Empirical likelihood for spatial dynamic panel data models with spatial errors and endogenous initial observations

Jie Tang Yongsong Qin[1]

Guangxi Key Lab of Multi-source Information Mining & Security and Department of Statistics, Guangxi Normal University, Guilin, Guangxi 541004, China

Abstract In this article, we study the construction of confidence regions for the parameters in spatial dynamic panel data (SDPD) models with spatial errors and endogenous initial observations by using the empirical likelihood (EL) method. It is shown that the EL ratio statistics are asymptotically χ^2-type distributed, which are used to obtain EL based confidence regions for the parameters in SDPD models. A simulation study is conducted to compare the performances of the EL and the quasi maximum likelihood (QML) methods.

Keywords: SDPD model; empirical likelihood; confidence region

AMS 2020 subject classification: Primary 62G05 Secondary 62E20

[1]Corresponding author.

Email address: ysqin@gxnu.edu.cn; 1243711126@qq.com

1. Introduction

Since Cliff and Ord (1973), the spatial data has been applied broadly to different fields, such as economics, sociology, biology and so on. The evolution of model with spatial element is divided into three generations: non-panel spatial data (SD) models, static spatial panel data (SPD) models and spatial dynamic panel data (SDPD) models. Firstly, the SD model deals with spatial interaction and spatial heterogeneity primarily in cross-section data. For the research developments of SD models, refer to Cliff and Ord (1973), Cressien (1993), Kelejian and Prucha (1998, 1999, 2006), Kelejian et al. (2004) and Liu et al. (2010), etc. Secondly, the SPD model studied in Anselin (1988), where data are observed both across sectional units and over time, is becoming increasingly attractive in empirical economic research. See, Elhorst (2003), Baltagi et al. (2003), Anselin et al. (2008), Parent and LeSage (2011), Lee and Yu (2016) and Baltagi et al. (2013) for an overview on the SPD models. Thirdly, by adding a dynamic element into SPD model, Anselin (2001) has proposed SDPD models, increasing greatly flexibility of the SPD model. See, Yang et al. (2006), Lee and Yu (2010), Elhorst (2012), Qu et al. (2017) and Li and Qin (2022) and so on.

SDPD models are distinguished into two categories due to the ways of initial observation are given: exogenously and endogenously. In this paper, we consider the latter type of SDPD model. In particular, the dynamic panel data model with spatial error and endogenous initial observations. Let y_{it} denote the observation at $i(=1, 2, \cdots, N)$ spatial unit and $t(=1, \cdots, T)$ time period, and then exogenous means that the initial observations $y_{i0}, i = 1, 2, \cdots, N$ are set to be a fixed constant under the assumption that they contain no information about the model parameters, while endogenous means initial observations are not exogenously given as they contain useful information of m periods before the 0th period and hence should be utilized in the model estimation. Su and Yang (2007) has found it is important to treat the initial period observations. A Monte Carlo study in Parent and LeSage (2012) shows that when the initial cross-sectional observations are endogenous, but incorrectly treated as exogenous, estimates and inferences can be biased, especially in the case of small T. More references on treatments of initial period observations can be seen in Yu et al. (2008), Parent and LeSage (2012), Su and Yang (2015), among others.

There are two major estimation approaches for the corresponding parameters in spatial models. One is the maximum likelihood (ML) method (e.g., Anselin 1988). The other is the generalized method of moments (GMMs) by Kelejian and Prucha (1999). The asymptotic properties of the ML estimator and the GMM estimator for the spatial models are researched by Anselin (1988) and Kelejian and Prucha (1999), respectively. These methods may be readily applied to the SDPD model. However, it may not be easy to use these normal approximation results to construct confidence region for the parameters in the SDPD model as the asymptotic covariance in the asymptotic distribution is unknown. More importantly, the accuracy of the normal approximation-based confidence region of the parameters in the model may be affected by estimating the asymptotic covariance.

Therefore, we propose to use the empirical likelihood (EL) method introduced by Owen (1988, 1990) to construct confidence regions for the parameters in the SDPD models. The shape and orientation of the EL confidence region are determined by data, and confidence region is obtained without covariance estimation. These features of the EL confidence region are the major motivations for our current proposal. At first, EL method was used for non-spatial models, for example,Owen (1991) used the EL method to construct confidence regions for a linear model with independence errors, Kolaczyk (1994) popularized EL method for generalized linear models, and Qin (1999) constructed EL ratio confidence regions in a partly linear model, etc. A comprehensive review on EL for regressions can be found in Chen and Keilegom (2009). More references on EL methods can be found in Chen and Qin (1993), Qin and Lawless (1994), Zhong and Rao (2000), Owen (2001) and Wu (2004), among others. Until recently, Jin and Lee (2019) and Qin (2021) independently found and successfully constructed the empirical likelihood ratio statistics for the parameters in SD models, which greatly promoted the development of empirical likelihood in spatial models. More references on EL methods in spatial models can be found in Nordman (2008), Bandyopadhyay et al. (2015), Li et al. (2020), Rong et al. (2021), Qin and Lei (2021), Li and Qin (2022), among others. As the research of the EL method on spatial models is still in its infancy, there is no research work on the EL method for SDPD models with endogenous initial observations.

It is worthwhile to mention the work related to this article. Li and Qin (2022) uses the EL method for the SDPD model in the exogenous case (or conditional specification). Rong et al. (2021) extends the model in Li and Qin (2022), still focusing on the exogenous case. In other words, SDPD models in Li and Qin (2022) belong to models in Rong et al. (2021), and these two articles only study the exogenous case for SDPD models. However, Nerlove (1999, p. 139) points out that conditioning on those initial values is an undesirable measure, especially when the time dimension of the panel is short. The initial observations may convey useful information because the within-sample observations and pre-sample values are generated by the same process, which makes it necessary to do some research on the endogenous (or unconditional) cases. Unfortunately, some time-varying exogenous explanatory variables or others in the pre-periods are unobservable resulting that the unconditional likelihood function does not exist and further the estimating equations for the EL method can not be derived. Hsiao (2003) suggests an alternative estimation procedure for SDPD models which has been extensively used in many endogenous cases. Su and Yang (2015) further develops the method in Hsiao (2003).

In this article, we focus on the endogenous situation and apply the EL method to construct confidence regions for SDPD models with spatial errors. We conduct simulations to compare the performances of LR based and EL based confidence regions. Our results show that, on the one hand, the EL can be computationally faster than the LR method to implement in practice; on the other hand, the confidence regions based on EL method are closer the nominal level than LR method, whether the error term is normally distributed or not.

The article is organized as follows. Section 2 gives the main results. Section 3 presents results from a simulation study. Section 4 concludes the article. Some lemmas and proofs are collected in Appendix A. Appendix B indicates where to obtain the R codes in the simulation study.

We conclude this section by introducing some notations which will be used later on. For a positive integer κ, let I_κ denote $\kappa \times \kappa$ identity matrix, $\mathbf{1}_\kappa$ a $\kappa \times 1$ vector of ones, and $\mathbf{0}_\kappa$ a $\kappa \times 1$ vector of zeros. Let $A_1 \otimes A_2$ denote the Kronecker product of two matrices A_1 and A_2. Let $|\cdot|$, $\|\cdot\|$, and $\mathrm{tr}(\cdot)$ denote, respectively, the determinant, the Frobenius norm, and the trace of a matrix.

2. Main results

Suppose that there are n individual units and T time periods, the following SDPD model with spatial errors is investigated:

$$y_{it} = \rho y_{i,t-1} + x_{it}'\beta + z_i'\gamma + u_{it}, \quad i = 1, 2, ..., n, \ t = 1, 2, ..., T, \tag{1}$$

where ρ is the scalar parameter with ($|\rho| < 1$) characterizing the dynamic effect, x_{it} is a $p \times 1$ vector of time-varying exogenous variables, and z_i is a $q \times 1$ vector of time-invariant exogenous variables that may include the constant term. The disturbance vector $u_t = (u_{1t}, u_{2t}, ..., u_{nt})'$ is assumed to show a spatially autocorrelated structure, i.e.,

$$u_t = \lambda W_n u_t + v_t, \quad t = 1, 2, ..., T, \tag{2}$$

where λ is the scalar autoregressive parameters with $|\lambda| < 1$, W_n is known $n \times n$ spatial weighting matrix of constants whose diagonal elements are zero, $v_t = (v_{1t}, v_{2t}, ..., v_{nt})'$ is an $n \times 1$ vector of model errors, and $\{v_{it}\}$ are i.i.d. across i and t with mean zero and variance σ^2.

Denoting $y_t = (y_{1t}, y_{2t}, ..., y_{nt})'$, $x_t = (x_{1t}, x_{2t}, ..., x_{nt})'$, and $z = (z_1, z_2, ..., z_n)'$, the model (1) has the following reduced-form representation,

$$y_t = \rho y_{t-1} + x_t\beta + z\gamma + u_t, \quad t = 1, 2, ..., T. \tag{3}$$

To describe clearly the differences between the models in this article and others, we make the following comparison. On the one hand, to compare the models (2.2)-(2.4) in Su and Yang (2015) and the model (2)-(3) in this article, we re-write the models (2.2)-(2.4) in Su and Yang (2015) as follows:

$$y_t = \rho y_{t-1} + x_t\beta + z\gamma + u_t,$$

$$u_t = \mu + \lambda W_n u_t + v_t, t = 1, 2, ..., T,$$

where $\mu = (\mu_1, \cdots, \mu_n)'$ represents the space-specific effects and other notations are the same as the models (2)-(3). Clearly, there is no space-specific effects μ in the model (2). 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, usual EL method cannot be directly used to construct confidence regions for this model, but some adjusted EL method may work, which is left for our future investigation.

On the other hand, to compare the models (1)-(2) in Rong et al. (2021) and the models (2)-(3) in this article, we re-write the models (1)-(2) in Rong et al. (2021) as follows:

$$y_t = \rho y_{t-1} + \kappa M_n y_t + x_t\beta + z\gamma + u_t$$
$$u_t = \lambda W_n u_t + v_t, t = 1, 2, ..., T,$$

where κ is the spatial autoregressive coefficients with the absolute values being less than 1, M_n is $n \times n$ spatial weighting matrices of constants and other notations are the same as the models (2)-(3). Clearly, the models in Rong et al. (2021) contain an additional term of spatial autoregressive $\kappa M_n y_t$. Although there is still no research for the models in Rong et al. (2021) in endogenous case, the method in this article provides a way to solve this problem.

Now, we continue with the model (2)-(3). With $t = 1, 2, ..., T$, model (2)-(3) can be written into a matrix form as follows:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{pmatrix} = \rho \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix} + \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{pmatrix} \beta + \begin{pmatrix} z \\ z \\ \vdots \\ z \end{pmatrix} \gamma + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{pmatrix}$$

with

$$\begin{pmatrix} B_n & 0 & \cdots & 0 \\ 0 & B_n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_n \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_T \end{pmatrix}$$

or

$$Y = \rho Y_{-1} + X\beta + Z\gamma + u, \tag{4}$$

with

$$(I_T \otimes B_n)u = v, \tag{5}$$

where $Y = (y_1', y_2', ..., y_T')'$, $Y_{-1} = (y_0', y_1', ..., y_{T-1}')'$, $X = (x_1', x_2', ..., x_T')'$, $Z = \mathbf{1}_T \otimes z$, $u = (u_1', u_2', ..., u_T')'$, $v = (v_1', v_2', ..., v_T')'$, and $B_n = B_n(\lambda) = I_n - \lambda W_n$.

The following derivations of the major part in Sections 2.1-2.3 can be found in Su and Yang (2015), which are stated here mainly for the consideration of completeness of notations.

2.1 Endogeneity of $\mathbf{y_0}$

We study the EL method for the above model (2)-(3) when y_0 is endogenous. To explain endogenous, we adopt a similar framework as Hsiao et al.(2002): (i) data collection starts from the 0th period; the process starts from the $-m$th period, i.e., m period before the start of data collection, $m = 0, 1, ...$, and then evolves according to the model specified by model (3); (ii) the starting position of the process y_{-m} is treated as exogenous; hence the exogenous variables (x_t, z) and the errors u_t start to have impact on the response from period $-m+1$ onwards; (iii) all exogenous quantities (y_{-m}, x_t, z) are considered as random and inferences proceed by conditioning on them, and (iv) variances of elements of y_{-m} are constant. Thus, when $m = 0, y_0 = y_{-m}$ is exogenous, when $m \geq 1$, y_0 becomes endogenous, and when $m = \infty$, the process has reached stationarity.

The initial observation y_0 is taken as endogenous intuitively that it is generated from the process specified by the model (3). The data collection starts m periods before the 0th period, and then y_0 contains significant information about the model parameters and hence should be considered in the model estimation. In this case, x_0 is needed, and the estimation makes use of $T + 1$ periods of data. By successive backward substitutions using (3), we have

$$y_0 = \rho^m y_{-m} + \sum_{j=0}^{m-1} \rho^j x_{-j}\beta + z\gamma \frac{1-\rho^m}{1-\rho} + \sum_{j=0}^{m-1} \rho^j B_n^{-1} v_{-j}. \tag{6}$$

Letting η_0 and ζ_0 be, respectively, the exogenous and endogenous components of y_0, we have

$$\eta_0 = \rho^m y_{-m} + \sum_{j=0}^{m-1} \rho^j x_{-j}\beta + z\gamma \frac{1-\rho^m}{1-\rho} = \eta_m + x_0\beta + z_m(\rho)\gamma, \tag{7}$$

where $\eta_m = \rho^m y_{-m} + \sum_{j=1}^{m-1} \rho^j x_{-j}\beta$, $z_m(\rho) = za_m$, and $a_m = a_m(\rho) = \frac{1-\rho^m}{1-\rho}$; and

$$\zeta_0 = \sum_{j=0}^{m-1} \rho^j B_n^{-1} v_{-j}, \tag{8}$$

where $E(\zeta_0) = 0$ and $\mathrm{Var}(\zeta_0) = \sigma^2 b_m (B_n' B_n)^{-1}$ and $b_m = b_m(\rho) = \frac{1-\rho^{2m}}{1-\rho^2}$.

Since both $\{x_{-j}, j = 1, 2, ..., m-1\}$ for $m \geq 2$ and y_{-m} for $m \geq 1$ in η_m are unobserved, which renders (7) cannot be used as a model for η_0. Some approximations are necessary. In this paper, we follow Bhargava and Sargan (1983) (see also Hsiao, 2003, p.76) and propose a model for the initial observations based on the following fundamental assumptions as Su and Yang (2015). Let $\mathbf{x} \equiv (x_0, x_1, ..., x_T)$.

Assumption A0: (i) Conditional on the observables $\mathbf{x}$ and z, the optimal predictors for $x_{-j}, j \geq 1$, are $\mathbf{x}$ and the optimal predictors for $E(y_{-m}), m \geq 1$, are $\mathbf{x}$ and z; and (ii) The error resulted from predicting η_m using $\mathbf{x}$ and z is ζ such that $\zeta \sim (0, \sigma_\zeta^2 I_n)$ and is independent of u, $\mathbf{x}$ and z.

These assumptions lead immediately to the following model for η_m:

$$\eta_m = \mathbf{1}_n \pi_1 + \mathbf{x}\pi_2 + z\pi_3 + \zeta \equiv \widetilde{\mathbf{x}}\pi + \zeta, \tag{9}$$

where $\widetilde{\mathbf{x}} = (\mathbf{1}_n, \mathbf{x}, z)$ and $\pi = (\pi_1, \pi_2', \pi_3')'$. Hence we have the following model for y_0 based on (6) – (9):

$$y_0 = \widetilde{\eta}_0 + u_0, \quad u_0 = \zeta + \zeta_0 \tag{10}$$

where $\widetilde{\eta}_0 = \widetilde{\mathbf{x}}\pi + x_0\beta + z_m(\rho)\gamma$. For the cases where ρ^m is not negligible, one can easily show that, under strict exogeneity of $\mathbf{x}$ and z, $E(u_0) = 0$, $E(u_0 u_0') = \sigma_\zeta^2 I_n + \sigma^2 b_m (B_n' B_n)^{-1}$, $E(u_0 u') = 0$ and $E(uu') = \sigma^2 I_T \otimes (B_n' B_n)^{-1}$.

2.2 Log-likelihood function

We continue with the model (4) – (5), and (10). The model of endogenous y_0 can be written into a matrix form as follows:

$$\begin{pmatrix} y_0 \\ Y - \rho Y_{-1} \end{pmatrix} = \begin{pmatrix} x_0\beta & + & z_m(\rho)\gamma & + & \widetilde{\mathbf{x}}\pi \\ X\beta & + & Z\gamma & + & \mathbf{0}_{nT\times k} \end{pmatrix} + \begin{pmatrix} u_0 \\ u \end{pmatrix},$$

with

$$\begin{pmatrix} u_0 \\ u \end{pmatrix} = \begin{pmatrix} \zeta + \zeta_0 \\ (I_T \otimes B_n^{-1})v \end{pmatrix} = \begin{pmatrix} D_1 \widetilde{v} \\ D_2 v \end{pmatrix} = \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix} \begin{pmatrix} \widetilde{v} \\ v \end{pmatrix} \widehat{=} D \begin{pmatrix} \widetilde{v} \\ v \end{pmatrix},$$

where $\widetilde{v} = (\zeta', v'_{-(m-1)}, v'_{-(m-2)}, ..., v_0')'$, $D_1 = (I_n, \rho^{m-1}B_n^{-1}, \rho^{m-2}B_n^{-1}, ..., \rho^0 B_n^{-1})$ and $D_2 = I_T \otimes B_n^{-1}$, or

$$Y^* = X^*\theta + u^*, \tag{11}$$

with

$$u^* = Dv^*, \tag{12}$$

where $Y^* = Y^*(\rho) = \begin{pmatrix} y_0 \\ Y - \rho Y_{-1} \end{pmatrix}$, $\theta = (\beta', \gamma', \pi')'$, $u^* = (u_0', u')'$, $v^* = (\widetilde{v}', v')'$ and $X^* = X^*(\rho) = \begin{pmatrix} x_0 & z_m(\rho) & \widetilde{\mathbf{x}} \\ X & Z & \mathbf{0}_{nT\times k} \end{pmatrix}$, $k = (T+1)p + q + 1$. Let $\epsilon^* = P^{-1}v^*$, $\epsilon^* = (\epsilon_1', \epsilon_2', ..., \epsilon'_{(m+T+1)})'$, model (12) also can be written into a new form as follow:

$$u^* = DP\epsilon^*, \tag{13}$$

where $P = \begin{pmatrix} \sqrt{\phi_\zeta} I_n & \mathbf{0}_{n\times n(m+T)} \\ \mathbf{0}_{n(m+T)\times n} & I_{n(m+T)} \end{pmatrix}$ and $\epsilon^* = (\epsilon_1', \epsilon_2', ..., \epsilon_{m+T+1}')'$, ϵ_i is $n \times 1$ vector, $i = 1, 2, \cdots, m+T+1$, and $E(\epsilon^*) = 0$ and $\mathrm{Var}(\epsilon^*) = \sigma^2 I_{n(m+T+1)}$.

Denote

$$\Omega^* \equiv \Omega^*(\rho, \lambda, \phi_\zeta) = \begin{pmatrix} \phi_\zeta I_n + b_m(B_n'B_n)^{-1} & 0 \\ 0 & I_T \otimes (B_n'B_n)^{-1} \end{pmatrix} \tag{14}$$

where $\phi_\zeta = \sigma_\zeta^2/\sigma^2$. Based on models (11) and (12), we adopt the quasi maximum likelihood method (QMLE) to estimate $\psi = (\theta', \sigma^2, \rho, \lambda, \phi_\zeta)'$. At this moment, suppose that Y^* is normally distributed, which is used to derive the EL statistic only and not employed in our main results. At this stage, we may say that we are using the so called QMLE. Moreover, the EL estimator derived by this method is asymptotically efficient as the parametric likelihood estimator when data are normally distributed. Then the log-likelihood function (apart from a constant term) is

$$L(\psi) = -\frac{n(T+1)}{2}\log\sigma^2 - \frac{1}{2}\log|\Omega^*| - \frac{1}{2\sigma^2}u^*(\rho,\theta)'\Omega^{*-1}u^*(\rho,\theta).$$

where $u^*(\rho, \theta) = \{(y_0 - x_0\beta - z_m(\rho)\gamma - \tilde{\mathbf{x}}\pi)', (Y - \rho Y_{-1} - X\beta - Z\gamma)'\}' \equiv Y^* - X^*\theta$.

2.3 Estimating equations

It can be shown that

$$\partial L(\psi)/\partial\theta = \frac{1}{\sigma^2}X^{*\prime}\Omega^{*-1}u^*,$$

$$\partial L(\psi)/\partial\sigma^2 = \frac{1}{2\sigma^4}\{u^{*\prime}\Omega^{*-1}u^* - \sigma^2 n(T+1)\},$$

$$\partial L(\psi)/\partial\rho = -\frac{1}{\sigma^2}u_\rho^{*\prime}\Omega^{*-1}u^* + \frac{1}{2\sigma^2}u^{*\prime}P_\rho^* u^* - \frac{1}{2}tr(P_\rho^*\Omega^*),$$

$$\partial L(\psi)/\partial\lambda = \frac{1}{2\sigma^2}u^{*\prime}P_\lambda^* u^* - \frac{1}{2}tr(P_\lambda^*\Omega^*),$$

$$\partial L(\psi)/\partial\phi_\zeta = \frac{1}{2\sigma^2}u^{*\prime}P_{\phi_\zeta}^* u^* - \frac{1}{2}tr(P_{\phi_\zeta}^*\Omega^*),$$

where $u_\rho^* = \frac{\partial}{\partial\rho}u^*$, $P_\omega^* = \Omega^{*-1}\Omega_\omega^*\Omega^{*-1}$, and $\Omega_\omega^* = \frac{\partial}{\partial\omega}\Omega^*$, for $\omega = \rho, \lambda$, and ϕ_ζ, given as $u_\rho^* = -\begin{pmatrix} \dot{a}_m z\gamma \\ Y_{-1} \end{pmatrix}$, $\Omega_\rho^* = \begin{pmatrix} \dot{b}_m(B_n'B_n)^{-1} & \mathbf{0}_{n\times nT} \\ \mathbf{0}_{nT\times n} & \mathbf{0}_{nT\times nT} \end{pmatrix}$, $\Omega_\lambda^* = \begin{pmatrix} b_m & \mathbf{0}_T' \\ \mathbf{0}_T & \mathbf{0}_{T\times T} \end{pmatrix} \otimes A$, and $\Omega_{\phi_\zeta}^* = \begin{pmatrix} 1 & \mathbf{0}_T' \\ \mathbf{0}_T & \mathbf{0}_{T\times T} \end{pmatrix} \otimes I_n$, where $A = \frac{\partial}{\partial\lambda}(B_n'B_n)^{-1} = (B_n'B_n)^{-1}(W_n'B_n + B_n'W_n)(B_n'B_n)^{-1}$, $\dot{a}_m = \frac{\partial}{\partial\rho}a_m$ and $\dot{b}_m = \frac{\partial}{\partial\rho}b_m$, and $\dot{a}_m$ or $\dot{b}_m$ expressions can easily be obtained.

Substituting (13) into above derivatives and letting above derivatives be 0, we obtain the following estimating equations:

$$X^{*\prime}\Omega^{*-1}DP\epsilon^{*}=0, \tag{15}$$

$$\epsilon^{*\prime}(DP)'\Omega^{*-1}DP\epsilon^{*}-\sigma^{2}n(T+1)=0, \tag{16}$$

$$-2u_{\rho}^{*\prime}\Omega^{*-1}DP\epsilon^{*}+\epsilon^{*\prime}(DP)'P_{\rho}^{*}DP\epsilon^{*}-\sigma^{2}tr(P_{\rho}^{*}(DP)(DP)')=0, \tag{17}$$

$$\epsilon^{*\prime}(DP)'P_{\lambda}^{*}DP\epsilon^{*}-\sigma^{2}tr(P_{\lambda}^{*}(DP)(DP)')=0, \tag{18}$$

$$\epsilon^{*\prime}(DP)'P_{\phi_\varsigma}^{*}DP\epsilon^{*}-\sigma^{2}tr(P_{\phi_\varsigma}^{*}(DP)(DP)')=0. \tag{19}$$

Considering that u_{ρ}^{*} contains Y_{-1} and Y_{-1} contains X, Z, and u^{*}, we need to separate out u^{*} from u_{ρ}^{*}. For this purpose, denote $Y_0=(\rho^0,\rho^1,...,\rho^{(T-1)})'\otimes\tilde{\eta}_0'$, $F=\Delta_{T-2}\otimes I_n$,

$$\Delta_t=\Delta_t(\rho)=\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0\\ \rho^0 & 0 & 0 & \cdots & 0 & 0\\ \rho^1 & \rho^0 & 0 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ \rho^{t-1} & \rho^{t-2} & \rho^{t-3} & \cdots & 0 & 0\\ \rho^{t} & \rho^{t-1} & \rho^{t-2} & \cdots & \rho^0 & 0 \end{pmatrix}_{(t+2)\times(t+2)},$$

and $F^{*}=\Delta_{T-1}\otimes I_n$. Then u_{ρ}^{*} can be expressed as

$$u_{\rho}^{*}=-\eta_{-1}^{*}-F^{*}u^{*},$$

where $\eta_{-1}^{*}=\begin{pmatrix}\dot{a}_m z\gamma\\ FX\beta+FZ\gamma+Y_0\end{pmatrix}$, then $-2u_{\rho}^{*\prime}\Omega^{*-1}DP\epsilon^{*}$ of (17) can be rewritten into

$$2{\eta_{-1}^{*}}'\Omega^{*-1}DP\epsilon^{*}+2\epsilon^{*\prime}(DP)'F^{*\prime}DP\epsilon^{*}, \tag{20}$$

For convenience, let $T_1=m+T+1$, $e=\epsilon^{*}$, i.e.,

$$e=e_{(nT_1)\times1}=\begin{pmatrix} e_1\\ e_2\\ \vdots\\ e_{nT_1}\end{pmatrix}=\begin{pmatrix} \epsilon_1\\ \epsilon_2\\ \vdots\\ \epsilon_{(m+T+1)}\end{pmatrix}, \tag{21}$$

By (20) and (21), (15)-(19) can be rewritten as

$$X^{*\prime}\Omega^{*-1}DPe=0, \tag{22}$$

$$e'(DP)'\Omega^{*-1}DPe-\sigma^{2}n(T+1)=0, \tag{23}$$

$$2\eta_{-1}^{*}{}'\Omega^{*-1}DPe+2e'(DP)'F^{*}{}'DPe+e'(DP)'P_{\rho}^{*}DPe-\sigma^{2}tr((DP)'P_{\rho}^{*}DP)=0, \quad (24)$$

$$e'(DP)'P_{\lambda}^{*}DPe-\sigma^{2}tr((DP)'P_{\lambda}^{*}DP))=0, \quad (25)$$

$$e'(DP)'P_{\phi_{\varsigma}}^{*}DPe-\sigma^{2}tr((DP)'P_{\phi_{\varsigma}}^{*}DP))=0. \quad (26)$$

2.4 Score function and EL ratio statistic

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 $H_1=(DP)'\Omega^{*-1}DP$, $H_2=(DP)'(F^{*}+F^{*}{}')DP$, $H_3=(DP)'P_{\rho}^{*}DP$, $H_4=(DP)'P_{\lambda}^{*}DP$, and $H_5=(DP)'P_{\phi_{\varsigma}}^{*}DP$. We use $h_{ij,\iota}$, $a_{i,1}$, and $a_{i,2}$ to denote the (i,j) element of the matrix $H_{\iota}(\iota=1,2,3,4,5)$, the i-th column of the matrix $X^{*}{}'\Omega^{*-1}$ DP and the i-th element of the vector $2\eta_{-1}^{*}{}'\Omega^{*-1}DP$, respectively, and adapt the convention that any sum with an upper index of less than one is zero. To deal with the quadratic form in (23)-(26), we follow Kelejian and Prucha (2001) to introduce a martingale difference array.

Define the σ-fields: $\mathcal{F}_0=\{\emptyset,\Omega\},\mathcal{F}_i=\sigma(e_1,e_2,...,e_i),1\le i\le nT_1$. Let

$$\widetilde{\mathcal{M}}_{i\iota}=h_{ii,\iota}(e_i^2-\sigma^2)+2e_i\sum_{j=1}^{i-1}h_{ij,\iota}e_j,\ \ \iota=1,2,3,4,5. \quad (27)$$

Then $\mathcal{F}_{i-1}\subseteq\mathcal{F}_i,\widetilde{\mathcal{M}}_{i\iota}$ is $\mathcal{F}_i-$measurable and $E(\widetilde{\mathcal{M}}_{i\iota}|\mathcal{F}_{i-1})=0$. Thus $\{\widetilde{\mathcal{M}}_{i\iota},\mathcal{F}_i,1\le i\le nT_1\}$ forms a martingale difference array and

$$e'H_{\iota}e-tr(H_{\iota})=\sum_{i=1}^{nT_1}\widetilde{\mathcal{M}}_{i\iota},\ \ \iota=1,2,3,4,5. \quad (28)$$

By (22)-(28), we propose the following profile EL ratio statistic for $\psi\in\mathbf{R}^{p+q+k+4}$:

$$R(\psi)=\sup_{p_i,1\le i\le nT_1}\prod_{i=1}^{nT_1}(nT_1)p_i, \quad (29)$$

where $\{p_i\}$ satisfy

$$p_i\ge 0,1\le i\le nT_1,\sum_{i=1}^{nT_1}p_i=1,$$

$$\sum_{i=1}^{nT_1}p_i a_{i,1}e_i=0,$$

$$\sum_{i=1}^{nT_1}p_i\left\{h_{ii,1}(e_i^2-\sigma^2)+2e_i\sum_{j=1}^{i-1}h_{ij,1}e_j\right\}=0,$$

$$\sum_{i=1}^{nT_1} p_i \left\{ a_{i,2}e_i + h_{ii,2}(e_i^2-\sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,2}e_j + h_{ii,3}(e_i^2-\sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,3}e_j \right\} = 0,$$

$$\sum_{i=1}^{nT_1} p_i \left\{ h_{ii,4}(e_i^2-\sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,4}e_j \right\} = 0,$$

$$\sum_{i=1}^{nT_1} p_i \left\{ h_{ii,5}(e_i^2-\sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,5}e_j \right\} = 0,$$

Let

$$\omega_i(\psi) = \begin{pmatrix} a_{i,1}e_i \\ h_{ii,1}(e_i^2-\sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,1}e_j \\ a_{i,2}e_i + h_{ii,2}(e_i^2-\sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,2}e_j + h_{ii,3}(e_i^2-\sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,3}e_j \\ h_{ii,4}(e_i^2-\sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,4}e_j \\ h_{ii,5}(e_i^2-\sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,5}e_j \end{pmatrix}_{(p+q+k+4)\times 1}$$

where e_i is the i-th component of ϵ^* which satisfied with $Y^* - X^*\theta \equiv DP\epsilon^*$.

Following Owen (1990), one can show that

$$p_i = \frac{1}{nT_1} \cdot \frac{1}{1+\lambda'(\psi)\omega_i(\psi)}, 1 \le i \le nT_1, \tag{30}$$

and

$$\ell(\psi) = -2\log R(\psi) = 2\sum_{i=1}^{nT_1} \log\{1+\lambda'(\psi)\omega_i(\psi)\}, \tag{31}$$

where $\lambda(\psi)$ is the solution of the following equation:

$$\frac{1}{nT_1} \sum_{i=1}^{nT_1} \frac{\omega_i(\psi)}{1+\lambda'(\psi)\omega_i(\psi)} = 0. \tag{32}$$

2.5 Main Result

Let $\vartheta_j = Ev_{11}^j, j = 3, 4$. Use $Vec(Diag(A))$ to denote the column vector formed with the diagonal elements of A. To obtain the asymptotical distribution of $\ell(\psi)$, we also need following assumptions.

A1. $\{v_{it}, 1 \le i \le n, 1 \le t \le T\}$ are iid for all i and t with $E(v_{it}) = 0$, $Var(v_{it}) = \sigma^2$, and $E|v_{it}|^{4+\eta_1} < \infty$ for some $\eta_1 > 0$.

A2. (i) v_{jt} are mutually independent, and they are independent of $x_{\iota s}$ and z_ι for all i, j, ι, t, s; (ii) All elements in (x_{it}, z_i) have $4+\eta_1$ moments for some $\eta_1 > 0$; (iii) $\{x_{it}, t = \cdots, -1, 0, 1, \cdots\}$ and $\{z_i\}$ are strictly exogenous and independent across i; (v) $\max\{|\rho|, |\lambda|\} < 1$.

A3. (i) The row and column sums of W_n are uniformly bounded in absolute value; (ii) $\{B_n^{-1}\}$ are uniformly bounded in either row or column sums, uniformly in λ in a compact parameter space $\mathbf{\Lambda}$, and $c_1 \leq \inf_{\lambda\in\mathbf{\Lambda}} \lambda_{max}(B_n' B_n) \leq \sup_{\lambda\in\mathbf{\Lambda}} \lambda_{max}(B_n' B_n) \leq c_2 \leq \infty$.

A4. There is a constants $c_j > 0, j = 3, 4$, such that $0 < c_3 \leq \lambda_{min}(n^{-1}\Sigma_{p+q+k+4}) \leq \lambda_{max}(n^{-1}\Sigma_{p+q+k+4}) \leq c_4 < \infty$, where $\lambda_{min}(A)$ and $\lambda_{max}(A)$ denote the minimum and maximum eigenvalues of a matrix A, respectively,

$$\Sigma_{p+q+k+4} = \Sigma'_{p+q+k+4} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \Sigma_{14} & \Sigma_{15} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} & \Sigma_{24} & \Sigma_{25} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} & \Sigma_{34} & \Sigma_{35} \\ \Sigma_{41} & \Sigma_{42} & \Sigma_{43} & \Sigma_{44} & \Sigma_{45} \\ \Sigma_{51} & \Sigma_{52} & \Sigma_{53} & \Sigma_{54} & \Sigma_{55} \end{pmatrix} \tag{33}$$

where

$$\Sigma_{11} = \sigma^2 E\{X^{*\prime}\Omega^{*-1}X^*\}, \quad \Sigma_{12} = \vartheta_3 E\{X^{*\prime}\Omega^{*-1}DP\}Vec(diag(H_1)),$$

$$\begin{aligned}\Sigma_{13} \quad = \quad & 2\sigma^2 E\{X^{*\prime}\Omega^{*-1}\eta^*_{-1}\} + \vartheta_3 E\{X^{*\prime}\Omega^{*-1}DP\}Vec(diag(H_2)) \\ & +\vartheta_3 E\{X^{*\prime}\Omega^{*-1}DP\}Vec(diag(H_3)),\end{aligned}$$

$$\Sigma_{14} = \vartheta_3 E\{X^{*\prime}\Omega^{*-1}DP\}Vec(diag(H_4)), \quad \Sigma_{15} = \vartheta_3 E\{X^{*\prime}\Omega^{*-1}DP\}Vec(diag(H_5)),$$

$$\Sigma_{22} = (\vartheta_4 - 3\sigma^4)\|Vec(diag(H_1))\|^2 + 2\sigma^4 tr(H_1^2),$$

$$\begin{aligned}\Sigma_{23} \quad = \quad & (\vartheta_4 - 3\sigma^4)Vec'(diag(H_1))Vec(diag(H_2)) + 2\sigma^4 tr(H_1 H_2) \\ & +(\vartheta_4 - 3\sigma^4)Vec'(diag(H_1))Vec(diag(H_3)) + 2\sigma^4 tr(H_1 H_3) \\ & +\vartheta_3 E\{2\eta^{*\prime}_{-1}\Omega^{*-1}DP\}Vec(diag(H_1)),\end{aligned}$$

$$\Sigma_{24} = (\vartheta_4 - 3\sigma^4)Vec'(diag(H_1))Vec(diag(H_4)) + 2\sigma^4 tr(H_1 H_4),$$

$$\Sigma_{25} = (\vartheta_4 - 3\sigma^4)Vec'(diag(H_1))Vec(diag(H_5)) + 2\sigma^4 tr(H_1 H_5),$$

$$\begin{aligned}\Sigma_{33} \quad = \quad & 4\sigma^2 E\{\eta^{*\prime}_{-1}\Omega^{*-1}\eta^*_{-1}\} + (\vartheta_4 - 3\sigma^4)\|Vec(diag(H_2))\|^2 + 2\sigma^4 tr(H_2^2), \\ & +(\vartheta_4 - 3\sigma^4)\|Vec(diag(H_3))\|^2 + 2\sigma^4 tr(H_3^2) \\ & +\vartheta_3 E\{2\eta^{*\prime}_{-1}\Omega^{*-1}DP\}Vec(diag(H_2)) + \vartheta_3 E\{2\eta^{*\prime}_{-1}\Omega^{*-1}DP\}Vec(diag(H_3)) \\ & +(\vartheta_4 - 3\sigma^4)Vec'(diag(H_2))Vec(diag(H_3)) + 2\sigma^4 tr(H_2 H_3),\end{aligned}$$

$$\Sigma_{34} = \vartheta_3 E\{2\eta_{-1}^{*}{}'\Omega^{*-1}DP\}Vec(diag(H_4)) + (\vartheta_4 - 3\sigma^4)Vec'(diag(H_2))Vec(diag(H_4)) + 2\sigma^4 tr(H_2H_4) + (\vartheta_4 - 3\sigma^4)Vec'(diag(H_3))Vec(diag(H_4)) + 2\sigma^4 tr(H_3H_4),$$

$$\Sigma_{35} = \vartheta_3 E\{2\eta_{-1}^{*}{}'\Omega^{*-1}DP\}Vec(diag(H_5)) + (\vartheta_4 - 3\sigma^4)Vec'(diag(H_2))Vec(diag(H_5)) + 2\sigma^4 tr(H_2H_5) + (\vartheta_4 - 3\sigma^4)Vec'(diag(H_3))Vec(diag(H_5)) + 2\sigma^4 tr(H_3H_5),$$

$$\Sigma_{44} = (\vartheta_4 - 3\sigma^4)\|Vec(diag(H_4))\|^2 + 2\sigma^4 tr(H_4^2),$$

$$\Sigma_{45} = (\vartheta_4 - 3\sigma^4)Vec'(diag(H_4))Vec(diag(H_5)) + 2\sigma^4 tr(H_4H_5),$$

$$\Sigma_{55} = (\vartheta_4 - 3\sigma^4)\|Vec(diag(H_5))\|^2 + 2\sigma^4 tr(H_5^2).$$

A5. $n \to \infty$ but $T \geq 2$ fixed.

REMARK 1 *(33) is verified in the proof of Lemma 3. Conditions A0-A4 are common assumptions for spatial models. For example, A0-A4 are used in Assumptions R0, G1, G2, R(i)-(iii) in Su and Yang (2015), and the analog of $0 < c_3 \leq \lambda_{min}(n^{-1}\Sigma_{p+q+k+4})$ is employed in the assumption of Theorem 1 in Kelejian and Prucha (2001). From Conditions A0-A3, one can see that $\lambda_{max}(n^{-1}\Sigma_{p+q+k+4}) \leq c_4 < \infty$.*

We now state the main results.

THEOREM 1 *Suppose that Assumptions A0-A5 are satisfied and the initial observation y_0 are endogenously given. Then under model (1) and (2), as n$\to \infty$,*

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

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

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

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

If one is interested to obtain the confidence region for some components of ψ, please refer to use the method of Corollary 5 in Qin and Lawless (1994) or the approach of Theorem 2 in Qin (2021) to treat this issue.

3. Simulations

According to Anselin(1988), when the error term $\{v_{it}\}$ is normally distributed, the likelihood ratio (LR): $LR(\psi_0) = 2(L(\hat{\psi}) - L(\psi_0))$ is asymptotically distributed as $\chi^2_{p+q+k+4}$ under the null hypothesis: $\psi = \psi_0$, where L is the corresponding log-likelihood and $\hat{\psi}$ is the MLE. It follows that LR-based confidence region for ψ with asymptotically correct coverage probability α can be constructed as follows:

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

We note that the LR method requires to know the form of the distribution of the population in study, while the EL method does not. This fact implies that the EL method performs better than LR method theoretically when the population distribution is not normal. Our following simulation results do confirm this conclusion.

We conduct a small simulation study to compare the finite sample performances of the confidence regions based on the EL and LR methods with confidence level $\alpha = 0.95$, and report the proportions of $\ell(\psi_0) \leq z_{0.95}(p + q + k + 4)$ and $LR(\psi_0) \leq z_{0.95}(p + q + k + 4)$ respectively in 500 replications, where ψ_0 is the true value of ψ. The results of simulations are reported in Tables 1-6.

In these simulations, we use the following model to generate data: $y_t = \rho y_{t-1} + x_t\beta + z\gamma + u_t, u_t = \lambda M_n u_t + v_t, t = 1, 2, ..., T$, $y_0 = \widetilde{\eta}_0 + u_0$ where y_t, y_{t-1}, x_t and z 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 et al. (2001), and the elements of z were randomly generated form Bernoulli(0.5). We chose $m = 6, \beta = 1, \gamma = 1, \pi = \mathbf{1}_k$, ζ is taken from $N(0, I_n)$, (ρ, λ) were taken as (-0.85,-0.8), (0.85,0.8), (-0.15,-0.2), and (0.15,0.2), respectively, and $v'_{it}s$ are taken from $N(0, 1)$, $t(5)$ and $\chi^2_4 - 4$, respectively.

For the contiguity weight matrix $W_n = (W_{ij})$, we took $W_{ij} = 1$ if spatial units i and j are neighbors by queen contiguity rule (namely, they share common border or vertex), $W_{ij} = 0$ otherwise (Anselin, 1988, p.18). We first considered four ideal cases of spatial units: $n = \tilde{n} \times \tilde{n}$ regular grid with $\tilde{n} = 7, 10, 13$ and 16, denoting W_n as $grid_{49}, grid_{100}, grid_{169}$ and $grid_{256}$, respectively. The sample size n is taken as $49, 100, 169$ and 256, respectively. A transformation is often used in applications to convert the matrix W_n to the unity of row sums.

From tables 1-6, we can see that the confidence regions based on LR performs well with coverage probabilities very close to the nominal level 0.95 when the error term v_{it} is normally distributed and n is large, but not well in other cases. The coverage probabilities of the confidence regions based on LR fall to the ranges [0.815, 0.910] (at $T = 3$) and [0.835, 0.898] (at $T = 4$) for t distribution , which are far from the nominal level 0.95, or [0.865, 0.930] (at $T = 3$) and [0.825, 0.946] (at $T = 4$) for χ^2

Table 1: Coverage probabilities of the LR and EL confidence regions with $e_i \sim N(0, 1)$ and $T = 3$.

(ρ, λ)	W_n	LR	EL	(ρ, λ)	W_n	LR	EL
(-0.85, -0.8)	$grid_{49}$	0.945	0.890	(-0.15, -0.2)	$grid_{49}$	0.926	0.849
	$grid_{100}$	0.952	0.940		$grid_{100}$	0.931	0.915
	$grid_{169}$	0.940	0.945		$grid_{169}$	0.923	0.950
	$grid_{256}$	0.941	0.9341		$grid_{256}$	0.960	0.947
(0.85, 0.8)	$grid_{49}$	0.951	0.910	(0.15, 0.2)	$grid_{49}$	0.942	0.847
	$grid_{100}$	0.959	0.948		$grid_{100}$	0.940	0.915
	$grid_{169}$	0.967	0.936		$grid_{169}$	0.950	0.950
	$grid_{256}$	0.978	0.944		$grid_{256}$	0.925	0.944

Table 2: Coverage probabilities of the EL and LR confidence regions with $e_i \sim N(0, 1)$ and $T = 4$.

(ρ, λ)	W_n	LR	EL	(ρ, λ)	W_n	LR	EL
(-0.85, -0.8)	$grid_{49}$	0.929	0.890	(-0.15, -0.2)	$grid_{49}$	0.934	0.835
	$grid_{100}$	0.942	0.945		$grid_{100}$	0.936	0.894
	$grid_{169}$	0.940	0.938		$grid_{169}$	0.937	0.946
	$grid_{256}$	0.950	0.948		$grid_{256}$	0.970	0.955
(0.85, 0.8)	$grid_{49}$	0.949	0.905	(0.15, 0.2)	$grid_{49}$	0.933	0.825
	$grid_{100}$	0.948	0.915		$grid_{100}$	0.944	0.910
	$grid_{169}$	0.973	0.950		$grid_{169}$	0.930	0.944
	$grid_{256}$	0.945	0.942		$grid_{256}$	0.910	0.936

Table 3: Coverage probabilities of the LR and EL confidence regions with $e_i \sim t(5)$ and $T = 3$.

(ρ, λ)	W_n	LR	EL	(ρ, λ)	W_n	LR	EL
(-0.85, -0.8)	$grid_{49}$	0.896	0.805	(-0.15, -0.2)	$grid_{49}$	0.847	0.728
	$grid_{100}$	0.859	0.884		$grid_{100}$	0.848	0.839
	$grid_{169}$	0.844	0.930		$grid_{169}$	0.830	0.925
	$grid_{256}$	0.850	0.935		$grid_{256}$	0.825	0.955
(0.85, 0.8)	$grid_{49}$	0.892	0.865	(0.15, 0.2)	$grid_{49}$	0.858	0.723
	$grid_{100}$	0.889	0.885		$grid_{100}$	0.860	0.848
	$grid_{169}$	0.898	0.910		$grid_{169}$	0.815	0.893
	$grid_{256}$	0.910	0.945		$grid_{256}$	0.845	0.924

Table 4: Coverage probabilities of the EL and LR confidence regions with $e_i \sim t(5)$ and $T = 4$.

(ρ, λ)	W_n	LR	EL	(ρ, λ)	W_n	LR	EL
(-0.85, -0.8)	$grid_{49}$	0.898	0.840	(-0.15, -0.2)	$grid_{49}$	0.844	0.780
	$grid_{100}$	0.871	0.901		$grid_{100}$	0.849	0.825
	$grid_{169}$	0.855	0.921		$grid_{169}$	0.845	0.889
	$grid_{256}$	0.835	0.936		$grid_{256}$	0.845	0.915
(0.85, 0.8)	$grid_{49}$	0.894	0.820	(0.15, 0.2)	$grid_{49}$	0.849	0.718
	$grid_{100}$	0.889	0.915		$grid_{100}$	0.858	0.838
	$grid_{169}$	0.860	0.920		$grid_{169}$	0.835	0.898
	$grid_{256}$	0.870	0.934		$grid_{256}$	0.855	0.930

Table 5: Coverage probabilities of the LR and EL confidence regions with $e_i + 4 \sim \chi^2(4)$ and $T = 3$.

(ρ, λ)	W_n	LR	EL	(ρ, λ)	W_n	LR	EL
(-0.85, -0.8)	$grid_{49}$	0.930	0.824	(-0.15, -0.2)	$grid_{49}$	0.880	0.740
	$grid_{100}$	0.913	0.902		$grid_{100}$	0.866	0.864
	$grid_{169}$	0.920	0.917		$grid_{169}$	0.910	0.910
	$grid_{256}$	0.910	0.933		$grid_{256}$	0.865	0.925
(0.85, 0.8)	$grid_{49}$	0.930	0.830	(0.15, 0.2)	$grid_{49}$	0.892	0.733
	$grid_{100}$	0.930	0.912		$grid_{100}$	0.887	0.882
	$grid_{169}$	0.915	0.932		$grid_{169}$	0.875	0.914
	$grid_{256}$	0.905	0.940		$grid_{256}$	0.865	0.930

Table 6: Coverage probabilities of the EL and LR confidence regions with $e_i + 4 \sim \chi^2(4)$ and $T = 4$.

(ρ, λ)	W_n	LR	EL	(ρ, λ)	W_n	LR	EL
(-0.85, -0.8)	$grid_{49}$	0.919	0.828	(-0.15, -0.2)	$grid_{49}$	0.876	0.732
	$grid_{100}$	0.914	0.902		$grid_{100}$	0.886	0.848
	$grid_{169}$	0.930	0.918		$grid_{169}$	0.825	0.902
	$grid_{256}$	0.935	0.934		$grid_{256}$	0.880	0.923
(0.85, 0.8)	$grid_{49}$	0.946	0.866	(0.15, 0.2)	$grid_{49}$	0.883	0.732
	$grid_{100}$	0.926	0.922		$grid_{100}$	0.904	0.844
	$grid_{169}$	0.935	0.932		$grid_{169}$	0.885	0.910
	$grid_{256}$	0.945	0.936		$grid_{256}$	0.855	0.927

distribution, which are close to the nominal level 0.95 at times. 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 v_{it} is normally distributed or not. In other words, the distribution of the error terms affects both the LR and EL methods, but the EL method performs much better than the LR method. This can be explained as follows: the EL method is a nonparametric method, which does not require to specify distribution of the data, but the LR 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 LR method than the EL method.

Our simulation results recommend EL method when we are not sure whether the errors are normally distributed. R codes related to this article are available online indicated in Appendix B.

4. Concluding remarks

In this article we consider a specification for spatial dynamic panel data models with spatial errors and without space-specific effects, which is studied in detail under the framework that the cross-sectional dimension n is large and the time dimension T is fixed. This specification of models are close to SDPD models in Su and Yang (2015) and in Rong et al. (2021). Su and Yang (2015) have developed and studied QML estimation procedures that treat the data generating process as endogenous versus exogenous. Rong et al. (2021) have considered the EL approach to spatial dynamic panel data models with spatial lags and spatial errors when the initial period cross-sectional observations are exogenous. As Parent and LeSage (2011) pointed out that correct treatment of the initial period observations are important, especially in the case of small T. We set forth the EL method to treat endogenous initial observations.

In endogenous case, it may not be easy to derive the initial period observations. In this paper, we follow Bhargava and Sargan (1983) to predict those observations under the assumption that the underlying process generating the data starts in the distant past. Our simulation shows that the confidence regions based on EL method closer to the nominal level 0.95 than LR method as the number of spatial units n is large enough, whether the error term v_{it} is normally distributed or not. On the other hand, calculation based on EL method simpler than LR method. In other words, the EL method performs much better than the LR method. The reason of this result may be that the EL method is a nonparametric method, but the LR method is based on the QML, and QML is a parametric method that depends to specify distribution of the data. The EL method is recommended when we are not sure the distributions of the data.

Future research might involve adding a spatial lag also as well a space-specific effect in the models, which may need to use an adjusted EL method.

Acknowledgements This work was partially supported by the National Natural Science Foundation of China (12061017, 12161009), and the Research Fund of Guangxi Key Lab of Multi-source Information Mining & Security (22-A-01-01).

Appendix A. Proof of Theorem 1

To prove the main results, we need following lemmas. Let

$$\widetilde{Q}_n = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{nij}\epsilon_{ni}\epsilon_{nj} + \sum_{i=1}^{n} b_{ni}\epsilon_{ni},$$

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

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

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

$$\mu_{\widetilde{Q}_n} = \sum_{i=1}^{n} a_{nii}\sigma_{ni}^2,$$

$$\sigma_{\widetilde{Q}_n}^2 = 2\sum_{i=1}^{n}\sum_{j=1}^{n} a_{nij}^2\sigma_{ni}^2\sigma_{nj}^2 + \sum_{i=1}^{n} b_{ni}^2\sigma_{ni}^2 + \sum_{i=1}^{n}\{a_{nii}^2(\mu_{ni}^{(4)} - 3\sigma_{ni}^4) + 2b_{ni}a_{nii}\mu_{ni}^{(3)}\}, \qquad (34)$$

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 $n^{-1}\sigma_{\widetilde{Q}_n}^2 \ge c$ for some constant $c > 0$. Then*

$$\frac{\widetilde{Q}_n - \mu_{\widetilde{Q}_n}}{\sigma_{\widetilde{Q}_n}} \xrightarrow{d} N(0,1).$$

Proof. See Theorem 1 and the remark 12 in Kelejian and Prucha (2001).

LEMMA 2 *Let $\eta_1, \eta_2, \cdots, \eta_n$ be a sequence of stationary random variables, with $E|\eta_1|^s < \infty$ for some constants $s > 0$ and $C > 0$. Then*

$$\max_{1\le i\le n} |\eta_i| = o(n^{1/s}), \quad a.s.$$

Proof. It is straightforward.

Lemma 3 *Suppose that Assumptions A1-A3 are satisfied, then as* $n \to \infty$,

$$Z_n = \max_{1 \le i \le nT_1} \|\omega_i(\psi)\| = o_p((nT_1)^{2/(4+\eta_1)}) \quad a.s., \tag{35}$$

$$\Sigma_{p+q+k+4}^{-1/2} \sum_{i=1}^{nT_1} \omega_i(\psi) \xrightarrow{d} N(0, I_{p+q+k+4}), \tag{36}$$

$$(nT_1)^{-1} \sum_{i=1}^{nT_1} \omega_i(\psi)\omega_i^{\tau}(\psi) = (nT_1)^{-1}\Sigma_{p+q+k+4} + o_p(1), \tag{37}$$

$$\sum_{i=1}^{nT_1} \|\omega_i(\psi)\|^3 = O_p(nT_1). \tag{38}$$

where $\Sigma_{p+q+k+4}$ *is given in 33.*

Proof. As

$$\begin{aligned}
Z_n \le & \max_{1\le i\le nT_1} \|a_{i,1}e_i\| + \max_{1\le i\le nT_1} \left| h_{ii,1}(e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,1}e_j \right| \\
& + \max_{1\le i\le nT_1} \left| a_{i,2}e_i + h_{ii,2}(e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,2}e_j + h_{ii,3}(e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,3}e_j \right| \\
& + \max_{1\le i\le nT_1} \left| h_{ii,4}(e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,4}e_j \right| + \max_{1\le i\le nT_1} \left| h_{ii,5}(e_i^2 - \sigma^2) + 2e_i \sum_{j=1}^{i-1} h_{ij,5}e_j \right| \\
\le & \max_{1\le i\le nT_1} \|a_{i,1}e_i\| + \max_{1\le i\le nT_1} |a_{i,2}e_i| + \sum_{\iota=1}^{5} \left\{ \max_{1\le i\le nT_1} \left| h_{ii,\iota}(e_i^2 - \sigma^2) \right| + \max_{1\le i\le nT_1} \left| 2e_i \sum_{j=1}^{i-1} h_{ij,\iota}e_j \right| \right\},
\end{aligned}$$

and by conditions A0-A3 and Lemma 2,

$$\max_{1\le i\le nT_1} \|a_{i,\iota}e_i\| = \max_{1\le i\le nT_1} \|a_{i,\iota}\| o_p((nT_1)^{1/(4+\eta_1)}) = o_p((nT_1)^{1/(4+\eta_1)}), \iota = 1, 2,$$

$$\max_{1\le i\le nT_1} |e_i^2 - \sigma^2| = o_p((nT_1)^{2/(4+\eta_1)}).$$

In addition, by Lemma B.2 in Su and Yang (2015), $(DP)'\Omega^{*-1}DP$, $(DP)'P_\rho^*DP$, $(DP)'P_\lambda^*DP$ and $(DP)'P_{\phi_\varsigma}^*DP$ are uniformly bounded in both row and column sums. It follows that

$$\begin{aligned}
\max_{1\le i\le nT_1} \left| e_i \sum_{j=1}^{i-1} h_{ij,\iota}e_j \right| &= (\max_{1\le i\le nT_1} |e_i|)^2 \cdot \max_{1\le i\le nT_1} \left(\sum_{j=1}^{i-1} |h_{ij,\iota}| \right) \\
&= o_p((nT_1)^{2/(4+\eta_1)}), \iota = 1, 2, 3, 4, 5,
\end{aligned}$$

$$\begin{aligned}\max_{1\le i\le nT_1}|h_{ii,\iota}(e_i^2-\sigma^2)| &= \max_{1\le i\le nT_1}|h_{ii,\iota}|o_p((nT_1)^{2/(4+\eta_1)})\\ &= o_p((nT_1)^{2/(4+\eta_1)}), \iota=1,2,3,4,5,\end{aligned}$$

Thus $Z_n=o_p((nT_1)^{2/(4+\eta_1)})$, and (35) is proved.

For any given $l=(l_1', l_2, l_3, l_4, l_5)'\in R^{p+q+k+4}$ with $\|l\|=1$, where $l_1\in R^{p+q+k}$, $l_j\in R, j=2,3,4,5$. Then

$$\begin{aligned}l'\omega_i(\psi) &= l_1'a_{i,1}e_i+l_2\left\{h_{ii,1}(e_i^2-\sigma^2)+2e_i\sum_{j=1}^{i-1}h_{ij,1}e_j\right\}\\ &\quad +l_3\left\{a_{i,2}e_i+h_{ii,2}(e_i^2-\sigma^2)+2e_i\sum_{j=1}^{i-1}h_{ij,2}e_j+h_{ii,3}(e_i^2-\sigma^2)+2e_i\sum_{j=1}^{i-1}h_{ij,3}e_j\right\}\\ &\quad +l_4\left\{h_{ii,4}(e_i^2-\sigma^2)+2e_i\sum_{j=1}^{i-1}h_{ij,4}e_j\right\}\\ &\quad +l_5\left\{h_{ii,5}(e_i^2-\sigma^2)+2e_i\sum_{j=1}^{i-1}h_{ij,5}e_j\right\}\\ &= (l_2h_{ii,1}+l_3h_{ii,2}+l_3h_{ii,3}+l_4h_{ii,4}+l_5h_{ii,5})(e_i^2-\sigma^2)\\ &\quad +2e_i\sum_{j=1}^{i-1}(l_2h_{ij,1}+l_3h_{ij,2}+l_3h_{ij,3}+l_4h_{ij,4}+l_5h_{ij,5})e_j+(l_1'a_{i,1}+l_3a_{i,2})e_i\end{aligned}$$

Denote

$$\sum_{i=1}^{nT_1}l'\omega_i(\psi)=\sum_{i=1}^{nT_1}\left\{c_{ii}(e_i^2-\sigma^2)+2e_i\sum_{j=1}^{i-1}c_{ij}e_j+d_ie_i\right\}$$

where

$$c_{ij}=l_2h_{ij,1}+l_3h_{ij,2}+l_3h_{ij,3}+l_4h_{ij,4}+l_5h_{ij,5},\ d_i=l_1'a_{i,1}+l_3a_{i,2}.$$

Let

$$Q_n=\sum_{i=1}^{nT_1}\sum_{j=1}^{nT_1}c_{ij}e_ie_j+\sum_{i=1}^{nT_1}d_ie_i,$$

Then

$$Q_n^*=Q_n-E(Q_n)=\sum_{i=1}^{nT_1}l'\omega_i(\psi)$$

We now derive the variance of Q_n^* given x_0, X and Z. It can be show that

$$\begin{aligned}
\sum_{i=1}^{nT_1}\sum_{j=1}^{nT_1} c_{ij}^2 &= \sum_{i=1}^{nT_1}\sum_{j=1}^{nT_1}(l_2h_{ij,1}+l_3h_{ij,2}+l_3h_{ij,3}+l_4h_{ij,4}+l_5h_{ij,5})^2 \\
&= l_2^2tr(H_1^2)+l_3^2tr(H_2^2)+l_3^2tr(H_3^2)+l_4^2tr(H_4^2)+l_5^2tr(H_5^2) \\
&\quad +2l_2l_3tr(H_1H_2)+2l_2l_3tr(H_1H_3)+2l_2l_4tr(H_1H_4)+2l_2l_5tr(H_1H_5) \\
&\quad +2l_3^2tr(H_2H_3)+2l_3l_4tr(H_2H_4)+2l_3l_5tr(H_2H_5)+2l_3l_4tr(H_3H_4) \\
&\quad +2l_3l_5tr(H_3H_5)+2l_4l_5tr(H_4H_5) \qquad (39)
\end{aligned}$$

$$\begin{aligned}
\sum_{i=1}^{nT_1} c_{ii}^2 &= \sum_{i=1}^{nT_1}(l_2h_{ii,1}+l_3h_{ii,2}+l_3h_{ii,3}+l_4h_{ii,4}+l_5h_{i,5})^2 \\
&= 2l_2l_3Vec'(diag(H_1))Vec(diag(H_2))+2l_2l_3Vec'(diag(H_1))Vec(diag(H_3)) \\
&\quad +2l_2l_4Vec'(diag(H_1))Vec(diag(H_4))+2l_2l_5Vec'(diag(H_1))Vec(diag(H_5)) \\
&\quad +2l_3^2Vec'(diag(H_2))Vec(diag(H_3))+2l_3l_4Vec'(diag(H_2))Vec(diag(H_4)) \\
&\quad +2l_3l_5Vec'(diag(H_2))Vec(diag(H_5))+2l_3l_4Vec'(diag(H_3))Vec(diag(H_4)) \\
&\quad +2l_3l_5Vec'(diag(H_3))Vec(diag(H_5))+2l_4l_5Vec'(diag(H_4))Vec(diag(H_5)) \\
&\quad +l_2^2\|Vec(diag(H_1))\|^2+l_3^2\|Vec(diag(H_2))\|^2+l_3^2\|Vec(diag(H_3))\|^2 \\
&\quad +l_4^2\|Vec(diag(H_4))\|^2+l_5^2\|Vec(diag(H_5))\|^2 \qquad (40)
\end{aligned}$$

$$\begin{aligned}
\sum_{i=1}^{nT_1} d_i^2 &= \sum_{i=1}^{nT_1}(l_1'a_{i,1}+l_3a_{i,2})^2 \\
&= l_1'\left(\sum_{i=1}^{nT_1}a_{i,1}a_{i,1}'\right)l_1+l_3^2\sum_{i=1}^{nT_1}a_{i,2}^2+2l_1'\left(\sum_{i=1}^{nT_1}a_{i,1}a_{i,2}\right)l_3 \\
&= l_1'\{X^{*\prime}\Omega^{*-1}X^*\}l_1+l_3^2\{4\eta_{-1}^{*}{}'\Omega^{*-1}\eta_{-1}^*\}+2l_1'l_3\{2X^{*\prime}\Omega^{*-1}\eta_{-1}^*\} \qquad (41)
\end{aligned}$$

$$\begin{aligned}
\sum_{i=1}^{nT_1} c_{ii}d_i &= \sum_{i=1}^{nT_1}(l_2h_{ii,1}+l_3h_{ii,2}+l_3h_{ii,3}+l_4h_{ii,4}+l_5h_{i,5})(l_1'a_{i,1}+l_3a_{i,2}) \\
&= l_1'l_2\sum_{i=1}^{nT_1}a_{i,1}h_{ii,1}+l_1'l_3\sum_{i=1}^{nT_1}a_{i,1}h_{ii,2}+l_1'l_3\sum_{i=1}^{nT_1}a_{i,1}h_{ii,3}+l_1'l_4\sum_{i=1}^{nT_1}a_{i,1}h_{ii,4} \\
&\quad +l_1'l_5\sum_{i=1}^{nT_1}a_{i,1}h_{ii,5}+l_2l_3\sum_{i=1}^{nT_1}a_{i,2}h_{ii,1}+l_3^2\sum_{i=1}^{nT_1}a_{i,2}h_{ii,2}+l_3^2\sum_{i=1}^{nT_1}a_{i,2}h_{ii,3} \\
&\quad +l_3l_4\sum_{i=1}^{nT_1}a_{i,2}h_{ii,4}+l_3l_5\sum_{i=1}^{nT_1}a_{i,2}h_{ii,5}
\end{aligned}$$

$$
\begin{aligned}
&= l_1' l_2 \{X^{*\prime}\Omega^{*-1}DP\}Vec(diag(H_1)) + l_2 l_3\{2\eta_{-1}^{*}{}'\Omega^{*-1}DP\}Vec(diag(H_1)) \\
&\quad + l_1' l_3 \{X^{*\prime}\Omega^{*-1}DP\}Vec(diag(H_2)) + l_3^2\{2\eta_{-1}^{*}{}'\Omega^{*-1}DP\}Vec(diag(H_2)) \\
&\quad + l_1' l_3 \{X^{*\prime}\Omega^{*-1}DP\}Vec(diag(H_3)) + l_3^2\{2\eta_{-1}^{*}{}'\Omega^{*-1}DP\}Vec(diag(H_3)) \\
&\quad + l_1' l_4 \{X^{*\prime}\Omega^{*-1}DP\}Vec(diag(H_4)) + l_3 l_4\{2\eta_{-1}^{*}{}'\Omega^{*-1}DP\}Vec(diag(H_4)) \\
&\quad + l_1' l_5 \{X^{*\prime}\Omega^{*-1}DP\}Vec(diag(H_5)) + l_3 l_5\{2\eta_{-1}^{*}{}'\Omega^{*-1}DP\}Vec(diag(H_5))
\end{aligned} \tag{42}
$$

It is easy to show (e.g., Su and Yang, 2015) that, $n^{-1}X^{*\prime}\Omega^{*-1}X^*$, $n^{-1}\eta_{-1}^{*}{}'\Omega^{*-1}\eta_{-1}^*$, $n^{-1}X^{*\prime}\Omega^{*-1}\eta_{-1}^*$, $n^{-1}X^{*\prime}\Omega^{*-1}DP$, and $n^{-1}\eta_{-1}^{*}{}'\Omega^{*-1}DP$ converge in probability to their expectations. Therefore, by (41)-(42), we have

$$
\begin{aligned}
\sum_{i=1}^{nT_1} d_i^2 &= l_1' E\{X^{*\prime}\Omega^{*-1}X^*\}l_1 + l_3^2 E\{4\eta_{-1}^{*}{}'\Omega^{*-1}\eta_{-1}^*\} + 2l_1' l_3 E\{2X^{*\prime}\Omega^{*-1}\eta_{-1}^*\} \\
&\quad + o_p(n)
\end{aligned} \tag{43}
$$

$$
\begin{aligned}
\sum_{i=1}^{nT_1} c_{ii}d_i &= l_1' l_2 E\{X^{*\prime}\Omega^{*-1}DP\}Vec(diag(H_1)) + l_2 l_3 E\{2\eta_{-1}^{*}{}'\Omega^{*-1}DP\}Vec(diag(H_1)) \\
&\quad + l_1' l_3 E\{X^{*\prime}\Omega^{*-1}DP\}Vec(diag(H_2)) + l_3^2 E\{2\eta_{-1}^{*}{}'\Omega^{*-1}DP\}Vec(diag(H_2)) \\
&\quad + l_1' l_3 E\{X^{*\prime}\Omega^{*-1}DP\}Vec(diag(H_3)) + l_3^2 E\{2\eta_{-1}^{*}{}'\Omega^{*-1}DP\}Vec(diag(H_3)) \\
&\quad + l_1' l_4 E\{X^{*\prime}\Omega^{*-1}DP\}Vec(diag(H_4)) + l_3 l_4 E\{2\eta_{-1}^{*}{}'\Omega^{*-1}DP\}Vec(diag(H_4)) \\
&\quad + l_1' l_5 E\{X^{*\prime}\Omega^{*-1}DP\}Vec(diag(H_5)) + l_3 l_5 E\{2\eta_{-1}^{*}{}'\Omega^{*-1}DP\}Vec(diag(H_5)) \\
&\quad + o_p(n)
\end{aligned} \tag{44}
$$

It follows from (34) that the variance of Q_n^* when the conditional expectation and variance given x_0, X and Z. Let

$$
\sigma_{Q^*}^2 = l'\Sigma_{p+q+k+4}l
$$

where $\Sigma_{p+q+k+4}$ is given in (33). The conditional expectation and variance given x_0, X and Z are denoted as E^* and Var^*, respectively. Then from (21) and note that $E(v^*) = 0$ and $E(\zeta) = 0$, we know that the variance of $\sigma_{Q^*|x_0,X,Z}^2$, substituting (39)-(40) and (43)-(44) into the following, is

$$
\begin{aligned}
\sigma_{Q^*|x_0,X,Z}^2 &= Var^*\left(\sum_{i=1}^{nT_1} l'\omega_i(\psi)\right) \\
&= 2\sum_{i=1}^{nT_1}\sum_{j=1}^{nT_1} c_{ij}^2\sigma^4 + \sum_{i=1}^{nT_1} d_i^2\sigma^2 + \sum_{i=1}^{nT_1}\{c_{ii}^2(\vartheta_4 - 3\sigma^4) + 2c_{ii}d_i\vartheta_3\}
\end{aligned}
$$

$$
\begin{aligned}
&= 2\sigma^4\left\{\sum_{i=1}^{nT_1}\sum_{j=1}^{nT_1}c_{ij}^2\right\}+\sigma^2\left\{\sum_{i=1}^{nT_1}d_i^2\right\}+(\vartheta_4-3\sigma^4)\left\{\sum_{i=1}^{nT_1}c_{ii}^2\right\}+2\vartheta_3\left\{\sum_{i=1}^{nT_1}c_{ii}d_i\right\}\\
&= l'\Sigma_{p+q+k+4}l+o_p(n) \qquad (45)
\end{aligned}
$$

From Condition A4, one can see that $(nT_1)^{-1}\sigma^2_{Q^*|x_0,X,Z}\geq c_1>0$. From Lemma 1, we have

$$\frac{Q_n^*-E^*(Q_n^*)}{\sigma_{Q^*|x_0,X,Z}}\xrightarrow{d^*}N(0,1), \tag{46}$$

where d^* stands for convergence in distribution given x_0, X, and Z. Noting that $(nT_1)^{-1}\sigma^2_{Q^*|x_0,X,Z}\geq c_1>0$ and

$$\sigma^2_{Q^*|x_0,X,Z}=\sigma^2_{Q^*}+o_p(n),$$

one can show that

$$\frac{\sigma^2_{Q^*}}{\sigma^2_{Q^*|x_0,X,Z}}\xrightarrow{p}1. \tag{47}$$

Combing $E^*(Q_n^*)=0$, (46) and (47), we have

$$\frac{Q_n^*}{\sigma_{Q^*}}\xrightarrow{d}N(0,1).$$

Thus proof of (36) is completed.

Next we will prove (37), i. e.

$$(nT_1)^{-1}\sum_{i=1}^{nT_1}\left\{l'\omega_i(\psi)\right\}^2=(nT_1)^{-1}\sigma^2_{Q^*}+o_p(1). \tag{48}$$

Let

$$Y_{in}=l'\omega_i(\psi)=c_{ii}(e_i^2-\sigma^2)+2e_i\sum_{j=1}^{i-1}c_{ij}e_j+d_ie_i=c_{ii}(e_i^2-\sigma^2)+Z_ie_i, \tag{49}$$

where $Z_i=2\sum_{j=1}^{i-1}c_{ij}e_j+d_i$. Let $\mathcal{F}_0=\{\emptyset,\Omega\},\mathcal{F}_i=\sigma(e_1,e_2,...,e_i),1\leq i\leq nT_1$. Then $\mathcal{F}_{i-1}\subseteq\mathcal{F}_i,Y_{in}$ is $\mathcal{F}_i-$measurable and $E(Y_{in}|\mathcal{F}_{i-1})=0$. Thus $\{Y_{in},\mathcal{F}_i,1\leq i\leq nT_1\}$ forms a martingale difference array. From (45), one can see that

$$\sigma^2_{Q^*}=\sum_{i=1}^{nT_1}E^*(Y_{in}^2)+o_p(n).$$

Thus

$$
\begin{aligned}
& (nT_1)^{-1}\sum_{i=1}^{nT_1}\{l'\omega_i(\psi)\}^2 - (nT_1)^{-1}\sigma_{Q^*}^2 \\
= & (nT_1)^{-1}\sum_{i=1}^{nT_1} Y_{in}^2 - (nT_1)^{-1}\sum_{i=1}^{nT_1} E^*(Y_{in}^2) + o_p(1) \\
= & (nT_1)^{-1}\sum_{i=1}^{nT_1}\{Y_{in}^2 - E^*(Y_{in}^2|\mathcal{F}_{i-1})\} \\
& +(nT_1)^{-1}\sum_{i=1}^{nT_1}\{E^*(Y_{in}^2|\mathcal{F}_{i-1}) - E^*Y_{in}^2\} + o_p(1) \\
= & (nT_1)^{-1}S_{n1} + (nT_1)^{-1}S_{n2} + o_p(1),
\end{aligned}
$$

where

$$
S_{n1} = \sum_{i=1}^{nT_1} N_{i,1} = \sum_{i=1}^{nT_1}\{Y_{in}^2 - E^*(Y_{in}^2|\mathcal{F}_{i-1})\},
$$

$$
S_{n2} = \sum_{i=1}^{nT_1} N_{i,2} = \sum_{i=1}^{nT_1}\{E^*(Y_{in}^2|\mathcal{F}_{i-1}) - E^*Y_{in}^2\}.
$$

Let $\mathcal{F}_0 = \{\emptyset, \Omega\}, \mathcal{F}_i = \sigma(e_1, e_2, ..., e_i), 1 \le i \le nT_1$. Then $\mathcal{F}_{i-1} \subseteq \mathcal{F}_i, N_{i,\iota}$ is $\mathcal{F}_i-$measurable and $E^*(N_{i,\iota}|\mathcal{F}_{i-1}) = 0, \iota = 1, 2$. Thus $\{N_{i,\iota}, \mathcal{F}_i, 1 \le i \le nT_1\}, \iota = 1, 2$ forms a martingale difference array. Next we will prove

$$
(nT_1)^{-1}S_{n1} = o_p(1), \tag{50}
$$

and

$$
(nT_1)^{-1}S_{n2} = o_p(1). \tag{51}
$$

In order to prove (50) and (51), we just need to prove $(nT_1)^{-2}ES_{n1}^2 \to 0$ and $(nT_1)^{-2}ES_{n2}^2 \to 0$ respectively. Obviously,

$$
Y_{in}^2 = c_{ii}^2(e_i^2 - \sigma^2)^2 + Z_i^2 e_i^2 + 2c_{ii}Z_i(e_i^2 - \sigma^2)e_i, \tag{52}
$$

then

$$
E^*(Y_{in}^2|\mathcal{F}_{i-1}) = c_{ii}^2 E(e_i^2 - \sigma^2)^2 + Z_i^2\sigma^2 + 2c_{ii}Z_i\vartheta_3, \tag{53}
$$

and

$$
\begin{aligned}
E^*Y_{in}^2 &= E^*\left\{E^*(Y_{in}^2|\mathcal{F}_{i-1})\right\} = c_{ii}^2 E(e_i^2 - \sigma^2)^2 + \sigma^2 E^*(Z_i^2) + 2c_{ii}\vartheta_3 E^*(Z_i) \\
&= c_{ii}^2 E(e_i^2 - \sigma^2)^2 + \sigma^2(4\sum_{j=1}^{i-1} c_{ij}^2\sigma^2 + d_i^2) + 2c_{ii}\vartheta_3 d_i. \qquad (54)
\end{aligned}
$$

By (52)-(53), it follows that

$$
\begin{aligned}
&(nT_1)^{-2}ES_{n1}^2 = (nT_1)^{-2}E\left\{\sum_{i=1}^{nT_1} N_{i,1}\right\}^2 = (nT_1)^{-2}\sum_{i=1}^{nT_1} EN_{i,1}^2 \\
= \;& (nT_1)^{-2}\sum_{i=1}^{nT_1} E\left\{Y_{in}^2 - E^*(Y_{in}^2|\mathcal{F}_{i-1})\right\}^2 \\
= \;& (nT_1)^{-2}\sum_{i=1}^{nT_1} E\Big\{c_{ii}^2\left[(e_i^2-\sigma^2)^2 - E(e_i^2-\sigma^2)^2\right] + Z_i^2(e_i^2-\sigma^2) \\
& +2c_{ii}Z_i(\epsilon_i^3-\sigma^2 e_i-\vartheta_3)\Big\}^2 \\
\leq \;& C(nT_1)^{-2}\sum_{i=1}^{nT_1} E\left\{c_{ii}^4\left[(e_i^2-\sigma^2)^2 - E(e_i^2-\sigma^2)^2\right]^2\right\} \\
& +C(nT_1)^{-2}\sum_{i=1}^{nT_1} E\left\{Z_i^4(e_i^2-\sigma^2)^2\right\} \\
& +C(nT_1)^{-2}\sum_{i=1}^{nT_1} E\left\{c_{ii}^2 Z_i^2(\epsilon_i^3-\sigma^2 e_i-\vartheta_3)^2\right\}. \qquad (55)
\end{aligned}
$$

By condition A0-A3, we have

$$
\begin{aligned}
&(nT_1)^{-2}\sum_{i=1}^{nT_1} E\left\{c_{ii}^4\left[(e_i^2-\sigma^2)^2 - E(e_i^2-\sigma^2)^2\right]^2\right\} \leq C(nT_1)^{-2}\sum_{i=1}^{nT_1} c_{ii}^4 \\
\leq \;& C(nT_1)^{-2}\sum_{i=1}^{nT_1} |l_2h_{ii,1}+l_3h_{ii,2}+l_3h_{ii,3}+l_4h_{ii,4}+l_5h_{ii,5}|^4 \\
\leq \;& C(nT_1)^{-1}\to 0, \qquad (56)
\end{aligned}
$$

and

$$
\begin{aligned}
&(nT_1)^{-2}\sum_{i=1}^{nT_1} E\left\{Z_i^4(e_i^2-\sigma^2)^2\right\} \leq C(nT_1)^{-2}\sum_{i=1}^{nT_1} E\Big(\sum_{j=1}^{i-1} c_{ij}e_j + d_i\Big)^4 \\
\leq \;& C(nT_1)^{-2}\sum_{i=1}^{nT_1} E\Big(\sum_{j=1}^{i-1} c_{ij}e_j\Big)^4 + C(nT_1)^{-2}\sum_{i=1}^{nT_1} Ed_i^4 \\
\leq \;& C(nT_1)^{-2}\sum_{i=1}^{nT_1}\sum_{j=1}^{i-1} c_{ij}^4\vartheta_4 + C(nT_1)^{-2}\sum_{i=1}^{nT_1}\Big(\sum_{j=1}^{i-1} c_{ij}^2\sigma^2\Big)^2 + C(nT_1)^{-2}\sum_{i=1}^{nT_1} Ed_i^4 \\
\leq \;& C(nT_1)^{-2}\sum_{i=1}^{nT_1}\sum_{j=1}^{i-1} |l_2h_{ij,1}+l_3h_{ij,2}+l_3h_{ij,3}+l_4h_{ij,4}+l_5h_{ij,5}|^4
\end{aligned}
$$

$$
\begin{aligned}
&+C(nT_1)^{-2}\sum_{i=1}^{nT_1}\Big(\sum_{j=1}^{i-1}|l_2h_{ij,1}+l_3h_{ij,2}+l_3h_{ij,3}+l_4h_{ij,4}+l_5h_{ij,5}|^2\Big)^2\\
&+C(nT_1)^{-2}\sum_{i=1}^{nT_1}E(l_1'a_{i,1}+l_3a_{i,2})^4\\
\leq\;& C(nT_1)^{-1}\to 0. \qquad (57)
\end{aligned}
$$

Similarly, one can show that

$$
(nT_1)^{-2}\sum_{i=1}^{nT_1}E\Big\{c_{ii}^2Z_i^2(\epsilon_i^3-\sigma^2e_i-\vartheta_3)^2\Big\}\to 0 \qquad (58)
$$

From (55)-(58), we have $(nT_1)^{-2}ES_{n1}^2\to 0$. Next to prove $(nT_1)^{-2}ES_{n2}^2\to 0$, by (53)-(54), we have

$$
\begin{aligned}
&(nT_1)^{-2}ES_{n2}^2=(nT_1)^{-2}E\left\{\sum_{i=1}^{nT_1}N_{i,2}\right\}^2=(nT_1)^{-2}\sum_{i=1}^{nT_1}EN_{i,2}^2\\
=\;&(nT_1)^{-2}\sum_{i=1}^{nT_1}E\Big\{E^*(Y_{in}^2|\mathcal{F}_{i-1})-E^*Y_{in}^2\Big\}^2\\
=\;&(nT_1)^{-2}\sum_{i=1}^{nT_1}E\Big\{\sigma^2(Z_i^2-4\sum_{j=1}^{i-1}c_{ij}^2\sigma^2-d_i^2)+2c_{ii}\vartheta_3(Z_i-d_i)\Big\}^2\\
=\;&(nT_1)^{-2}\sum_{i=1}^{nT_1}E\Big\{\sigma^2\Big[(2\sum_{j=1}^{i-1}c_{ij}e_j)^2-4\sum_{j=1}^{i-1}c_{ij}^2\sigma^2+4\sum_{j=1}^{i-1}c_{ij}e_jd_i\Big]+2c_{ii}\vartheta_3(2\sum_{j=1}^{i-1}c_{ij}e_j)\Big\}^2\\
\leq\;&C(nT_1)^{-2}\sum_{i=1}^{nT_1}E\Big\{\sigma^2\Big[(2\sum_{j=1}^{i-1}c_{ij}e_j)^2-4\sum_{j=1}^{i-1}c_{ij}^2\sigma^2\Big]\Big\}^2+C(nT_1)^{-2}\sum_{i=1}^{nT_1}E\Big\{\sigma^2(4\sum_{j=1}^{i-1}c_{ij}e_jd_i)\Big\}^2\\
&+C(nT_1)^{-2}\sum_{i=1}^{nT_1}E\Big\{2c_{ii}\vartheta_3(2\sum_{j=1}^{i-1}c_{ij}e_j)\Big\}^2. \qquad (59)
\end{aligned}
$$

Note that

$$
\begin{aligned}
&(nT_1)^{-2}\sum_{i=1}^{nT_1}E\Big\{\sigma^2\Big[(2\sum_{j=1}^{i-1}c_{ij}e_j)^2-4\sum_{j=1}^{i-1}c_{ij}^2\sigma^2\Big]\Big\}^2\leq C(nT_1)^{-2}\sigma^4\sum_{i=1}^{nT_1}E\Big(\sum_{j=1}^{i-1}c_{ij}e_j\Big)^4\\
\leq\;&C(nT_1)^{-2}\sum_{i=1}^{nT_1}\sum_{j=1}^{i-1}c_{ij}^4\vartheta_4+C(nT_1)^{-2}\sum_{i=1}^{nT_1}(\sum_{j=1}^{i-1}c_{ij}^2\sigma^2)^2\leq C(nT_1)^{-1}\to 0, \qquad (60)
\end{aligned}
$$

$$
(nT_1)^{-2}\sum_{i=1}^{nT_1}E\Big\{\sigma^2(4\sum_{j=1}^{i-1}c_{ij}e_jd_i)\Big\}^2\leq (nT_1)^{-2}\sigma^6\sum_{i=1}^{nT_1}d_i^2\sum_{j=1}^{i-1}c_{ij}^2\leq C(nT_1)^{-2}\to 0, \qquad (61)
$$

and

$$(nT_1)^{-2}\sum_{i=1}^{nT_1} E\Big\{2c_{ii}\vartheta_3(2\sum_{j=1}^{i-1} c_{ij}e_j)\Big\}^2 = 4\vartheta_3^2\sigma^2(nT_1)^{-2}\sum_{i=1}^{nT_1} c_{ii}^2\sum_{j=1}^{i-1} c_{ij}^2 \leq C(nT_1)^{-1} \to 0, \quad (62)$$

where we use conditions A1 and A2. From (59)-(62), we have $(nT_1)^{-2}ES_{n2}^2 \to 0$. Then (48) can be proved. The proof of (37) is thus complete.

Finally, we will prove (38). Note that

$$\begin{aligned}
&\sum_{i=1}^{nT_1} E\|\omega_i(\psi)\|^3 \leq \sum_{i=1}^{nT_1} E\|a_{i,1}e_i\|^3 + \sum_{i=1}^{nT_1} E\Big|h_{ii,1}(e_i^2-\sigma^2) + 2e_i\sum_{j=1}^{i-1} h_{ij,1}e_j\Big|^3 \\
&+\sum_{i=1}^{nT_1} E\Big|a_{i,2}e_i + h_{ii,2}(e_i^2-\sigma^2) + 2e_i\sum_{j=1}^{i-1} h_{ij,2}e_j + h_{ii,3}(e_i^2-\sigma^2) + 2e_i\sum_{j=1}^{i-1} h_{ij,3}e_j\Big|^3 \\
&+\sum_{i=1}^{nT_1} E\Big|h_{ii,4}(e_i^2-\sigma^2) + 2e_i\sum_{j=1}^{i-1} h_{ij,4}e_j\Big|^3 \\
&+\sum_{i=1}^{nT_1} E\Big|h_{ii,5}(e_i^2-\sigma^2) + 2e_i\sum_{j=1}^{i-1} h_{ij,5}e_j\Big|^3. \quad (63)
\end{aligned}$$

By Conditions A2 and A3,

$$\sum_{i=1}^{nT_1} E\|a_{i,1}e_i\|^3 = O(nT_1), \quad (64)$$

$$\begin{aligned}
&\sum_{i=1}^{nT_1} E\Big|h_{ii,\iota}(e_i^2-\sigma^2) + 2e_i\sum_{j=1}^{i-1} h_{ij,\iota}e_j\Big|^3 \\
\leq\ & C\sum_{i=1}^{nT_1} E\Big|h_{ii,\iota}(e_i^2-\sigma^2)\Big|^3 + C\sum_{i=1}^{nT_1} E\Big|2e_i\sum_{j=1}^{i-1} h_{ij,\iota}e_j\Big|^3 \\
\leq\ & C\sum_{i=1}^{nT_1} E\Big|h_{ii,\iota}(e_i^2-\sigma^2)\Big|^3 + C\sum_{i=1}^{nT_1} E|e_i|^3\sum_{j=1}^{i-1} E\Big|h_{ij,\iota}e_j\Big|^3 \\
&+C\sum_{i=1}^{nT_1} E|e_i|^3\Big\{\sum_{j=1}^{i-1} E\Big(h_{ij,\iota}e_j)^2\Big\}^{3/2} = O(nT_1), \iota = 1, 4, 5 \quad (65)
\end{aligned}$$

The last inequality of (65) is obtained by Rosenthal inequality. Similarly,

$$\sum_{i=1}^{nT_1} E\Big|a_{i,2}e_i + h_{ii,2}(e_i^2-\sigma^2) + 2e_i\sum_{j=1}^{i-1} h_{ij,2}e_j + h_{ii,3}(e_i^2-\sigma^2) + 2e_i\sum_{j=1}^{i-1} h_{ij,3}e_j\Big|^3$$

$$
\begin{aligned}
\leq \ & \sum_{i=1}^{nT_1} E\Big|a_{i,2}e_i\Big|^3 + \sum_{i=1}^{nT_1} E\Big|h_{ii,2}(e_i^2-\sigma^2)\Big|^3 + \sum_{i=1}^{nT_1} E\Big|2e_i\sum_{j=1}^{i-1} h_{ij,2}e_j\Big|^3 \\
& + \sum_{i=1}^{nT_1} E\Big|h_{ii,3}(e_i^2-\sigma^2)\Big|^3 + \sum_{i=1}^{nT_1} E\Big|2e_i\sum_{j=1}^{i-1} h_{ij,3}e_j\Big|^3 = O(nT_1), \qquad (66)
\end{aligned}
$$

From (63)-(66) we have

$$
\sum_{i=1}^{nT_1} E\|\omega_i(\psi)\|^3 = O(nT_1), \qquad (67)
$$

Further, using (67) and Markov inequality, we obtain

$$
\sum_{i=1}^{nT_1} \|\omega_i(\psi)\|^3 = O_p(nT_1).
$$

Thus (37) 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.

Appendix B. Supplementary material

Through https://github.com/Tang-Jay/Empirical-Likelihood-for-SDPD.git, one can find R codes related to this article online.

References

Anselin, L. (1988). *Spatial econometrics: methods and models*. Amsterdam: Kluwer Academic Publishers.

Anselin, L. (2001). Spatial econometrics. In: Baltagi, B. H. (Ed.), *A companion to theoretical econometrics*. Blackwell Publishers Ltd., Massachusetts, 310-330.

Anselin, L., Le Gallo, J., & Jayet, H. (2008). Spatial panel econometrics. In: Mátyás, L., Sevestre, P. (Eds.), *The econometrics of panel data: fundamentals and recent developments in theory and practice*. SpringerVerlag, Berlin Heidelberg, 625-660.

Baltagi, B.H., Song, S.H., & Koh, W. (2003). Testing panel data regression models with spatial error correlation. *Journal of Econometrics, 117*, 123-150.

Baltagi, B.H., Egger, P., & Pfaffermayr, M. (2013). A generalized spatial panel data model with random effects. *Econometric Reviews, 32*, 650-685.

Bandyopadhyay, S., Lahiri, N., & Nordman, D. J. (2015). A frequency domain empirical likelihood method for irregularly spaced spatial data. *Annals of Statistics, 43*, 519-545.

Bhargava, A., & Sargan, J.D. (1983). Estimating dynamic random effects models from panel data covering short time periods. *Econometrica, 51*, 1635-1659.

Chen, S. X., & Keilegom, I. V. (2009). A review on empirical likelihood methods for regression. *Test, 18*, 415-447.

Chen, J., & Qin, J. (1993). Empirical likelihood estimation for finite populations and the effective usage of auxiliary information. *Biometrika, 80*, 107-116.

Cliff, A. D., & Ord, J. K. (1973). *Spatial autocorrelation.* London: Pion Ltd.

Cressien, N. (1993). *Statistics for spatial data.* New York: Wiley.

Elhorst, J.P. (2003). Specification and estimation of spatial panel data models. *International Regional Science Review , 26*, 244-268.

Elhorst, J.P. (2012). Dynamic spatial panels: models, methods, and inferences. *Journal of geographical systems, 14*, 5-28.

Hsiao, C. (2003). *Analysis of Panel Data, second ed.*. Cambridge University Press, Cambridge.

Hsiao, C., Pesaran, M.H., & Tahmiscioglu, A.K. (2002). Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods. *Econometrics, 109*, 107-150.

Jin, F., Lee, & L. F. (2019). GEL estimation and tests of spatial autoregressive models. *Journal of Econometrics, 208*, 585-612.

Kelejian, H. H., & Prucha, L. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics, 17*, 99-121.

Kelejian, H. H., & Prucha, L. R. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review, 40*, 509-533.

Kelejian, H.H., & Prucha, I.R. (2001). On the asymptotic distribution of the Moran I test statistic with applications. *Econometrics, 104*, 219-257.

Kelejian, H. H., Prucha, I. R., & Yuzefovich, Y. (2004). Instrumental variable estimation of a spatial autoregressive model with autoregressive disturbances: large and small sample results. *Spatial and Spatiotemporal Econometrics, 18*, 163-198.

Kelejian, H. H., & Prucha, I. R. (2006). HAC estimation in a spatial framework. *Journal of Econometrics, 140*, 131-154.

Kolaczyk, E. D. (1994). Emprical likelihood for generalized linear models. *Statistica Sinica, 4*, 199-218.

Lee, L.F., & Yu, J. (2010). A spatial dynamic panel data model with both time and individual fixed effects. *Econometric Theory, 26*, 564-597.

Lee, L.F., & Yu, J. (2016). Identification of Spatial Durbin Panel Models. *Applied Economics, 31*, 133-162.

Li, Y. H., Li, Y., & Qin, Y. S. (2020). Empirical likelihood for panel data models with spatial errors. *Communications in Statistics-Theory and Methods, 51*, 2838-2857.

Li, Y. H., & Qin, Y. S. (2022). Empirical likelihood for spatial dynamic panel data models, *Journal of the Korean Statistical Society, 51*, 500-525.

Liu, X., Lee, L. F., & Bollinger, C. R. (2010). An efficient GMM estimator of spatial autoregressive models. *Journal of Econometrics, 159*, 303-319.

Nerlove, M. (1999). *Properties of alternative estimators of dynamic panel models: an empirical analysis of cross-country data for the study of economic growth*. In: Hsiao C, Lahiri K, Lee LF, Pesaran MH (eds) Analysis of panels and limited dependent variable models. Cambridge University Press,Cambridge, pp 136-170

Nordman, D. J. (2008). An empirical likelihood method for spatial regression. *Metrika, 68*, 351-363.

Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika, 75*, 237-249.

Owen, A. B. (1990). Empirical likelihood ratio confidence regions. *Annals of Statistics, 18*, 90-120.

Owen, A. B. (1991). Empirical likelihood for linear models. *Annals of Statistics, 19*, 1725-1747.

Owen, A. B. (2001). *Empirical Likelihood*. New York: Chapman & Hall.

Parent, O., & LeSage, J.P. (2011). A space-time filter for panel data models containing random effects. *Computational Statistics & Data Analysis, 55*, 475-490.

Parent, O., & LeSage, J.P. (2012). Spatial dynamic panel data models with random effects. *Regional Science and Urban Economics, 42*, 727-738.

Qin, Y. S. (1999). Empirical likelihood ratio confidence regions in a partly linear model. *Chinese Journal of applied probability and statistics, 15*, 363-369.

Qin, J., & Lawless, J. (1994). Empirical likelihood and general estimating equations. *Annals of Statistics, 22*, 300-325.

Qin, Y. S. (2021). Empirical likelihood for spatial autoregressive models with spatial autoregressive disturbances. *Sankhyā A: Indian J. Statist., 83*, 1-25.

Qin, Y. S., & Lei, Q. Z. (2021). Empirical likelihood for mixed regressive, spatial autoregressive model based on GMM. *Sankhyā A: Indian J. Statist., 83*, 353-378.

Rong, J. R., Liu, Y., & Qin, Y. S. (2021). Empirical likelihood for spatial dynamic panel data models with spatial lags and spatial errors. *Communications in Statistics-Theory and Methods*. https://doi.org/10.1080/03610926.2022.2032172.

Qu, X., Lee, L.F., & Yu, J. (2017). QML estimation of spatial dynamic panel data models with endogenous time varying spatial weights matrices. *Econometrics, 197*, 173-201.

Su, L., & Yang, Z. (2007). QML estimation of dynamic panel data models with spatial errors. *Singapore Management University manuscript.*

Su, L., & Yang, Z. (2015). QML estimation of dynamic panel data models with spatial errors. *Econometrics, 185*, 230-258.

Wu, C.B. (2004). Weighted empirical likelihood inference. *Statistics & Probability Letters, 66*, 67-79.

Yang, Z., Li, C., & Tse, Y.K. (2006). Functional form and spatial dependence in dynamic panels. *Economics Letters, 91*, 138-145.

Yu, J., R, de Jong., & Lee, L.F. (2008). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. *Journal of Econometrics, 146*, 118-134.

Zhong, B., & Rao, J. N. K. (2000). Empirical likelihood inference under stratified random sampling using auxiliary population information. *Biometrika, 87*, 929-938.