ being finite (unknown) or infinite.
Note:

1. When \( m = 0 \), \( y_0 \) is exogenous; when \( m \geq 1 \), \( y_0 \) is endogenous, and when \( m = \infty \), processes have become stationary.

2. Model (1) can be extended to allow for higher-order spatial lags;

3. It can also be reduced to some specific models:
   - Model with \( \lambda_1 \) and \( \lambda_2 \) only: Yu et al. (2008), large \( n \) and large \( T \);
   - Model with \( \lambda_3 \) only: Su and Yang (2015), large \( n \) and small \( T \);
   - Model with \( \lambda_1 \) only: Elhorst (2010), large \( n \) and small \( T \).

4. The methods introduced in this chapter can be applied to all of the above models, and are not restricted to small \( T \).

The methods provide a unified framework for estimating short SDPD models, free from the initial conditions and robust against nonnormality.
Challenges in QML-type estimation of FE-SDPD model with small $T$

1. **Incidental parameters problem**: number of parameters increases with sample size;

2. **Initial values problem**: the distribution of the vector of initial observations depends on the past unobservables.

**Pros and cons with QML and GMM:**

- QML estimation is more efficient than GMM estimation;
- QML estimation needs the distribution of $y_0$, or $\Delta y_1$ for setting up the unconditional (quasi) likelihood;
- The distribution of $y_0$ or $\Delta y_1$ involves unobservables (e.g., $X_{-1}, X_{-2}, \ldots, y_{-m}$), ⇒ a proper ‘model’ for $y_0$ or $\Delta y_1$ is required;
- Such a model may depend on unknown process start time ‘$-m$’, and may need stronger conditions on $X_t$;
- The traditional modeling strategy for $y_0$ or $\Delta y_1$ may not work for models with spatial lags.
Literature related to SDPD models with small $T$

This chapter,

1. introduces a unified method for estimating the FE-SDPD models with short panels, the *M-estimation*, which is free from the specifications of the initial conditions and robust against nonnormality of errors;

2. presents results for consistency and asymptotic normality of the M-estimators;

3. introduces robust method of estimating the VC matrix of the M-estimators, free initial-conditions and allowing nonnormal errors;

4. presents results for consistency of the VC-matrix estimator;

5. presents Monte Carlo results for the finite sample performance of the methods introduced;

6. presents an empirical application to illustrate the proposed method.
Recall: M-Estimator or Zero-Estimator. The term M-estimation was coined by Huber (1964) to mean the maximum likelihood type estimation. It can be defined as either

1. the solution of a maximization problem:

\[ \hat{\psi}_n = \arg \max \{ Q_n(\psi) \}; \]

2. or the root of a set of estimating equations:

\[ \hat{\psi}_n = \arg \{ S_n(\psi) = 0 \}. \]

The latter is also called zero estimator as it makes the estimating equation zero. See van der Vaart (1998, *Asymptotic Statistics*).

See also Huber (1981, *Robust Statistics*).
Conditional QML estimation

Taking first-difference of Model (1) to eliminate the fixed effects $\mu$:

$$
\Delta y_t = \rho \Delta y_{t-1} + \lambda_1 W_1 \Delta y_t + \lambda_2 W_2 \Delta y_{t-1} + \Delta X_t \beta + \Delta u_t,
$$

$$
\Delta u_t = \lambda_3 W_3 \Delta u_t + \Delta v_t, \quad t = 2, 3, \cdots, T. \tag{2}
$$

**Note:** time-invariant variables $Z$ are also eliminated, and the terms corresponding to $\alpha_t$ are merged into $X_t$, as $T$ is fixed.

Stacking these vectors, the model is written in matrix form:

$$
\Delta Y = \rho \Delta Y_{-1} + \lambda_1 W_1 \Delta Y + \lambda_2 W_2 \Delta Y_{-1} + \Delta X \beta + \Delta u,
$$

$$
\Delta u = \lambda_3 W_3 \Delta u + \Delta v, \tag{3}
$$

- $W_r = I_{T-1} \otimes W_r, \ r = 1, 2, 3; \ \otimes$: Kronecker product; $I_k$: identity matrix.

Denote: $\psi = (\beta', \sigma_v^2, \rho, \lambda')', \ \lambda = (\lambda_1, \lambda_2, \lambda_3)'$, $\theta = (\beta', \rho, \lambda_1, \lambda_2)'$.

$$
B_1(\lambda_1) = I_n - \lambda_1 W_1, \quad B_2(\rho, \lambda_2) = \rho I_n + \lambda_2 W_2, \quad B_3(\lambda_3) = I_n - \lambda_3 W_3.
$$

$$
B_r(\lambda_r) = I_{T-1} \otimes B_r(\lambda_r), \ r = 1, 3, \quad B_2(\rho, \lambda_2) = I_{T-1} \otimes B_2(\rho, \lambda_2).
$$
The quasi Gaussian loglikelihood of $\psi$, as if $\Delta y_1$ is exogenous is:

$$
\ell(\psi) = -\frac{n(T-1)}{2} \log(\sigma^2_v) - \frac{1}{2} \log |\Omega(\lambda_3)| + \log |B_1(\lambda_1)| - \frac{1}{2\sigma^2_v} \Delta u(\theta)'\Omega(\lambda_3)^{-1}\Delta u(\theta),
$$

(4)

- $\Delta u(\theta) = B_1(\lambda_1)\Delta Y - B_2(\rho, \lambda_2)\Delta Y_{-1} - \Delta X\beta$,
- $\Omega(\lambda_3) = \frac{1}{\sigma^2_v} \text{Var}(\Delta u) = C \otimes [B'_3(\lambda_3)B_3(\lambda_3)]^{-1}$, and

$$
C = \begin{pmatrix}
2 & -1 & 0 & \cdots & 0 & 0 & 0 \\
-1 & 2 & -1 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & -1 & 2 & -1 \\
0 & 0 & 0 & \cdots & 0 & -1 & 2 \\
\end{pmatrix}.
$$

- Maximizing $\ell(\psi)$ in (4) leads to the conditional QMLE $\tilde{\psi}$ of $\psi$.
- However, $\tilde{\psi}$ cannot be consistent if $T$ is small and fixed,
- as the information about parameters contained in $\Delta y_1$ is ignored.
Conditional QML estimation

Let \( \theta_1 = (\beta', \rho, \lambda_2)' \). Given \( \lambda_1 \) and \( \lambda_3 \), (4) is maximized at

\[
\tilde{\theta}_1(\lambda_1, \lambda_3) = (\Delta \mathbf{X}' \Omega^{-1} \Delta \mathbf{X})^{-1} \Delta \mathbf{X}' \Omega \mathbf{B}_1(\lambda_1) \Delta \mathbf{Y},
\]

(5)

\[
\tilde{\sigma}_V^2(\lambda_1, \lambda_3) = \frac{1}{n(T-1)} \Delta \tilde{u}'(\lambda_1, \lambda_3) \Omega^{-1} \Delta \tilde{u}(\lambda_1, \lambda_3),
\]

(6)

• \( \tilde{u}(\lambda_1, \lambda_3) = \mathbf{B}_1(\lambda_1) \Delta \mathbf{Y} - \Delta \mathbf{X} \tilde{\theta}(\lambda_1, \lambda_3); \Delta \mathbf{X} = (\Delta \mathbf{X}, \Delta \mathbf{Y}_{-1}, \mathbf{W}_2 \Delta \mathbf{Y}_{-1}) \).

Substituting \( \tilde{\theta}_1(\lambda_1, \lambda_3) \) and \( \tilde{\sigma}_V^2(\lambda_1, \lambda_3) \) back into (4) gives the concentrated conditional quasi loglikelihood of \( (\lambda_1, \lambda_3) \), ignoring the constant term,

\[
\ell_{\text{STLE}}^c(\lambda_1, \lambda_3) = -\frac{n(T-1)}{2} \log[\tilde{\sigma}_V^2(\lambda_1, \lambda_3)] - \frac{1}{2} \log |\Omega(\lambda_3)| + \log |\mathbf{B}_1(\lambda_1)|. \quad (7)
\]

Maximizing \( \ell_{\text{STLE}}^c(\lambda_1, \lambda_3) \) gives the conditional QML (CQML) estimators \( \tilde{\lambda}_1 \) and \( \tilde{\lambda}_3 \) of \( \lambda_1 \) and \( \lambda_3 \), and thus the CQML estimators of \( \theta_1 \) and \( \sigma_V^2 \) as

\( \tilde{\theta}_1 \equiv \tilde{\theta}_1(\tilde{\lambda}_1, \tilde{\lambda}_3) \) and \( \tilde{\sigma}_V^2 \equiv \tilde{\sigma}_V^2(\tilde{\lambda}_1, \tilde{\lambda}_3) \).
Inconsistency of CQMLE under fixed $T$

**Note:** the $\ell_{\text{STLE}}(\psi)$ is a quasi Gaussian loglikelihood both in
- the traditional sense that $\{v_{it}\}$ are not exactly Gaussian but Gaussian likelihood is used,
- and the sense that $\Delta y_1$ is not exogenous but is treated as exogenous.
- The latter causes inconsistency of the CQMLEs when $T$ is small.

Furthermore,
- we see from the results presented below that even if $T$ increases with $n$, the CQMLEs may encounter an asymptotic bias;
- we introduce a method that not only gives a consistent estimator of the model parameters when $T$ is small, but also eliminates the asymptotic bias when $T$ is large.

Details on these important points follow.
First, to simplify the notation,

- a parametric quantity (scalar, vector or matrix) evaluated at the general values of the parameters is denoted by dropping its arguments, e.g., $B_1 \equiv B_1(\lambda_1)$, $B_1 \equiv B_1(\lambda_1)$, $\Omega \equiv \Omega(\lambda_3)$, and similarly for $B_r$ and $B_r$, $r = 2, 3$;

- a parametric quantity evaluated at the true values of the parameters is denoted by dropping its argument and then adding a subscript 0, e.g., $B_{10} \equiv B_1(\lambda_{10})$, $\Omega_0 \equiv \Omega(\lambda_{30})$.

- Let $C = C \otimes I_n$.

- Denote $\Delta u \equiv \Delta u(\theta_0)$.

- The usual expectation, variance and covariance operators, ‘$E$’, ‘$Var$’ and ‘$Cov$’, correspond to the true parameter values.
The conditional quasi score function $S(\psi) = \frac{\partial}{\partial \psi} \ell(\psi)$ has the form:

$$S(\psi) = \begin{cases} 
\frac{1}{\sigma_v^2} \Delta X' \Omega^{-1} \Delta u(\theta), \\
\frac{1}{2\sigma_v^4} \Delta u(\theta)' \Omega^{-1} \Delta u(\theta) - \frac{n(T-1)}{2\sigma_v^2}, \\
\frac{1}{\sigma_v^2} \Delta u(\theta)' \Omega^{-1} \Delta Y_{-1}, \\
\frac{1}{\sigma_v^2} \Delta u(\theta)' \Omega^{-1} W_1 \Delta Y - \text{tr}(B_1^{-1} W_1), \\
\frac{1}{\sigma_v^2} \Delta u(\theta)' \Omega^{-1} W_2 \Delta Y_{-1}, \\
\frac{1}{2\sigma_v^2} \Delta u(\theta)' (C^{-1} \otimes A_3) \Delta u(\theta) - (T-1) \text{tr}(G_3),
\end{cases} \quad (8)$$

where $A_3 = W_3' B_3 + B_3' W_3$ and $G_3 = W_3 B_3^{-1}$.

- Under mild conditions, maximizing the conditional loglikelihood $\ell_{STLE}(\psi)$ is equivalent to solving the estimating equation $S_{STLE}(\psi) = 0$;
- The QML type estimation is special case of $M$-estimation;
- A necessary condition for the $M$-estimators to be consistent is that the probability limit of the estimating function at the true parameter value is zero (see, e.g., van der Vaart, 1998).
Inconsistency of CQMLE under fixed $T$

For the estimation of the FE-SDPD model, this condition becomes,

$$\lim_{n \to \infty} \frac{1}{nT} S_{STLE}(\psi_0) \xrightarrow{p} 0.$$ 

However, as shown below this is not the case. Thus,

1. CQMLEs are not consistent unless $T \to \infty$.
2. Further, even if $T$ goes to infinity with $n$ (proportionally), the CQMLEs encounter a bias of order $O(T^{-1})$, giving the asymptotic bias:
   - if $\frac{1}{nT} E[S_{STLE}(\psi_0)] = O(\frac{1}{T})$, then $\frac{1}{\sqrt{nT}} E[S_{STLE}(\psi_0)] = O((\frac{n}{T})^{\frac{1}{2}})$,
   - implying $E[\sqrt{nT}(\tilde{\psi} - \psi_0)] = O((\frac{n}{T})^{\frac{1}{2}})$, and
   - $\sqrt{nT}(\tilde{\psi} - \psi_0)$ converges to a non-centered normal if $\frac{n}{T} \to c > 0$.
3. If $\frac{n}{T} \to 0$ (large $T$ case), the asymptotic bias vanishes, but this would not be a case of interest to a spatial panel model.
4. To overcome this major problem, we first derive $E[S_{STLE}(\psi_0)]$, and then adjust $S_{STLE}(\psi)$ so that the adjusted quasi score (AQS) vector, say $S_{STLE}^*(\psi)$, is such that $\operatorname{plim}_{n \to \infty} \frac{1}{nT} S_{STLE}^*(\psi_0) = 0$. 

Z. L. Yang, SMU
In contrast with Hsiao et al. (2002), Elhorst (2010), and Su and Yang (2015), we only need to have very minimum knowledge about the processes in the past.

**Assumption A:** (i) the processes started \( m \) periods before the start of data collection, the 0th period, and (ii) if \( m \geq 1 \), \( \Delta y_0 \) is independent of future errors \( \{v_t, t \geq 1\} \); if \( m = 0 \), \( y_0 \) is independent of future errors \( \{v_t, t \geq 1\} \).

- The proposed method requires neither \( \{y_s, s = -m, \ldots, -1\} \) to follow the same processes as \( \{y_t, t = 0, 1, \ldots, T\} \), nor \( \{x_{it}\} \) to be trend-stationary or first-difference stationary;
- has a much weaker requirement on processes starting positions \( y_m \).

To derive the results, the reduced form of (2) is important:

\[
\Delta y_t = B_0 \Delta y_{t-1} + B_{10}^{-1} \Delta X_t \beta_0 + B_{10}^{-1} B_{30}^{-1} \Delta v_t, \quad t = 2, \ldots, T, \tag{9}
\]

where \( B \equiv B(\rho, \lambda_1, \lambda_2) = B_1^{-1}(\lambda_1)B_2(\rho, \lambda_2) \).
Lemma. Suppose Assumption A holds. Assume further that, for \( i = 1, \ldots, n \) and \( t = 0, 1, \ldots, T \), (i) the idiosyncratic errors \( \{v_{it}\} \) are iid across \( i \) and \( t \) with mean 0 and variance \( \sigma_{v0}^2 \), (ii) the time-varying regressors \( X_t \) are exogenous, and (iii) both \( B_{10}^{-1} \) and \( B_{30}^{-1} \) exist. Then,

\[
E(\Delta Y_{-1} \Delta v') = -\sigma_{v0}^2 D_{-10} B_{30}^{-1},
\]

\[
E(\Delta Y \Delta v') = -\sigma_{v0}^2 D_0 B_{30}^{-1},
\]

where \( D_{-1} = D_{-1}(\rho, \lambda_1, \lambda_2) \) and \( D = D(\rho, \lambda_1, \lambda_2) \), given as,

\[
D_{-1} = \begin{pmatrix}
I_n, & 0, & \cdots & 0, & 0 \\
B - 2I_n, & I_n, & \cdots & 0, & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
B^{T-4}(I_n - B)^2, & B^{T-5}(I_n - B)^2, & \cdots & B - 2I_n, & I_n
\end{pmatrix} B_{10}^{-1},
\]

\[
D = \begin{pmatrix}
B - 2I_n, & I_n, & \cdots & 0 \\
(I_n - B)^2, & B - 2I_n, & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
B^{T-3}(I_n - B)^2, & B^{T-4}(I_n - B)^2, & \cdots & B - 2I_n
\end{pmatrix} B_{10}^{-1}.
\]
The results of the Lemma lead immediately to

\[
E(\Delta u'\Omega_0^{-1}\Delta Y_{-1}) = -\sigma_v^2 v_0 \text{tr}(C^{-1}D_{-10}), \quad (12)
\]

\[
E(\Delta u'\Omega_0^{-1}W_1\Delta Y) = -\sigma_v^2 v_0 \text{tr}(C^{-1}D_0W_1), \quad (13)
\]

\[
E(\Delta u'\Omega_0^{-1}W_2\Delta Y_{-1}) = -\sigma_v^2 v_0 \text{tr}(C^{-1}D_{-10}W_2), \quad (14)
\]

\[\Rightarrow\] the \((\rho, \lambda_1, \lambda_2)\) elements of \(E[S(\psi_0)]\) are of order \(O(n)\), and hence

\[\text{plim}_{n \to \infty} \frac{1}{n(T-1)} S(\psi_0) \neq 0,\]

\[\Rightarrow\] at least, the conditional QMLEs \(\tilde{\rho}, \tilde{\lambda}_1,\) and \(\tilde{\lambda}_2\) are inconsistent.

For example, when \(\lambda_1 = \lambda_2 = 0\), i.e., for the FE-SDPD model with SE only,

\[
\text{plim} \frac{1}{n(T-1)} \frac{\partial}{\partial \rho} \ell_{\text{STLE}}(\psi_0) = \frac{1-\rho_0^T}{T^2(1-\rho_0)^2} - \frac{1}{T(1-\rho_0)},
\]

- which is not zero, and thus \(\tilde{\rho}\) is not consistent;
- even if \(T \to \infty\) with \(n\), \(\text{Bias}(\tilde{\rho}) = O(\frac{1}{T})\).
It is expected that this result would hold for the general model, and the bias in $\hat{\rho}$ would spill over to the other CQMLEs.

The results (12)-(14) give a set of adjusted quasi score (AQS) functions:

$$S^*(\psi) = \begin{cases} 
\frac{1}{\sigma^2} \Delta X' \Omega^{-1} \Delta u(\theta), \\
\frac{1}{2\sigma^4} \Delta u(\theta)' \Omega^{-1} \Delta u(\theta) - \frac{n(T-1)}{2\sigma^2}, \\
\frac{1}{\sigma^2} \Delta u(\theta)' \Omega^{-1} \Delta Y_{-1} + \text{tr}(C^{-1}D_{-1}), \\
\frac{1}{\sigma^2} \Delta u(\theta)' \Omega^{-1} W_1 \Delta Y + \text{tr}(C^{-1}DW_1), \\
\frac{1}{\sigma^2} \Delta u(\theta)' \Omega^{-1} W_2 \Delta Y_{-1} + \text{tr}(C^{-1}D_{-1}W_2), \\
\frac{1}{2\sigma^2} \Delta u(\theta)' (C^{-1} \otimes A_3) \Delta u(\theta) - (T - 1)\text{tr}(G_3),
\end{cases}$$

which are a set of unbiased and consistent estimating functions, i.e., $E[S^*(\psi_0)] = 0$, and $\text{plim}_{n \to \infty} \frac{1}{n(T-1)} S^*(\psi_0) = 0$, even when $T$ is fixed.

Solving $S^*(\psi) = 0$ gives the M-estimator $\hat{\psi}_M$ of $\psi$!
This root-finding process can be simplified: given \( \delta = (\rho, \lambda')' \), the constrained \( M \)-estimators of \( \beta \) and \( \sigma^2_V \) are:

\[
\begin{align*}
\hat{\beta}(\delta) &= \left( \Delta X' \Omega^{-1} \Delta X \right)^{-1} \Delta X' \Omega^{-1} \left( B_1 \Delta Y - B_2 \Delta Y_{-1} \right), \\
\hat{\sigma}^2_V(\delta) &= \frac{1}{n(T-1)} \Delta \hat{u}(\delta)' \Omega^{-1} \Delta \hat{u}(\delta),
\end{align*}
\]

where \( \Delta \hat{u}(\delta) = \Delta u(\hat{\beta}(\delta), \rho, \lambda_1, \lambda_2) \);

\[
S^*_C(\delta) = \begin{cases} \\
\frac{1}{\hat{\sigma}^2_V, M(\delta)} \Delta \hat{u}(\delta)' \Omega^{-1} \Delta Y_{-1} + \text{tr}(C^{-1}D_{-1}), \\
\frac{1}{\hat{\sigma}^2_V, M(\delta)} \Delta \hat{u}(\delta)' \Omega^{-1} W_1 \Delta Y + \text{tr}(C^{-1}D W_1), \\
\frac{1}{\hat{\sigma}^2_V, M(\delta)} \Delta \hat{u}(\delta)' \Omega^{-1} W_2 \Delta Y_{-1} + \text{tr}(C^{-1}D_{-1} W_2), \\
\frac{1}{2\hat{\sigma}^2_V, M(\delta)} \Delta \hat{u}(\delta)'(C^{-1} \otimes A_3) \Delta \hat{u}(\delta) - (T - 1)\text{tr}(G_3).
\end{cases}
\]

Solving \( S^*_C(\delta) = 0 \) gives the \( M \)-estimators \( \hat{\delta}_M \) of \( \delta \), and hence the \( M \)-estimators of \( \beta \) and \( \sigma^2_V \): \( \hat{\beta}_M \equiv \hat{\beta}(\hat{\delta}_M) \) and \( \hat{\sigma}^2_V, M \equiv \hat{\sigma}^2_V(\hat{\delta}_M) \).
10.3. Asymptotic Properties of the M-Estimators

**Notation:** Recall: $\psi_0 = \text{true value of } \psi$, $B_1 = B_1(\lambda_1)$, $B_{10} = B_0(\lambda_{10})$, $\Omega \equiv \Omega(\lambda_3)$, $\Omega_0 \equiv \Omega(\lambda_{30})$, etc.; $\Delta u \equiv \Delta u(\theta_0)$. Further,

(i) $\Delta$ is the parameter space of $\delta = (\rho, \lambda')'$;
(ii) ‘E’ and ‘Var’ correspond to the true parameter values $\psi_0$;
(iii) $\text{tr}(\cdot)$, $|\cdot|$, $\|\cdot\|$: trace, determinant, Frobenius norm;
(iv) $\gamma_{\text{max}}(A)$, $\gamma_{\text{min}}(A)$: largest and smallest eigenvalues of a real symmetric matrix $A$.
(v) $\text{diag}(a_k)$ forms a diagonal matrix using the elements $\{a_k\}$,
     $\text{blkdiag}(A_k)$ forms a block-diagonal matrix using matrices $\{A_k\}$. 
**Assumption B:** The innovations $v_{it}$ are iid for all $i$ and $t$ with $E(v_{it}) = 0$, $\text{Var}(v_{it}) = \sigma_v^2$, and $E|v_{it}|^{4+\epsilon_0} < \infty$ for some $\epsilon_0 > 0$.

**Assumption C:** The space $\Delta$ is compact, and the true parameter $\delta_0$ lies in its interior.

**Assumption D:** The time-varying regressors $\{X_t, \ t = 0, 1, \ldots, T\}$ are exogenous, their values are uniformly bounded, and $\lim_{n \to \infty} \frac{1}{nT} \Delta X' \Delta X$ exists and is nonsingular.

**Assumption E:** (i) For $r = 1, 2, 3$, the elements $w_{r,ij}$ of $W_r$ are at most of order $h_n^{-1}$, uniformly in all $i$ and $j$, and $w_{r,ii} = 0$ for all $i$; (ii) $h_n/n \to 0$ as $n \to \infty$; (iii) $\{W_r, r = 1, 2, 3\}$ and $\{B_{r0}^{-1}, r = 1, 3\}$ are uniformly bounded in both row and column sums; (iv) For $r = 1, 3$, $\{B_r^{-1}\}$ are uniformly bounded in either row or column sums, uniformly in $\lambda_r$ in a compact parameter space $\Lambda_r$, and $0 < c_r \leq \inf_{\lambda_r \in \Lambda_r} \gamma_{\min}(B_r'B_r) \leq \sup_{\lambda_r \in \Lambda_r} \gamma_{\max}(B_r'B_r) \leq \bar{c}_r < \infty$. 

Z. L. Yang, SMU ECON6002, Term II 2020-21 22/71
Assumption F: For an $n \times n$ matrix $\Phi$ uniformly bounded in either row or column sums, with elements of uniform order $h_n^{-1}$, and an $n \times 1$ vector $\phi$ with elements of uniform order $h_n^{-1/2}$,

(i) $\frac{h_n}{n} \Delta y_1' \Phi \Delta y_1 = O_p(1)$ and $\frac{h_n}{n} \Delta y_1' \Phi \Delta v_2 = O_p(1)$;  
(ii) $\frac{h_n}{n} (\Delta y_1 - E(\Delta y_1))' \phi = o_p(1)$;  
(iii) $\frac{h_n}{n} [\Delta y_1' \Phi \Delta y_1 - E(\Delta y_1' \Phi \Delta y_1)] = o_p(1)$, and  
(iv) $\frac{h_n}{n} [\Delta y_1' \Phi \Delta v_2 - E(\Delta y_1' \Phi \Delta v_2)] = o_p(1)$.

Define $\tilde{S}^*(\psi) = E[S_{STLE}^*(\psi)]$, the population counterpart of $S^*(\psi)$ given in (15). Given $\delta$, $\tilde{S}_{STLE}^*(\psi) = 0$ is partially solved at

$$
\bar{\beta}(\delta) = (\Delta X' \Omega^{-1} \Delta X)^{-1} \Delta X' \Omega^{-1} (B_1 E \Delta Y - B_2 E \Delta Y_{-1}), 
$$

(19)

$$
\bar{\sigma}_{\psi}^2(\delta) = \frac{1}{n(T-1)} E[\Delta \bar{u}(\delta)' \Omega^{-1} \Delta \bar{u}(\delta)],
$$

(20)

where $\Delta \bar{u}(\delta) = \Delta u(\theta)|_{\beta = \bar{\beta}(\delta)} = B_1 \Delta Y - B_2 \Delta Y_{-1} - \Delta X \bar{\beta}(\delta)$. 

Z. L. Yang, SMU ECON6002, Term II 2020-21
Asymptotic properties

These lead to the population counter part of $S_c^*(\psi)$ given in (18), upon substituting $\bar{\beta}(\delta)$ and $\bar{\sigma}_\psi^2(\delta)$ back into the $\delta$-component of $\bar{S}_c^*(\psi)$:

$$
\bar{S}_c^*(\delta) = \begin{cases}
\frac{1}{\bar{\sigma}_\psi^2(\delta)} E[\Delta \bar{u}(\delta)' \Omega^{-1} \Delta Y_{-1}] + \text{tr}(C^{-1}D_{-1}), \\
\frac{1}{\bar{\sigma}_\psi^2(\delta)} E[\Delta \bar{u}(\delta)' \Omega^{-1} W_1 \Delta Y] + \text{tr}(C^{-1}D W_1), \\
\frac{1}{\bar{\sigma}_\psi^2(\delta)} E[\Delta \bar{u}(\delta)' \Omega^{-1} W_1 \Delta Y_{-1}] + \text{tr}(C^{-1}D_{-1} W_1), \\
\frac{1}{2\bar{\sigma}_\psi^2(\delta)} E[\Delta \bar{u}(\delta)'(C^{-1} \otimes A_3) \Delta \bar{u}(\delta)] - (T - 1)\text{tr}(G_3).
\end{cases}
$$

(21)

Idea for consistency: By Theorem 5.9 of van der Vaart (1998), $\hat{\delta}_M$ will be consistent for $\delta_0$ if $\sup_{\delta \in \Delta} \frac{1}{\sqrt{n(T-1)}} \|S_c^*(\delta) - \bar{S}_c^*(\delta)\|_p \to 0$, and the following identification condition holds.

Assumption G: $\inf_{\delta: \ d(\delta, \delta_0) \geq \epsilon} \|\bar{S}_c^*(\delta)\| > 0$ for every $\epsilon > 0$, where $d(\delta, \delta_0)$ is a measure of distance between $\delta_0$ and $\delta$. 

Z. L. Yang, SMU, ECON6002, Term II 2020-21
Theorem

Suppose Assumptions A-G hold. Assume further that

(i) \( \gamma_{\text{max}}[\text{Var}(\Delta Y)] \) and \( \gamma_{\text{max}}[\text{Var}(\Delta Y_{-1})] \) are bounded, and

(ii) \( \inf_{\delta \in \Delta} \gamma_{\text{min}}(\text{Var}(B_1 \Delta Y - B_2 \Delta Y_{-1})) \geq c_y > 0. \)

We have, as \( n \to \infty, \hat{\psi}_M \xrightarrow{p} \psi_0. \)

For asymptotic normality, let \( \Delta y_1 = 1_{T-1} \otimes \Delta y_1. \) We have,

\[
\Delta Y = R \Delta y_1 + \eta + S \Delta \nu, \quad (22)
\]

\[
\Delta Y_{-1} = R_{-1} \Delta y_1 + \eta_{-1} + S_{-1} \Delta \nu, \quad (23)
\]

where \( R = \text{blkdiag}(B_0, B_0^2, \ldots, B_0^{T-1}), R_{-1} = \text{blkdiag}(I, B_0, \ldots, B_0^{T-2}), \)

\( \eta = B B_{10}^{-1} \Delta X_0 \beta_0, \eta_{-1} = B_{-1} B_{10}^{-1} \Delta X_0 \beta_0, S = B B_{10}^{-1} B_{30}^{-1}, S_{-1} = B_{-1} B_{10}^{-1} B_{30}^{-1}, \)
Asymptotic properties

\[ \mathbf{B} = \begin{pmatrix} l_n & 0 & \cdots & 0 & 0 \\ B_0 & l_n & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ B_0^{T-2} & B_0^{T-3} & \cdots & B_0 & l_n \end{pmatrix}, \mathbf{B}^{-1} = \begin{pmatrix} l_n & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ B_0^{T-3} & B_0^{T-4} & \cdots & l_n & 0 \end{pmatrix}. \]

Using (22) and (23), the AQS function can be represented as follows:

\[
S^*(\psi_0) = \begin{cases} \\
\Delta \nu' \Pi_1, \\
\Delta \nu' \Phi_1 \Delta \nu - \frac{n(T-1)}{2\sigma^2 v_0}, \\
\Delta \nu' \Psi_1 \Delta y_1 + \Delta \nu' \Pi_2 + \Delta \nu' \Phi_2 \Delta \nu + \text{tr}(\mathbf{C}^{-1} \mathbf{D}_{-10}) , \\
\Delta \nu' \Psi_2 \Delta y_1 + \Delta \nu' \Pi_3 + \Delta \nu' \Phi_3 \Delta \nu + \text{tr}(\mathbf{C}^{-1} \mathbf{D}_0 \mathbf{W}_1) , \\
\Delta \nu' \Psi_3 \Delta y_1 + \Delta \nu' \Pi_4 + \Delta \nu' \Phi_4 \Delta \nu + \text{tr}(\mathbf{C}^{-1} \mathbf{D}_{-10} \mathbf{W}_2) , \\
\Delta \nu' \Phi_5 \Delta \nu - (T-1) \text{tr}(G_{30}), \\
\end{cases}
\]

\[
\Pi_1 = \frac{1}{\sigma^2 v_0} \mathbf{C}_b \mathbf{A} \mathbf{X}, \quad \Pi_2 = \frac{1}{\sigma^2 v_0} \mathbf{C}_b \mathbf{Y}_{-1}, \quad \Pi_3 = \frac{1}{\sigma^2 v_0} \mathbf{C}_b \mathbf{W}_1 \mathbf{Y}_{-1}, \quad \Pi_4 = \frac{1}{\sigma^2 v_0} \mathbf{C}_b \mathbf{W}_2 \mathbf{Y}_{-1}, \quad \Phi_1 = \frac{1}{2\sigma^4 v_0} \mathbf{C}^{-1}, \\
\Phi_2 = \frac{1}{\sigma^2 v_0} \mathbf{C}_b \mathbf{S}_{-1}, \quad \Phi_3 = \frac{1}{\sigma^2 v_0} \mathbf{C}_b \mathbf{W}_1 \mathbf{S}, \quad \Phi_4 = \frac{1}{\sigma^2 v_0} \mathbf{C}_b \mathbf{W}_2 \mathbf{S}_{-1}, \quad \Phi_5 = \frac{1}{\sigma^2 v_0} [\mathbf{C}^{-1} \otimes (G'_{30} + G_{30})], \\
\Psi_1 = \frac{1}{\sigma^2 v_0} \mathbf{C}_b \mathbf{R}_{-1}, \quad \Psi_2 = \frac{1}{\sigma^2 v_0} \mathbf{C}_b \mathbf{W}_1 \mathbf{R}, \quad \Psi_3 = \frac{1}{\sigma^2 v_0} \mathbf{C}_b \mathbf{W}_2 \mathbf{R}_{-1}, \quad \text{and} \quad \mathbf{C}_b = \mathbf{C}^{-1} \otimes \mathbf{B}_{30}. 
\]
Theorem

Assume Assumptions A-G hold. We have, as \( n \to \infty \),

\[
\sqrt{n(T-1)}(\hat{\psi}_M - \psi_0) \xrightarrow{D} N[0, \lim_{n \to \infty} \Sigma^{-1}(\psi_0)\Gamma(\psi_0)\Sigma^{-1}(\psi_0)],
\]

where \( \Sigma^*(\psi_0) = -\frac{1}{n(T-1)}E[\frac{\partial}{\partial \psi'} S^*(\psi_0)] \) and \( \Gamma^*(\psi_0) = \frac{1}{n(T-1)}\text{Var}[S^*(\psi_0)] \), both assumed to exist and \( \Sigma^*(\psi_0) \) to be positive definite, for sufficiently large \( n \).

- In practical applications, one needs to estimate \( \Sigma^*(\psi_0) \) and \( \Gamma^*(\psi_0) \) to get standard errors of the \( M \)-estimators.
- As \( \Sigma^*_{\text{STLE}}(\psi_0) \) is the expected negative modified Hessian, its observed counterpart immediately offers a consistent estimate, i.e.,

\[
\Sigma^*(\hat{\psi}_M) = -\frac{1}{n(T-1)} \frac{\partial}{\partial \psi'} S^*(\psi)\bigg|_{\psi=\hat{\psi}_M}.
\]
10.4. OPMD Estimation of Robust VC Matrix

However,

- Estimation of $\Gamma^*(\psi_0)$ runs into problems.
- From (24), AQS function $S^*(\psi_0)$ contains three types of elements:
  \[ \Pi' \Delta \nu, \quad \Delta \nu' \Phi \Delta \nu, \quad \Delta \nu' \Psi \Delta y_1. \]
- $\text{Var}(\Delta \nu' \Psi \Delta y_1)$, etc, depend on unobservables.

**Idea of the Proposed Method:** write the above quantities as sums of martingale difference (MD) sequences. Then, *outer-product-of-MDs* (OPMD) gives a consistent estimate of the VC matrix.

- For a square matrix $A$, decompose $A = A^u + A^l + A^d$, sum of upper-triangular, lower-triangular, and diagonal matrix.
- Denote by $\Pi_t$, $\Phi_{ts}$ and $\Psi_{ts}$ the submatrices of $\Pi$, $\Phi$ and $\Psi$ partitioned according to $t, s = 2, \ldots, T$. 
First, for the terms linear in $\Delta v$:

\[
\Pi' \Delta v = \sum_{t=2}^{T} \Pi_t' \Delta v_t \\
= \sum_{t=2}^{T} \sum_{i=1}^{n} \Pi_{it}' \Delta v_{it} \\
= \sum_{i=1}^{n} \sum_{t=2}^{T} \Pi_{it}' \Delta v_{it} \equiv \sum_{i=1}^{n} g_{1i}.
\]

Then, for the terms quadratic in $\Delta v$: \(E(\Delta v' \Phi \Delta v) = \sigma_{v0}^2 \text{tr}(C\Phi)\), and

\[
\Delta v' \Phi \Delta v - E(\Delta v' \Phi \Delta v) \\
= \sum_{t} \sum_{s} \Delta v'_t \Phi_{ts} \Delta v_s - \sigma_{v0}^2 \text{tr}(C\Phi) \\
= \sum_{t} \sum_{s} \Delta v'_t (\Phi_{ts}', + \Phi_{ts}' + \Phi_{ts}' d) \Delta v_s - \sigma_{v0}^2 \text{tr}(C\Phi) \\
= \sum_{t} \sum_{s} \Delta v'_t (\Phi_{st}', + \Phi_{ts}' + \Phi_{ts}' d) \Delta v_s - \sigma_{v0}^2 \text{tr}(C\Phi) \\
= \sum_{t} \Delta v'_t \Delta \xi_t + \sum_{t} \Delta v'_t \Delta v^*_t - \sigma_{v0}^2 \text{tr}(C\Phi) \\
= \sum_{i=1}^{n} \sum_{(t} (\Delta v_{it} \Delta \xi_{it} + \Delta v_{it} \Delta v^*_{it} - \sigma_{v0}^2 d_{it}) \equiv \sum_{i=1}^{n} g_{2i},
\]

- \(\Delta \xi_t = \sum_{s=2}^{T} (\Phi_{st}' + \Phi_{ts}') \Delta v_s; \Delta v^*_{t} = \sum_{s=2}^{T} \Phi_{ts}' \Delta v_s; \{d_{it}\} = \text{diag}(C\Phi)\).
Finally, for bilinear terms: $\Delta v' \Psi \Delta y_1$, define

- $\psi_{t+} = \sum_{s=2}^{T} \psi_{ts}, \ t = 2, \ldots, T$,
- $\Delta y_1^\circ = B_{30} B_{10} \Delta y_1; \ \Delta y_{1t}^* = \psi_{t+} \Delta y_1$,
- $\Theta = \psi_{2+} (B_{30} B_{10})^{-1}; \ \{\Theta_{ii}\} = \text{diag}(\Theta)$,
- $\{\Delta \zeta_{1i}\} = \Delta v'_2 \Theta^u; \ \{\Delta \zeta_{2i}\} = \Theta^l \Delta y_1^\circ$.

We obtain:

$$
\Delta v' \Psi \Delta y_1 - E(\Delta v' \Psi \Delta y_1) = \sum_{i=1}^{n} (\Delta \zeta_{1i} \Delta y_{1i}^\circ + \Delta v_{2i} \Delta \zeta_{2i} + \Theta_{ii} (\Delta v_{2i} \Delta y_{1i}^\circ + \sigma_{v0}^2) + \sum_{t=3}^{T} \Delta v_{it} \Delta y_{1it}^*)
\equiv \sum_{i=1}^{n} g_{3i}
$$

The $\{g_{ri}\}$ are MD sequences, for $r = 1, 2, 3$!
Applying the above results to the elements of $S^*(\psi_0)$ in (24):

- For each $\Pi_r$ term in (24), define $g_{1ri}, \ r = 1, 2, 3, 4,$
- For each $\Phi_r$ term in (24), define $g_{2ri}, \ r = 1, 2, 3, 4, 5,$
- For each $\Psi_r$ term in (24), define $g_{3ri}, \ r = 1, 2, 3.$

Define

$$g_i = \begin{cases} 
  g_{11i}, \\
  g_{21i}, \\
  g_{31i} + g_{12i} + g_{22i}, \\
  g_{32i} + g_{13i} + g_{23i}, \\
  g_{33i} + g_{14i} + g_{24i}, \\
  g_{25i}.
\end{cases}$$

- Then, $S^*(\psi_0) = \sum_{i=1}^n g_i,$
- $\{g_i\}$ form a vector M.D. sequence, and hence
- $\text{Var}[S^*(\psi_0)] = \sum_{i=1}^n E(g_i g'_i).$
The ‘average’ of the outer products of the estimated $g_i'$s, i.e.,

$$\hat{\Gamma}^* = \frac{1}{n(T-1)} \sum_{i=1}^{n} \hat{g}_i\hat{g}_i',$$  \hspace{1cm} (26)

thus gives a consistent estimator of the variance of $\Gamma_{STLE}^*(\psi_0)$, where $\hat{g}_i$ is obtained by replacing $\psi_0$ in $g_i$ by $\hat{\psi}_M$ and $\Delta \nu$ in $g_i$ by its observed counterpart $\hat{\Delta} \nu$. Noting that $\Delta y_1$ is observed, we have:

**Theorem**

*Under the assumptions of Theorem (1), we have, as $n \to \infty$,*

$$\hat{\Gamma}^* - \Gamma^* (\psi_0) = \frac{1}{n(T-1)} \sum_{i=1}^{n} [\hat{g}_i\hat{g}_i' - E(g_ig_i')] \xrightarrow{p} 0,$$

*and hence, $\Sigma^{-1}(\hat{\psi}_M)\hat{\Gamma}^*\Sigma^{-1}(\hat{\psi}_M) - \Sigma^{-1}(\psi_0)\Gamma^*(\psi_0)\Sigma^{-1}(\psi_0) \xrightarrow{p} 0$.***
Certain submodels deserve some special attention. We concentrate on the submodels that contain spatial dependence, namely,

- the FE-DPD model with only $SE$ dependence,
- the FE-DPD model with only $SL$ dependence,
- the FE-DPD model with both $SL$ and $STL$ dependence, and
- the FE-DPD model with both $SL$ and $SE$ dependence.

We are particularly interested in comparing our approach with the standard small $T$ or large $T$ approaches, to demonstrate that

- when $T$ is small our approach provides results that are comparable with the standard full QML approach when the initial model is correctly specified.
- However, our approach provides results that are more robust against misspecification of the initial model than does the full QML approach.
- When $T$ is large, our approach provides results that are less biased compared with the conditional QML approach.
Setting $\lambda_1 = \lambda_2 = 0$, Model (2) reduces to an FE-SDPD with only $\text{SE}$ dependence of a SAR form,

- which has been rigorously treated in Su and Yang (2015) based on a full QML approach where the initial differences are modeled.
- It would be certainly interesting to see how the proposed approach compares with this full QML approach.

The conditional quasi Gaussian loglikelihood (4) simplifies to:

$$
\ell_{\text{SE}}(\psi) = -\frac{n(T-1)}{2} \log(\sigma_v^2) - \frac{1}{2} \log |\Omega| - \frac{1}{2\sigma_v^2} \Delta u(\theta)'\Omega^{-1}\Delta u(\theta),
$$

(27)

where $\psi = \{\beta', \sigma_v^2, \rho, \lambda_3\}'$ and $\theta = (\beta', \rho)'$ and $u(\theta) = \Delta Y - \rho \Delta Y_{-1} - \Delta X \beta$. Given $\lambda_3$, $\ell_{\text{SE}}(\psi)$ is maximized at

$$
\tilde{\theta}(\lambda_3) = (\Delta X'\Omega^{-1}\Delta X)^{-1}\Delta X'\Omega\Delta Y,
\tilde{\sigma}_v^2(\lambda_3) = \frac{1}{n(T-1)}\Delta \tilde{u}'(\lambda_3)\Omega^{-1}\Delta \tilde{u}(\lambda_3),
$$

where $\Delta \tilde{u}(\lambda_3) = \Delta Y - \Delta X \tilde{\theta}(\lambda_3)$, and $\Delta X = (\Delta X, \Delta Y_{-1})$. 

Substituting $\tilde{\theta}(\lambda_3)$ and $\tilde{\sigma}_v^2(\lambda_3)$ back into $\ell_{SE}(\psi)$ gives the concentrated quasi loglikelihood function of $\lambda_3$,

$$\ell_{SE}^c(\lambda_3) = -\frac{n(T-1)}{2} \log(\tilde{\sigma}_v^2(\lambda_3)) - \frac{1}{2} \log |\Omega|.$$  \hspace{1cm} (28)

Maximizing $\ell_{SE}^c(\lambda_3)$ gives the CQMLE $\tilde{\lambda}_3$ of $\lambda_3$, and thus the CQMLEs $\tilde{\theta} \equiv \tilde{\theta}(\tilde{\lambda}_3)$ and $\tilde{\sigma}_v^2 \equiv \tilde{\sigma}_v^2(\tilde{\lambda}_3)$ of $\beta$ and $\sigma_v^2$, respectively.

Now, the quasi score $S_{SE}(\psi) = \frac{\partial}{\partial \psi} \ell_{SE}(\psi)$ is:

$$S_{SE}(\psi) = \begin{cases} 
\frac{1}{\sigma_v^2} \Delta X'\Omega^{-1} \Delta u(\theta), \\
\frac{1}{2\sigma_v^4} \Delta u(\theta)'\Omega^{-1} \Delta u(\theta) - \frac{n(T-1)}{2\sigma_v^2}, \\
\frac{1}{\sigma_v^2} \Delta u(\theta)'\Omega^{-1} \Delta Y_{-1}, \\
\frac{1}{2\sigma_v^2} \Delta u(\theta)'(C^{-1} \otimes A_3) \Delta u(\theta) - (T - 1)\text{tr}(G_3).
\end{cases}$$

Only the $\rho$-element of $E[S_{SE}(\psi_0)]$ is non-zero,

$$\sigma_{v_0}^{-2} E(\Delta u'\Omega^{-1} Y_{-1}) = -n\text{tr}[C^{-1} D(\rho_0)],$$  \hspace{1cm} (29)
where the \((T - 1) \times (T - 1)\) matrix \(D(\rho)\) has the following expression, noting that \(D_{-1}\) in the Lemma reduces to \(D(\rho) \otimes I_n\):

\[
D(\rho) = \\
\begin{pmatrix}
1 & 0 & \cdots & 0 & 0 \\
\rho - 2 & 1 & \cdots & 0 & 0 \\
(1 - \rho)^2 & \rho - 2 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots \\
\rho^{T-5}(1 - \rho)^2 & \rho^{T-6}(1 - \rho)^2 & \cdots & 1 & 0 \\
\rho^{T-4}(1 - \rho)^2 & \rho^{T-5}(1 - \rho)^2 & \cdots & \rho - 2 & 1
\end{pmatrix}.
\]

It is easy to see that, when \(|\rho| < 1\),

\[
\text{tr}[C^{-1} D(\rho)] = \frac{1}{1-\rho} - \frac{1-\rho^T}{T(1-\rho)^2},
\]

which is a result that has appeared in the literature of non-spatial dynamic panel data models (e.g., Nickell, 1981; Lancaster, 2002; and Alvarez and Arellano, 2004), and was derived from different angles.
The result suggests that
- the $\rho$-element of the conditional quasi score function is such that
  \[ \lim_{n \to \infty} \frac{1}{nT\sigma_v^2} \Delta u'\Omega^{-1}\Delta Y_{-1} \neq 0, \] unless $T$ also approaches $\infty$.
- A necessary condition for consistency is violated, and hence the conditional QMLE of $\rho$ is inconsistent when $T$ is fixed.
- It also suggests that even under the large $n$ and large $T$ set up, the conditional QMLE of $\rho$ would incur a bias of order $O(T^{-1})$ as shown in Hahn and Kuersteiner (2002) for the regular DPD model.

With (29) and the fact that other score elements have zero expectation, the adjusted quasi score becomes

\[
S_{SE}^*(\psi) = \begin{cases} 
\frac{1}{\sigma_v^2} \Delta X'\Omega^{-1}\Delta u(\theta), \\
\frac{1}{2\sigma_v^4} \Delta u(\theta)'\Omega^{-1}\Delta u(\theta) - \frac{n(T-1)}{2\sigma_v^2}, \\
\frac{1}{\sigma_v^2} \Delta u(\theta)'\Omega^{-1}\Delta Y_{-1} + n\text{tr}(C^{-1}D(\rho)), \\
\frac{1}{2\sigma_v^2} \Delta u(\theta)'(C^{-1} \otimes A_3)\Delta u(\theta) - (T - 1)\text{tr}(G_3). 
\end{cases} 
\] (30)
Solving $S_{SE}^*(\psi) = 0$ leads to the $M$-estimator $\hat{\psi}_M$ of $\psi$.

This root-finding process can be simplified by first solving the equations for $\beta$ and $\sigma^2_v$, given $\delta = (\rho, \lambda_3)'$, resulting in the constrained $M$-estimators of $\beta$ and $\sigma^2_v$ as

$$
\hat{\beta}(\delta) = (\Delta X'\Omega^{-1}\Delta X)^{-1}\Delta X'\Omega^{-1}\Delta Y(\rho)
$$

$$
\hat{\sigma}^2_v(\delta) = \frac{1}{n(T-1)}\Delta \hat{u}(\delta)'\Omega^{-1}\Delta \hat{u}(\delta),
$$

where $\Delta Y(\rho) = \Delta Y - \rho \Delta Y_{-1}$ and $\Delta \hat{u}(\delta) = \Delta u(\hat{\beta}(\delta), \rho)$.

Substituting $\hat{\beta}(\delta)$ and $\hat{\sigma}^2_v(\delta)$ into the last two components of the AQS function in (30) gives the concentrated AQS functions:

$$
S_{SE}^{*c}(\delta) = \begin{cases}
\frac{1}{\hat{\sigma}^2_v,M(\delta)}\Delta \hat{u}(\delta)'\Omega^{-1}\Delta Y_{-1} + n\text{tr}(C^{-1}D(\rho)), \\
\frac{1}{2\hat{\sigma}^2_v,M(\delta)}\Delta \hat{u}(\delta)'(C^{-1} \otimes A_3)\Delta \hat{u}(\delta) - (T - 1)\text{tr}(G_3).
\end{cases}
$$

Solving the resulted concentrated estimating equations, $S_{SE}^{*c}(\delta) = 0$, we obtain the unconstrained $M$-estimators $\hat{\delta}_M = (\hat{\rho}_M, \hat{\lambda}_{3,M})'$ of $\delta$. 

Z. L. Yang, SMU
The unconstrained $M$-estimators of $\beta$ and $\sigma^2_v$ are thus $\hat{\beta}_M \equiv \hat{\beta}(\hat{\delta}_M)$ and $\hat{\sigma}^2_{v,M} \equiv \hat{\sigma}^2_v(\hat{\delta}_M)$.

Compared with the full QML estimation of Su and Yang (2015), the proposed $M$-estimation, though slightly less efficient, is much simpler as it is free from the specification of the initial conditions, and is thus robust against misspecifications of initial conditions.

In contrast, the full QML estimation requires that the process starting time $m$ is known a priori and that the processes evolve in the same manner before and after the data collection.

Our Monte Carlo results and those in Su and Yang (2015) confirm these points.
Setting \( \lambda_2 = \lambda_3 = 0 \) gives the FE-SDPD model with only \( SL \). Now, \( \psi = (\beta', \sigma^2_v, \rho, \lambda_1)' \). The conditional quasi loglikelihood of \( \psi \) reduces to:

\[
\ell_{SL}(\psi) = -\frac{n(T-1)}{2} \log(\sigma^2_v) + \log |B_1| - \frac{1}{2} \log |C| - \frac{1}{2\sigma^2_v} \Delta \nu(\theta)'C^{-1}\Delta \nu(\theta), \tag{32}
\]

where \( \theta = (\theta_1', \lambda_1)' \), \( \theta_1 = (\beta', \rho)' \), and \( \nu(\theta) = B_1 \Delta Y - \rho \Delta Y_{-1} - \Delta X \beta \).

- Given \( \lambda_1 \), \( \ell_{SL}(\psi) \) is maximized at
  
  \[
  \tilde{\theta}_1(\lambda_1) = (\Delta X'C^{-1}\Delta X)^{-1}\Delta X'C^{-1}B_1 \Delta Y
  \]
  
  \[
  \tilde{\sigma}^2_v(\lambda_1) = \frac{1}{n(T-1)} \Delta \tilde{\nu}(\lambda_1)'C^{-1}\Delta \tilde{\nu}(\lambda_1),
  \]

  where \( \Delta \tilde{\nu}(\lambda_1) = B_1 \Delta Y - \Delta X \tilde{\theta}(\lambda_1) \), and \( \Delta X = (\Delta X, \Delta Y_{-1}) \).

- Substituting \( \tilde{\theta}_1(\lambda_1) \) and \( \tilde{\sigma}^2_v(\lambda_1) \) back into \( \ell_{SL}(\psi) \) gives the concentrated conditional loglikelihood function of \( \lambda_1 \),

\[
\ell^c_{SL}(\lambda_1) = \log |B_1| - \frac{n(T-1)}{2} \log(\tilde{\sigma}^2_v(\lambda_1)) - \frac{1}{2} \log |C|. \tag{33}
\]
Maximizing $\ell^c_{SL}(\lambda_1)$ gives the CQMLE $\tilde{\lambda}_1$ of $\lambda_1$, and thus the CQMLEs $\tilde{\theta} \equiv \tilde{\theta}(\tilde{\lambda}_1)$ and $\tilde{\sigma}_v^2 \equiv \tilde{\sigma}_v^2(\tilde{\lambda}_1)$ of $\theta$ and $\sigma_v^2$, respectively.

Now, the CQS function $S_{SL}(\psi)$ has the form:

$$S_{SL}(\psi) = \begin{cases} 
\frac{1}{\sigma_v^2} \Delta X' C^{-1} \Delta v(\theta), \\
\frac{1}{2\sigma_v^4} \Delta v(\theta)' C^{-1} \Delta v(\theta) - \frac{n(T-1)}{2\sigma_v^2}, \\
\frac{1}{\sigma_v^2} \Delta v(\theta)' C^{-1} \Delta Y_{-1}, \\
\frac{1}{\sigma_v^2} \Delta v(\theta)' C^{-1} W_1 \Delta Y - \text{tr}(B_1^{-1} W_1). 
\end{cases}$$

The expectations of the first two components of $S_{SL}(\psi_0)$ are zero, but these of the last two are not as by the Lemma,

$$E(\Delta v' C^{-1} \Delta Y_{-1}) = -\sigma_v^2 \text{tr}(C^{-1} D_{-10}), \quad \text{and} \quad (34)$$
$$E(\Delta v' C^{-1} W_1 \Delta Y) = -\sigma_v^2 \text{tr}(C^{-1} D_0 W_1), \quad (35)$$

where $D_{-1}$ and $D$ are from the general model, but $B$ simplifies to $\rho B_1^{-1}$. 
FE-SDPD model with SL dependence

- These show that the last two elements of \( \lim_{n \to \infty} \frac{1}{nT} S_{SL}(\psi_0) \) are not zero, showing that the CQMLEs of the SL model are inconsistent.
- Even when \( T \) grows with \( n \), it can be shown that the CQMLE of \( \rho \) has a bias of order \( O(T^{-1}) \) instead of the desired order \( O((nT)^{-1}) \).
- Some modifications are thus necessary whether \( T \) is fixed or not.

The adjusted quasi score function is,

\[
S_{SL}^*(\psi) = \begin{cases} 
\frac{1}{\sigma_v^2} \Delta X' C^{-1} \Delta \nu(\theta), \\
\frac{1}{2\sigma_v^4} \Delta \nu(\theta)' C^{-1} \Delta \nu(\theta) - \frac{n(T-1)}{2\sigma_v^2}, \\
\frac{1}{\sigma_v^2} \Delta \nu(\theta)' C^{-1} \Delta Y_{-1} + \text{tr}(C^{-1}D_{-1}), \\
\frac{1}{\sigma_v^2} \Delta \nu(\theta)' C^{-1} W_1 \Delta Y + \text{tr}(C^{-1}DW_1). 
\end{cases}
\]

The \( M \)-estimator for the FE-SDPD-SLD model is thus defined as

\[
\hat{\psi}_M = \arg\{S_{SL}^*(\psi) = 0\}.
\]
The root-finding process can be simplified by first solving the equations for $\beta$ and $\sigma^2_v$, given $\delta = (\rho, \lambda_1)'$, leading to the constrained $M$-estimators for $\beta$ and $\sigma^2_v$:

$$
\hat{\beta}(\delta) = (\Delta X' C^{-1} \Delta X)^{-1} \Delta X' C^{-1} \Delta Y(\delta),
$$
$$
\hat{\sigma}^2_v(\delta) = \frac{1}{n(T-1)} \Delta \hat{v}(\delta)' C^{-1} \Delta \hat{v}(\delta),
$$

where $\Delta Y(\delta) = B_1 \Delta Y - \rho \Delta Y_{-1}$ and $\Delta \hat{v}(\delta) = \Delta v(\hat{\beta}(\delta), \delta)$.

Substituting $\hat{\beta}(\delta)$ and $\hat{\sigma}^2_v(\delta)$ into the last two components of (36) gives the concentrated AQS function of $\delta$:

$$
S^{*c}_{SL}(\delta) = \begin{cases} 
\frac{1}{\hat{\sigma}^2_{v,M}(\delta)} \Delta \hat{v}(\delta)' C^{-1} \Delta Y_{-1} + \text{tr}(C^{-1} D_{-1}), \\
\frac{1}{\hat{\sigma}^2_{v,M}(\delta)} \Delta \hat{v}(\delta)' C^{-1} W_1 \Delta Y + \text{tr}(C^{-1} D W_1) 
\end{cases}
$$

(37)

Solving the concentrated equations, $S^{*c}_{SL}(\delta) = 0$, gives the unconstrained $M$-estimator $\hat{\delta}_M$ of $\delta$. The unconstrained $M$-estimators of $\beta$ and $\sigma^2_v$ are thus $\hat{\beta}_M \equiv \hat{\beta}(\hat{\delta}_M)$ and $\hat{\sigma}^2_{v,M} \equiv \hat{\sigma}^2_v(\hat{\delta}_M)$. 

Setting $\lambda_3 = 0$ gives the FE-SDPD model with $SL$ and $STL$ dependence.

Now, $\psi = (\beta', \sigma_v^2, \rho, \lambda_1, \lambda_2)'$, and the conditional quasi loglikelihood of $\psi$:

$$
l_{STL}(\psi) = -\frac{n(T-1)}{2} \log(\sigma_v^2) + \log |B_1| - \frac{1}{2} \log |C| - \frac{1}{2\sigma_v^2} \Delta v(\theta)'C^{-1}\Delta v(\theta), \quad (38)$$

where $\theta = (\beta', \rho, \lambda_1, \lambda_2)'$, and $v(\theta) = B_1 \Delta Y - (\rho I_n + \lambda_2 W_2) \Delta Y_{-1} - \Delta X \beta$.

- Let $\theta_1 = (\beta', \rho, \lambda_2)'$. Given $\lambda_1$, $l_{STL}(\psi)$ is maximized at

$$
\tilde{\theta}_1(\lambda_1) = (\Delta X' C^{-1} \Delta X)^{-1} \Delta X' C^{-1} B_1 \Delta Y,
\tilde{\sigma}_v^2(\lambda_1) = \frac{1}{n(T-1)} \Delta \tilde{v}'(\lambda_1) C^{-1} \Delta \tilde{v}(\lambda_1),
$$

where $\Delta X = (\Delta X, \Delta Y_{-1}, W_2 \Delta Y_{-1})$ and $\Delta \tilde{v}(\lambda_1) = B_1 \Delta Y - \Delta X \tilde{\theta}(\lambda_1)$.

- The concentrated conditional quasi loglikelihood function of $\lambda_1$ is,

$$
l_{STL}^c(\lambda_1) = + \log |B_1| - \frac{n(T-1)}{2} \log(\tilde{\sigma}_v^2(\lambda_1)) - \frac{1}{2} \log |C|. \quad (39)$$

- Maximizing $l_{STL}^c(\lambda_1)$ gives the CQMLE $\tilde{\lambda}_1$, and thus the CQMLEs $\tilde{\theta}_1 \equiv \tilde{\theta}_1(\tilde{\lambda}_1)$ and $\tilde{\sigma}_v^2 \equiv \tilde{\sigma}_v^2(\tilde{\lambda}_1)$.
The CQS function $S_{\text{STL}}(\psi)$ becomes:

$$S_{\text{STL}}(\psi) = \begin{cases} 
\frac{1}{\sigma_v^2} \Delta X' C^{-1} \Delta v(\theta), \\
\frac{1}{2\sigma_v^4} \Delta v(\theta)' C^{-1} \Delta v(\theta) - \frac{n(T-1)}{2\sigma_v^2}, \\
\frac{1}{\sigma_v^2} \Delta v(\theta)' C^{-1} \Delta Y_{-1}, \\
\frac{1}{\sigma_v^2} \Delta v(\theta)' C^{-1} W_1 \Delta Y - \text{tr}(B_1^{-1} W_1), \\
\frac{1}{\sigma_v^2} \Delta v(\theta)' C^{-1} W_2 \Delta Y_{-1}. 
\end{cases}$$

It is easy to see that the last components of $E[S_{\text{STL}}(\psi_0)]$ are not zero, and are obtained from (12)-(14). The adjusted quasi score function is,

$$S^*_{\text{STL}}(\psi) = \begin{cases} 
\frac{1}{\sigma_v^2} \Delta X' C^{-1} \Delta v(\theta), \\
\frac{1}{2\sigma_v^4} \Delta v(\theta)' C^{-1} \Delta v(\theta) - \frac{n(T-1)}{2\sigma_v^2}, \\
\frac{1}{\sigma_v^2} \Delta v(\theta)' C^{-1} \Delta Y_{-1} + \text{tr}(C^{-1} D_{-1}), \\
\frac{1}{\sigma_v^2} \Delta v(\theta)' C^{-1} W_1 \Delta Y + \text{tr}(C^{-1} D W_1), \\
\frac{1}{\sigma_v^2} \Delta v(\theta)' C^{-1} W_2 \Delta Y_{-1} + \text{tr}(C^{-1} D_{-1} W_2). 
\end{cases}$$ (40)
The $M$-estimator for the FE-SDPD-SLD model is thus defined as
\[ \hat{\psi}_{SL} = \arg\{ S_M^*(\psi) = 0 \}. \]

It can be found by first solving the equations for $\beta$ and $\sigma^2_v$, given $\delta = (\rho, \lambda_1, \lambda_2)'$, leading to the constrained $M$-estimators
\[
\hat{\beta}(\delta) = (\Delta X' C^{-1} \Delta X)^{-1} \Delta X' C^{-1} \Delta Y(\delta),
\]
\[
\hat{\sigma}^2_v(\delta) = \frac{1}{n(T-1)} \Delta \hat{\nu}(\delta)' C^{-1} \Delta \hat{\nu}(\delta),
\]
where $\Delta Y(\delta) = B_1 \Delta Y - (\rho I_n + \lambda_2 W_2) \Delta Y_{-1}$ and $\Delta \hat{\nu}(\delta) = \Delta \nu(\hat{\beta}(\delta), \delta)$;

and then solving the concentrated estimating equations, $S^*_c(\delta) = 0$, where the concentrated AQS function of $\delta$ has the form:
\[
S^*_c(\delta) = \begin{cases} 
\frac{1}{\tilde{\sigma}^2_{v,M}(\delta)} \tilde{\nu}(\delta)' C^{-1} \Delta \Delta Y_{-1} + \text{tr}(C^{-1} D_{-1}), \\
\frac{1}{\tilde{\sigma}^2_{v,M}(\delta)} \Delta \tilde{\nu}(\delta)' C^{-1} W_1 \Delta Y + \text{tr}(W_1 C^{-1} D), \\
\frac{1}{\tilde{\sigma}^2_{v,M}(\delta)} \Delta \tilde{\nu}(\delta)' C^{-1} W_2 \Delta Y_{-1} + \text{tr}(C^{-1} D_{-1} W_2). 
\end{cases}
\]

The $M$-estimators of $\beta$ and $\sigma^2_v$ are, thus, $\hat{\beta}_M \equiv \hat{\beta}(\hat{\delta}_M)$ and $\hat{\sigma}^2_{v,M} \equiv \hat{\sigma}^2_v(\hat{\delta}_M)$. 
Setting $\lambda_2 = 0$ gives the FE-SDPD model with SL and SE dependence. The conditional quasi loglikelihood of $\psi = (\beta', \sigma_v^2, \rho, \lambda_1, \lambda_3)'$ is,

$$
\ell_{SLE}(\psi) = -\frac{n(T-1)}{2} \log(\sigma_v^2) + \log |B_1| - \frac{1}{2} \log |\Omega| - \frac{1}{2\sigma_v^2} \Delta u(\theta)'\Omega^{-1}\Delta u(\theta), \quad (42)
$$

where $\theta = (\beta', \rho, \lambda_1)'$ and $\Delta u(\theta) = B_1 \Delta Y - \rho \Delta Y_{-1} - \Delta X \beta$.

- $\ell_{SLE}(\psi)$ is partially maximized at

  $$
  \tilde{\theta}(\lambda) = (\Delta X'\Omega^{-1}\Delta X)^{-1} \Delta X'\Omega^{-1} B_1 \Delta Y,
  \tilde{\sigma}_v^2(\lambda) = \frac{1}{n(T-1)} \Delta \tilde{u}'(\lambda)\Omega^{-1}\Delta \tilde{u}(\lambda),
  $$

  where $\Delta \tilde{u}(\lambda) = B_1 \Delta Y - \Delta X \tilde{\theta}(\lambda)$, and $\Delta X = (\Delta X, \Delta Y_{-1})$.

- Maximizing the concentrated loglikelihood function of $\lambda$,

  $$
  \ell_{SLE}^c(\lambda) = \log |B_1| - \frac{n(T-1)}{2} \log(\tilde{\sigma}_v^2(\lambda)) - \frac{1}{2} \log |\Omega|. \quad (43)
  $$

  gives the CQMLE $\tilde{\lambda}$, and thus the CQMLEs $\tilde{\theta} \equiv \tilde{\theta}(\tilde{\lambda})$ and $\tilde{\sigma}_v^2 \equiv \tilde{\sigma}_v^2(\tilde{\lambda})$. 
The CQS function $S_{\text{SLE}}(\psi)$ has the form:

$$S_{\text{SLE}}(\psi) = \begin{cases} \frac{1}{\sigma_v^2} \Delta X' \Omega^{-1} \Delta u(\theta), \\ \frac{1}{2\sigma_v^4} \Delta u(\theta)' \Omega^{-1} \Delta u(\theta) - \frac{n(T-1)}{2\sigma_v^2}, \\ \frac{1}{\sigma_v^2} \Delta u(\theta)' \Omega^{-1} \Delta Y_{-1}, \\ \frac{1}{\sigma_v^2} \Delta u(\theta)' \Omega^{-1} W_1 \Delta Y - \text{tr}(B_1^{-1} W_1), \\ \frac{1}{2\sigma_v^2} \Delta u(\theta)'(C^{-1} \otimes A_3) \Delta u(\theta) - (T-1)\text{tr}(G_3). \end{cases}$$

The $\rho$ and $\lambda_1$ components of $E[S_{\text{SLE}}(\psi_0)]$ are not zero, as seen from the Lemma given for the general model:

- $E(\Delta u' \Omega^{-1} \Delta Y_{-1}) = -\sigma_{v0}^2 \text{tr}(C^{-1}D_{-10})$,
- $E(\Delta u' \Omega^{-1} W_1 \Delta Y) = -\sigma_{v0}^2 \text{tr}(C^{-1}D_0 W_1)$,

which are of identical forms as those for the SLD model.

- The existence of SE does not effect the adjustments on the conditional quasi score!
The results show that the CQMLEs are not consistent unless $T$ is also large. The conditional quasi score function should be modified as:

$$
S^*_{\text{SLE}}(\psi) = \begin{cases} 
\frac{1}{\sigma^2_v} \Delta X' \Omega^{-1} \Delta u(\theta), \\
\frac{1}{2\sigma^4_v} \Delta u(\theta)' \Omega^{-1} \Delta u(\theta) - \frac{n(T-1)}{2\sigma^2_v}, \\
\frac{1}{\sigma^2_v} \Delta u(\theta)' \Omega^{-1} \Delta Y_{-1} + \text{tr}(C^{-1} D_{-1}), \\
\frac{1}{2\sigma^4_v} \Delta u(\theta)' \Omega^{-1} W_1 \Delta Y + \text{tr}(C^{-1} D W_1), \\
\frac{1}{2\sigma^2_v} \Delta u(\theta)' (C^{-1} \otimes A_3) \Delta u(\theta) - (T - 1) \text{tr}(G_3),
\end{cases}
$$

(44)

The $M$-estimator of the FE-DPD-SLE model is defined as

$$
\hat{\psi}_M = \arg\{S^*_{\text{SLE}}(\psi) = 0\}.
$$

which can be found by first solving for $\beta$ and $\sigma^2_v$ to give:

$$
\hat{\beta}(\delta) = (\Delta X' \Omega^{-1} \Delta X)^{-1} \Delta X' \Omega^{-1} \Delta Y(\rho, \lambda_1),
$$

$$
\hat{\sigma}^2_v(\delta) = \frac{1}{n(T-1)} \Delta \hat{u}(\delta)' \Omega^{-1} \Delta \hat{u}(\delta),
$$

where $\Delta Y(\rho, \lambda_1) = B_1 \Delta Y - \rho \Delta Y_{-1}$ and $\Delta \hat{u}(\delta) = \Delta u(\hat{\beta}(\delta), \rho, \lambda_1)$. 

Then, solve the concentrated estimating equations $S_{SLE}(\delta) = 0$ to give the $M$-estimator $\hat{\delta}_M$ of $\delta$, where the concentrated AQS function is:

$$S_{SLE}^*(\delta) = \begin{cases} 
\frac{1}{\hat{\sigma}^2_v, M(\delta)} \Delta \hat{u}(\delta)' \Omega^{-1} \Delta Y_{-1} + \text{tr}(C^{-1}D_{-1}), \\
\frac{1}{\hat{\sigma}^2_v, M(\delta)} \Delta \hat{u}(\delta)' \Omega^{-1} W_1 \Delta Y + \text{tr}(C^{-1}DW_1), \\
\frac{1}{2\hat{\sigma}^2_v, M(\delta)} \Delta \hat{u}(\delta)' (C^{-1} \otimes A_3) \Delta \hat{u}(\delta) - (T - 1) \text{tr}(G_3),
\end{cases}$$

(45)

obtained by substituting $\hat{\beta}(\delta)$ and $\hat{\sigma}^2_v(\delta)$ into the last three components of $S_{SLE}^*(\psi)$ given by (44). The $M$-estimators of $\beta$ and $\sigma^2_v$ are thus:

$$\hat{\beta}_M \equiv \hat{\beta}(\hat{\delta}_M) \text{ and } \hat{\sigma}^2_{v, M} \equiv \hat{\sigma}^2_v(\hat{\delta}_M).$$

The consistency and asymptotic normality of $\hat{\psi}_M$ for the SLE model are implied by the results for the general model. The estimate of robust VC matrix is obtained using the relevant submatrices for the general model.
Monte Carlo experiments are carried out to investigate

(i) the finite sample performance of the $M$-estimators of the FE-SDPD models,

(ii) the finite sample performance of the proposed OPMD estimates of the robust standard errors,

(iii) the performance of proposed methods relative to existing ones.

We use the following five models:

\[ \text{SE : } y_t = \rho y_{t-1} + \beta_0 \nu_n + X_t \beta_1 + Z \gamma + \mu + u_t, \quad u_t = \lambda_3 W_3 u_t + v_t, \]
\[ \text{SL : } y_t = \rho y_{t-1} + \lambda_1 W_1 y_t + \beta_0 \nu_n + X_t \beta_1 + Z \gamma + \mu + v_t, \]
\[ \text{SLE : } y_t = \rho y_{t-1} + \lambda_1 W_1 y_t + \beta_0 \nu_n + X_t \beta_1 + Z \gamma + \mu + u_t, \]
\[ \quad u_t = \lambda_3 W_3 u_t + v_t, \]
\[ \text{STL : } y_t = \rho y_{t-1} + \lambda_1 W_1 y_t + \lambda_2 W_2 y_{t-1} + \beta_0 \nu_n + X_t \beta_1 + Z \gamma + \mu + v_t, \]
\[ \text{STLE : } y_t = \rho y_{t-1} + \lambda_1 W_1 y_t + \lambda_2 W_2 y_{t-1} + \beta_0 \nu_n + X_t \beta_1 + Z \gamma + \mu + u_t, \]
\[ \quad u_t = \lambda_2 W_3 u_t + v_t. \]
<table>
<thead>
<tr>
<th>err ψ</th>
<th>CQMLE</th>
<th>FQMLE</th>
<th>M-Est</th>
<th>CQMLE</th>
<th>FQMLE</th>
<th>M-Est</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 1</td>
<td>1.0152(.096)</td>
<td>1.0017(.100)</td>
<td>1.0015(.100)</td>
<td>1.0109(.050)</td>
<td>1.0021(.052)</td>
<td>1.0020(.053)</td>
</tr>
<tr>
<td>1</td>
<td>.9154(.135)</td>
<td>.9678(.148)</td>
<td>.9719(.154)</td>
<td>.9080(.065)</td>
<td>.9960(.079)</td>
<td>.9962(.080)</td>
</tr>
<tr>
<td>.5</td>
<td>.3605(.055)</td>
<td>.4995(.065)</td>
<td>.5015(.066)</td>
<td>.2869(.033)</td>
<td>.5009(.043)</td>
<td>.5013(.044)</td>
</tr>
<tr>
<td>.5</td>
<td>.4702(.107)</td>
<td>.4761(.093)</td>
<td>.4793(.105)</td>
<td>.4775(.073)</td>
<td>.4877(.060)</td>
<td>.4907(.070)</td>
</tr>
<tr>
<td>2 1</td>
<td>1.0142(.098)</td>
<td>1.0007(.102)</td>
<td>1.0002(.102)</td>
<td>1.0099(.050)</td>
<td>1.0015(.053)</td>
<td>1.0014(.053)</td>
</tr>
<tr>
<td>1</td>
<td>.9176(.266)</td>
<td>.9662(.284)</td>
<td>.9785(.307)</td>
<td>.9045(.128)</td>
<td>.9920(.152)</td>
<td>.9935(.155)</td>
</tr>
<tr>
<td>.5</td>
<td>.3610(.066)</td>
<td>.4975(.069)</td>
<td>.5023(.078)</td>
<td>.2876(.041)</td>
<td>.5002(.047)</td>
<td>.5018(.052)</td>
</tr>
<tr>
<td>.5</td>
<td>.4701(.106)</td>
<td>.4770(.092)</td>
<td>.4803(.104)</td>
<td>.4741(.075)</td>
<td>.4844(.063)</td>
<td>.4883(.072)</td>
</tr>
<tr>
<td>3 1</td>
<td>1.0133(.099)</td>
<td>1.0001(.103)</td>
<td>.9997(.103)</td>
<td>1.0090(.047)</td>
<td>1.0003(.049)</td>
<td>1.0003(.049)</td>
</tr>
<tr>
<td>1</td>
<td>.9192(.198)</td>
<td>.9678(.212)</td>
<td>.9771(.227)</td>
<td>.9060(.099)</td>
<td>.9938(.119)</td>
<td>.9947(.121)</td>
</tr>
<tr>
<td>.5</td>
<td>.3585(.059)</td>
<td>.4953(.066)</td>
<td>.4992(.071)</td>
<td>.2881(.036)</td>
<td>.5018(.046)</td>
<td>.5029(.048)</td>
</tr>
<tr>
<td>.5</td>
<td>.4681(.110)</td>
<td>.4736(.093)</td>
<td>.4786(.106)</td>
<td>.4741(.075)</td>
<td>.4852(.062)</td>
<td>.4884(.073)</td>
</tr>
</tbody>
</table>

**Note:** Par = ψ = (β, σ², ρ, λ³)'; err=1 (normal), 2 (normal mixture), and 3 (chi-square).

*χ* values are generated with \( θ_x = (g, φ_1, φ_2, σ_1, σ_2) = (.01, .5, .5, 1, .5) \), as in Footnote 1.

*W₃* is generated according to Group Interaction scheme as in Footnote 2.
Table 10.1a. Cont’d, $T = 7$

<table>
<thead>
<tr>
<th>err $\psi$</th>
<th>$n = 50$</th>
<th>$n = 100$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQMLE</td>
<td>FQMLE</td>
</tr>
<tr>
<td>1 1</td>
<td>1.0248(.044)</td>
<td>1.0015(.044)</td>
</tr>
<tr>
<td>1</td>
<td>.9771(.081)</td>
<td>.9888(.083)</td>
</tr>
<tr>
<td>.5</td>
<td>.4456(.028)</td>
<td>.4987(.029)</td>
</tr>
<tr>
<td>.5</td>
<td>.4928(.057)</td>
<td>.4920(.055)</td>
</tr>
<tr>
<td>2 1</td>
<td>1.0247(.045)</td>
<td>1.0012(.045)</td>
</tr>
<tr>
<td>1</td>
<td>.9776(.183)</td>
<td>.9887(.186)</td>
</tr>
<tr>
<td>.5</td>
<td>.4461(.028)</td>
<td>.4988(.029)</td>
</tr>
<tr>
<td>.5</td>
<td>.4919(.058)</td>
<td>.4914(.055)</td>
</tr>
<tr>
<td>3 1</td>
<td>1.0250(.044)</td>
<td>1.0015(.044)</td>
</tr>
<tr>
<td>1</td>
<td>.9751(.130)</td>
<td>.9863(.133)</td>
</tr>
<tr>
<td>.5</td>
<td>.4458(.028)</td>
<td>.4986(.029)</td>
</tr>
<tr>
<td>.5</td>
<td>.4903(.057)</td>
<td>.4896(.056)</td>
</tr>
</tbody>
</table>
Table 10.1b. Empirical sd and average of estimated standard errors of M-Estimator

SE Model, $T = 3$, $m = 5$, Parameter configurations as in Table 1a.

<table>
<thead>
<tr>
<th>dgp $\psi$</th>
<th>sd</th>
<th>$\tilde{se}$</th>
<th>$\tilde{se}$</th>
<th>rse</th>
<th>sd</th>
<th>$\tilde{se}$</th>
<th>$\tilde{se}$</th>
<th>rse</th>
<th>sd</th>
<th>$\tilde{se}$</th>
<th>$\tilde{se}$</th>
<th>rse</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 1</td>
<td>.100</td>
<td>.112</td>
<td>.099</td>
<td>.096</td>
<td>.071</td>
<td>.073</td>
<td>.070</td>
<td>.069</td>
<td>.053</td>
<td>.053</td>
<td>.051</td>
<td>.051</td>
</tr>
<tr>
<td>1 1</td>
<td>.154</td>
<td>.165</td>
<td>.150</td>
<td>.146</td>
<td>.113</td>
<td>.114</td>
<td>.110</td>
<td>.109</td>
<td>.080</td>
<td>.081</td>
<td>.079</td>
<td>.080</td>
</tr>
<tr>
<td>.5 1</td>
<td>.066</td>
<td>.068</td>
<td>.064</td>
<td>.065</td>
<td>.059</td>
<td>.054</td>
<td>.054</td>
<td>.056</td>
<td>.044</td>
<td>.040</td>
<td>.042</td>
<td>.044</td>
</tr>
<tr>
<td>.5 1</td>
<td>.105</td>
<td>.111</td>
<td>.099</td>
<td>.096</td>
<td>.083</td>
<td>.086</td>
<td>.081</td>
<td>.080</td>
<td>.070</td>
<td>.070</td>
<td>.068</td>
<td>.068</td>
</tr>
<tr>
<td>2 1</td>
<td>.102</td>
<td>.124</td>
<td>.099</td>
<td>.093</td>
<td>.071</td>
<td>.078</td>
<td>.069</td>
<td>.068</td>
<td>.053</td>
<td>.055</td>
<td>.051</td>
<td>.050</td>
</tr>
<tr>
<td>1 1</td>
<td>.307</td>
<td>.117</td>
<td>.152</td>
<td>.263</td>
<td>.209</td>
<td>.076</td>
<td>.110</td>
<td>.198</td>
<td>.155</td>
<td>.079</td>
<td>.147</td>
<td>.147</td>
</tr>
<tr>
<td>.5 1</td>
<td>.078</td>
<td>.071</td>
<td>.064</td>
<td>.070</td>
<td>.065</td>
<td>.053</td>
<td>.054</td>
<td>.063</td>
<td>.052</td>
<td>.037</td>
<td>.042</td>
<td>.051</td>
</tr>
<tr>
<td>.5 1</td>
<td>.104</td>
<td>.126</td>
<td>.099</td>
<td>.090</td>
<td>.089</td>
<td>.095</td>
<td>.082</td>
<td>.078</td>
<td>.072</td>
<td>.074</td>
<td>.068</td>
<td>.067</td>
</tr>
<tr>
<td>3 1</td>
<td>.103</td>
<td>.117</td>
<td>.099</td>
<td>.095</td>
<td>.070</td>
<td>.075</td>
<td>.069</td>
<td>.069</td>
<td>.049</td>
<td>.053</td>
<td>.051</td>
<td>.051</td>
</tr>
<tr>
<td>1 1</td>
<td>.227</td>
<td>.133</td>
<td>.151</td>
<td>.203</td>
<td>.162</td>
<td>.089</td>
<td>.110</td>
<td>.153</td>
<td>.121</td>
<td>.061</td>
<td>.079</td>
<td>.113</td>
</tr>
<tr>
<td>.5 1</td>
<td>.071</td>
<td>.070</td>
<td>.064</td>
<td>.066</td>
<td>.062</td>
<td>.053</td>
<td>.054</td>
<td>.060</td>
<td>.048</td>
<td>.039</td>
<td>.042</td>
<td>.047</td>
</tr>
<tr>
<td>.5 1</td>
<td>.106</td>
<td>.118</td>
<td>.099</td>
<td>.093</td>
<td>.088</td>
<td>.091</td>
<td>.082</td>
<td>.079</td>
<td>.073</td>
<td>.072</td>
<td>.068</td>
<td>.067</td>
</tr>
</tbody>
</table>
Table 10.2a. Empirical Mean(sd) of CQMLE and M-Estimator, SL Model, $T = 3$, $m = 5$, $rho = 0.5$

<table>
<thead>
<tr>
<th>err</th>
<th>$n = 50$</th>
<th>$n = 100$</th>
<th>$n = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQMLE</td>
<td>M-Est</td>
<td>CQMLE</td>
</tr>
<tr>
<td>1</td>
<td>1.0190(.053)</td>
<td>.9992(.055)</td>
<td>1.0024(.035)</td>
</tr>
<tr>
<td>1</td>
<td>.9365(.133)</td>
<td>.9695(.143)</td>
<td>1.0024(.035)</td>
</tr>
<tr>
<td>.5</td>
<td>.4279(.042)</td>
<td>.5015(.047)</td>
<td>.4467(.024)</td>
</tr>
<tr>
<td>.2</td>
<td>.2331(.060)</td>
<td>.1933(.064)</td>
<td>.2114(.049)</td>
</tr>
<tr>
<td>1</td>
<td>1.0193(.052)</td>
<td>.9992(.054)</td>
<td>1.0005(.034)</td>
</tr>
<tr>
<td>1</td>
<td>.9391(.260)</td>
<td>.9743(.280)</td>
<td>.9558(.184)</td>
</tr>
<tr>
<td>.5</td>
<td>.4289(.045)</td>
<td>.5031(.048)</td>
<td>.4474(.027)</td>
</tr>
<tr>
<td>.2</td>
<td>.2335(.061)</td>
<td>.1938(.065)</td>
<td>.2124(.050)</td>
</tr>
<tr>
<td>3</td>
<td>1.0180(.055)</td>
<td>.9980(.056)</td>
<td>1.0019(.035)</td>
</tr>
<tr>
<td>1</td>
<td>.9388(.203)</td>
<td>.9730(.218)</td>
<td>.9581(.147)</td>
</tr>
<tr>
<td>.5</td>
<td>.4277(.043)</td>
<td>.5018(.047)</td>
<td>.4461(.027)</td>
</tr>
<tr>
<td>.2</td>
<td>.2354(.060)</td>
<td>.1960(.064)</td>
<td>.2121(.050)</td>
</tr>
</tbody>
</table>

**Note:** Par = $\psi = (\beta, \sigma_v^2, \rho, \lambda_1)'$; err = 1 (normal), 2 (normal mixture), and 3 (chi-square).

$X_t$ values are generated with $\theta_x = (g, \phi_1, \phi_2, \sigma_1, \sigma_2) = (.01, .5, .5, 2, 1)$, as in Footnote 1.

$W_1$ is generated according to Queen Contiguity scheme.
Table 10.2a. Cont’d, $\rho = -0.5$

<table>
<thead>
<tr>
<th>err $\psi$</th>
<th>$n = 50$</th>
<th></th>
<th></th>
<th>$n = 100$</th>
<th></th>
<th></th>
<th>$n = 200$</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQMLE</td>
<td>M-Est</td>
<td></td>
<td>CQMLE</td>
<td>M-Est</td>
<td></td>
<td>CQMLE</td>
<td>M-Est</td>
<td></td>
</tr>
<tr>
<td>1 1</td>
<td>1.0261(.057)</td>
<td>.9978(.059)</td>
<td></td>
<td>1.0188(.037)</td>
<td>.9993(.037)</td>
<td></td>
<td>1.0226(.027)</td>
<td>.9992(.027)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.9600(.136)</td>
<td>.9753(.141)</td>
<td></td>
<td>.9761(.099)</td>
<td>.9891(.102)</td>
<td></td>
<td>.9797(.068)</td>
<td>.9917(.070)</td>
<td></td>
</tr>
<tr>
<td>-.5</td>
<td>-.5577(.046)</td>
<td>-.4976(.050)</td>
<td></td>
<td>-.5543(.031)</td>
<td>-.4989(.034)</td>
<td></td>
<td>-.5513(.022)</td>
<td>-.4991(.024)</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>.1898(.096)</td>
<td>.1841(.097)</td>
<td></td>
<td>.1855(.059)</td>
<td>.1978(.059)</td>
<td></td>
<td>.1824(.044)</td>
<td>.1975(.044)</td>
<td></td>
</tr>
<tr>
<td>2 1</td>
<td>1.0254(.056)</td>
<td>.9971(.058)</td>
<td></td>
<td>1.0179(.037)</td>
<td>.9985(.037)</td>
<td></td>
<td>1.0228(.028)</td>
<td>.9993(.028)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.9557(.281)</td>
<td>.9719(.290)</td>
<td></td>
<td>.9745(.201)</td>
<td>.9878(.206)</td>
<td></td>
<td>.9857(.140)</td>
<td>.9980(.143)</td>
<td></td>
</tr>
<tr>
<td>-.5</td>
<td>-.5557(.048)</td>
<td>-.4958(.052)</td>
<td></td>
<td>-.5542(.033)</td>
<td>-.4990(.035)</td>
<td></td>
<td>-.5518(.023)</td>
<td>-.4995(.024)</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>.1985(.093)</td>
<td>.1920(.094)</td>
<td></td>
<td>.1853(.058)</td>
<td>.1974(.058)</td>
<td></td>
<td>.1815(.043)</td>
<td>.1966(.043)</td>
<td></td>
</tr>
<tr>
<td>3 1</td>
<td>1.0261(.057)</td>
<td>.9978(.059)</td>
<td></td>
<td>1.0184(.037)</td>
<td>.9989(.037)</td>
<td></td>
<td>1.0225(.028)</td>
<td>.9990(.028)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.9487(.198)</td>
<td>.9640(.204)</td>
<td></td>
<td>.9752(.155)</td>
<td>.9884(.159)</td>
<td></td>
<td>.9808(.105)</td>
<td>.9929(.108)</td>
<td></td>
</tr>
<tr>
<td>-.5</td>
<td>-.5562(.048)</td>
<td>-.4971(.052)</td>
<td></td>
<td>-.5547(.034)</td>
<td>-.4992(.037)</td>
<td></td>
<td>-.5514(.023)</td>
<td>-.4991(.025)</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>.1938(.097)</td>
<td>.1877(.098)</td>
<td></td>
<td>.1824(.058)</td>
<td>.1943(.058)</td>
<td></td>
<td>.1827(.043)</td>
<td>.1974(.044)</td>
<td></td>
</tr>
</tbody>
</table>
Table 10.2b. Empirical sd and average of estimated standard errors of M-Estimator

SL Model, $T = 3$, $m = 5$, Parameter configurations as in Table 2a.

<table>
<thead>
<tr>
<th>err $\psi$</th>
<th>$n = 50$</th>
<th>$n = 100$</th>
<th>$n = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>sd</td>
<td>$\hat{se}$</td>
<td>$\hat{se}$</td>
</tr>
<tr>
<td>1 1</td>
<td>.055</td>
<td>.060</td>
<td>.054</td>
</tr>
<tr>
<td>1 .143</td>
<td>.158</td>
<td>.142</td>
<td>.138</td>
</tr>
<tr>
<td>.5 .047</td>
<td>.049</td>
<td>.044</td>
<td>.044</td>
</tr>
<tr>
<td>.2 .064</td>
<td>.070</td>
<td>.063</td>
<td>.061</td>
</tr>
<tr>
<td>2 1</td>
<td>.054</td>
<td>.066</td>
<td>.054</td>
</tr>
<tr>
<td>1 .280</td>
<td>.105</td>
<td>.143</td>
<td>.255</td>
</tr>
<tr>
<td>.5 .048</td>
<td>.052</td>
<td>.044</td>
<td>.046</td>
</tr>
<tr>
<td>.2 .065</td>
<td>.077</td>
<td>.063</td>
<td>.061</td>
</tr>
<tr>
<td>3 1</td>
<td>.056</td>
<td>.062</td>
<td>.054</td>
</tr>
<tr>
<td>1 .218</td>
<td>.122</td>
<td>.143</td>
<td>.196</td>
</tr>
<tr>
<td>.5 .047</td>
<td>.050</td>
<td>.044</td>
<td>.045</td>
</tr>
<tr>
<td>.2 .064</td>
<td>.074</td>
<td>.063</td>
<td>.060</td>
</tr>
</tbody>
</table>
Table 10.3a. Empirical Mean(sd) of CQMLE and M-Estimator, SLE Model, $T = 3$, $m = 5$

<table>
<thead>
<tr>
<th>err ψ</th>
<th>CQMLE</th>
<th>M-Est</th>
<th>CQMLE</th>
<th>M-Est</th>
<th>CQMLE</th>
<th>M-Est</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n = 50$</td>
<td></td>
<td>$n = 100$</td>
<td></td>
<td>$n = 200$</td>
<td></td>
</tr>
<tr>
<td>1 1</td>
<td>1.0028(.051)</td>
<td>1.0012(.052)</td>
<td>.9894(.035)</td>
<td>.9990(.036)</td>
<td>1.0130(.028)</td>
<td>1.0001(.028)</td>
</tr>
<tr>
<td>1</td>
<td>.9268(.133)</td>
<td>.9542(.141)</td>
<td>.9491(.093)</td>
<td>.9800(.100)</td>
<td>.9629(.070)</td>
<td>.9911(.075)</td>
</tr>
<tr>
<td>.5</td>
<td>.4332(.041)</td>
<td>.4993(.044)</td>
<td>.4285(.029)</td>
<td>.4999(.032)</td>
<td>.4337(.019)</td>
<td>.5010(.021)</td>
</tr>
<tr>
<td>.2</td>
<td>.2064(.078)</td>
<td>.1930(.083)</td>
<td>.2190(.069)</td>
<td>.1905(.075)</td>
<td>.1871(.063)</td>
<td>.1938(.066)</td>
</tr>
<tr>
<td>.2</td>
<td>.1383(.185)</td>
<td>.1489(.183)</td>
<td>.1561(.148)</td>
<td>.1582(.146)</td>
<td>.1781(.122)</td>
<td>.1723(.120)</td>
</tr>
<tr>
<td>2 1</td>
<td>1.0006(.050)</td>
<td>.9989(.051)</td>
<td>.9908(.036)</td>
<td>1.0005(.037)</td>
<td>1.0123(.027)</td>
<td>.9995(.028)</td>
</tr>
<tr>
<td>1</td>
<td>.9219(.263)</td>
<td>.9505(.280)</td>
<td>.9495(.188)</td>
<td>.9813(.200)</td>
<td>.9626(.138)</td>
<td>.9910(.146)</td>
</tr>
<tr>
<td>.5</td>
<td>.4355(.043)</td>
<td>.5011(.045)</td>
<td>.4291(.032)</td>
<td>.5006(.034)</td>
<td>.4332(.022)</td>
<td>.5001(.023)</td>
</tr>
<tr>
<td>.2</td>
<td>.2016(.078)</td>
<td>.1881(.083)</td>
<td>.2199(.067)</td>
<td>.1921(.072)</td>
<td>.1881(.064)</td>
<td>.1940(.067)</td>
</tr>
<tr>
<td>.2</td>
<td>.1411(.175)</td>
<td>.1525(.171)</td>
<td>.1597(.148)</td>
<td>.1634(.145)</td>
<td>.1778(.122)</td>
<td>.1733(.120)</td>
</tr>
<tr>
<td>3 1</td>
<td>1.0001(.052)</td>
<td>.9984(.053)</td>
<td>.9920(.036)</td>
<td>1.0015(.037)</td>
<td>1.0121(.028)</td>
<td>.9993(.028)</td>
</tr>
<tr>
<td>1</td>
<td>.9247(.199)</td>
<td>.9527(.212)</td>
<td>.9461(.143)</td>
<td>.9771(.152)</td>
<td>.9596(.102)</td>
<td>.9875(.108)</td>
</tr>
<tr>
<td>.5</td>
<td>.4345(.042)</td>
<td>.5006(.046)</td>
<td>.4287(.031)</td>
<td>.4996(.034)</td>
<td>.4324(.021)</td>
<td>.4991(.022)</td>
</tr>
<tr>
<td>.2</td>
<td>.2038(.080)</td>
<td>.1901(.086)</td>
<td>.2209(.071)</td>
<td>.1925(.076)</td>
<td>.1898(.063)</td>
<td>.1955(.066)</td>
</tr>
<tr>
<td>.2</td>
<td>.1394(.186)</td>
<td>.1510(.182)</td>
<td>.1598(.148)</td>
<td>.1629(.145)</td>
<td>.1745(.121)</td>
<td>.1695(.118)</td>
</tr>
</tbody>
</table>

Note: Par = $\psi = (\beta, \sigma_\nu^2, \rho, \lambda_1, \lambda_3)'$; err = 1 (normal), 2 (normal mixture), and 3 (chi-square).

$X_t$ values are generated with $\theta_x = (g, \phi_1, \phi_2, \sigma_1, \sigma_2) = (.01, .5, .5, 2, 1)$, as in Footnote 1.

$W_1$ and $W_3$ are from Group Interaction scheme, and not equal; see Footnote 2.
<table>
<thead>
<tr>
<th>err $\psi$</th>
<th>$n = 50$</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>$n = 100$</th>
<th></th>
<th></th>
<th></th>
<th>$n = 200$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>sd</td>
<td>$\tilde{se}$</td>
<td>$\hat{se}$</td>
<td>$\tilde{rse}$</td>
<td>sd</td>
<td>$\tilde{se}$</td>
<td>$\hat{se}$</td>
<td>$\tilde{rse}$</td>
<td>sd</td>
<td>$\tilde{se}$</td>
<td>$\hat{se}$</td>
<td>$\tilde{rse}$</td>
<td></td>
</tr>
<tr>
<td>1 1</td>
<td>.052</td>
<td>.056</td>
<td>.051</td>
<td>.050</td>
<td>.036</td>
<td>.038</td>
<td>.036</td>
<td>.035</td>
<td>.028</td>
<td>.029</td>
<td>.028</td>
<td>.028</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.141</td>
<td>.157</td>
<td>.140</td>
<td>.137</td>
<td>.100</td>
<td>.108</td>
<td>.102</td>
<td>.100</td>
<td>.075</td>
<td>.075</td>
<td>.072</td>
<td>.072</td>
<td></td>
</tr>
<tr>
<td>.5</td>
<td>.044</td>
<td>.045</td>
<td>.042</td>
<td>.042</td>
<td>.032</td>
<td>.032</td>
<td>.031</td>
<td>.031</td>
<td>.021</td>
<td>.021</td>
<td>.021</td>
<td>.021</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>.083</td>
<td>.072</td>
<td>.080</td>
<td>.098</td>
<td>.075</td>
<td>.065</td>
<td>.072</td>
<td>.086</td>
<td>.066</td>
<td>.056</td>
<td>.064</td>
<td>.076</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>.183</td>
<td>.179</td>
<td>.160</td>
<td>.163</td>
<td>.146</td>
<td>.143</td>
<td>.135</td>
<td>.138</td>
<td>.120</td>
<td>.116</td>
<td>.114</td>
<td>.117</td>
<td></td>
</tr>
<tr>
<td>2 1</td>
<td>.051</td>
<td>.062</td>
<td>.050</td>
<td>.049</td>
<td>.037</td>
<td>.041</td>
<td>.036</td>
<td>.035</td>
<td>.028</td>
<td>.030</td>
<td>.028</td>
<td>.027</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.280</td>
<td>.107</td>
<td>.140</td>
<td>.247</td>
<td>.200</td>
<td>.067</td>
<td>.102</td>
<td>.192</td>
<td>.146</td>
<td>.043</td>
<td>.072</td>
<td>.141</td>
<td></td>
</tr>
<tr>
<td>.5</td>
<td>.045</td>
<td>.049</td>
<td>.041</td>
<td>.042</td>
<td>.034</td>
<td>.033</td>
<td>.031</td>
<td>.033</td>
<td>.023</td>
<td>.021</td>
<td>.021</td>
<td>.022</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>.083</td>
<td>.080</td>
<td>.079</td>
<td>.095</td>
<td>.072</td>
<td>.070</td>
<td>.072</td>
<td>.085</td>
<td>.067</td>
<td>.059</td>
<td>.064</td>
<td>.074</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>.171</td>
<td>.208</td>
<td>.160</td>
<td>.153</td>
<td>.145</td>
<td>.155</td>
<td>.134</td>
<td>.134</td>
<td>.120</td>
<td>.122</td>
<td>.113</td>
<td>.114</td>
<td></td>
</tr>
<tr>
<td>3 1</td>
<td>.053</td>
<td>.058</td>
<td>.051</td>
<td>.050</td>
<td>.037</td>
<td>.039</td>
<td>.036</td>
<td>.035</td>
<td>.028</td>
<td>.029</td>
<td>.028</td>
<td>.028</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.212</td>
<td>.123</td>
<td>.140</td>
<td>.190</td>
<td>.152</td>
<td>.080</td>
<td>.101</td>
<td>.144</td>
<td>.108</td>
<td>.054</td>
<td>.072</td>
<td>.105</td>
<td></td>
</tr>
<tr>
<td>.5</td>
<td>.046</td>
<td>.046</td>
<td>.041</td>
<td>.043</td>
<td>.034</td>
<td>.031</td>
<td>.031</td>
<td>.033</td>
<td>.022</td>
<td>.021</td>
<td>.021</td>
<td>.022</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>.086</td>
<td>.076</td>
<td>.080</td>
<td>.097</td>
<td>.076</td>
<td>.067</td>
<td>.072</td>
<td>.085</td>
<td>.066</td>
<td>.057</td>
<td>.064</td>
<td>.075</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>.182</td>
<td>.191</td>
<td>.160</td>
<td>.159</td>
<td>.145</td>
<td>.147</td>
<td>.134</td>
<td>.136</td>
<td>.118</td>
<td>.119</td>
<td>.114</td>
<td>.115</td>
<td></td>
</tr>
</tbody>
</table>
Table 10.4a. Empirical Mean(sd) of CQMLE and M-Estimator, STL Model, $T = 7$, $m = 5$

<table>
<thead>
<tr>
<th>err $\psi$</th>
<th>$n = 50$</th>
<th></th>
<th>$n = 100$</th>
<th></th>
<th>$n = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CQMLE</td>
<td>M-Est</td>
<td>CQMLE</td>
<td>M-Est</td>
<td>CQMLE</td>
<td>M-Est</td>
</tr>
<tr>
<td>1 1</td>
<td>1.0045 (.025)</td>
<td>1.0000 (.025)</td>
<td>1.0073 (.017)</td>
<td>.9998 (.017)</td>
<td>1.0078 (.012)</td>
</tr>
<tr>
<td>1</td>
<td>.9809 (.079)</td>
<td>.9855 (.080)</td>
<td>.9897 (.058)</td>
<td>.9940 (.058)</td>
<td>.9923 (.041)</td>
</tr>
<tr>
<td>.5</td>
<td>.4780 (.018)</td>
<td>.4996 (.019)</td>
<td>.4807 (.012)</td>
<td>.5001 (.012)</td>
<td>.4812 (.008)</td>
</tr>
<tr>
<td>.2</td>
<td>.1904 (.046)</td>
<td>.1950 (.046)</td>
<td>.1971 (.031)</td>
<td>.1994 (.031)</td>
<td>.1974 (.023)</td>
</tr>
<tr>
<td>.2</td>
<td>.2280 (.041)</td>
<td>.2043 (.042)</td>
<td>.2208 (.030)</td>
<td>.2006 (.030)</td>
<td>.2200 (.021)</td>
</tr>
<tr>
<td>2 1</td>
<td>1.0040 (.026)</td>
<td>.9994 (.026)</td>
<td>1.0078 (.017)</td>
<td>1.0003 (.018)</td>
<td>1.0074 (.012)</td>
</tr>
<tr>
<td>1</td>
<td>.9920 (.182)</td>
<td>.9968 (.184)</td>
<td>.9884 (.130)</td>
<td>.9927 (.131)</td>
<td>.9934 (.090)</td>
</tr>
<tr>
<td>.5</td>
<td>.4780 (.018)</td>
<td>.4999 (.019)</td>
<td>.4805 (.013)</td>
<td>.4999 (.013)</td>
<td>.4816 (.009)</td>
</tr>
<tr>
<td>.2</td>
<td>.1908 (.046)</td>
<td>.1954 (.046)</td>
<td>.1962 (.032)</td>
<td>.1985 (.032)</td>
<td>.1986 (.023)</td>
</tr>
<tr>
<td>.2</td>
<td>.2276 (.042)</td>
<td>.2038 (.042)</td>
<td>.2216 (.031)</td>
<td>.2014 (.031)</td>
<td>.2186 (.022)</td>
</tr>
<tr>
<td>3 1</td>
<td>1.0050 (.025)</td>
<td>1.0006 (.025)</td>
<td>1.0076 (.018)</td>
<td>1.0001 (.018)</td>
<td>1.0075 (.012)</td>
</tr>
<tr>
<td>1</td>
<td>.9815 (.135)</td>
<td>.9861 (.137)</td>
<td>.9912 (.095)</td>
<td>.9955 (.096)</td>
<td>.9903 (.067)</td>
</tr>
<tr>
<td>.5</td>
<td>.4783 (.018)</td>
<td>.4999 (.018)</td>
<td>.4805 (.012)</td>
<td>.4999 (.012)</td>
<td>.4810 (.008)</td>
</tr>
<tr>
<td>.2</td>
<td>.1905 (.048)</td>
<td>.1950 (.048)</td>
<td>.1954 (.031)</td>
<td>.1977 (.031)</td>
<td>.1978 (.023)</td>
</tr>
<tr>
<td>.2</td>
<td>.2278 (.042)</td>
<td>.2041 (.042)</td>
<td>.2226 (.030)</td>
<td>.2024 (.030)</td>
<td>.2198 (.022)</td>
</tr>
</tbody>
</table>

Note: Par = $\psi = (\beta, \sigma_v^2, \rho, \lambda_1, \lambda_2)'$; err = 1 (normal), 2 (normal mixture), and 3 (chi-square).

$X_t$ values are generated with $\theta_x = (g, \phi_1, \phi_2, \sigma_1, \sigma_2) = (.01, .5, .5, 2, 1)$, as in Footnote 1.

$W_1$ and $W_2$ are from Queen Contiguity, and equal.
Table 10.4b. Empirical sd and average of estimated standard errors of M-Estimator

STL Model, $T = 7$, $m = 5$, Parameter configurations as in Table 4a.

<table>
<thead>
<tr>
<th>err $\psi$</th>
<th>$n = 50$</th>
<th>$n = 100$</th>
<th>$n = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$</td>
<td>$\tilde{\sigma}$</td>
<td>$\hat{\sigma}$</td>
<td>$\tilde{\sigma}$</td>
</tr>
<tr>
<td>1 1</td>
<td>.025</td>
<td>.027</td>
<td>.025</td>
</tr>
<tr>
<td>1</td>
<td>.080</td>
<td>.088</td>
<td>.081</td>
</tr>
<tr>
<td>.5</td>
<td>.019</td>
<td>.019</td>
<td>.018</td>
</tr>
<tr>
<td>.2</td>
<td>.046</td>
<td>.049</td>
<td>.047</td>
</tr>
<tr>
<td>.2</td>
<td>.042</td>
<td>.044</td>
<td>.042</td>
</tr>
<tr>
<td>2 1</td>
<td>.026</td>
<td>.029</td>
<td>.025</td>
</tr>
<tr>
<td>1</td>
<td>.184</td>
<td>.046</td>
<td>.082</td>
</tr>
<tr>
<td>.5</td>
<td>.019</td>
<td>.021</td>
<td>.019</td>
</tr>
<tr>
<td>.2</td>
<td>.046</td>
<td>.052</td>
<td>.047</td>
</tr>
<tr>
<td>.2</td>
<td>.042</td>
<td>.047</td>
<td>.042</td>
</tr>
<tr>
<td>3 1</td>
<td>.025</td>
<td>.028</td>
<td>.025</td>
</tr>
<tr>
<td>1</td>
<td>.137</td>
<td>.060</td>
<td>.081</td>
</tr>
<tr>
<td>.5</td>
<td>.018</td>
<td>.020</td>
<td>.018</td>
</tr>
<tr>
<td>.2</td>
<td>.048</td>
<td>.050</td>
<td>.046</td>
</tr>
<tr>
<td>.2</td>
<td>.042</td>
<td>.045</td>
<td>.042</td>
</tr>
<tr>
<td>err ψ</td>
<td>CQMLE (n=50)</td>
<td>M-Est (n=50)</td>
<td>CQMLE (n=100)</td>
</tr>
<tr>
<td>-------</td>
<td>---------------</td>
<td>--------------</td>
<td>---------------</td>
</tr>
<tr>
<td>1 1.0076(.031)</td>
<td>.9995(.031)</td>
<td>1.0092(.024)</td>
<td>.9999(.025)</td>
</tr>
<tr>
<td>1 .9287(.133)</td>
<td>.9437(.138)</td>
<td>.9622(.099)</td>
<td>.9746(.102)</td>
</tr>
<tr>
<td>.3 .2578(.033)</td>
<td>.3000(.035)</td>
<td>.2663(.021)</td>
<td>.2992(.022)</td>
</tr>
<tr>
<td>.2 .1966(.073)</td>
<td>.1957(.075)</td>
<td>.1877(.059)</td>
<td>.1967(.060)</td>
</tr>
<tr>
<td>.2 .2209(.084)</td>
<td>.2055(.091)</td>
<td>.2240(.047)</td>
<td>.2037(.049)</td>
</tr>
<tr>
<td>.2 .1463(.185)</td>
<td>.1459(.191)</td>
<td>.1846(.132)</td>
<td>.1745(.135)</td>
</tr>
<tr>
<td>2 1.0084(.031)</td>
<td>1.0003(.032)</td>
<td>1.0091(.025)</td>
<td>.9998(.025)</td>
</tr>
<tr>
<td>1 .9288(.264)</td>
<td>.9448(.273)</td>
<td>.9591(.195)</td>
<td>.9717(.201)</td>
</tr>
<tr>
<td>.3 .2564(.035)</td>
<td>.2986(.036)</td>
<td>.2661(.022)</td>
<td>.2988(.022)</td>
</tr>
<tr>
<td>.2 .1989(.073)</td>
<td>.1983(.076)</td>
<td>.1895(.060)</td>
<td>.1985(.060)</td>
</tr>
<tr>
<td>.2 .2168(.082)</td>
<td>.2008(.089)</td>
<td>.2211(.047)</td>
<td>.2008(.050)</td>
</tr>
<tr>
<td>.2 .1367(.184)</td>
<td>.1358(.190)</td>
<td>.1805(.131)</td>
<td>.1702(.134)</td>
</tr>
<tr>
<td>3 1.0066(.031)</td>
<td>.9985(.031)</td>
<td>1.0084(.025)</td>
<td>.9991(.025)</td>
</tr>
<tr>
<td>1 .9318(.201)</td>
<td>.9476(.208)</td>
<td>.9644(.148)</td>
<td>.9769(.152)</td>
</tr>
<tr>
<td>.3 .2590(.034)</td>
<td>.3014(.036)</td>
<td>.2679(.022)</td>
<td>.3008(.022)</td>
</tr>
<tr>
<td>.2 .1983(.071)</td>
<td>.1978(.073)</td>
<td>.1879(.060)</td>
<td>.1969(.060)</td>
</tr>
<tr>
<td>.2 .2193(.081)</td>
<td>.2035(.087)</td>
<td>.2207(.047)</td>
<td>.2003(.049)</td>
</tr>
<tr>
<td>.2 .1412(.184)</td>
<td>.1403(.190)</td>
<td>.1852(.133)</td>
<td>.1750(.135)</td>
</tr>
</tbody>
</table>

**Note:** Par = ψ = (β, β^2, ρ, λ_1, λ_2, λ_3)^'; err = 1 (normal), 2 (normal mixture), and 3 (chi-square). X_t values are generated with θ_x = (g, φ_1, φ_2, σ_1, σ_2) = (.01, .5, .5, 3, 1), as in Footnote 1. W_1, W_2 and W_3 are all from Queen Contiguity, and equal.
<table>
<thead>
<tr>
<th>errψ</th>
<th>n = 50</th>
<th>n = 100</th>
<th>n = 200</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>sd</td>
<td>se</td>
<td>se</td>
</tr>
<tr>
<td>1</td>
<td>.031</td>
<td>.035</td>
<td>.031</td>
</tr>
<tr>
<td>1</td>
<td>.138</td>
<td>.156</td>
<td>.136</td>
</tr>
<tr>
<td>.3</td>
<td>.035</td>
<td>.038</td>
<td>.035</td>
</tr>
<tr>
<td>.2</td>
<td>.075</td>
<td>.080</td>
<td>.071</td>
</tr>
<tr>
<td>.2</td>
<td>.091</td>
<td>.079</td>
<td>.085</td>
</tr>
<tr>
<td>.2</td>
<td>.191</td>
<td>.209</td>
<td>.184</td>
</tr>
<tr>
<td>2</td>
<td>.032</td>
<td>.039</td>
<td>.031</td>
</tr>
<tr>
<td>1</td>
<td>.273</td>
<td>.109</td>
<td>.137</td>
</tr>
<tr>
<td>.3</td>
<td>.036</td>
<td>.042</td>
<td>.034</td>
</tr>
<tr>
<td>.2</td>
<td>.076</td>
<td>.089</td>
<td>.070</td>
</tr>
<tr>
<td>.2</td>
<td>.089</td>
<td>.088</td>
<td>.083</td>
</tr>
<tr>
<td>.2</td>
<td>.190</td>
<td>.243</td>
<td>.184</td>
</tr>
<tr>
<td>3</td>
<td>.031</td>
<td>.037</td>
<td>.031</td>
</tr>
<tr>
<td>1</td>
<td>.208</td>
<td>.124</td>
<td>.137</td>
</tr>
<tr>
<td>.3</td>
<td>.036</td>
<td>.040</td>
<td>.035</td>
</tr>
<tr>
<td>.2</td>
<td>.073</td>
<td>.084</td>
<td>.070</td>
</tr>
<tr>
<td>.2</td>
<td>.087</td>
<td>.084</td>
<td>.084</td>
</tr>
<tr>
<td>.2</td>
<td>.190</td>
<td>.224</td>
<td>.183</td>
</tr>
</tbody>
</table>
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).
- 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) for illustrating the bias-correction and refined inferences for fixed effects spatial panel data 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
- \( \text{Unemp} \) is the state unemployment rate.

This model is now extended by adding the dynamic effect and one or more spatial 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 normalised. For models with more than one spatial term, the corresponding \( W^\prime \)s are taken to be the same.
Each of the five models discussed in the chapter is estimated using
- (a) full data,
- (b) data from the last six years \((T + 1 = 6)\),
- and (c) data from first six years.

Table 10.6a summarizes the CQMLE, FQMLE, M-Est, and the standard error of the M-Est for the \(SE\) model, as for this model the full QMLE is available (Su and Yang, 2015). From the results we see that
- (i) the dynamic and \(SE\) effects are highly significant in all models,
- (ii) three methods give quite different estimates of dynamic effect,
- and (iii) the FQMLE of \(\rho\) improves over CQMLE in that it is much closer to the M-estimate in particular when \(T\) is small.

FQMLE uses \(m = 6\), and the time mean of the regressors as the predictor for the initial differences. The results are quite robust to the value of \(m\), but not quite to the choice of the predictors.
Table 10.6a. CQMLE, FQMLE, M-Est and its t-Ratio based on Munnell Data: SE Model

<table>
<thead>
<tr>
<th></th>
<th>Full Data</th>
<th>Last 6 Years</th>
<th>First 6 Years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQM</td>
<td>FQM</td>
<td>M-Est</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-.0433</td>
<td>-.0234</td>
<td>-.0467</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>-.0393</td>
<td>-.0309</td>
<td>-.0702</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>-.0024</td>
<td>-.0026</td>
<td>-.0028</td>
</tr>
<tr>
<td>$\sigma_v^2$</td>
<td>.0001</td>
<td>.0001</td>
<td>.0001</td>
</tr>
<tr>
<td>$\rho$</td>
<td>.7772</td>
<td>.8283</td>
<td>.9140</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>.7592</td>
<td>.7550</td>
<td>.7697</td>
</tr>
</tbody>
</table>

Table 10.6b summarize the results for the other four models. The results show that, for any model estimated and data used,

- (i) the dynamic effect is always significant,
- (ii) there is always at least one spatial effect that is significant,
- and (iii) the CQMLE is always significantly smaller than the corresponding M-estimate.

The empirical results are consistent with the theories.
### Table 10.6b. CQMLE, M-Est and its $t$-Ratio based on Munnell Data: Other Models

<table>
<thead>
<tr>
<th></th>
<th>Full Data</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQMLE</td>
<td>M-Est</td>
<td>$t$-ratio</td>
<td>CQMLE</td>
<td>M-Est</td>
<td>$t$-ratio</td>
<td>CQMLE</td>
<td>M-Est</td>
</tr>
<tr>
<td><strong>SL Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-0.0620</td>
<td>-0.0598</td>
<td>-1.8194</td>
<td>-0.1850</td>
<td>-0.0169</td>
<td>-2.5069</td>
<td>-0.0165</td>
<td>-0.0079</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.0296</td>
<td>0.0105</td>
<td>0.3514</td>
<td>-0.0365</td>
<td>-0.1850</td>
<td>-2.5069</td>
<td>-0.1081</td>
<td>-0.2194</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.3045</td>
<td>0.2480</td>
<td>3.1542</td>
<td>0.9917</td>
<td>0.3916</td>
<td>10.4729</td>
<td>0.3916</td>
<td>0.2369</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>-0.0025</td>
<td>-0.0027</td>
<td>-4.0988</td>
<td>-0.0016</td>
<td>-0.0018</td>
<td>-2.5384</td>
<td>-0.0018</td>
<td>-0.0018</td>
</tr>
<tr>
<td>$\sigma^2_v$</td>
<td>0.0001</td>
<td>0.0001</td>
<td>9.5094</td>
<td>0.0001</td>
<td>0.0001</td>
<td>8.6974</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.5333</td>
<td>0.6132</td>
<td>7.0194</td>
<td>0.1625</td>
<td>0.2849</td>
<td>4.4754</td>
<td>0.2849</td>
<td>0.4801</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.2131</td>
<td>0.2046</td>
<td>4.3797</td>
<td>0.2077</td>
<td>0.3767</td>
<td>4.4475</td>
<td>0.3767</td>
<td>0.4134</td>
</tr>
<tr>
<td><strong>SLE Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-0.0412</td>
<td>-0.0454</td>
<td>-1.6237</td>
<td>-0.0888</td>
<td>-0.1023</td>
<td>-1.8749</td>
<td>-0.1023</td>
<td>-0.0829</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>-0.0364</td>
<td>-0.0675</td>
<td>-1.3981</td>
<td>-0.0197</td>
<td>0.4341</td>
<td>-0.8777</td>
<td>0.4341</td>
<td>0.0429</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.2649</td>
<td>0.1685</td>
<td>1.4418</td>
<td>0.7585</td>
<td>0.4201</td>
<td>5.3430</td>
<td>0.4201</td>
<td>0.3343</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>-0.0024</td>
<td>-0.0027</td>
<td>-3.9247</td>
<td>-0.0021</td>
<td>-0.0025</td>
<td>-3.5416</td>
<td>-0.0025</td>
<td>-0.0031</td>
</tr>
<tr>
<td>$\sigma^2_v$</td>
<td>0.0001</td>
<td>0.0001</td>
<td>5.1623</td>
<td>0.0000</td>
<td>0.0000</td>
<td>4.8590</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.7752</td>
<td>0.9092</td>
<td>5.9496</td>
<td>0.4515</td>
<td>0.3754</td>
<td>8.0173</td>
<td>0.3754</td>
<td>0.6123</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>-0.0235</td>
<td>-0.0123</td>
<td>-0.3143</td>
<td>-0.0804</td>
<td>-0.3615</td>
<td>-0.8565</td>
<td>-0.3615</td>
<td>-0.1289</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.7753</td>
<td>0.7757</td>
<td>17.6446</td>
<td>0.7800</td>
<td>0.8878</td>
<td>10.7070</td>
<td>0.8878</td>
<td>0.7789</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>0.7753</td>
<td>0.7757</td>
<td>17.6446</td>
<td>0.7800</td>
<td>0.8878</td>
<td>10.7070</td>
<td>0.8878</td>
<td>0.7789</td>
</tr>
</tbody>
</table>
### Table 10.6b. Cont’d

<table>
<thead>
<tr>
<th></th>
<th>Full Data</th>
<th>Last 6 Years</th>
<th>First 6 Years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CQMLE M-Est t-ratio</td>
<td>CQMLE M-Est t-ratio</td>
<td>CQMLE M-Est t-ratio</td>
</tr>
<tr>
<td><strong>STL Model</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-0.0383 -0.0343 -1.2882</td>
<td>-0.1367 -0.1072 -3.0105</td>
<td>-0.0791 -0.0727 -0.8560</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.0215 0.0040 0.1641</td>
<td>-0.0158 -0.0262 -0.6303</td>
<td>0.1456 0.0937 0.8758</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.2414 0.1844 2.9434</td>
<td>0.7215 0.5669 5.5058</td>
<td>0.4769 0.4040 4.3346</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>-0.0011 -0.0012 -3.4687</td>
<td>-0.0014 -0.0017 -2.8457</td>
<td>-0.0017 -0.0018 -3.1086</td>
</tr>
<tr>
<td>$\sigma_v^2$</td>
<td>0.0001 0.0001 6.1872</td>
<td>0.0000 0.0000 5.0666</td>
<td>0.0000 0.0000 4.9172</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.7547 0.8474 12.1490</td>
<td>0.4757 0.6365 7.2715</td>
<td>0.4258 0.5700 4.6003</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.6662 0.681 15.2637</td>
<td>0.4890 0.5409 7.9038</td>
<td>0.5533 0.5565 10.9247</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>-0.6350 -0.6747 -11.3723</td>
<td>-0.466 -0.5797 -6.4991</td>
<td>-0.5343 -0.5775 -4.5748</td>
</tr>
<tr>
<td><strong>STLE Model</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-0.0399 -0.0432 -1.7639</td>
<td>-0.1255 -0.1071 -2.8461</td>
<td>-0.0657 -0.0322 -0.2979</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>-0.0370 -0.0617 -1.3938</td>
<td>-0.0180 -0.0264 -0.5836</td>
<td>0.1254 0.0584 0.5115</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.2146 0.1353 1.2129</td>
<td>0.7684 0.5690 3.7925</td>
<td>0.4517 0.3512 2.5418</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>-0.0023 -0.0026 -3.5825</td>
<td>-0.0017 -0.0017 -2.3548</td>
<td>-0.0015 -0.0012 -1.1755</td>
</tr>
<tr>
<td>$\sigma_v^2$</td>
<td>0.0000 0.0001 4.5221</td>
<td>0.0000 0.0000 5.0517</td>
<td>0.0000 0.0000 4.2264</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.7973 0.9164 6.2388</td>
<td>0.4484 0.6349 5.3390</td>
<td>0.4367 0.6001 3.8399</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>-0.5538 -0.5566 -5.3667</td>
<td>0.4137 0.5381 3.6888</td>
<td>0.5976 0.6711 3.9109</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.4985 0.5331 4.8853</td>
<td>-0.4138 -0.5770 -3.6064</td>
<td>-0.5514 -0.6536 -3.4999</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>0.9074 0.9059 31.9162</td>
<td>0.2058 0.0078 0.0237</td>
<td>-0.1215 -0.3409 -0.6752</td>
</tr>
</tbody>
</table>
References


