Empirical likelihood for spatial dynamic panel data models with spatial lags and spatial errors

Jianrong Rong, Yan Liu, and Yongsong Qin

College of Mathematics and Statistics, Guangxi Normal University, Guilin, Guangxi, China

ABSTRACT
We study spatial dynamic panel data models with both spatial lags and spatial errors. The empirical likelihood (EL) ratio statistics are constructed for the parameters of the models. It is shown that the limiting distributions of the EL ratio statistics are chi-square distributions, which are used to construct confidence regions for the parameters of the models. A simulation study is conducted to compare the performances of the EL based and the normal approximation (NA) based confidence regions. Simulation results show that the EL can be computationally easier than the NA method to implement in practice.

1. Introduction
Owen (1988, 1990, 1991) proposed the empirical likelihood (EL) method to construct confidence regions in non-parametric problems. There are many literatures addressing the EL method for non-spatial models such as Qin and Lawleses (1994), Chen and Keilegom (2009) and Owen (2001), among others. On the other hand, there only exist a few publications studying the EL method for spatial models. For example, Nordman (2008), Nordman and Caragea (2008), and Bandyopadhyay, Lahiri, and Nordman (2015) used the block-wise EL (BEL) proposed by Kitamura (1997) to spatial cross-sectional data. Recently, by exploring inherent martingale structures, Qin (2021) and Jin and Lee (2019) used the EL method to construct confidence intervals/regions for spatial cross-sectional data models. Further, Li, Li, and Qin (2020) extended the EL method proposed by Qin (2021) and Jin and Lee (2019) to spatial panel data models without any dynamic element. The advantages of the EL method include that the shape of confidence interval is determined by the sample automatically and the confidence interval is obtained without covariance estimation. These features are the major motivations for our current research.

Spatial panel data (SPD) models without any dynamic element are studied intensively. Refer to Anselin (1988), Elhorst (2003), Baltagi, Song, and Koh (2003), Anselin, Le Gallo, and Jayet (2008), Kapoor, Kelejian, and Prucha (2007), Lee and Yu (2010a), Lee
and Yu (2010b), Lee and Yu (2016), Fingleton (2008), Parent and LeSage (2011), among others. By taking into account the dynamic elements, Anselin (2001) proposed a spatial dynamic panel data (SDPD) model, which significantly increases the flexibility of the SPD model. For the research developments of SDPD models, refer to Elhorst (2012), Baltagi, Song, and Koh (2003), Elhorst (2005), Yang, Li, and Tse (2006), Qu, Lee, and Yu (2017), Lee and Yu (2010c), Baltagi, Egger, and Pfaffermayr (2013) and Elhorst (2010), and so on.

In above literatures for SPD and SDPD models, two major estimation approaches are employed. One is the quasi-maximum likelihood (QML) method (e.g. Yang, Li, and Tse 2006). The other is computationally more efficient method, the generalized method of moments (GMM) (e.g. Mutl 2006). To construct confidence intervals for the parameters in the models when the above methods are employed, the normal approximation (NA) approach is usually used, where one has to obtain a consistent estimator of the asymptotic covariance of estimators, which may reduce the accuracy of the interval estimation. In this article, we apply the EL method to construct confidence intervals for SDPD models with both spatial lags and spatial errors and conduct simulations to compare the performances of NA based and EL based confidence intervals. Our results show that the EL can be computationally easier than the NA method to implement in practice.

The article is organized as follows. Section 2 presents the main results. Results from a simulation study are reported in Sec. 3. All the technical details are presented in Sec. 4.

2. Main results

Suppose that there are \( n \) individual units and \( T \) time periods. In this article, the following SDPD model with both spatial lag and spatial error is considered:

\[
y_t = \rho y_{t-1} + \kappa M_n y_t + x_t \beta + z_t \gamma + \varepsilon_t, \tag{1}
\]

\[
\varepsilon_t = \lambda W_n \varepsilon_t + \nu_t, t = 1, 2, \ldots, T, \tag{2}
\]

where \( y_t = (y_{1t}, \ldots, y_{nt})' \) is an \( n \)-dimensional column vector of observed dependent variables, \( x_t = (x_{1t}, \ldots, x_{nt})' \) is an \( n \times p \) matrix of time-varying exogenous variables, \( z = (z_1, \ldots, z_n)' \) is an \( n \times q \) matrix of time-invariant exogenous variables. The scalar parameter \( \rho (|\rho| < 1) \) characterizes the dynamic effect, \( \kappa \) and \( \lambda \) are the spatial autoregressive coefficients with the absolute values being less than 1, \( \beta \) and \( \gamma \) are \( p \times 1 \) and \( q \times 1 \) regression coefficients, respectively, and \( W_n \) and \( M_n \) are \( n \times n \) spatial weighting matrices of constants. The disturbance vector \( \varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{nt})' \) is an \( n \times 1 \) vector of errors, \( \nu_t = (\nu_{1t}, \ldots, \nu_{nt})' \) is an \( n \times 1 \) column vector, and \( \{\nu_{it}\} \) are i.i.d. across \( t \) and \( i \) with zero mean and variance \( \sigma^2 \). To compare the model (2.1)-(2.3) in Su and Yang (2015) and the model (1) and (2) in this article, we re-write the model (2.1)-(2.3) in Su and Yang (2015) as follows:

\[
y_t = \rho y_{t-1} + x_t \beta + z_t \gamma + \mu + \varepsilon_t, \tag{1}
\]

\[
\varepsilon_t = \lambda W_n \varepsilon_t + \nu_t, t = 1, 2, \ldots, T, \tag{2}
\]

where \( \mu = (\mu_1, \ldots, \mu_n)' \) represents the space-specific effects and other notations are the same as the model (1) and (2). In model (1) and (2), the spatial lag term \( \kappa M_n y_t \) of the
dependent variable is introduced, but there is no space-specific effects \( \mu \). Our investigation shows that the EL statistic for the SDPD model with space-specific effects has the asymptotic distribution of a weighted sum of independent chi-squared random variables with one degree of freedom. Therefore, to use the EL method to construct confidence regions for this model, one needs to use an adjusted EL method, which is left for our future research.

We study the EL method for the above model (1) and (2) when \( y_0 \) is exogenous. In this case, we set \( y_0 \) to be a fixed constant vector as it contains no information about the model parameters. The case that \( y_0 \) is endogenous can be treated similarly and is left for our future study. For convenience, for any positive integer \( k \), we let \( \mathbf{1}_k \) be a \( k \times 1 \) vector of ones, and \( I_k \) be a \( k \times k \) identity matrix, where \( \otimes \) is the Kronecker product.

Let \( Y = (y_1', y_2', ..., y_T')', \ Y_{-1} = (y_0', y_1', ..., y_{T-1}')', \ X = (x_1', x_2', ..., x_T')', \ Z = 1_T \otimes \mathbf{z}, \ \nu = (\nu_1', \nu_2', ..., \nu_T')', \ A = A(\kappa) = I_n - \kappa M_n, \ B = B(\lambda) = I_n - \lambda W_n \) and \( \varepsilon = (\varepsilon_1', \varepsilon_2', ..., \varepsilon_T')' \). Model (1) and (2) can be rewritten into a matrix form as follows:

\[
\begin{pmatrix}
A & 0 & \cdots & 0 \\
0 & A & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & A
\end{pmatrix}
\begin{pmatrix}
y_1 \\
y_2 \\
\vdots \\
y_T
\end{pmatrix} = \begin{pmatrix}
x_1 \\
x_2 \\
\vdots \\
x_T
\end{pmatrix} \beta + \begin{pmatrix}
z \\
z \\
\vdots \\
z
\end{pmatrix} \gamma + \begin{pmatrix}
y_0 \\
y_1 \\
\vdots \\
y_{T-1}
\end{pmatrix} \rho + \begin{pmatrix}
\varepsilon_1 \\
\varepsilon_2 \\
\vdots \\
\varepsilon_T
\end{pmatrix},
\]

with

\[
\begin{pmatrix}
B & 0 & \cdots & 0 \\
0 & B & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & B
\end{pmatrix}
\begin{pmatrix}
\varepsilon_1 \\
\varepsilon_2 \\
\vdots \\
\varepsilon_T
\end{pmatrix} = \begin{pmatrix}
\nu_1 \\
\nu_2 \\
\vdots \\
\nu_T
\end{pmatrix}.
\]

or

\[\left( I_T \otimes A \right) Y = \rho Y_{-1} + X \beta + Z \gamma + \varepsilon, \quad (3)\]

with

\[\left( I_T \otimes B \right) \varepsilon = \nu. \quad (4)\]

Denote

\[\Omega(\lambda) = I_T \otimes \left( B' B \right)^{-1}, \ \Omega(\kappa) = I_T \otimes \left( A' A \right)^{-1}.\]

Let \( \theta = (\beta', \gamma', \rho)' \) and \( \psi = (\theta', \lambda, \sigma^2, \kappa)' \). We adopt the QML method to derive the estimating equations for the EL method. Under the assumption of normality (which is only used at this moment), based on (3) and (4) the log-likelihood function is

\[\tilde{L}(\psi) = -\frac{nT}{2} \log (2\pi) - \frac{nT}{2} \log \sigma^2 - \frac{1}{2} |\Omega(\kappa)| - \frac{1}{2} |\Omega(\lambda)| - \frac{1}{2 \sigma^2} \nu' \nu,\]

where \( \nu = \left( I_T \otimes B \right) \left[ (I_T \otimes A) Y - \rho Y_{-1} - X \beta - Z \gamma \right] \). It can be shown that
\[
\frac{\partial \hat{L}(\psi)}{\partial \theta} = \sigma^{-2} \hat{X} (I_T \otimes B') \nu,
\]
\[
\frac{\partial \hat{L}(\psi)}{\partial \phi} = -\frac{1}{2} \text{tr}(I_T \otimes (BCB')) + \frac{1}{2\sigma^2} \nu' (I_T \otimes (BCB')) \nu,
\]
\[
\frac{\partial \hat{L}(\psi)}{\partial \sigma^2} = -\frac{nT}{2\sigma^2} + \frac{1}{2\sigma^4} \nu',
\]
\[
\frac{\partial \hat{L}(\psi)}{\partial \kappa} = -\frac{1}{2} \text{tr}(I_T \otimes (ADA')) + \frac{1}{\sigma^2} [(I_T \otimes (BM_nA^{-1})) \hat{X} \hat{\theta}]' \nu
\]
\[
+ \frac{1}{2\sigma^4} \nu' (I_T \otimes (BADA'B^{-1})) \nu,
\]
where \( \hat{X} = (X, Z, Y_{-1}) \), \( C = (B'B)^{-1} (W'B + B'W_n)(B'B)^{-1} \), and \( D = (A'A)^{-1} (M_n'A + A'M_n)(A'A)^{-1} \). Letting above derivatives be 0, we obtain the following estimating equations of the QML method:
\[
\hat{X}' (I_T \otimes B') \nu = 0, \tag{5}
\]
\[-\sigma^2 \text{tr}(I_T \otimes (BCB')) + \nu' (I_T \otimes (BCB')) \nu = 0, \tag{6}
\]
\[-nT \nu^2 + \nu = 0, \tag{7}
\]
\[-\sigma^2 \text{tr}(I_T \otimes (ADA')) + 2 [(I_T \otimes (BM_nA^{-1})) \hat{X} \hat{\theta}]' \nu
\]
\[+ \nu' (I_T \otimes (BADA'B^{-1})) \nu = 0. \tag{8}
\]
Noting that \( \hat{X} = (X, Z, Y_{-1}) \) and \( Y_{-1} \) contains \( X, Z, \) and \( \nu \), we need to separate out \( \nu \) from \( \hat{X} \). For this purpose, denote
\[
Y_0 = (Y_{0,0}', Y_{0,1}', \ldots, Y_{0,T-1}'), Y_{0,t} = \rho'(A^{-1})y_0,
\]
\[
F = \begin{pmatrix} 0 & A^{-1} & \rho(A^2)^{-1} & \ldots & \rho^{T-2}(A^{T-1})^{-1} \\ 0 & 0 & A^{-1} & \ldots & \rho^{T-3}(A^{T-2})^{-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & A^{-1} \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix},
\]
and \( G = F \otimes B^{-1} \). Then \( Y_{-1} \) can be expressed as
\[
Y_{-1} = F'X\beta + F'Z\gamma + G'\nu + Y_0.
\]
Let \( \hat{X}_1 = (X, Z) \), \( \hat{\theta}_1 = (\beta', \gamma)' \) and \( \hat{X}_2 = F'X\beta + F'Z\gamma + Y_0 \), then (5) can be decomposed into
\[
\hat{X}_1' (I_T \otimes B') \nu = 0, \tag{9}
\]
\[
\hat{X}_2' (I_T \otimes B') \nu + \nu' G (I_T \otimes B') \nu = 0, \tag{10}
\]
and (8) can be rewritten as
\(-\sigma^2 \text{tr} \begin{pmatrix} I_T \otimes (ADA') \\ I_T \otimes (BM_nA^{-1}) \end{pmatrix} + 2 \begin{pmatrix} \sigma^2 \text{tr} \begin{pmatrix} I_T \otimes (BM_nA^{-1}) \end{pmatrix} \end{pmatrix}' \nu + \nu' \begin{pmatrix} I_T \otimes (BADA'B^{-1}) \end{pmatrix} + 2\rho \{ \begin{pmatrix} I_T \otimes (BM_nA^{-1}) \end{pmatrix} \}_{k} \}' \nu + 2\rho \nu' G [I_T \otimes (BM_nA_{-1})]' \nu = 0. \)

(11)

For convenience, let \( e = \nu, \) i.e.,

\[
e = e_{(nT) \times 1} = \begin{pmatrix} 1 \\ 2 \\ \vdots \\ nT \end{pmatrix} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_{nT} \\ \nu_1 \\ \nu_2 \\ \vdots \\ \nu_T \end{pmatrix}.
\]

(12)

Then (6, 7), and (9)–(11) can be rewritten as

\[
\begin{align*}
\tilde{X}_1'I_T \otimes B' e &= 0, \\
\tilde{X}_2'I_T \otimes B' e + e'G(I_T \otimes B') e &= 0, \\
-\sigma^2 \text{tr} \begin{pmatrix} I_T \otimes (BCB') \\ I_T \otimes (BM_nA^{-1}) \end{pmatrix} e &= 0, \\
-\sigma^2 \text{tr} \begin{pmatrix} I_T \otimes (BCB') \\ I_T \otimes (BM_nA^{-1}) \end{pmatrix} + e' \begin{pmatrix} I_T \otimes (BCB') \\ I_T \otimes (BM_nA^{-1}) \end{pmatrix}' e &= 0, \\
2 \begin{pmatrix} I_T \otimes (BM_nA^{-1}) \end{pmatrix} \tilde{X}_1 \tilde{\theta}_1' e &= 0, \\
2(\nu_1 \text{tr} \begin{pmatrix} I_T \otimes (BM_nA^{-1}) \end{pmatrix} \tilde{X}_2)' e &= 0.
\end{align*}
\]

(13)–(17)

To use the EL method, we need to represent the quadratic forms of \( e \) in above estimating equations into the linear forms of a well behaved random variables. To this end, we let \( H_1 = \frac{1}{2}(G(I_T \otimes B') + (I_T \otimes B)G'), \) \( H_2 = I_T \otimes (BCB'), \) \( H_3 = I_T \otimes (BADA'B^{-1}) \) and \( H_4 = \rho G(I_T \otimes (BM_nA^{-1})]' + \rho (I_T \otimes (BM_nA^{-1})')G'. \) Use \( h_{ij,k}, a_{i,1}, a_{i,2}, a_{i,3}, \) and \( a_{i,4} \) to denote the \((i, j)\) element of matrix \( H_k(k = 1, 2, 3, 4)\), the \(i\)-th column of the matrix \( \tilde{X}_1'I_T \otimes B' \), the \(i\)-th element of the vector \( \tilde{X}_2'I_T \otimes B' \), the \(i\)-th element of the vector \( 2 \begin{pmatrix} I_T \otimes (BM_nA^{-1}) \end{pmatrix} \tilde{X}_1 \tilde{\theta}_1' \) and the \(i\)-th element of the vector \( 2\rho \begin{pmatrix} I_T \otimes (BM_nA^{-1}) \end{pmatrix} \tilde{X}_2' \), respectively, and adapt the convention that any sum with an upper index of less than one is zero. To deal with quadratic form in (14, 15) and (17), we follow Kelejian and Prucha (2001) to introduce a martingale difference array.

Define the \( \sigma \)-fields: \( F_0 = \{ \phi, \Omega \}, \) \( F_i = \sigma (e_1, e_2, \ldots, e_i), 1 \leq i \leq nT. \) Let

\[
\tilde{M}_{ik} = h_{ii,k}(e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,k} e_j, k = 1, 2, 3, 4.
\]

(18)

Then, \( F_{i-1} \subseteq F_i, \) \( \tilde{M}_{ik} \) is \( F_i \)-measurable and \( E(\tilde{M}_{ik} | F_{i-1}) = 0. \) Thus \( \{ \tilde{M}_{ik}, F_i, 1 \leq i \leq nT \} \) form a martingale difference array and

\[
ed' H_k e - \sigma^2 \text{tr}(H_k) = \sum_{i=1}^{nT} \tilde{M}_{ik}, k = 1, 2, 3, 4.
\]

(19)

Based on (13)–(19), we propose the following EL ratio statistic for \( \psi \in R^{d+q+4} \).
\[
\begin{align*}
L(\psi) &= \sup_{p_i, 1 \leq i \leq nT} \prod_{iT = 1}^{nT} (nT p_i), \\
\end{align*}
\]
where \(\{p_i\}\) satisfy
\[
\begin{align*}
p_i &\geq 0, 1 \leq i \leq nT, \sum_{iT = 1}^{nT} p_i = 1, \\
\sum_{iT = 1}^{nT} p_i a_{i,1} e_i &= 0, \\
\sum_{iT = 1}^{nT} p_i \left\{ a_{i,2} e_i + h_{ii,1} \left( e_i^2 - \sigma^2 \right) + 2 e_i \sum_{iT = 1}^{i-1} h_{ij,1} e_j \right\} &= 0, \\
\sum_{iT = 1}^{nT} p_i \left\{ h_{ii,2} \left( e_i^2 - \sigma^2 \right) + 2 e_i \sum_{iT = 1}^{i-1} h_{ij,2} e_j \right\} &= 0, \\
\sum_{iT = 1}^{nT} p_i (e_i^2 - \sigma^2) &= 0, \\
\sum_{iT = 1}^{nT} p_i \left\{ a_{i,3} e_i + a_{i,4} e_i + h_{ii,3} \left( e_i^2 - \sigma^2 \right) + 2 e_i \sum_{iT = 1}^{i-1} h_{ij,3} e_j \\
&+ h_{ii,4} \left( e_i^2 - \sigma^2 \right) + 2 e_i \sum_{iT = 1}^{i-1} h_{ij,4} e_j \right\} &= 0.
\end{align*}
\]

Let
\[
\omega_i(\psi) = \begin{pmatrix}
a_{i,1} e_i \\
a_{i,2} e_i + h_{ii,1} \left( e_i^2 - \sigma^2 \right) + 2 e_i \sum_{iT = 1}^{i-1} h_{ij,1} e_j \\
h_{ii,2} \left( e_i^2 - \sigma^2 \right) + 2 e_i \sum_{iT = 1}^{i-1} h_{ij,2} e_j \\
e_i^2 - \sigma^2 \\
a_{i,3} e_i + a_{i,4} e_i + h_{ii,3} \left( e_i^2 - \sigma^2 \right) + 2 e_i \sum_{iT = 1}^{i-1} h_{ij,3} e_j \\
+ h_{ii,4} \left( e_i^2 - \sigma^2 \right) + 2 e_i \sum_{iT = 1}^{i-1} h_{ij,4} e_j \\
\end{pmatrix}_{(p+q+4) \times 1}
\]
where \(e_i\) is the \(i\)-th component of \((I_T \otimes B)((I_T \otimes A)Y - \rho Y_{-1} - X\beta - Z\gamma)\). Following Owen (1990), we can show that
\[
\ell(\psi) \doteq -2 \log L(\psi) = 2 \sum_{iT = 1}^{nT} \log \left\{1 + \lambda'(\psi)\omega_i(\psi)\right\},
\]
where \(\lambda(\psi) \in R^{p+q+4}\) is the solution of following equation:
\[
\frac{1}{nT} \sum_{iT = 1}^{nT} \frac{\omega_i(\psi)}{1 + \lambda'(\psi)\omega_i(\psi)} = 0.
\]

Let \(\psi_j = Ev_j^{1/1}, j = 3, 4\). Use \(vec_D(A)\) to denote the vector formed by the diagonal elements of a matrix \(A\) and \(\|a\|\) to denote the \(L_2\)-norm of a vector \(a\). To obtain the asymptotic distribution of \(\ell(\psi)\), we need following assumptions:
A1. (i) \( \nu_{jt} \) are mutually independent, and they are independent of \( x_{ks} \) and \( z_k \) for all \( j, k, t, s \).

(ii) All elements in \( (x_{it}, z_i) \) have \( 4 + \eta_1 \) moments for some \( \eta_1 > 0 \).

A2. (i) \( \{ \nu_{jt}, t = 1, ..., T, i = 1, ..., n \} \) are independent and identically distributed for all \( i \) and \( t \) with mean 0, variance \( \sigma^2 > 0 \) and \( E|\nu_{jt}|^{4+\eta_1} < \infty \) for some \( \eta_1 > 0 \).

(ii) \( \{x_{it}, t = ..., -1, 0, 1, ... \} \) and \( \{z_i\} \) are strictly exogenous and independent across \( i \).

(iii) \( \max \{|\rho|, |\kappa|, |\lambda|\} < 1 \).

A3. Let \( W_n, M_n, \{B^{-1}\} \) and \( \{A^{-1}\} \) be as described above. They satisfy the following conditions:

(i) The row and column sums of \( W_n \) are uniformly bounded in absolute value.

(ii) \( \{B^{-1}\} \) are uniformly bounded in either row or column sums, uniformly in \( \lambda \) in a compact parameter space \( \Lambda \), and \( 0 < c_1 \leq \inf_{\lambda \in \Lambda} \lambda_{\text{max}}(B'B) \leq \sup_{\lambda \in \Lambda} \lambda_{\text{max}}(B'B) \leq c_2 < \infty \) for two constants \( c_1 \) and \( c_2 \).

(ii) The row and column sums of \( M_n \) and \( \{A^{-1}\} \) are uniformly bounded in absolute value.

A4. There are constants \( c_j > 0, j = 3, 4 \), such that

\[
0 < c_3 \leq \lambda_{\text{min}} \left( (nT)^{-1} \Sigma_{p+q+4} \right) \leq \lambda_{\text{max}} \left( (nT)^{-1} \Sigma_{p+q+4} \right) \leq c_4 < \infty,
\]

where \( \lambda_{\text{min}}(H) \) and \( \lambda_{\text{max}}(H) \) denote the minimum and maximum eigenvalues of a matrix \( H \), respectively.

\[
\Sigma_{p+q+4} = \Sigma'_{p+q+4} = \text{Cov} \left\{ \sum_{i=1}^{nT} \omega_i(\psi) \right\}
\]

\[
= \begin{pmatrix}
\Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \Sigma_{14} & \Sigma_{15} \\
* & \Sigma_{22} & \Sigma_{23} & \Sigma_{24} & \Sigma_{25} \\
* & * & \Sigma_{33} & \Sigma_{34} & \Sigma_{35} \\
* & * & * & \Sigma_{44} & \Sigma_{45} \\
* & * & * & * & \Sigma_{55}
\end{pmatrix}_{(p+q+4) \times (p+q+4)}
\]

where

\[
\Sigma_{11} = \sigma^2 E \left[ \tilde{X}_1' (I_T \otimes (B'B)) \tilde{X}_1 \right],
\]

\[
\Sigma_{12} = \sigma^2 E \left[ \tilde{X}_1' (I_T \otimes (B'B)) \tilde{X}_2 \right] + \vartheta_3 E(\tilde{X}_1')(I_T \otimes B') \text{vec}_D(H_1),
\]

\[
\Sigma_{13} = \vartheta_3 E(\tilde{X}_1')(I_T \otimes B') \text{vec}_D(H_2), \Sigma_{14} = \vartheta_3 E(\tilde{X}_1')(I_T \otimes B') 1_{nT},
\]

\[
\Sigma_{15} = \vartheta_3 E(\tilde{X}_1')(I_T \otimes B') \text{vec}_D(H_3) + \vartheta_3 E(\tilde{X}_1')(I_T \otimes B') \text{vec}_D(H_4)
\]

\[
+ 2\sigma^2 E \left\{ \tilde{X}_1' (I_T \otimes B') \left[ (I_T \otimes (B \tilde{M}_n A^{-1})) \tilde{X}_1 \tilde{\theta}_1 \right] \right\}
\]

\[
+ 2\sigma^2 E \left\{ \tilde{X}_1' (I_T \otimes B') \left[ (I_T \otimes (B \tilde{M}_n A^{-1})) \tilde{X}_2 \cdot \rho \right] \right\},
\]

\[
\Sigma_{22} = (\vartheta_4 - 3\sigma^4) \| \text{vec}_D(H_1) \|^2 + 2\sigma^4 \text{tr}(H_1^2)
\]

\[
+ \sigma^2 E \left[ \tilde{X}_2' (I_T \otimes (B'B)) \tilde{X}_2 \right] + 2\vartheta_3 E(\tilde{X}_2')(I_T \otimes B') \text{vec}_D(H_1),
\]

\[
\Sigma_{23} = (\vartheta_4 - 3\sigma^4) \text{vec}_D(H_1) \text{vec}_D(H_2)
\]

\[
+ 2\sigma^4 \text{tr}(H_1H_2) + \vartheta_3 E(\tilde{X}_2')(I_T \otimes B') \text{vec}_D(H_2),
\]

\[
\Sigma_{24} = (\vartheta_4 - 3\sigma^4) \text{vec}_D(H_1) 1_{nT},
\]

\[
\Sigma_{25} = \vartheta_3 E(\tilde{X}_2')(I_T \otimes B') \text{vec}_D(H_3) + \vartheta_3 E(\tilde{X}_2')(I_T \otimes B') \text{vec}_D(H_4)
\]

\[
+ 2\sigma^2 E \left\{ \tilde{X}_2' (I_T \otimes B') \left[ (I_T \otimes (B \tilde{M}_n A^{-1})) \tilde{X}_2 \cdot \rho \right] \right\},
\]

\[
\Sigma_{33} = (\vartheta_4 - 3\sigma^4) \text{vec}_D(H_1) \text{vec}_D(H_2)
\]

\[
+ 2\sigma^4 \text{tr}(H_1H_2) + \vartheta_3 E(\tilde{X}_2')(I_T \otimes B') \text{vec}_D(H_2),
\]

\[
\Sigma_{34} = (\vartheta_4 - 3\sigma^4) \text{vec}_D(H_1) 1_{nT},
\]

\[
\Sigma_{35} = \vartheta_3 E(\tilde{X}_2')(I_T \otimes B') \text{vec}_D(H_3) + \vartheta_3 E(\tilde{X}_2')(I_T \otimes B') \text{vec}_D(H_4)
\]

\[
+ 2\sigma^2 E \left\{ \tilde{X}_2' (I_T \otimes B') \left[ (I_T \otimes (B \tilde{M}_n A^{-1})) \tilde{X}_2 \cdot \rho \right] \right\},
\]

\[
\Sigma_{44} = (\vartheta_4 - 3\sigma^4) \text{vec}_D(H_1) \text{vec}_D(H_2)
\]

\[
+ 2\sigma^4 \text{tr}(H_1H_2) + \vartheta_3 E(\tilde{X}_2')(I_T \otimes B') \text{vec}_D(H_2),
\]

\[
\Sigma_{45} = (\vartheta_4 - 3\sigma^4) \text{vec}_D(H_1) 1_{nT},
\]

\[
\Sigma_{55} = \vartheta_3 E(\tilde{X}_2')(I_T \otimes B') \text{vec}_D(H_3) + \vartheta_3 E(\tilde{X}_2')(I_T \otimes B') \text{vec}_D(H_4)
\]

\[
+ 2\sigma^2 E \left\{ \tilde{X}_2' (I_T \otimes B') \left[ (I_T \otimes (B \tilde{M}_n A^{-1})) \tilde{X}_2 \cdot \rho \right] \right\}.
\]
\[ \Sigma_{24} = (\vartheta_4 - 3\sigma^4)tr(H_1) + 2\sigma^2tr(H_1) + \vartheta_3 E(\tilde{X}_2')(I_T \otimes B')1_{nt}, \]
\[ \Sigma_{25} = (\vartheta_4 - 3\sigma^4)\text{vec}'(H_1)\text{vec}(H_3) + 2\sigma^4tr(H_1H_3) + \vartheta_3 E(\tilde{X}_2')(I_T \otimes B')\text{vec}(H_3) + \vartheta_3 E(\tilde{X}_2')(I_T \otimes B')\text{vec}(H_4) + 2\sigma^4tr(H_1H_4) + \vartheta_3 E(\tilde{X}_2')(I_T \otimes B')\text{vec}(H_4), \]
\[ \Sigma_{33} = (\vartheta_4 - 3\sigma^4)\|\text{vec}(H_2)\|^2 + 2\sigma^4tr(H_2^2), \]
\[ \Sigma_{34} = (\vartheta_4 - 3\sigma^4)tr(H_2) + 2\sigma^4tr(H_2), \]
\[ \Sigma_{35} = (\vartheta_4 - 3\sigma^4)\text{vec}'(H_2)\text{vec}(H_3) + 2\sigma^4tr(H_2H_3) + \vartheta_3 E(\tilde{X}_2')(I_T \otimes (BM_nA^{-1}))\tilde{X}_1\tilde{\theta}_1 \]
\[ + 2\sigma^2E\left\{ \left[(I_T \otimes (BM_nA^{-1}))\tilde{X}_2\tilde{\theta}_1 \right] \right\}\text{vec}(H_2) + 2\sigma^2E\left\{ \left[(I_T \otimes (BM_nA^{-1}))\tilde{X}_2\cdot\rho \right] \right\}\text{vec}(H_2), \]
\[ \Sigma_{44} = nT(\vartheta_4 - \sigma^4), \]
\[ \Sigma_{45} = (\vartheta_4 - 3\sigma^4)tr(H_3) + 2\sigma^4tr(H_3) + (\vartheta_4 - 3\sigma^4)tr(H_4) + 2\sigma^4tr(H_4) \]
\[ + 2\vartheta_3 E\left\{ \left[(I_T \otimes (BM_nA^{-1}))\tilde{X}_1\tilde{\theta}_1 \right] \right\}1_{nt} \]
\[ + 2\vartheta_3 E\left\{ \left[(I_T \otimes (BM_nA^{-1}))\tilde{X}_2\cdot\rho \right] \right\}1_{nt}, \]
\[ \Sigma_{55} = (\vartheta_4 - 3\sigma^4)\|\text{vec}(H_3)\|^2 + 2\sigma^4tr(H_3^2) + (\vartheta_4 - 3\sigma^4)\|\text{vec}(H_4)\|^2 + 2\sigma^4tr(H_4^2) \]
\[ + 4\sigma^2E\left\{ \left[(I_T \otimes (BM_nA^{-1}))\tilde{X}_1\tilde{\theta}_1 \right] \right\}\left\{ \left[(I_T \otimes (BM_nA^{-1}))\tilde{X}_1\tilde{\theta}_1 \right] \right\} \]
\[ + 4\sigma^2E\left\{ \left[(I_T \otimes (BM_nA^{-1}))\tilde{X}_2\cdot\rho \right] \right\}\left\{ \left[(I_T \otimes (BM_nA^{-1}))\tilde{X}_2\cdot\rho \right] \right\} \]
\[ + 2(\vartheta_4 - 3\sigma^4)\text{vec}'(H_3)\text{vec}(H_4) + 4\sigma^4tr(H_3H_4) \]
\[ + 8\sigma^2E\left\{ \left[(I_T \otimes (BM_nA^{-1}))\tilde{X}_1\tilde{\theta}_1 \right] \right\}\left\{ \left[(I_T \otimes (BM_nA^{-1}))\tilde{X}_2\cdot\rho \right] \right\} \]
\[ + 4\vartheta_3 E\left\{ \left[(I_T \otimes (BM_nA^{-1}))\tilde{X}_1\tilde{\theta}_1 \right] \right\}\text{vec}(H_3) \]
\[ + 4\vartheta_3 E\left\{ \left[(I_T \otimes (BM_nA^{-1}))\tilde{X}_2\cdot\rho \right] \right\}\text{vec}(H_3) \]
\[ + 4\vartheta_3 E\left\{ \left[(I_T \otimes (BM_nA^{-1}))\tilde{X}_1\tilde{\theta}_1 \right] \right\}\text{vec}(H_4) \]
\[ + 4\vartheta_3 E\left\{ \left[(I_T \otimes (BM_nA^{-1}))\tilde{X}_2\cdot\rho \right] \right\}\text{vec}(H_4). \]

A5. \( n \to \infty \) but \( T \) is fixed.
Remark 1. Conditions A1-A5 are common assumptions for spatial models. For example, A1, A2, A3 (i) and A3 (ii), which are used in Su and Yang (2015), A3 (iii) is used in assumptions 4 and 5 in Lee (2004), and the analog of $0 < c_3 \leq \hat{c}_{\min}(nT)^{-1} \sum_{p+q+4}$ is employed in the assumption of Theorem 1 in Kelejian and Prucha (2001).

Remark 2. Assumption A5 corresponds to traditional panel data models with large $n$ and small $T$, which is also employed in Su and Yang (2015). It is interesting to extend the EL method to panels with large $n$ and large $T$, which is left for our future study.

We now state the main results.

**Theorem 1.** Suppose that Assumptions A1-A5 are satisfied. Then under model (1) and (2), as $n \to \infty$,

$$
\ell(\psi) \xrightarrow{d} \chi^2_{p+q+4},
$$

where $\chi^2_{p+q+4}$ is a chi-squared distributed random variable with $p + q + 4$ degrees of freedom.

Let $z_\alpha(p + q + 4)$ satisfy $P(\chi^2_{p+q+4} \leq z_\alpha(p + q + 4)) = \alpha$ for $0 < \alpha < 1$. It follows from Theorem 1 that an EL based confidence region for $\psi$ with asymptotically correct coverage probability $\alpha$ can be constructed as

$$
\{ \psi : \ell(\psi) \leq z_\alpha(p + q + 4) \}.
$$

It is noted that one does not need to estimate $\Sigma_{p+q+4}$ in constructing confidence regions for the parameters in the model studied. In practice, R package “emplik” is recommended in implementing the EL method (http://www.ms.uky.edu/mai/EmpLik.html).

### 3. Simulations

In our simulations, we first take $y_0$ as 0. To reduce the effect of the choice of $y_0$, we generate data as follows. Based on $y_0$, generate $y_1, y_2, \ldots, y_{100+T}$ and use $y_{101}, y_{102}, \ldots, y_{100+T}$ as the final sample for $y$ with sample size $T$. Let $\theta = (b', \gamma', \rho)', \psi = (\theta', \lambda, \sigma^2, \kappa)'$. Denote $\Omega_{kk} = 2(A'A)^{-1}[(M'_{10}A + A'M_{10}) \cdot D - M'_{10}M_{10}(A'A)^{-1}], \Omega_{\lambda \lambda} = 2(B'B)^{-1}[(W_{10}B + B'W_{10}) \cdot D - W_{10}W_{10}(B'B)^{-1}], \psi_l = -W_{10}CB + B\Omega_{kk}B - Bcw_{10}^l, P_k = -M_{10}DA' + A\Omega_{kk}A' - ADM_{10}', P_Y = (I_T \otimes A)Y - X\theta$. The QMLE of $\psi$ is asymptotically normal with a limiting variance that depends on the second order partial derivatives of $\hat{L}(\psi)$. It can be shown that

$$
\frac{\partial^2 \hat{L}(\psi)}{\partial \theta \partial \theta'} = -\frac{1}{\sigma^2} \hat{X}'(I_T \otimes (B'B))\hat{X},
$$

$$
\frac{\partial^2 \hat{L}(\psi)}{\partial \theta \partial \lambda} = -\frac{1}{\sigma^2} \hat{X}'(I_T \otimes W_n)\nu - \frac{1}{\sigma^2} \hat{X}'(I_T \otimes (B'W_n))P_Y,
$$

$$
\frac{\partial^2 \hat{L}(\psi)}{\partial \theta \partial \sigma^2} = -\frac{1}{\sigma^2} \hat{X}'(I_T \otimes B')\nu, \quad \frac{\partial^2 \hat{L}(\psi)}{\partial \theta \partial \kappa} = -\frac{1}{\sigma^2} \hat{X}'(I_T \otimes (B'BM_n))Y,
$$

$$
\frac{\partial^2 \hat{L}(\psi)}{\partial \lambda \partial \nu} = \frac{1}{\sigma^2} \hat{X}'(I_T \otimes (B'BM_n))Y,
$$

$$
\frac{\partial^2 \hat{L}(\psi)}{\partial \lambda \partial \kappa} = \frac{1}{\sigma^2} \hat{X}'(I_T \otimes (B'BM_n))Y,
$$

$$
\frac{\partial^2 \hat{L}(\psi)}{\partial \sigma^2 \partial \kappa} = \frac{1}{\sigma^2} \hat{X}'(I_T \otimes (B'BM_n))Y,
$$

$$
\frac{\partial^2 \hat{L}(\psi)}{\partial \kappa \partial \nu} = \frac{1}{\sigma^2} \hat{X}'(I_T \otimes (B'BM_n))Y.
$$
\[ \frac{\partial^2 \hat{L}(\psi)}{\partial \psi^2} = -\frac{1}{2} \text{tr}(I_T \otimes P_\lambda) - \frac{1}{\sigma^2} \{(I_T \otimes W_n)Y'[(I_T \otimes (BCB'))\nu \\
- \frac{1}{2\sigma^2} \nu'(I_T \otimes P_\lambda)\nu, \]
\[ \frac{\partial^2 \hat{L}(\psi)}{\partial \lambda^2} = -\frac{1}{\sigma^4} \nu'(I_T \otimes (BCB'))\nu, \]
\[ \frac{\partial^2 \hat{L}(\psi)}{\partial \lambda \partial \kappa} = \frac{1}{\sigma^2} [(I_T \otimes (BM_n))Y'(I_T \otimes (BCB'))\nu, \]
\[ \frac{\partial^2 \hat{L}(\psi)}{\partial \sigma^2 \partial \sigma^2} = \frac{1}{\sigma^4} \nu' + \frac{nT}{2\sigma^4}, \]
\[ \frac{\partial^2 \hat{L}(\psi)}{\partial \sigma^2 \partial \kappa} = -\frac{1}{\sigma^4} \nu'[(I_T \otimes (BM_n))Y], \]
\[ \frac{\partial^2 \hat{L}(\psi)}{\partial \kappa^2} = -\frac{1}{2} \text{tr}(I_T \otimes P_\kappa) - \frac{1}{\sigma^2} \nu'(I_T \otimes (BADA'B^{-1}))[(I_T \otimes (BM_n))Y] \\
+ \frac{1}{2\sigma^2} \nu'(I_T \otimes (BADA'B^{-1}))\nu + \frac{1}{\sigma^2} [(I_T \otimes (BM_nA^{-1}M_nA^{-1}))\tilde{X}\theta]'\nu \\
- \frac{1}{\sigma^2} [(I_T \otimes (BM_nA^{-1}))\tilde{X}\theta]'[(I_T \otimes (BM_n))Y]. \]

According to Su and Yang (2015), the QMLE \( \hat{\psi} \) of \( \psi \) satisfies:
\[ \sqrt{nT}(\hat{\psi} - \psi) \overset{d}{\rightarrow} N(0, -\Sigma^{-1}), \]
where \( \Sigma = \lim_{n \to \infty} \frac{1}{nT} E[\Sigma_n(\psi)] \) and \( \Sigma_n(\psi) = \frac{\partial^2}{\partial \psi \partial \psi'} \hat{L}(\psi) \).

Based on the above asymptotic result, we can obtain the NA based confidence region for \( \psi \) as follows:
\[ \{ \psi : (\hat{\psi} - \psi)'\left(-\tilde{\Sigma}/(nT)\right)(\hat{\psi} - \psi) \leq z_x(p + q + 4) \}, \]
where \( z_x(p + q + 4) \) is defined at the end of Sec. 2 and \( \tilde{\Sigma} \) is a consistent estimator of \( \Sigma \). It can be shown that \( \frac{1}{nT}\Sigma_n(\hat{\psi}) \) can be served as \( \tilde{\Sigma} \) (e.g. Su and Yang 2015). We note that the NA method depends on the availability of a consistent estimator of the asymptotic covariance matrix in practical applications, while the EL method does not.

We conducted a small simulation study to compare the finite sample performances of the confidence regions based on EL and NA methods with confidence level \( \alpha = 0.95 \), and report the proportions of \( \ell(\psi) \leq z_{0.95}(p + q + 4) \) and \( (\hat{\psi} - \psi)'\left(-\tilde{\Sigma}/(nT)\right)(\hat{\psi} - \psi) \leq z_{0.95}(p + q + 4) \) respectively in 1,000 replications.

In these simulations, we considered the following model to generate data: \( \gamma_t = \rho \gamma_{t-1} + \kappa W_n y_t + x_i \beta + z_i \varepsilon_t + \varepsilon_t = \lambda M_n \varepsilon_t + \nu_t, t = 1, 2, 3, \) where \( y_t, y_{t-1}, x_t \) and \( z_t \) were all \( n \times 1 \) vectors. The elements of \( x_t \) were generated from \( N(0, 4) \), alternatively, the elements of \( x_t \) could be randomly generated in a similar fashion as in Hsiao, Pesaran, and Tahmiscioglu (2002), and the elements of \( z_t \) were randomly generated from Bernoulli(0.5). We chose \( \beta = 1, \gamma = 1, \rho = 0.15 \) or 0.85, and \( (\kappa, \lambda) \) were taken as \(( -0.8, -0.7), ( -0.2, -0.1), (0.2, 0.1), (0.8, 0.7) \) respectively, and \( \nu_t \)'s were i.i.d. from \( N(0, 1) \), \( t(5) \) and \( \chi^2_{j-4} \), respectively. For the contiguity weight matrix \( W_n = (W_{ij}) \), we took \( W_{ij} = 1 \), if spatial units \( i \) and \( j \) are neighbors, \( W_{ij} = 0 \) otherwise by queen contiguity rule (Anselin 1988, P.18). In addition, we took \( W_n = M_n \) in the simulations. We considered five ideal cases of spatial units: \( n = m \times m \) regular grid with \( m = \)
The error term behaves well with coverage probabilities very close to the nominal level 0.95 when the distribution of the error terms affects both the NA and EL methods, but the EL method converges to the nominal level 0.95 as the number of spatial units is large, but not well in other cases. The coverage probabilities of the confidence regions based on NA fall to the ranges [0.784, 0.939] (at \( \rho = 0.15 \)) and [0.752, 0.875] (at \( \rho = 0.15 \)) for the \( \chi^2 \) distribution, which are far from the nominal level 0.95. On the other hand, the confidence regions based on EL method converge to the nominal level 0.95 as the number of spatial units \( n \) is large enough, whether the error term \( \nu_{it} \) is normally distributed or not. In other words, the distribution of the error terms affects both the NA and EL methods, but the EL method performs much better than the NA method. This can be explained as follows: the EL

### Table 1. Coverage probabilities of the NA and EL confidence regions with \( \nu_{it} \sim N(0, 1) \) and \( \rho = 0.15 \).

| \((\kappa, \lambda)\) | \(W_n\) | NA   | EL   | \((\kappa, \lambda)\) | \(W_n\) | NA   | EL   |
|-----------------------|--------|------|------|-----------------------|--------|------|------|
| \((-0.8, -0.7)\)      | grid_{49} | 0.958 | 0.902 | \((-0.2, -0.1)\)     | grid_{49} | 0.908 | 0.913 |
| grid_{100}           | 0.942  | 0.927 |      | grid_{100}           | 0.956  | 0.930 |
| grid_{169}           | 0.954  | 0.934 |      | grid_{169}           | 0.956  | 0.936 |
| grid_{256}           | 0.972  | 0.924 |      | grid_{256}           | 0.960  | 0.958 |
| grid_{400}           | 0.950  | 0.928 |      | grid_{400}           | 0.970  | 0.943 |

### Table 2. Coverage probabilities of the NA and EL confidence regions with \( \nu_{it} \sim N(0, 1) \) and \( \rho = 0.85 \).

| \((\kappa, \lambda)\) | \(W_n\) | NA   | EL   | \((\kappa, \lambda)\) | \(W_n\) | NA   | EL   |
|-----------------------|--------|------|------|-----------------------|--------|------|------|
| \((-0.8, -0.7)\)      | grid_{49} | 0.930 | 0.928 | \((-0.2, -0.1)\)     | grid_{49} | 0.922 | 0.934 |
| grid_{100}           | 0.946  | 0.942 |      | grid_{100}           | 0.934  | 0.964 |
| grid_{169}           | 0.964  | 0.938 |      | grid_{169}           | 0.954  | 0.932 |
| grid_{256}           | 0.965  | 0.976 |      | grid_{256}           | 0.965  | 0.946 |
| grid_{400}           | 0.960  | 0.952 |      | grid_{400}           | 0.975  | 0.952 |

### Table 3. Coverage probabilities of the NA and EL confidence regions with \( \nu_{it} \sim t(5) \) and \( \rho = 0.15 \).

| \((\kappa, \lambda)\) | \(W_n\) | NA   | EL   | \((\kappa, \lambda)\) | \(W_n\) | NA   | EL   |
|-----------------------|--------|------|------|-----------------------|--------|------|------|
| \((-0.8, -0.7)\)      | grid_{49} | 0.814 | 0.798 | \((-0.2, -0.1)\)     | grid_{49} | 0.835 | 0.816 |
| grid_{100}           | 0.844  | 0.868 |      | grid_{100}           | 0.842  | 0.902 |
| grid_{169}           | 0.844  | 0.910 |      | grid_{169}           | 0.858  | 0.916 |
| grid_{256}           | 0.835  | 0.922 |      | grid_{256}           | 0.876  | 0.886 |
| grid_{400}           | 0.828  | 0.916 |      | grid_{400}           | 0.848  | 0.930 |

### Table 4. Coverage probabilities of the NA and EL confidence regions with \( \nu_{it} \sim t(5) \) and \( \rho = 0.85 \).

| \((\kappa, \lambda)\) | \(W_n\) | NA   | EL   | \((\kappa, \lambda)\) | \(W_n\) | NA   | EL   |
|-----------------------|--------|------|------|-----------------------|--------|------|------|
| \((-0.8, -0.7)\)      | grid_{49} | 0.798 | 0.814 | \((-0.2, -0.1)\)     | grid_{49} | 0.822 | 0.813 |
| grid_{100}           | 0.844  | 0.872 |      | grid_{100}           | 0.858  | 0.907 |
| grid_{169}           | 0.816  | 0.932 |      | grid_{169}           | 0.822  | 0.909 |
| grid_{256}           | 0.865  | 0.924 |      | grid_{256}           | 0.840  | 0.912 |
| grid_{400}           | 0.784  | 0.939 |      | grid_{400}           | 0.812  | 0.925 |

7, 10, 13, 16, 20, denoting \( W_n \) as \( grid_{49}, grid_{100}, grid_{169}, grid_{256}, \) and \( grid_{400} \), respectively. A transformation is often used in applications to convert the matrix \( W_n \) to the unity of row sums. The results of simulations are reported in Tables 1–6.

From the simulation results, we can see that the confidence regions based on NA behave well with coverage probabilities very close to the nominal level 0.95 when the error term \( \nu_{it} \) is normally distributed and \( n \) is large, but not well in other cases. The coverage probabilities of the confidence regions based on NA fall to the ranges [0.784, 0.876] (at \( \rho = 0.15 \)) and [0.752, 0.875] (at \( \rho = 0.85 \)) for \( t \) distribution, and [0.734, 0.885] (at \( \rho = 0.15 \)) and [0.767, 0.888] (at \( \rho = 0.85 \)) for \( \chi^2 \) distribution, which are far from the nominal level 0.95. On the other hand, the confidence regions based on EL method converge to the nominal level 0.95 as the number of spatial units \( n \) is large enough, whether the error term \( \nu_{it} \) is normally distributed or not. In other words, the distribution of the error terms affects both the NA and EL methods, but the EL method performs much better than the NA method. This can be explained as follows: the EL
method is a nonparametric method, which does not require to specify the distribution of the data, but the NA method is based on the QML, which is close to a parametric method. So the distribution of the error term has a bigger impact on the NA method than the EL method. Our simulation results recommend EL method when we are not sure whether the errors are normally distributed.

4. Proofs

To prove the main results, we need to use Theorem 1 in Kelejian and Prucha (2001). We now state this result. Let

\[ \tilde{Q}_n = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{nij} \epsilon_i \epsilon_j + \sum_{i=1}^{n} b_{ni} \epsilon_i, \]

Table 4. Coverage probabilities of the NA and EL confidence regions with \( \nu_{10} \sim t(5) \) and \( \rho = 0.85 \).

| \((\kappa, \lambda)\) | \(W_n\) | NA | EL | \((\kappa, \lambda)\) | \(W_n\) | NA | EL |
|---------------------|--------|----|----|---------------------|--------|----|----|
| \((-0.8, -0.7)\)    | grid_{49} | 0.800 | 0.928 | \((-0.2, -0.1)\)   | grid_{49} | 0.836 | 0.924 |
|                     | grid_{100} | 0.850 | 0.944 |                     | grid_{100} | 0.868 | 0.932 |
|                     | grid_{69} | 0.860 | 0.954 |                     | grid_{69} | 0.816 | 0.938 |
|                     | grid_{56} | 0.865 | 0.960 |                     | grid_{56} | 0.865 | 0.938 |
|                     | grid_{35} | 0.820 | 0.966 |                     | grid_{35} | 0.809 | 0.949 |
|                     | grid_{39} | 0.792 | 0.902 | \((0.2, 0.1)\)     | grid_{39} | 0.832 | 0.932 |
|                     | grid_{100} | 0.762 | 0.948 |                     | grid_{100} | 0.846 | 0.940 |
|                     | grid_{69} | 0.770 | 0.950 |                     | grid_{69} | 0.848 | 0.952 |
|                     | grid_{56} | 0.805 | 0.934 |                     | grid_{56} | 0.875 | 0.958 |
|                     | grid_{400} | 0.752 | 0.932 |                     | grid_{400} | 0.840 | 0.948 |

Table 5. Coverage probabilities of the NA and EL confidence regions with \( \nu_{10} + 4 \sim \chi^2_4 \) and \( \rho = 0.15 \).

| \((\kappa, \lambda)\) | \(W_n\) | NA | EL | \((\kappa, \lambda)\) | \(W_n\) | NA | EL |
|---------------------|--------|----|----|---------------------|--------|----|----|
| \((-0.8, -0.7)\)    | grid_{49} | 0.872 | 0.832 | \((-0.2, -0.1)\)   | grid_{49} | 0.786 | 0.836 |
|                     | grid_{100} | 0.865 | 0.904 |                     | grid_{100} | 0.843 | 0.906 |
|                     | grid_{69} | 0.874 | 0.896 |                     | grid_{69} | 0.826 | 0.922 |
|                     | grid_{56} | 0.875 | 0.932 |                     | grid_{56} | 0.885 | 0.926 |
|                     | grid_{35} | 0.876 | 0.933 |                     | grid_{35} | 0.856 | 0.927 |
|                     | grid_{39} | 0.740 | 0.830 | \((0.2, 0.1)\)     | grid_{39} | 0.822 | 0.847 |
|                     | grid_{100} | 0.746 | 0.888 |                     | grid_{100} | 0.812 | 0.901 |
|                     | grid_{69} | 0.734 | 0.918 |                     | grid_{69} | 0.854 | 0.918 |
|                     | grid_{56} | 0.796 | 0.926 |                     | grid_{56} | 0.828 | 0.925 |
|                     | grid_{400} | 0.760 | 0.916 |                     | grid_{400} | 0.845 | 0.928 |

Table 6. Coverage probabilities of the NA and EL confidence regions with \( \nu_{10} + 4 \sim \chi^2_4 \) and \( \rho = 0.85 \).

| \((\kappa, \lambda)\) | \(W_n\) | NA | EL | \((\kappa, \lambda)\) | \(W_n\) | NA | EL |
|---------------------|--------|----|----|---------------------|--------|----|----|
| \((-0.8, -0.7)\)    | grid_{49} | 0.864 | 0.882 | \((-0.2, -0.1)\)   | grid_{49} | 0.814 | 0.918 |
|                     | grid_{100} | 0.848 | 0.930 |                     | grid_{100} | 0.835 | 0.958 |
|                     | grid_{69} | 0.858 | 0.944 |                     | grid_{69} | 0.882 | 0.922 |
|                     | grid_{56} | 0.860 | 0.950 |                     | grid_{56} | 0.888 | 0.954 |
|                     | grid_{35} | 0.896 | 0.958 |                     | grid_{35} | 0.868 | 0.946 |
|                     | grid_{39} | 0.770 | 0.894 | \((0.2, 0.1)\)     | grid_{39} | 0.814 | 0.910 |
|                     | grid_{100} | 0.767 | 0.940 |                     | grid_{100} | 0.852 | 0.938 |
|                     | grid_{69} | 0.770 | 0.932 |                     | grid_{69} | 0.864 | 0.936 |
|                     | grid_{56} | 0.788 | 0.935 |                     | grid_{56} | 0.875 | 0.940 |
|                     | grid_{400} | 0.800 | 0.941 |                     | grid_{400} | 0.850 | 0.944 |
where $\epsilon_{ni}$ are real valued random variables, and the $a_{nij}$ and $b_{ni}$ denote the real valued coefficients of the linear-quadratic form. We need the following assumptions in Lemma 1.

(C1) $\{\epsilon_{ni}, 1 \leq i \leq n\}$ are independent random variables with mean 0 and $\sup_{1 \leq i \leq n, n \geq 1} E|\epsilon_{ni}|^{4+\eta_1} < \infty$ for some $\eta_1 > 0$;

(C2) For all $1 \leq i, j \leq n, n \geq 1$, $a_{nij} = a_{nji}$, $\sup_{1 \leq i \leq n, n \geq 1} \sum_{i=1}^{n} |a_{nij}| < \infty$, and $\sup_{n \geq 1} n^{-1} \sum_{i=1}^{n} |b_{ni}|^{2+\eta_2} < \infty$ for some $\eta_2 > 0$.

Given above assumptions (C1) and (C2), the mean and variance of $\tilde{Q}_n$ are given as (e.g. Kelejian and Prucha 2001):

\[ \nu_{\tilde{Q}_n} = \sum_{i=1}^{n} a_{nii} \sigma_{ni}^2, \]

\[ \sigma_{\tilde{Q}_n}^2 = 2 \sum_{i=1}^{n} \sum_{j=1}^{n} a_{nij} \sigma_{ni}^2 \sigma_{nj}^2 + \sum_{i=1}^{n} b_{ni}^2 \sigma_{ni}^2 + \sum_{i=1}^{n} \{a_{nii}^2 (\mu_{ni}^{(4)} - 3 \sigma_{ni}^4) + 2 b_{ni} a_{nii} \mu_{ni}^{(3)}\}, \]

with $\sigma_{ni}^2 = E(\epsilon_{ni}^2)$ and $\mu_{ni}^{(s)} = E(\epsilon_{ni}^s)$ for $s = 3, 4$.

**Lemma 1.** Suppose that Assumptions C1 and C2 hold true and $n^{-1} \sigma_{\tilde{Q}_n}^2 \geq c$ for some constant $c > 0$. Then

\[ \frac{\tilde{Q}_n - \nu_{\tilde{Q}_n}}{\sigma_{\tilde{Q}_n}} \xrightarrow{d} N(0, 1). \]

**Proof.** See Theorem 1 in Kelejian and Prucha (2001).

**Lemma 2.** Let $\xi_1, \xi_2, \ldots, \xi_n$ be a sequence of stationary random variables, with $E|\xi_1|^s < \infty$ for some constants $s > 0$. Then

\[ \max_{1 \leq i \leq n} |\xi_i| = o(n^{1/s}), \ a.s. \]

**Proof.** Straightforward.

**Lemma 3.** Suppose that Assumptions A1-A5 are satisfied. Then as $n \to \infty$,

\[ Z_n = \max_{1 \leq i \leq \lfloor nT \rfloor} \|\omega_i(\psi)\| = o_p((nT)^{2/(4+\eta_1)}) \ a.s., \]  \hspace{1cm} (23)

\[ \Sigma_{p+q+4}^{-1/2} \sum_{i=1}^{nT} \omega_i(\psi) \xrightarrow{d} N(0, I_{p+q+4}), \]  \hspace{1cm} (24)

\[ (nT)^{-1} \sum_{i=1}^{nT} \omega_i(\psi)\omega_i'(\psi) = (nT)^{-1} \Sigma_{p+q+4} + o_p(1), \]  \hspace{1cm} (25)
\[
\sum_{i=1}^{nT} \|o_i(\psi)\|^3 = O_p(nT).
\] (26)

**Proof.** The proof of this lemma is similar to that of Lemma 3 in Qin (2021). However, there are a few differences in detail. To make things clear, we here present the detailed proof of this lemma.

Note that
\[
Z_n \leq \max_{1 \leq i \leq nT} \left\{ \max_{1 \leq i \leq nT} |a_{i,1}e_i|, \max_{1 \leq i \leq nT} |a_{i,2}e_i + h_{ii,1}(e_i^2 - \sigma^2)| \right. \\
+ 2e_i \sum_{j=1}^{i-1} |h_{ij,1}e_j|, \max_{1 \leq i \leq nT} |h_{ii,2}(e_i^2 - \sigma^2)| \\
+ 2e_i \sum_{j=1}^{i-1} |h_{ij,1}e_j|, \max_{1 \leq i \leq nT} |e_i^2 - \sigma^2|, \max_{1 \leq i \leq nT} |a_{i,3}e_i + a_{i,4}e_i + h_{ii,3}(e_i^2 - \sigma^2)| \\
+ 2e_i \sum_{j=1}^{i-1} |h_{ij,3}e_j + h_{ii,4}(e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,4}e_j| \right\}.
\]

By Conditions A1-A3 and **Lemma 2**, we have
\[
\max_{1 \leq i \leq nT} |a_{i,k}e_i| = o_p\left((nT)^{1/(4+\eta_1)}\right), k = 1, 2, 3, 4,
\]
\[
\max_{1 \leq i \leq nT} |e_i^2 - \sigma^2| = o_p\left((nT)^{2/(4+\eta_1)}\right).
\]

In addition, by Lemma B.2 in Su and Yang (2015), \(G(I_T \otimes B'), (I_T \otimes (BCB')), (I_T \otimes (BADA'B^{-1}))\) and \(G(I_T \otimes (BM_nA^{-1}))'\) are uniformly bounded in both row and column sums. It follows that
\[
\max_{1 \leq i \leq nT} |h_{ii,k}(e_i^2 - \sigma^2)| = \max_{1 \leq i \leq nT} |h_{ii,k}|o_p\left((nT)^{2/(4+\eta_1)}\right) \\
= o_p\left((nT)^{2/(4+\eta_1)}\right), k = 1, 2, 3, 4,
\]
\[
\max_{1 \leq i \leq nT} |e_i \sum_{j=1}^{i-1} h_{ij,k}e_j| \leq \left( \max_{1 \leq i \leq nT} |e_i| \right)^2 \cdot \max_{1 \leq i \leq nT} \left( \sum_{j=1}^{i-1} |h_{ij,k}| \right) \\
= o_p\left((nT)^{2/(4+\eta_1)}\right), k = 1, 2, 3, 4.
\]

Thus \(Z_n = o_p((nT)^{2/(4+\eta_1)})\), and (23) is proved.

We now prove (24). For any given \(l = (l_1, l_2, l_3, l_4, l_5)' \in R^{p+q+4}\) with \(\|l\| = 1\), where \(l_1 \in R^{p+q}, l_2, l_3, l_4, l_5 \in R\), it is clear that
\[ l' \omega_i(\psi) = l'_1 a_{i,1} e_i + l_2 \left\{ a_{i,2} e_i + h_{ii,1} (e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,1} e_j \right\} + l_3 \left\{ h_{ii,2} (e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,2} e_j \right\} + l_4 (e_i^2 - \sigma^2) + l_5 \left\{ a_{i,3} e_i + a_{i,4} e_i + h_{ii,3} (e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,3} e_j + h_{ii,4} (e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,4} e_j \right\} = (l_2 h_{ii,1} + l_3 h_{ii,2} + l_4 + l_5 h_{ii,3} + l_5 h_{ii,4})(e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} (l_2 h_{jj,1} + l_3 h_{jj,2} + l_5 h_{jj,3} + l_5 h_{jj,4}) e_j + (l'_1 a_{i,1} + l_2 a_{i,2} + l_5 a_{i,3} + l_5 a_{i,4}) e_i. \]

Denote

\[ Q_n = \sum_{i=1}^{nT} l' \omega_i(\psi) = \sum_{i=1}^{nT} \sum_{j=1}^{nT} u_{ij} e_i e_j + \sum_{i=1}^{nT} b_i e_i = e' U_{1n} e + U_{2n} e, \]

where

\[ U_{1n} = (u_{ij})_{(nT) \times (nT)}, \quad U_{2n} = (b_i)_{1 \times (nT)}, \]

\[ u_{ii} = l_2 h_{ii,1} + l_3 h_{ii,2} + l_4 + l_5 h_{ii,3} + l_5 h_{ii,4}, \]

\[ u_{ij} = l_2 h_{jj,1} + l_3 h_{jj,2} + l_5 h_{jj,3} + l_5 h_{jj,4} (i \neq j), \]

\[ b_i = l'_1 a_{i,1} + l_2 a_{i,2} + l_5 a_{i,3} + l_5 a_{i,4}. \]

Note that

\[ U_{1n} = l'_1 \tilde{X}_1 (I_T \otimes B') + l_2 \tilde{X}_2 (I_T \otimes B') + 2l_5 [(I_T \otimes (BM_n A^{-1})) \tilde{X}_1 \tilde{t}^1] + 2l_5 [(I_T \otimes (BM_n A^{-1})) \tilde{X}_2 \cdot \rho]', \]

The conditional expectation and variance given \( X, Z \) are denoted as \( E^* \) and \( \text{Var}^* \), respectively. Then from (12) and note that \( E(\nu) = 0 \), we know that the variance of \( Q_n \) is

\[ \sigma^2_{Q_n} = \text{Var} \left( \sum_{i=1}^{nT} l' \omega_i(\psi) \right) = \text{Var}(\nu' U_{1n} \nu) + \text{Var}(U_{2n} \nu) + 2\text{Cov}(\nu' U_{1n} \nu, U_{2n} \nu), \]

and

\[ \text{Var}^*_{Q_n} = (\var_4 - 3\sigma^4) \| \text{vec}_D(U_{1n}) \|_2^2 + \sigma^4 [tr(U_{1n} U_{1n}') + tr(U_{2n} U_{2n}')] + \sigma^2 U_{2n} U_{2n}^t + 2\var_3 U_{2n} \text{vec}_D(U_{1n}). \]

Further
\[
\| \text{vec}_D(U_{1n}) \|^2 = \| \text{vec}_D(l_2H_1 + l_3H_2 + l_4H_3 + l_5H_4) \|^2 \\
= \| l_2 \text{vec}_D(H_1) + l_3 \text{vec}_D(H_2) + l_4 \text{vec}_D(I_{nT}) \\
+ l_5 \text{vec}_D(H_3) + l_5 \text{vec}_D(H_4) \|^2 \\
= l_2^2 \| \text{vec}_D(H_1) \|^2 + l_3^2 \| \text{vec}_D(H_2) \|^2 \\
+ l_4^2 \text{tr}(H_2) + l_5^2 \| \text{vec}_D(H_3) \|^2 + l_5^2 \| \text{vec}_D(H_4) \|^2 \\
+ 2l_2l_3 \text{vec}_D^T(H_1) \text{vec}_D(H_2) + 2l_2l_4 \text{tr}(H_1) \\
+ 2l_2l_5 \text{vec}_D^T(H_1) \text{vec}_D(H_3) + 2l_2l_5 \text{vec}_D^T(H_1) \text{vec}_D(H_4) \\
+ 2l_3l_4 \text{tr}(H_2) + 2l_3l_5 \text{vec}_D^T(H_2) \text{vec}_D(H_3) \\
+ 2l_3l_5 \text{vec}_D^T(H_2) \text{vec}_D(H_4) + 2l_4l_5 \text{tr}(H_3) \\
+ 2l_4l_5 \text{tr}(H_4) + 2l_5^2 \text{vec}_D^T(H_3) \text{vec}_D(H_4) \\
= \tilde{l}' G_1 \tilde{l},
\]

where \( \tilde{l} = (l_2, l_3, l_4, l_5, l_5)' \),

\[
G_1 = \begin{pmatrix}
\| \text{vec}_D(H_1) \|^2 & \text{vec}_D^T(H_1) \text{vec}_D(H_2) & \text{tr}(H_1) & \text{vec}_D^T(H_1) \text{vec}_D(H_3) & \text{vec}_D^T(H_1) \text{vec}_D(H_4) \\
* & \| \text{vec}_D(H_2) \|^2 & \text{tr}(H_2) & \text{vec}_D^T(H_2) \text{vec}_D(H_3) & \text{vec}_D^T(H_2) \text{vec}_D(H_4) \\
* & * & nT & \text{tr}(H_3) & \text{tr}(H_3) \\
* & * & * & \| \text{vec}_D(H_3) \|^2 & \text{vec}_D^T(H_3) \text{vec}_D(H_4) \\
* & * & * & * & \| \text{vec}_D(H_4) \|^2 
\end{pmatrix}.
\]

And

\[
\text{tr}(U_{1n} U_{1n}') = l_2^2 \text{tr}(H_1^2) + 2l_2l_3 \text{tr}(H_1H_2) + 2l_2l_4 \text{tr}(H_1) \\
+ 2l_2l_5 \text{tr}(H_1H_3) + 2l_2l_5 \text{tr}(H_1H_4) \\
+ l_3^2 \text{tr}(H_2^2) + 2l_3l_4 \text{tr}(H_2) + 2l_3l_5 \text{tr}(H_2H_3) \\
+ 2l_3l_5 \text{tr}(H_2H_4) + l_4^2 nT \\
+ 2l_4l_5 \text{tr}(H_3) + 2l_4l_5 \text{tr}(H_4) + l_5^2 \text{tr}(H_3^2) \\
+ 2l_5^2 \text{tr}(H_4H_3) + l_5^2 \text{tr}(H_4^2) \\
= \tilde{l}' G_2 \tilde{l},
\]

where

\[
G_2 = \begin{pmatrix}
\text{tr}(H_1^2) & \text{tr}(H_1H_2) & \text{tr}(H_1) & \text{tr}(H_1H_3) & \text{tr}(H_1H_4) \\
* & \text{tr}(H_2^2) & \text{tr}(H_2) & \text{tr}(H_2H_3) & \text{tr}(H_2H_4) \\
* & * & nT & \text{tr}(H_3) & \text{tr}(H_4) \\
* & * & * & \text{tr}(H_3^2) & \text{tr}(H_3H_4) \\
* & * & * & * & \text{tr}(H_4^2) 
\end{pmatrix}.
\]

Moreover,
\[ U_{2n}U'_{2n} \]
\[ = I_1' \tilde{X}_1' (I_T \otimes (B'B)) \tilde{X}_1 I_1 + I_2' \tilde{X}_2' (I_T \otimes (B'B)) \tilde{X}_2 \]
\[ + 2I_1' \tilde{X}_1' (I_T \otimes (B'B)) \tilde{X}_2 I_2 \]
\[ + 4I_2' \left[ \left( I_T \otimes (BMI) \right) \tilde{X}_1 I_1 - \left( I_T \otimes (BMI) \right) \tilde{X}_2 I_2 \right] \]
\[ + 4I_2' \left[ \left( I_T \otimes (BMI) \right) \tilde{X}_2 \cdot \rho \left( I_T \otimes (BMI) \right) \tilde{X}_2 \cdot \rho \right] \]
\[ + 4I_2' \tilde{X}_2' (I_T \otimes B') \left( I_T \otimes (BMI) \right) \tilde{X}_1 I_1 \]
\[ + 4I_2' \tilde{X}_2' (I_T \otimes B') \left( I_T \otimes (BMI) \right) \tilde{X}_2 I_2 \]
\[ + 8I_2' \left( I_T \otimes (BMI) \right) \tilde{X}_1 I_1 \]
\[ + \left( I_T \otimes (BMI) \right) \tilde{X}_2 I_2 \]
\[ + o_p(n). \]

It is easy to show (e.g., Su and Yang 2015) that, \( n^{-1} \tilde{X}_1' (I_T \otimes (B'B)) \tilde{X}_1, \ n^{-1} \tilde{X}_2' (I_T \otimes (B'B)) \tilde{X}_2, \ n^{-1} [\tilde{X}_1' (I_T \otimes (BMI)) \tilde{X}_1 I_1 - \tilde{X}_2' (I_T \otimes (BMI)) \tilde{X}_2 I_2], \ n^{-1} [\tilde{X}_2' (I_T \otimes (BMI)) \tilde{X}_2 \cdot \rho], \ n^{-1} \tilde{X}_1' (I_T \otimes (B'B)) \tilde{X}_2, \ n^{-1} \tilde{X}_2' (I_T \otimes (B'B)) \tilde{X}_1 I_1, \ n^{-1} \tilde{X}_2' (I_T \otimes (B'B)) \tilde{X}_2 I_2, \ n^{-1} \tilde{X}_2' (I_T \otimes (BMI)) \tilde{X}_1 I_1, \ n^{-1} \tilde{X}_2' (I_T \otimes (BMI)) \tilde{X}_2 I_2 \]

converge in probability to their expectations. Therefore,
\[ U_{2n}U'_{2n} \]
\[ = I_1' E \left[ \tilde{X}_1' (I_T \otimes (B'B)) \tilde{X}_1 \right] I_1 + I_2' E \left[ \tilde{X}_2' (I_T \otimes (B'B)) \tilde{X}_2 \right] 
\[ + 2I_1' E \left[ \tilde{X}_1' (I_T \otimes (B'B)) \tilde{X}_2 \right] I_2 \]
\[ + 4I_2' E \left[ \left( I_T \otimes (BMI) \right) \tilde{X}_1 I_1 - \left( I_T \otimes (BMI) \right) \tilde{X}_2 I_2 \right] \]
\[ + 4I_2' E \left[ \left( I_T \otimes (BMI) \right) \tilde{X}_2 \cdot \rho \left( I_T \otimes (BMI) \right) \tilde{X}_2 \cdot \rho \right] \]
\[ + 4I_2' \tilde{X}_2' (I_T \otimes B') \left( I_T \otimes (BMI) \right) \tilde{X}_1 I_1 \]
\[ + 4I_2' \tilde{X}_2' (I_T \otimes B') \left( I_T \otimes (BMI) \right) \tilde{X}_2 I_2 \]
\[ + 8I_2' \left( I_T \otimes (BMI) \right) \tilde{X}_1 I_1 \]
\[ + \left( I_T \otimes (BMI) \right) \tilde{X}_2 I_2 \]
\[ + o_p(n). \]

And
\[
U_{2n} \text{vec}_D(U_{1n}) = l_1' \tilde{X}_1 (I_T \otimes B') \text{vec}_D(H_1) l_2 + l_1' \tilde{X}_1 (I_T \otimes B') \text{vec}_D(H_2) l_3 \\
+ l_1' \tilde{X}_1 (I_T \otimes B') \mathbf{1}_{nT} l_4 + l_1' \tilde{X}_1 (I_T \otimes B') \text{vec}_D(H_3) l_5 \\
+ l_2 \tilde{X}_2 (I_T \otimes B') \text{vec}_D(H_4) l_5 + l_2 \tilde{X}_2 (I_T \otimes B') \mathbf{1}_{nT} l_4 \\
+ l_2 \tilde{X}_2 (I_T \otimes B') \text{vec}_D(H_3) l_5 + l_2 \tilde{X}_2 (I_T \otimes B') \text{vec}_D(H_4) l_5 \\
+ 2l_5 \left[ (I_T \otimes (BM_n A^{-1})) \tilde{X}_1 \theta_1 \right]' \text{vec}_D(H_1) l_2 \\
+ 2l_5 \left[ (I_T \otimes (BM_n A^{-1})) \tilde{X}_1 \theta_1 \right]' \text{vec}_D(H_2) l_3 \\
+ 2l_5 \left[ (I_T \otimes (BM_n A^{-1})) \tilde{X}_1 \theta_1 \right]' \mathbf{1}_{nT} l_4 \\
+ 2l_5 \left[ (I_T \otimes (BM_n A^{-1})) \tilde{X}_1 \theta_1 \right]' \text{vec}_D(H_3) l_5 \\
+ 2l_5 \left[ (I_T \otimes (BM_n A^{-1})) \tilde{X}_1 \theta_1 \right]' \text{vec}_D(H_4) l_5 \\
+ 2l_5 \left[ (I_T \otimes (BM_n A^{-1})) \tilde{X}_2 \cdot \rho \right]' \text{vec}_D(H_1) l_2 \\
+ 2l_5 \left[ (I_T \otimes (BM_n A^{-1})) \tilde{X}_2 \cdot \rho \right]' \text{vec}_D(H_2) l_3 \\
+ 2l_5 \left[ (I_T \otimes (BM_n A^{-1})) \tilde{X}_2 \cdot \rho \right]' \mathbf{1}_{nT} l_4 \\
+ 2l_5 \left[ (I_T \otimes (BM_n A^{-1})) \tilde{X}_2 \cdot \rho \right]' \text{vec}_D(H_3) l_5 \\
+ 2l_5 \left[ (I_T \otimes (BM_n A^{-1})) \tilde{X}_2 \cdot \rho \right]' \text{vec}_D(H_4) l_5 \\
= l_1' \left( \text{vec}_D(H_1), \text{vec}_D(H_2), \mathbf{1}_{nT}, \text{vec}_D(H_3), \text{vec}_D(H_4) \right)' (I_T \otimes B) E(\tilde{X}_1) l_1 \\
+ l_1' \left( \text{vec}_D(H_1), \text{vec}_D(H_2), \mathbf{1}_{nT}, \text{vec}_D(H_3), \text{vec}_D(H_4) \right)' (I_T \otimes B) E(\tilde{X}_2) l_2 \\
+ 2l_5' \left( \text{vec}_D(H_1), \text{vec}_D(H_2), \mathbf{1}_{nT}, \text{vec}_D(H_3), \text{vec}_D(H_4) \right)' \\
\cdot E \left[ (I_T \otimes (BM_n A^{-1})) \tilde{X}_1 \theta_1 \right] l_5 \\
+ 2l_5' \left( \text{vec}_D(H_1), \text{vec}_D(H_2), \mathbf{1}_{nT}, \text{vec}_D(H_3), \text{vec}_D(H_4) \right)' \\
\cdot E \left[ (I_T \otimes (BM_n A^{-1})) \tilde{X}_2 \cdot \rho \right] l_5 + o_p(n).}

Combine (28)–(33), we have
\[
Var_{Q_n}^* = (\vartheta_4 - 3\sigma^4)||\text{vec}_D(U_{1n})||^2 + \sigma^4 [tr(U_{1n}U_{1n}'n) + tr(U_{1n}^2)] \\
+ \sigma^2 U_{2n}U_{2n}'n + 2\vartheta_3 U_{2n}\text{vec}_D(U_{1n}) \\
= (\vartheta_4 - 3\sigma^4)I^T G_1 I + 2\sigma^4 I^T G_2 I + \lambda_1\sigma^2 E \left[ \tilde{X}_1' (I_T \otimes (B'B)) \tilde{X}_1 \right]I_1 \\
+ 4\lambda_2^2 \sigma^2 E \left\{ \left[ (I_T \otimes (BM_nA^{-1})) \tilde{X}_1 \tilde{\theta}_1 \right]' \left[ (I_T \otimes (BM_nA^{-1})) \tilde{X}_1 \tilde{\theta}_1 \right] \right\} \\
+ 4\lambda_2^2 \sigma^2 E \left\{ \left[ (I_T \otimes (BM_nA^{-1})) \tilde{X}_2 \cdot \rho \right]' \left[ (I_T \otimes (BM_nA^{-1})) \tilde{X}_2 \cdot \rho \right] \right\} \\
+ 2\lambda_2' \sigma^2 E \left[ \tilde{X}_1' (I_T \otimes (B'B)) \tilde{X}_2 \right]I_2 + \lambda_2^2 \sigma^2 E \left[ \tilde{X}_2' (I_T \otimes (B'B)) \tilde{X}_2 \right]I_5 \\
+ 4\lambda_2' \sigma^2 E \left[ \tilde{X}_1' (I_T \otimes B') \left( (I_T \otimes (BM_nA^{-1})) \tilde{X}_1 \tilde{\theta}_1 \right) \right]I_5 \\
+ 4\lambda_2' \sigma^2 E \left[ \tilde{X}_2' (I_T \otimes B') \left( (I_T \otimes (BM_nA^{-1})) \tilde{X}_1 \tilde{\theta}_1 \right) \right]I_5 \\
+ 4\lambda_2' \sigma^2 E \left[ \tilde{X}_2' (I_T \otimes B') \left( (I_T \otimes (BM_nA^{-1})) \tilde{X}_2 \cdot \rho \right) \right]I_5 \\
+ 2\lambda_2' \vartheta_3 \left( \text{vec}_D(H_1), \text{vec}_D(H_2), 1_{nT}, \text{vec}_D(H_3), \text{vec}_D(H_4) \right)' \left( (I_T \otimes B) \text{E}(\tilde{X}_1) \right)I_1 \\
+ 2\lambda_2' \vartheta_3 \left( \text{vec}_D(H_1), \text{vec}_D(H_2), 1_{nT}, \text{vec}_D(H_3), \text{vec}_D(H_4) \right)' \left( (I_T \otimes B) \text{E}(\tilde{X}_2) \right)I_2 + 4\lambda_2' \vartheta_3 \left( \text{vec}_D(H_1), \text{vec}_D(H_2), 1_{nT}, \text{vec}_D(H_3), \text{vec}_D(H_4) \right)' \left( (I_T \otimes (BM_nA^{-1})) \tilde{X}_1 \tilde{\theta}_1 \right)I_5 \\
+ 4\lambda_2' \vartheta_3 \left( \text{vec}_D(H_1), \text{vec}_D(H_2), 1_{nT}, \text{vec}_D(H_3), \text{vec}_D(H_4) \right)' \left[ (I_T \otimes (BM_nA^{-1})) \tilde{X}_2 \cdot \rho \right]I_5 \\
+ 8\lambda_2 \sigma^2 E \left\{ \left[ (I_T \otimes (BM_nA^{-1})) \tilde{X}_1 \tilde{\theta}_1 \right]' \left[ (I_T \otimes (BM_nA^{-1})) \tilde{X}_2 \cdot \rho \right] \right\}I_5 \\
+ \lambda_2 \Sigma_{p+q+4} l + o_p(n) \\
= \lambda_2 \Sigma_{p+q+4} l + o_p(n),
\]

where \(\Sigma_{p+q+4}\) is given in (22). From Condition A4, one can see that \((nT)^{-1}Var^*(Q_n) \geq c_1 > 0\). From Lemma 1, we have

\[
\frac{Q_n - E^*(Q_n)}{\sqrt{Var^*(Q_n)}} \overset{d}{\rightarrow} N(0, 1),
\]

where \(d^*\) stands for convergence in distribution given \(X, Z\). Noting that \((nT)^{-1}Var^*(Q_n) \geq c_1 > 0\) and

\[
Var^*(Q_n) = \sigma_{Q_n}^2 + o_p(n),
\]

one can show that
Combing $E^*(Q_n) = 0$, (35) and (36), we thus have

$$
\frac{Q_n}{\sigma_{Q_n}} \overset{d}{\rightarrow} N(0, 1).
$$

Then (24) holds true.

Next we will prove (25), i.e.,

$$
(nT)^{-1} \sum_{i=1}^{nT} \left( \ell' \omega_i(\psi) \right)^2 = (nT)^{-1} \sigma^2_{Q_n} + o_p(1).
$$

Let

$$
N_{in} = \ell' \omega_i(\psi) = u_{ii} (\epsilon_i^2 - \sigma^2) + 2 \sum_{j=1}^{i-1} u_{ij} \epsilon_i \epsilon_j + b_i \epsilon_i = u_{ii} (\epsilon_i^2 - \sigma^2) + R_i \epsilon_i,
$$

where $R_i = 2 \sum_{j=1}^{i-1} u_{ij} \epsilon_j + b_i$. Let $\mathcal{F}_0 = \{ \phi, \Omega \}, \mathcal{F}_i = \sigma(e_1, ..., e_i), 1 \leq i \leq nT$, then $\{N_{in}, \mathcal{F}_i, 1 \leq i \leq nT\}$ form a martingale difference array given $X, Z$. From (27) and (34), one can see that

$$
\sigma^2_{Q_n} = \sum_{i=1}^{nT} E^* (N_{in}^2) + o_p(n).
$$

It follows that

$$
(nT)^{-1} \sum_{i=1}^{nT} \left( \ell' \omega_i(\psi) \right)^2 = (nT)^{-1} \sigma^2_{Q_n}
$$

$$
= (nT)^{-1} \sum_{i=1}^{nT} (N_{in}^2 - E^*(N_{in}^2)) + o_p(1)
$$

$$
= (nT)^{-1} \sum_{i=1}^{nT} \{N_{iin}^2 - E^*(N_{in}^2)\} + E^*(N_{in}^2|\mathcal{F}_{i-1})
$$

$$
+ E^*(N_{in}^2|\mathcal{F}_{i-1}) - E^*(N_{in}^2) + o_p(1)
$$

$$
= (nT)^{-1} S_{n1} + (nT)^{-1} S_{n2} + o_p(1),
$$

where

$$
S_{n1} = \sum_{i=1}^{nT} \{N_{in}^2 - E^*(N_{iin}^2)\},
$$

$$
S_{n2} = \sum_{i=1}^{nT} \{E^*(N_{in}^2|\mathcal{F}_{i-1}) - E^*(N_{in}^2)\}.
$$

Next we will show that (i)

$$
(nT)^{-1} S_{n1} = o_p(1),
$$

(40)
and (ii)
\[(nT)^{-1} S_{n2} = a_p(1). \quad (41)\]

To show (i) and (ii), it is sufficient to show that \((nT)^{-2} E(S_{n1}) \to 0\) and \((nT)^{-2} E(S_{n2}) \to 0\) respectively. Obviously,
\[N_{m2}^2 = u_{ii}^2 (e_i^2 - \sigma^2)^2 + R_i^2 (e_i^2 - \sigma^2)^2 + 2u_{ii}R_i (e_i^2 - \sigma^2)e_i.\]

Thus
\[E(N_{m2}^2 | F_{i-1}) = u_{ii}^2 E(e_i^2 - \sigma^2)^2 + R_i^2 \sigma^2 + 2u_{ii}R_i \theta_3.\]

It follows that
\[(nT)^{-2} E(S_{n1}^2) = (nT)^{-2} \sum_{i=1}^{nT} E\{N_{m2}^2 - E(N_{m2}^2 | F_{i-1})\}^2\]
\[= (nT)^{-2} \sum_{i=1}^{nT} E\left[u_{ii}^2 \left((e_i^2 - \sigma^2)^2 - E(e_i^2 - \sigma^2)^2\right)\right.\]
\[\quad + R_i^2 (e_i^2 - \sigma^2)^2 + 2u_{ii}R_i (e_i^2 - \sigma^2)e_i - \theta_3]^2\]
\[\leq C(nT)^{-2} \sum_{i=1}^{nT} E\left[u_{ii}^4 \left((e_i^2 - \sigma^2)^2 - E(e_i^2 - \sigma^2)^2\right)\right.\]
\[\quad + R_i^4 (e_i^2 - \sigma^2)^2\]
\[\quad + C(nT)^{-2} \sum_{i=1}^{nT} E\left[u_{ii}^2 R_i^2 (e_i^2 - \sigma^2 e_i - \theta_3)^2\right] \quad (42)\]

By Conditions A1-A3, we have the following
\[(nT)^{-2} \sum_{i=1}^{nT} E\left[u_{ii}^4 \left((e_i^2 - \sigma^2)^2 - E(e_i^2 - \sigma^2)^2\right)\right] \leq C(nT)^{-2} \sum_{i=1}^{nT} u_{ii}^4\]
\[\leq C(nT)^{-2} \sum_{i=1}^{nT} |l_2 h_{ii,1} + l_3 h_{ii,2} + l_4 + l_5 h_{ii,3} + l_6 h_{ii,4}|^4\]
\[\leq C(nT)^{-1} \to 0, \quad (43)\]

and
\[(nT)^{-2} \sum_{i=1}^{nT} E\left[R_i^4 (e_i^2 - \sigma^2)^2\right] = (nT)^{-2} \sum_{i=1}^{nT} E\left[R_i^4 (e_i^2 - \sigma^2)^2\right]\]
\[\leq C(nT)^{-2} \sum_{i=1}^{nT} E \left(\sum_{j=1}^{i-1} u_{ij}e_j + b_i\right)^4\]
\[\leq C(nT)^{-2} \sum_{i=1}^{nT} E \left(\sum_{j=1}^{i-1} u_{ij}e_j\right)^4 + C(nT)^{-2} \sum_{i=1}^{nT} Eb_i^4\]
From (42) to (45), we have

\[ \leq C(nT)^{-2} \sum_{i=1}^{nT} \sum_{j=1}^{i-1} u_{ij}^2 \vartheta_4 + C(nT)^{-2} \sum_{i=1}^{nT} \left( \sum_{j=1}^{i-1} u_{ij}^2 \sigma^2 \right)^2 \]

\[ + C(nT)^{-2} \sum_{i=1}^{nT} \left( \sum_{j=1}^{i-1} \left| l_i a_{i,1} + l_2 a_{i,2} + l_3 a_{i,3} + l_5 a_{i,4} \right|^4 \right) \]

\[ \leq C(nT)^{-2} \sum_{i=1}^{nT} \sum_{j=1}^{i-1} \left| l_2 h_{ij,1} + l_3 h_{ij,2} + l_5 h_{ij,3} + l_5 h_{ij,4} \right|^4 \]

\[ + C(nT)^{-2} \sum_{i=1}^{nT} \left( \sum_{j=1}^{i-1} \left| l_2 h_{ij,1} + l_3 h_{ij,2} + l_5 h_{ij,3} + l_5 h_{ij,4} \right|^2 \right)^2 \]

\[ + C(nT)^{-2} \sum_{i=1}^{nT} E \left( l_i a_{i,1} + l_2 a_{i,2} + l_3 a_{i,3} + l_5 a_{i,4} \right)^4 \leq Cn^{-1} \to 0. \]

Similarly, we can prove that

\[ (nT)^{-2} \sum_{i=1}^{nT} E \left\{ u_{ii}^2 R_i^2 \left( e_i^2 - \sigma^2 e_i - \vartheta_3 \right)^2 \right\} \to 0. \] (45)

From (42) to (45), we have \((nT)^{-2} E(S_{n1}^2) \to 0\). Furthermore,

\[ E^* \left( N_{in}^2 \right) = E^* \left\{} E^* \left( N_{in}^2 | F_{i-1} \right) \right\} \]

\[ = u_{ii}^2 E \left( e_i^2 - \sigma^2 \right)^2 + \sigma^2 E^* \left( R_i^2 \right) + 2 u_{ii} \vartheta_3 E^* \left( R_i \right) \]

\[ = u_{ii}^2 E \left( e_i^2 - \sigma^2 \right)^2 + \sigma^2 \left( 4 \sum_{j=1}^{i-1} u_{ij}^2 \sigma^2 + b_i^2 \right) + 2 u_{ii} \vartheta_3 b_i. \]

Thus,

\[ (nT)^{-2} E \left( S_{n2}^2 \right) = (nT)^{-2} E \left[ \sum_{i=1}^{nT} \left\{ E^* \left( N_{in}^2 | F_{i-1} \right) - E^* \left( N_{in}^2 \right) \right\} \right]^2 \]

\[ = (nT)^{-2} E \left[ \sum_{i=1}^{nT} \left\{ R_i^2 \sigma^2 - \sigma^2 \left( 4 \sum_{j=1}^{i-1} u_{ij}^2 \sigma^2 + b_i^2 \right) + 2 u_{ii} \vartheta_3 (R_i - b_i) \right\} \right]^2 \]

\[ = (nT)^{-2} \sum_{i=1}^{nT} E \left[ \sigma^2 \left\{ 4 \sum_{j=1}^{i-1} u_{ij}^2 \sigma^2 \right\} - \left( 2 \sum_{j=1}^{i-1} u_{ij} e_j \right)^2 - 4 \sum_{j=1}^{i-1} u_{ij}^2 \sigma^2 \right]^2 \]

\[ + 4 \left( \sum_{j=1}^{i-1} u_{ij} e_j \right) b_i \sigma^2 + 2 u_{ii} \vartheta_3 \left( 2 \sum_{j=1}^{i-1} u_{ij} e_j \right)^2 \]
\[
\leq C(nT)^{-2} \sum_{i=1}^{nT} E \left\{ \sigma^2 \left( \sum_{j=1}^{i-1} u_{ij} e_j \right)^2 - \sum_{j=1}^{i-1} u_{ij}^2 \sigma^2 \right\}^2 \\
+ C(nT)^{-2} \sum_{i=1}^{nT} E \left\{ \left( \sum_{j=1}^{i-1} u_{ij} e_j \right) b_i \sigma^2 \right\}^2 \\
+ C(nT)^{-2} \sum_{i=1}^{nT} E \left\{ 2u_{ii} \vartheta_3 \left( \sum_{j=1}^{i-1} u_{ij} e_j \right) \right\}^2 .
\]  

(46)

Note that

\[
(nT)^{-2} \sum_{i=1}^{nT} E \left[ \sigma^2 \left( \sum_{j=1}^{i-1} u_{ij} e_j \right)^2 - \sum_{j=1}^{i-1} u_{ij}^2 \sigma^2 \right] \leq (nT)^{-2} \sigma^4 \sum_{i=1}^{nT} E \left( \sum_{j=1}^{i-1} u_{ij} e_j \right)^4 
\]  

(47)

\[
\leq C(nT)^{-2} \sum_{i=1}^{nT} \sum_{j=1}^{i-1} u_{ij}^4 \vartheta_4 + C(nT)^{-2} \sum_{i=1}^{nT} \left( \sum_{j=1}^{i-1} u_{ij}^2 \sigma^2 \right)^2 \\
\leq Cn^{-1} \rightarrow 0,
\]

(48)

\[
(nT)^{-2} \sum_{i=1}^{nT} E \left\{ \left( \sum_{j=1}^{i-1} u_{ij} e_j \right) b_i \sigma^2 \right\}^2 = (nT)^{-2} \sigma^6 \sum_{i=1}^{nT} E(b_i^2) \sum_{j=1}^{i-1} u_{ij}^2 \\
\leq C(nT)^{-2} \rightarrow 0,
\]

(49)

and

\[
\leq 4\vartheta_3^2 \sigma^2 (nT)^{-2} \sum_{i=1}^{nT} u_i^2 \sum_{j=1}^{i-1} u_{ij}^2 \leq C(nT)^{-1} \rightarrow 0,
\]

where we have used Conditions A2 and A3. From (46) to (49), we have \((nT)^{-2} ES_{n2}^2 \rightarrow 0\). The proof of (25) is thus complete.
Finally, we will prove (26). Note that

\[
\sum_{i=1}^{nT} E\|\omega_i(\psi)\|^3 \leq \sum_{i=1}^{nT} E\|a_{i,1}e_i\|^3 \\
+ \sum_{i=1}^{nT} E|a_{i,2}e_i + h_{ii,1}(e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,1}e_j|^3 \\
+ \sum_{i=1}^{nT} E|h_{ii,2}(e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,2}e_j|^3 + \sum_{i=1}^{nT} E|e_i^2 - \sigma^2|^3 \\
+ \sum_{i=1}^{nT} E|a_{i,3}e_i + a_{i,4}e_i + h_{ii,3}(e_i^2 - \sigma^2) \\
+ 2e_i \sum_{j=1}^{i-1} h_{ij,3}e_j + h_{ii,4}(e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,4}e_j|^3.
\]

By Conditions A2 and A3, we have

\[
\sum_{i=1}^{nT} E\|a_{i,1}e_i\|^3 \leq CnT \left( \max_{1 \leq i \leq nT} E\|a_{i,1}\| \right)^3 E|e_i|^3 = O(nT), \tag{51}
\]

\[
\sum_{i=1}^{nT} E|a_{i,2}e_i + h_{ii,1}(e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,1}e_j|^3 \\
\leq C \sum_{i=1}^{nT} E|a_{i,2}e_i|^3 + C \sum_{i=1}^{nT} E|h_{ii,1}(e_i^2 - \sigma^2)|^3 + C \sum_{i=1}^{nT} E\left( 2e_i \sum_{j=1}^{i-1} h_{ij,1}e_j \right)^3 \\
\leq C \sum_{i=1}^{nT} E|a_{i,2}|^3 E|e_i|^3 + C \sum_{i=1}^{nT} E|h_{ii,1}(e_i^2 - \sigma^2)|^3 \\
+ C \sum_{i=1}^{nT} E|e_i|^3 \sum_{j=1}^{i-1} E|h_{ij,1}e_j|^3 + C \sum_{i=1}^{nT} E|e_i|^3 \left\{ \sum_{j=1}^{i-1} E(h_{ij,1}e_j)^2 \right\}^{3/2} \\
= O(nT),
\]

and

\[
\sum_{i=1}^{nT} E|e_i^2 - \sigma^2|^3 = O(nT). \tag{53}
\]

Similarly,

\[
\sum_{i=1}^{nT} E\left| h_{ii,2}(e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,2}e_j \right|^3 = O(nT), \tag{54}
\]

\[
\sum_{i=1}^{nT} \left| a_{i,3}e_i + a_{i,4}e_i + h_{ii,3}(e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,3}e_j \\
+ h_{ii,4}(e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,4}e_j \right|^3 = O(nT). \tag{55}
\]
From (50) to (55), we have
\[ \sum_{i=1}^{nT} E|\omega_i(\psi)|^3 = O(nT). \] (56)

Further, using (56) and Markov inequality, we obtain
\[ \sum_{i=1}^{nT} |\omega_i(\psi)|^3 = O_p(nT^2). \]

Thus, (26) is proved.

**Proof of Theorem 1.** Using Lemma 3 and following the proof of Theorem 1 in Qin (2021), we can easily show that Theorem 1 holds true.

**Acknowledgments**

The authors are thankful to the referees for constructive suggestions.

**Funding**

This work was partially supported by the National Natural Science Foundation of China (12061017, 12161009) and the Innovation Project of Guangxi Graduate Education (YCSW2021073).

**References**

Anselin, L. 1988. *Spatial econometrics: Methods and models.* The Netherlands: Kluwer Academic Publishers.

Anselin, L. 2001. Spatial econometrics. In: *A companion to theoretical econometrics,* ed. B. H. Baltagi, 310–30. Massachusetts: Blackwell Publishers Ltd.

Anselin, L., J. Le Gallo, and H. Jayet. 2008. Spatial panel econometrics. In: *The econometrics of panel data: Fundamentals and recent developments in theory and practice,* eds. L. Mátyás and P. Sevstrø, 625–60. Berlin Heidelberg: Springer-Verlag.

Baltagi, B. H., P. Egger, and M. Pfaffermayr. 2013. A generalized spatial panel data model with random effects. *Econometric Reviews* 32 (5-6):650–85. doi:10.1080/07474938.2012.742342.

Baltagi, B. H., S. H. Song, and W. Koh. 2003. Testing panel data regression models with spatial error correlation. *Journal of Econometrics* 117 (1):123–50. doi:10.1016/S0304-4076(03)00120-9.

Bandyopadhyay, S., S. N. Lahiri, and D. J. Nordman. 2015. A frequency domain empirical likelihood method for irregularly spaced spatial data. *Annals of Statistics* 43:519–45.

Chen, S. X., and I. V. Keilegom. 2009. A review on empirical likelihood methods for regression (with discussions). *Test* 18 (3):415–47. doi:10.1007/s11749-009-0159-5.

Elhorst, J. P. 2003. Specification and estimation of spatial panel data models. *International Regional Science Review* 26 (3):244–68. doi:10.1177/0160017603253791.

Elhorst, J. P. 2005. Unconditional maximum likelihood estimation of linear and log-linear dynamic models for spatial panels. *Geographical Analysis* 37 (1):85–106. doi:10.1111/j.1538-4632.2005.00577.x.

Elhorst, J. P. 2010. Dynamic panels with endogenous interaction effects when t is small. *Regional Science and Urban Economics* 40 (5):272–82. doi:10.1016/j.regsciurbeco.2010.03.003.

Elhorst, J. P. 2012. Dynamic spatial panels: Models, methods, and inferences. *Journal of Geographical Systems* 14 (1):5–28. doi:10.1007/s10109-011-0158-4.
Fingleton, B. 2008. 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 (1): 27–44. doi:10.1080/17421770701774922.

Hsiao, C., M. H. Pesaran, and A. K. Tahmiscioglu. 2002. Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods. *Journal of Econometrics* 109 (1):107–50. doi:10.1016/S0304-4076(01)00143-9.

Jin, F., and L. F. Lee. 2019. GEL estimation and tests of spatial autoregressive models. *Journal of Econometrics* 208 (2):585–612. doi:10.1016/j.jeconom.2018.07.007.

Kapoor, M., H. H. Kelejian, and I. R. Prucha. 2007. Panel data models with spatially correlated error components. *Journal of Econometrics* 140 (1):97–130. doi:10.1016/j.jeconom.2006.09.004.

Kelejian, H. H., and I. R. Prucha. 2001. On the asymptotic distribution of the Moran I test statistic with applications. *Journal of Econometrics* 104 (2):219–57. doi:10.1016/S0304-4076(01)00064-1.

Kitamura, Y. 1997. Empirical likelihood methods with weakly dependent processes. *Annals of Statistics* 25:2084–102.

Lee, L. F. 2004. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* 72 (6):1899–925. doi:10.1111/j.1468-0262.2004.00558.x.

Lee, L. F., and J. Yu. 2010a. Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics* 154 (2):165–85. doi:10.1016/j.jeconom.2009.08.001.

Lee, L. F., and J. Yu. 2010b. Some recent developments in spatial panel data models. *Regional Science and Urban Economics* 40 (5):255–71. doi:10.1016/j.regsciurbeco.2009.09.002.

Lee, L. F., and J. Yu. 2010c. A spatial dynamic panel data model with both time and individual fixed effects. *Econometric Theory* 26 (2):564–97. doi:10.1017/S0266466909000997.

Lee, L. F., and J. Yu. 2016. Identification of Spatial Durbin Panel Models. *Journal of Applied Econometrics* 31 (1):133–62. doi:10.1002/jae.2450.

Li, Y., Y. Li, and Y. Qin. 2020. Empirical likelihood for panel data models with spatial errors. *Communications in Statistics - Theory and Methods* 1–20. doi:10.1080/03610926.2020.1780449.

Mutl, J. 2006. Dynamic panel data models with spatially correlated disturbances. Ph. D. diss. University of Maryland, College Park.

Nordman, D. J. 2008. An empirical likelihood method for spatial regression. *Metrika* 68 (3):351–63. doi:10.1007/s00184-007-0167-y.

Nordman, D. J., and P. C. Caragea. 2008. Point and interval estimation of variogram models using spatial empirical likelihood. *Journal of the American Statistical Association* 103 (481):350–61. doi:10.1198/016214507000001391.

Owen, A. B. 1988. Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* 75 (2):237–49. doi:10.1093/biomet/75.2.237.

Owen, A. B. 1990. Empirical likelihood ratio confidence regions. *The Annals of Statistics* 18 (1):90–120. doi:10.1214/aos/1176347494.

Owen, A. B. 1991. Empirical likelihood for linear models. *The Annals of Statistics* 19:1725–47.

Owen, A. B. 2001. *Empirical likelihood*. London: Chapman & Hall.

Parent, O., and J. P. LeSage. 2011. A space-time filter for panel data models containing random effects. *Computational Statistics & Data Analysis* 55 (1):475–90. doi:10.1016/j.csda.2010.05.016.

Qin, J., and J. Lawless. 1994. Empirical likelihood and general estimating equations. *The Annals of Statistics* 22 (1):300–25. doi:10.1214/aos/1176325370.

Qin, Y. 2021. Empirical likelihood for spatial autoregressive models with spatial autoregressive disturbances. *Sankhya A* 83 (1):1–25. doi:10.1007/s13171-019-00166-3.

Qu, X., L. F. Lee, and J. Yu. 2017. QML estimation of spatial dynamic panel data models with endogenous time varying spatial weights matrices. *Journal of Econometrics* 197 (2):173–201. doi:10.1016/j.jeconom.2016.11.004.

Su, L., and Z. Yang. 2015. QML estimation of dynamic panel data models with spatial errors. *Journal of Econometrics* 185 (1):230–58. doi:10.1016/j.jeconom.2014.11.002.

Yang, Z., C. Li, and Y. K. Tse. 2006. Functional form and spatial dependence in dynamic panels. *Economics Letters* 91 (1):138–45. doi:10.1016/j.econlet.2005.11.011.