Principal component empirical likelihood method for spatial data with a diverging number of parameters

Jie Tang   Yunlong Zou   Yongsong Qin[1]

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

Abstract The statistical inference for high-dimensional data faces many difficulties and challenges. Traditional empirical likelihood (EL) method do not work when the dimensions of the moment restrictions and the parameters diverge along with the sample size. We propose the principal component empirical likelihood (PCEL) method to address this problem. It is shown that the asymptotically distributions of the PCEL ratio statistics are chi-squared distributions, which are used to obtain confidence regions for the possibly high-dimensional parameters in the spatial autoregressive models with spatial autoregressive disturbances (SARSAR models). A simulation study is conducted to compare the performances of the PCEL with the usual EL method.

Keywords: SARSAR model; empirical likelihood; confidence region; high-dimensional statistical inference;

AMS 2020 subject classification: Primary 62G05   Secondary 62E20

[1]Corresponding author.

Email address: ysqin@gxnu.edu.cn

1. Introduction

Obviously, two geographical units can affect each other mutually, so the spatial data models dealing with spatial interaction effects among geographical units was proposed in Cliff and Ord (1973) and developed into a specialized subject of spatial econometrics after 50 years. Spatial econometric models can also be used to explain the behavior of economic agents other than geographical units, such as individuals, firms, or governments, if they are related to each other through networks. It is necessary to conduct relevant research on the spatial data models because it has been applied broadly to different fields, such as economics, sociology, biology and so on.

In this article, we consider the following spatial autoregressive model with spatial autoregressive disturbances (SARSAR model):

$$Y_n = \rho_1 W_n Y_n + X_n \beta + u_{(n)}, u_{(n)} = \rho_2 M_n u_{(n)} + \epsilon_{(n)}, \tag{1}$$

where n is spatial sample size, $Y_n = (y_1, y_2, ..., y_n)^\tau$ is an $n \times 1$ vector of observations on the dependent variable, the matrix $X_n = (x_1, x_2, \cdots, x_n)^\tau$ with $x_i, i = 1, 2, \cdots, n$ is a $p \times 1$ exogenous vector of observations on the independent variable, β is the $p \times 1$ vector of regression parameters, the scalar parameters $\rho_j, j = 1, 2$, are spatial autoregressive coefficients with $|\rho_j| < 1, j = 1, 2$, W_n and M_n are known $n \times n$ spatial weight matrices whose diagonal elements are zero, the disturbance vector $u_{(n)} = (u_1, u_2, \cdots, u_n)^\tau$ contains a spatially autocorrelated setting, the vector $\epsilon_{(n)} = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n)^\tau$ represents the random innovations which satisfies $E\epsilon_{(n)} = 0, Var(\epsilon_{(n)}) = \sigma^2 I_n$. The $\rho_1 W_n$ term is a spacial lag in the dependent variable and its coefficient represents the spatial influence due to neighbors realized dependent variable. The $\rho_2 M_n$ term is a spacial lag in the disturbances and its coefficient represents the spacial effect of unobservables on neighboring units. This model belongs to the cross-sectional data models, containing three different types of interaction effects: endogenous interaction effects among the dependent variable Y_n, exogenous interaction effects among the independent variables X_n, and interaction effects among the error terms $\epsilon_{(n)}$, which is helpful to study the spatial panel models.

There are two major estimation approaches for the corresponding parameters in spatial cross-sectional models: the quasi-maximum likelihood (QML) and the generalized method of moments (GMMs). The asymptotic properties of the QML estimator and the GMM estimator for the spatial models are researched by Anselin (1988) and Kelejian and Prucha (1999), respectively. However, the accuracy of the normal approximation-based confidence regions of the parameters in the models based on the above methods may be affected by the estimation of asymptotic covariance. More researches on spatial non-panel models can refer to Cliff and Ord (1973), Cressien (1993), Kelejian and Prucha (1998, 1999, 2006), Kelejian et al. (2004) and Liu et al. (2010), etc. These parameter methods are widely extend to the spatial panel model. See, Elhorst (2003), Baltagi et al. (2003, 2013), Yang et al. (2006), Anselin (2001), Anselin et al. (2008), Lee and Yu (2010), Parent and LeSage (2011), Elhorst (2012), Lee and Yu (2016), Qu et al. (2017) and so on.

We propose to use the empirical likelihood (EL) method introduced by Owen (1988, 1990) to construct confidence regions for the parameters in the SARSAR model. 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. Originally, EL method is 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) for generalized linear models, and Qin (1999) for a partly linear model. 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 EL ratio statistics for the parameters in the SARSAR model by a martingale sequence, which greatly promoted the development of empirical likelihood in spatial models. Li et al. (2020) used the EL method to study a spatial static panel data model, Li and Qin (2022) to a spatial dynamic panel data model. More references on EL methods in spatial models can be found in Nordman (2008), Bandyopadhyay et al. (2015), Rong et al. (2021), Qin and Lei (2021), among others.

With the frequent occurrence of high-dimensional data, the curse of dimensionality make the traditional EL methods unable to analyze high-dimensional data. Therefore, the study of high-dimensional data has become a hot issue in statistics. To clearly state the existing idea in using the EL method under high dimensional case, for the moment, we use p and r to denote the dimension of the unknown parameters and the number of moment restrictions, respectively. Let n and ℓ_n be the sample size and the EL statistic, respectively. Under independent samples, for fixed p and r with $r \geq p$, it is shown that $\ell_n(\theta_0) \xrightarrow{d} \chi_p^2$ as $n \to \infty$, where θ_0 is the true value of the parameter vector θ (e.g., Qin and Lawless, 1994). Firstly, a natural substitute for this result is that $(2p)^{-1}(\ell_n(\theta_0) - p) \xrightarrow{d} N(0,1)$ as both n and p go to infinity. Chen et al. (2009) and Hjort et al. (2009) shown that this result holds true under the condition $p = o(n^{1/2})$ and $p = o(n^{1/3})$, respectively. Secondly, the penalized EL method is found based on the idea of reducing the dimension of the parameter vector. Tang and Leng (2010) investigated the context of the mean parameters by the penalized EL method for the first time, and Leng and Tang (2012) also studied the penalized empirical likelihood for general estimating equations with growing dimensionality, which extended the EL method to independent and high-dimensional but sparse cases. However, the extension of EL to the problem with diverging numbers of parameters, especially for dependent and non-sparse data, is still a challenging task. Up to now, there is no research on the EL method for spatial non-panel data models with the dimensions of the parameters diverge along with the sample size.

In this article, we focus on the case of the high-dimensional and non-sparse spatial cross-sectional model, and propose an innovate EL method, the principal component empirical likelihood (PCEL) method, which combines the idea of the principal component analysis and the usual EL method to reduce the dimensions of the moment restrictions. We prove the limit distribution of the PCEL statistics is chi-square distribution as well as the traditional methods, but the degree of freedom of the new distribution is given and independent of the data dimension, much less than the degree of freedom under the traditional method. Meanwhile, we conduct simulations to compare the performances of the usual EL-based and PCEL-based confidence regions, which shows that the PCEL is computationally faster than the EL method in practice, and that the coverage probabilities of the confidence regions based on the PCEL method are closer to the nominal level than the usual EL method, especially when the dimension of estimation equations is growing along with the sample size.

The article is organized as follows. Section 2 introduces the PCEL for estimation function models. Section 3 presents the the main results of PCEL for ASARSAR models. Section 4 presents results from a simulation study. Section 5 gives some lemmas and proofs.

2. The PCEL method and estimating equations

In this section, we introduce the concept of estimation equation, state a conclusion on the usual EL method for estimation equations, and present a result on the PCEL method for estimation equations.

2.1 The usual EL method for estimation equations

Suppose that $X \in R^d$ is a population and $X_1, X_2, \cdots, X_n$ are the i.i.d. observations of X. We further assume that there are r known functions $g_j(x, \theta), 1 \leq j \leq r$, such that

$$Eg_j(X, \theta) = 0, 1 \leq j \leq r, \tag{2}$$

where $\theta \in \Theta \subseteq R^p$.

Suppose that 0 is inside the convex hull of the $\{g(X_i, \theta), 1 \leq i \leq n\}$. Define the scoring function

$$g(x, \theta) = (g_1(x, \theta), g_2(x, \theta), \cdots, g_r(x, \theta))^\tau, x \in R^d, \theta \in \Theta,$$

and the EL statistic (e.g., Qin and Lawless, 1994):

$$\ell_E(\theta) = 2\sum_{i=1}^{n} \log\{1 + t^\tau(\theta)g(X_i, \theta)\},$$

where $t(\theta) \in R^r$ is the solution of the following equations:

$$\sum_{i=1}^{n} \frac{g(X_i, \theta)}{1 + t^\tau(\theta) g(X_i, \theta)} = 0.$$

Following Qin and Lawless (1994), one can obtain the following result which states the limiting distribution of $\ell_E(\theta)$.

Theorem 1. *Suppose that* $E||g(X, \theta_0)||^3 < \infty$ *and* $Cov(g(X, \theta_0))$ *is positive definite, where* $||a||$ *is the* L_2*-norm in* R^r *and* θ_0 *is the true value of* θ. *Then for fixed* p *and* r, *as* $n \to \infty$,

$$\ell_E(\theta_0) \xrightarrow{d} \chi_r^2,$$

where χ_r^2 *is a chi-squared distributed random variable with* r *degrees of freedom.*

2.2 The PCEL method for estimation equations

Large datasets are increasingly common and are often difficult to interpret. Principal component analysis is a technique for reducing the dimensionality of such datasets, increasing interpretability but at the same time minimizing information loss. We next apply the idea of principal component to the EL method.

Let $A(\theta_0) = E\{g(X, \theta_0) g^\tau(X, \theta_0)\} > 0$ and use $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r$ to denote the ordered eigenvalues of $A(\theta_0)$. Further, assume that ξ_i is the eigenvector associated with $\lambda_i, 1 \leq i \leq r$. Let $H(\theta_0) = (\xi_1, \xi_2, \cdots, \xi_s)^\tau$ is a $s \times r$ matrix and $\tilde{g}(x, \theta) = H(\theta_0) g(x, \theta)$, and define the new estimation equations:

$$E\{\tilde{g}(X, \theta)\} = 0. \tag{3}$$

We then define the PCEL statistic based on (3):

$$\tilde{\ell}_E^s(\theta) = 2 \sum_{i=1}^{n} \log\{1 + \tilde{t}^\tau(\theta) \tilde{g}(X_i, \theta)\},$$

where $\tilde{t}(\theta) \in R^s$ is the solution of the following equations:

$$\sum_{i=1}^{n} \frac{\tilde{g}(X_i, \theta)}{1 + \tilde{t}^\tau(\theta) \tilde{g}(X_i, \theta)} = 0.$$

By the Theorem 1 in Qin and Lawless (1994), asymptotic covariance of the EL estimator (denoted as $\hat{\theta}_{EL}$) of θ under (2) is

$$S = \left[E\left(\frac{\partial g(X, \theta_0)}{\partial \theta} \right)^\tau (E g(X, \theta_0) g^\tau(X, \theta_0))^{-1} E\left(\frac{\partial g(X, \theta_0)}{\partial \theta} \right) \right]^{-1}$$

and the total asymptotic variances of $\hat{\theta}_{EL}$ is the trace of S. It can be shown that the asymptotic covariance of the PCEL estimator $\hat{\theta}_{PCEL}$ of θ under (3) is equal to S if $s = r$, and larger than S if $s < r$. To reduce the dimensionality of the estimation equations, we need s to be small, and to minimize the information loss, s should not be too small. In general, one may choose s to satisfy

$$\min_s \frac{\sum_{i=1}^{s} \lambda_i}{\sum_{i=1}^{r} \lambda_i} \geq 75\%, s = 1, 2, \ldots, r.$$

Based on Theorem 1, one can show the limiting distribution of $\tilde{\ell}_E(\theta)$ presented in Theorem 2.

THEOREM 2. *Suppose that the conditions of Theorem 1 hold true. Then for fixed* p *and* r*, as* $n \to \infty$*,*

$$\tilde{\ell}_E^s(\theta_0) \xrightarrow{d} \chi_s^2,$$

where χ_s^2 *is a chi-squared distributed random variable with* s *degrees of freedom.*

REMARK 1. *In practice,* θ_0 *is unknown. Then we may use the plug-in method to replace* θ_0 *by* $\hat{\theta}_{EL}$ *in the expression of* $H(\theta_0)$ *and the result of Theorem 2 still holds true.*

3. The PCEL method for the SARSAR model

Let $A_n(\rho_1) = I_n - \rho_1 W_n$, $B_n(\rho_2) = I_n - \rho_2 M_n$ and suppose that $A_n(\rho_1)$ and $B_n(\rho_2)$ are nonsingular. Then (1) can be written as

$$Y_n = A_n^{-1}(\rho_1) X_n \beta + A_n^{-1}(\rho_1) B_n^{-1}(\rho_2) \epsilon_{(n)}.$$

At this moment, suppose that $\epsilon_{(n)}$ is normally distributed, which is firstly used to derive the usual EL statistic only and not employed in our main results. Then the log-likelihood function based on the response vector Y_n is

$$L = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 + \log|A_n(\rho_1)| + \log|B_n(\rho_2)| - \frac{1}{2\sigma^2}\epsilon_{(n)}^{\tau}\epsilon_{(n)},$$

where $\epsilon_{(n)} = B_n(\rho_2)\{A_n(\rho_1)Y_n - X_n\beta\}$. Let $G_n = B_n(\rho_2)W_n A_n^{-1}(\rho_1)B_n^{-1}(\rho_2), H_n = M_n B_n^{-1}(\rho_2), \tilde{G}_n = \frac{1}{2}(G_n + G_n^{\tau})$ and $\tilde{H}_n = \frac{1}{2}(H_n + H_n^{\tau})$. It can be shown that (e.g. Anselin, 1988, pp. 74-75)

$$\partial L/\partial\beta = \frac{1}{\sigma^2} X_n^{\tau} B_n^{\tau}(\rho_2)\epsilon_{(n)},$$

$$\partial L/\partial\rho_1 = \frac{1}{\sigma^2}\{B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n\beta\}^{\tau}\epsilon_{(n)} + \frac{1}{\sigma^2}\{\epsilon_{(n)}^{\tau}\tilde{G}_n\epsilon_{(n)} - \sigma^2 tr(\tilde{G}_n)\},$$

$$\partial L/\partial\rho_2 = \frac{1}{\sigma^2}\{\epsilon_{(n)}^{\tau}\tilde{H}_n\epsilon_{(n)} - \sigma^2 tr(\tilde{H}_n)\},$$

$$\partial L/\partial\sigma^2 = \frac{1}{2\sigma^4}\{\epsilon_{(n)}^{\tau}\epsilon_{(n)} - n\sigma^2\}.$$

Letting above derivatives be 0, we obtain the following estimating equations:

$$X_n^\tau B_n^\tau(\rho_2)\epsilon_{(n)} = 0, \tag{4}$$

$$\{B_n(\rho_2)W_nA_n^{-1}(\rho_1)X_n\beta\}^\tau\epsilon_{(n)} + \{\epsilon_{(n)}^\tau\tilde{G}_n\epsilon_{(n)} - \sigma^2 tr(\tilde{G}_n)\} = 0, \tag{5}$$

$$\epsilon_{(n)}^\tau\tilde{H}_n\epsilon_{(n)} - \sigma^2 tr(\tilde{H}_n) = 0, \tag{6}$$

$$\epsilon_{(n)}^\tau\epsilon_{(n)} - n\sigma^2 = 0. \tag{7}$$

Let $\tilde{g}_{ij}$, $\tilde{h}_{ij}$, b_i and s_i denote the (i,j) element of the matrix $\tilde{G}_n$, the (i,j) element of the matrix $\tilde{H}_n$, the i-th column of the matrix $X_n^\tau B_n^\tau(\rho_2)$ and i-th component of the vector $B_n(\rho_2)W_nA_n^{-1}(\rho_1)X_n\beta$, respectively, and adapt the convention that any sum with an upper index of less than one is zero. To deal with the quadratic forms in (4) and (5), we follow Kelejian and Prucha (2001) to introduce a martingale difference array. Define the σ-fields: $\mathcal{F}_0 = \{\emptyset, \Omega\}, \mathcal{F}_i = \sigma(\epsilon_1, \epsilon_2, \cdots, \epsilon_i), 1 \le i \le n$. Let

$$\tilde{Y}_{in} = \tilde{g}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i\sum_{j=1}^{i-1}\tilde{g}_{ij}\epsilon_j, \quad \tilde{Z}_{in} = \tilde{h}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i\sum_{j=1}^{i-1}\tilde{h}_{ij}\epsilon_j. \tag{8}$$

Then $\mathcal{F}_{i-1} \subseteq \mathcal{F}_i$, $\tilde{Y}_{in}$ is $\mathcal{F}_i$-measurable and $E(\tilde{Y}_{in}|\mathcal{F}_{i-1}) = 0$. Thus $\{\tilde{Y}_{in}, \mathcal{F}_i, 1 \le i \le n\}$ and $\{\tilde{Z}_{in}, \mathcal{F}_i, 1 \le i \le n\}$ form two martingale difference arrays and

$$\epsilon_{(n)}^\tau\tilde{G}_n\epsilon_{(n)} - \sigma^2 tr(\tilde{G}_n) = \sum_{i=1}^n\tilde{Y}_{in}, \quad \epsilon_{(n)}^\tau\tilde{H}_n\epsilon_{(n)} - \sigma^2 tr(\tilde{H}_n) = \sum_{i=1}^n\tilde{Z}_{in}. \tag{9}$$

Based on (4) to (9), we can get the score function:

$$\omega_i(\theta) = \begin{pmatrix} b_i\epsilon_i \\ \tilde{g}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i\sum_{j=1}^{i-1}\tilde{g}_{ij}\epsilon_j + s_i\epsilon_i \\ \tilde{h}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i\sum_{j=1}^{i-1}\tilde{h}_{ij}\epsilon_j \\ \epsilon_i^2 - \sigma^2 \end{pmatrix}_{(p+3)\times 1},$$

where ϵ_i is the i-th component of $\epsilon_{(n)} = B_n(\rho_2)\{A_n(\rho_1)Y_n - X_n\beta\}$.

Following Qin (2021), the following EL ration statistic for $\theta = (\beta^\tau, \rho_1, \rho_2, \sigma^2)^\tau \in R^{p+3}$ is defined as:

$$L_n(\theta) = \sup\left\{\prod_{i=1}^n(n\hat{p}_i) : \hat{p}_i \ge 0, \sum_{i=1}^n\hat{p}_i = 1, \sum_{i=1}^n\hat{p}_i\omega_i(\theta) = 0\right\}.$$

It can be shown that the profile empirical log-likelihood function of θ is defined as

$$\ell_n(\theta) \hat{=} -2\log L_n(\theta) = 2\sum_{i=1}^n\log\{1 + \hat{\lambda}^\tau(\theta)\omega_i(\theta)\}, \tag{10}$$

where $\hat{\lambda}(\theta) \in R^{p+3}$ is the solution of the following equation:

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\omega_i(\theta)}{1+\hat{\lambda}^{\tau}(\theta)\omega_i(\theta)} = 0. \tag{11}$$

In this article, we use the following PCEL ratio statistic for θ:

$$\tilde{L}_n^s(\theta) = \sup\left\{\prod_{i=1}^{n}(np_i) : p_i \geq 0, \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i\tilde{\omega}_i(\theta) = 0\right\},$$

where $\tilde{\omega}_i(\theta) = \Gamma_s\omega_i(\theta)$, $\Gamma_s = (\xi_1, \xi_2, \cdots, \xi_s)^{\tau}$, (λ_i, ξ_i) represents i-th eigenvalue and associated eigenvector of Σ_{p+3} in (14) with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{p+3}$ and $\|\xi_j\| = 1$, and $s \leq p+3$. Similar to Owen's framework, one can show that

$$\tilde{\ell}_n^s(\theta) \hat{=} -2\log\tilde{L}_n^s(\theta) = 2\sum_{i=1}^{n}\log\{1+\lambda^{\tau}(\theta)\tilde{\omega}_i(\theta)\}, \tag{12}$$

where $\lambda^{\tau}(\theta) \in R^s$ is the solution of the following equation:

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\tilde{\omega}_i(\theta)}{1+\lambda^{\tau}(\theta)\tilde{\omega}_i(\theta)} = 0. \tag{13}$$

Let $\mu_j = E(\epsilon_1^j), j = 3, 4$, use $Vec(diagA)$ to denote the vector formed by the diagonal elements of a matrix A and use $\|a\|$ to denote the L_2-norm of a vector a. Furthermore, Let $\mathbf{1}_n$ present the n-dimensional (column) vector with 1 as its components. To obtain the asymptotical distribution of $\tilde{\ell}_n^s(\theta)$, we need following assumptions.

A1. $\{\epsilon_i, 1 \leq i \leq n\}$ are independent and identically distributed random variables with mean 0, variance σ^2 and $E|\epsilon_1|^{4+\eta_1} < \infty$ for some $\eta_1 > 0$.

A2. $W_n, M_n, A_n^{-1}(\rho_1), B_n^{-1}(\rho_2)$ and $\{x_i\}$ be as described above. They satisfy the following conditions:

(i) The row and column sums of $W_n, M_n, A_n^{-1}(\rho_1)$ and $B_n^{-1}(\rho_2)$ are uniformly bounded in absolute value,

(ii) $\{x_i\}, i = 1, 2, ..., n$ are uniformly bounded.

A3. There is a constants $c_j > 0$, $j = 1, 2$, such that $0 < c_1 \leq \lambda_{min}(n^{-1}\Sigma_{p+3}) \leq \lambda_{max}(n^{-1}\Sigma_{p+3}) \leq c_2 < \infty$, where $\lambda_{min}(A)$ and $\lambda_{max}(A)$ denote the minimum and maximum eigenvalues of a matrix A, respectively,

$$\Sigma_{p+3} = \Sigma_{p+3}^{\tau} = Cov\left\{\sum_{i=1}^{n}\omega_i(\theta)\right\} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \Sigma_{14} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} & \Sigma_{24} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} & \Sigma_{34} \\ \Sigma_{41} & \Sigma_{42} & \Sigma_{43} & \Sigma_{44} \end{pmatrix}, \tag{14}$$

where

$$
\begin{aligned}
\Sigma_{11} &= \sigma^2\{B_n(\rho_2)X_n\}'B_n(\rho_2)X_n, \\
\Sigma_{12} &= \sigma^2\{B_n(\rho_2)X_n\}'B_n(\rho_2)W_nA_n^{-1}(\rho_1)X_n\beta + \mu_3\{B_n(\rho_2)X_n\}'Vec(diag\tilde{G}_n), \\
\Sigma_{13} &= \mu_3\{B_n(\rho_2)X_n\}'Vec(diag\tilde{H}_n), \quad \Sigma_{14} = \mu_3\{B_n(\rho_2)X_n\}'\mathbf{1}_n, \\
\Sigma_{22} &= 2\sigma^4tr(\tilde{G}_n^{\ 2}) + \sigma^2\{B_n(\rho_2)W_nA_n^{-1}(\rho_1)X_n\beta\}'B_n(\rho_2)W_nA_n^{-1}(\rho_1)X_n\beta \\
&\quad +(\mu_4-3\sigma^4)\|Vec(diag\tilde{G}_n)\|^2 + 2\mu_3\{B_n(\rho_2)W_nA_n^{-1}(\rho_1)X_n\beta\ \}'Vec(diag\tilde{G}_n), \\
\Sigma_{23} &= 2\sigma^4tr(\tilde{G}_n\tilde{H}_n) + (\mu_4-3\sigma^4)Vec'(diag\tilde{G}_n)Vec(diag\tilde{H}_n) \\
&\quad +\mu_3\{B_n(\rho_2)W_nA_n^{-1}(\rho_1)X_n\beta\ \}'Vec(diag\tilde{H}_n), \\
\Sigma_{24} &= (\mu_4-\sigma^4)tr(\tilde{G}_n) + \mu_3\{B_n(\rho_2)W_nA_n^{-1}(\rho_1)X_n\beta\ \}'\mathbf{1}_n, \\
\Sigma_{33} &= 2\sigma^4tr(\tilde{H}_n^{\ 2})(\mu_4-3\sigma^4)\|Vec(diag\tilde{H}_n)\|^2, \\
\Sigma_{34} &= (\mu_4-\sigma^4)tr(\tilde{H}_n), \quad \Sigma_{44} = n(\mu_4-\sigma^4).
\end{aligned}
$$

We now state the main results.

THEOREM 3. *Suppose that conditions A1-A3 hold. Let θ_0 be the true value of θ. As $n \to \infty$,*

$$\tilde{\ell}_n^s(\theta_0) \xrightarrow{d} \chi_s^2,$$

where χ_s^2 is a chi-squared distributed random variable with s degree of freedom.

Let $z_s(\alpha)$ satisfy $P(\chi_s^2 > z_s(\alpha)) = \alpha$ for $0 < \alpha < 1$. It follows from Theorem 3 that an PCEL-based confidence region for θ with asymptotically correct coverage probability $1-\alpha$ can be constructed as

$$\{\theta : \tilde{\ell}_n^s(\theta_0) \leq z_s(\alpha)\}. \tag{15}$$

4. Simulations

Let I_p denote a $p \times p$ identity matrix, and PCELs denote the PCEL method that reduces the initial $p+3$ dimension to s. According to Qin (2021), the usual empirical log-likelihood ratio $\ell_n(\theta_0)$ in (10) is asymptotically distributed as χ_{p+3}^2 under the null hypothesis: $\theta = \theta_0$. It follows that the usual EL-based confidence region for θ with asymptotically correct coverage probability $1-\alpha$ can be constructed as

$$\{\theta : \ell_n(\theta) \leq z_{p+3}(\alpha)\}, \tag{16}$$

where $z_{p+3}(\alpha)$ satisfies $P(\chi_{p+3}^2 > z_{p+3}(\alpha)) = \alpha$ for $0 < \alpha < 1$.

We conducted a small simulation study to compare the finite sample performances of the confidence regions based on EL and PCEL methods with confidence level $\alpha = 0.05$, and report the proportion of $\ell_n(\theta_0) \leq z_{p+3}(0.05)$ and $\tilde{\ell}_n^s(\theta_0) \leq z_s(0.05)$, $s = 1, 2, 3, 4$ respectively in our 2,000 simulations where θ_0 is the true value of θ. The results of simulations are reported in Tables 1-5.

In the simulations, we used the model: $Y_n = \rho_1 W_n Y_n + X_n \beta_0 + u_{(n)}, u_{(n)} = \rho_2 M_n u_{(n)} + \epsilon_{(n)}$, with $(\rho_1, \rho_2) = (0.85, 0.15)$, $\beta_0 = \mathbf{1}_p$, and $\{X_i\} \sim N(\mathbf{0}, I_p)$. To make p and n increase simultaneously, we considered one growth rates for p with respect to n: $p = [3n^{index}]$ and $[x]$ is the integral function for x. For the contiguity weight matrix $W_n = (w_{ij})$, we took $w_{ij} = 1$ if spatial units i and j are neighbours by queen contiguity rule (namely, they share common border or vertex), $w_{ij} = 0$ otherwise (Anselin, 1988, p.18). We considered five ideal cases of spatial units: $n = m \times m$, $m = 10, 15, 20, 25, 30$ and $M_n = W_n$.

Experiment 1

Experiment 1 is designed to compare the finite sample performances of the confidence regions between usual EL and PCEL method in low-dimensional case. We chose $index = 0, 0.1, 0.2, 0.3, 0.4$ and 0.5. The errors $\epsilon_i' s$ were taken from $N(0, 1)$, $N(0, 0.75)$, $t(5)$ and $\chi_4^2 - 4$, respectively, which implied $\sigma_0^2 = 1, 0.75, 2.5, 8$, $\mu_3 = 0, 0, 0, 32$, $\mu_4 = 3, \frac{9}{4}, 25, 384$, respectively. The results of simulations are reported in Tables 1-4.

Experiment 2

Experiment 2 only presents the performances of confidence regions based on PCEL tests in high-dimensional case, because the solution $\hat{\lambda}$ in (11) does not exist in the case of $index > 0.5$, while λ in (13) has numerical solution. We chose $index = 0.6, 0.7, 0.8, 0.9, 1, 1.1$ and the errors $\epsilon_i' s$ were taken from $N(0, 1)$. The results of simulations are reported in Table 5.

Experimental Analysis

Tables 1-4 show that the confidence regions based on the usual EL behave well with coverage probabilities very close to the nominal level 0.95 as n increase, but not well as p is growing. Especially when $index \geq 0.3$, the coverage probabilities of the confidence regions based on the usual EL fall to the range [0.022, 0.905] for $N(0, 1)$ distribution , [0.020, 0.909] for $N(0, 0.75)$, [0.011, 0.840] for $t(5)$ and [0.010, 0.849] for χ^2 distribution, which are far from the nominal level 0.95.

Tables 1-4 also show that the confidence regions based on PCEL behave well with coverage probabilities very close to the nominal level 0.95, as n increase, whether p is increasing. Although $index \geq 0.3$, the coverage probabilities of the confidence regions based on PCEL1 fall to the range [0.939, 0.959] for $N(0, 1)$ distribution, [0.942, 0.956] for $N(0, 0.75)$, [0.921, 0.953] for $t(5)$ and [0.913, 0.951] for χ^2 distribution, which are still close to the nominal level 0.95.

Table 5 shows that, when $index > 0.5$, the coverage probabilities of the confidence regions based on PCEL1 fall to the range [0.936, 0.957] for $N(0, 1)$ distribution, PCEL2 to [0.924, 0.961], PCEL3 to [0.907, 0.958] and PCEL4 to [0.905, 0.955], which are still close to the nominal level 0.95, even $p > n$, so the PCEL method is effective in some applications.

We can see, from Tables 1-5, the confidence regions based on PCEL method converge to the nominal level 0.95 as the number of spatial units n is large enough, whether the data dimension is low or not. Besides, the calculation speed under the PCEL method is much faster than that under the usual EL method, which has higher application value. One can download R codes related to this article at https://github.com/Tang-Jay/PCEL. Our simulation results recommend PCEL method when the dimensions of the moment restrictions diverge along with the sample size.

5. Proof

To prove the main results, we need following lemmas.

LEMMA 1. *Suppose that Assumptions A1-A3 are satisfied, then as $n \to \infty$,*

$$Z_n = \max_{1 \le i \le n} \|\omega_i(\theta)\| = o_p(n^{1/2}), \tag{17}$$

$$\Sigma_{p+3}^{-1/2} \sum_{i=1}^{n} \omega_i(\theta) \xrightarrow{d} N(0, I_{p+3}), \tag{18}$$

$$n^{-1} \sum_{i=1}^{n} \omega_i(\theta) \omega_i^{\tau}(\theta) = n^{-1} \Sigma_{p+3} + o_p(1), \tag{19}$$

$$\sum_{i=1}^{n} \|\omega_i\|^3 = O_p(n), \tag{20}$$

where Σ_{p+3} is given in (14).

Proof. See Lemma 3 in Qin (2021).

LEMMA 2. *Suppose that Assumptions A1-A3 are satisfied, then as $n \to \infty$,*

$$\tilde{Z}_n = \max_{1 \le i \le n} \|\tilde{\omega}_i(\theta)\| = o_p(n^{1/2}), \tag{21}$$

$$\Lambda_s^{-1/2} \sum_{i=1}^{n} \tilde{\omega}_i(\theta) \xrightarrow{d} N(0, I_s), \tag{22}$$

$$n^{-1} \sum_{i=1}^{n} \tilde{\omega}_i(\theta) \tilde{\omega}_i^{\tau}(\theta) = n^{-1} \Lambda_s + o_p(1), \tag{23}$$

$$\sum_{i=1}^{n} \|\tilde{\omega}_i(\theta)\|^3 = O_p(n), \tag{24}$$

where $\Lambda_s = diag(\lambda_1, \lambda_2, \cdots, \lambda_s)$.

Table 1: Coverage probabilities of the EL and PCELs (s = 1, 2, 3, 4) confidence regions with $\epsilon_i \sim N(0,1)$

n	$index$	p	EL	PCEL1	PCEL2	PCEL3	PCEL4
100	0	3	0.861	0.943	0.930	0.929	0.908
	0.1	5	0.824	0.943	0.935	0.925	0.910
	0.2	8	0.740	0.933	0.930	0.920	0.901
	0.3	12	0.548	0.939	0.936	0.931	0.913
	0.4	19	0.250	0.943	0.931	0.914	0.909
	0.5	30	0.022	0.947	0.938	0.923	0.910
225	0	3	0.920	0.948	0.943	0.940	0.935
	0.1	5	0.910	0.954	0.948	0.945	0.945
	0.2	9	0.863	0.946	0.944	0.936	0.919
	0.3	15	0.774	0.949	0.949	0.940	0.932
	0.4	26	0.508	0.947	0.936	0.933	0.930
	0.5	45	0.096	0.946	0.943	0.941	0.933
400	0	3	0.931	0.950	0.947	0.943	0.936
	0.1	5	0.929	0.941	0.940	0.936	0.929
	0.2	10	0.905	0.943	0.946	0.944	0.948
	0.3	18	0.856	0.950	0.946	0.943	0.935
	0.4	33	0.655	0.941	0.939	0.943	0.939
	0.5	60	0.163	0.956	0.952	0.944	0.944
625	0	3	0.948	0.948	0.947	0.949	0.948
	0.1	6	0.935	0.945	0.944	0.947	0.950
	0.2	11	0.931	0.943	0.941	0.948	0.945
	0.3	21	0.889	0.948	0.951	0.946	0.946
	0.4	39	0.739	0.954	0.948	0.951	0.948
	0.5	75	0.262	0.947	0.950	0.946	0.945
900	0	3	0.951	0.954	0.949	0.949	0.944
	0.1	6	0.942	0.953	0.939	0.950	0.946
	0.2	12	0.928	0.947	0.947	0.946	0.948
	0.3	23	0.905	0.947	0.950	0.943	0.942
	0.4	46	0.792	0.959	0.956	0.955	0.950
	0.5	90	0.353	0.956	0.943	0.943	0.948

Table 2: Coverage probabilities of the EL and PCELs (s = 1, 2, 3, 4) confidence regions with $\epsilon_i \sim N(0, 0.75)$

n	*index*	p	EL	PCEL1	PCEL2	PCEL3	PCEL4
100	0	3	0.872	0.946	0.936	0.923	0.909
	0.1	5	0.824	0.947	0.925	0.919	0.904
	0.2	8	0.718	0.941	0.929	0.922	0.911
	0.3	12	0.545	0.943	0.931	0.919	0.901
	0.4	19	0.266	0.949	0.939	0.927	0.909
	0.5	30	0.020	0.945	0.931	0.920	0.916
225	0	3	0.917	0.944	0.943	0.940	0.928
	0.1	5	0.907	0.953	0.943	0.933	0.931
	0.2	9	0.878	0.948	0.947	0.944	0.945
	0.3	15	0.783	0.950	0.946	0.939	0.931
	0.4	26	0.508	0.951	0.950	0.947	0.934
	0.5	45	0.084	0.942	0.941	0.930	0.936
400	0	3	0.929	0.950	0.944	0.945	0.946
	0.1	5	0.925	0.949	0.939	0.934	0.936
	0.2	10	0.911	0.944	0.940	0.945	0.943
	0.3	18	0.850	0.950	0.942	0.945	0.947
	0.4	33	0.655	0.951	0.950	0.942	0.935
	0.5	60	0.186	0.955	0.950	0.947	0.942
625	0	3	0.938	0.950	0.946	0.941	0.945
	0.1	6	0.924	0.958	0.950	0.951	0.946
	0.2	11	0.926	0.953	0.946	0.943	0.946
	0.3	21	0.890	0.945	0.945	0.939	0.937
	0.4	39	0.751	0.956	0.953	0.952	0.946
	0.5	75	0.267	0.943	0.940	0.944	0.936
900	0	3	0.935	0.946	0.939	0.935	0.940
	0.1	6	0.941	0.943	0.952	0.948	0.946
	0.2	12	0.936	0.942	0.940	0.939	0.940
	0.3	23	0.909	0.948	0.952	0.950	0.949
	0.4	46	0.798	0.947	0.945	0.947	0.942
	0.5	90	0.337	0.949	0.956	0.952	0.952

Table 3: Coverage probabilities of the EL and PCELs (s = 1, 2, 3, 4) confidence regions with $\epsilon_i \sim t(5)$

n	*index*	p	EL	PCEL1	PCEL2	PCEL3	PCEL4
100	0	3	0.778	0.904	0.879	0.860	0.842
	0.1	5	0.720	0.914	0.887	0.866	0.851
	0.2	8	0.591	0.923	0.885	0.866	0.839
	0.3	12	0.392	0.921	0.880	0.865	0.849
	0.4	19	0.132	0.926	0.884	0.861	0.826
	0.5	30	0.011	0.933	0.888	0.864	0.837
225	0	3	0.866	0.919	0.916	0.911	0.901
	0.1	5	0.857	0.913	0.898	0.897	0.889
	0.2	9	0.786	0.929	0.910	0.913	0.902
	0.3	15	0.614	0.945	0.913	0.902	0.896
	0.4	26	0.317	0.935	0.910	0.907	0.907
	0.5	45	0.025	0.938	0.909	0.902	0.898
400	0	3	0.895	0.926	0.928	0.920	0.921
	0.1	5	0.891	0.931	0.917	0.919	0.926
	0.2	10	0.856	0.934	0.914	0.923	0.922
	0.3	18	0.723	0.945	0.918	0.915	0.904
	0.4	33	0.459	0.941	0.920	0.918	0.907
	0.5	60	0.061	0.940	0.915	0.917	0.912
625	0	3	0.922	0.935	0.927	0.925	0.918
	0.1	6	0.906	0.929	0.928	0.931	0.924
	0.2	11	0.879	0.931	0.922	0.920	0.928
	0.3	21	0.811	0.947	0.929	0.922	0.917
	0.4	39	0.557	0.949	0.926	0.929	0.929
	0.5	75	0.085	0.940	0.911	0.920	0.921
900	0	3	0.921	0.932	0.928	0.925	0.924
	0.1	6	0.915	0.934	0.932	0.938	0.931
	0.2	12	0.906	0.944	0.938	0.934	0.931
	0.3	23	0.840	0.949	0.928	0.925	0.929
	0.4	46	0.624	0.952	0.936	0.935	0.934
	0.5	90	0.118	0.953	0.942	0.942	0.928

Table 4: Coverage probabilities of the EL and PCELs (s = 1, 2, 3, 4) confidence regions with $\epsilon_i \sim \chi^2(4) - 4$

n	*index*	p	EL	PCEL1	PCEL2	PCEL3	PCEL4
100	0	3	0.775	0.903	0.886	0.867	0.842
	0.1	5	0.730	0.918	0.889	0.869	0.855
	0.2	8	0.624	0.919	0.888	0.868	0.849
	0.3	12	0.452	0.913	0.878	0.850	0.833
	0.4	19	0.182	0.929	0.899	0.868	0.841
	0.5	30	0.010	0.923	0.883	0.865	0.844
225	0	3	0.877	0.938	0.922	0.904	0.903
	0.1	5	0.861	0.931	0.910	0.899	0.896
	0.2	9	0.788	0.937	0.914	0.896	0.883
	0.3	15	0.665	0.946	0.928	0.913	0.905
	0.4	26	0.384	0.932	0.919	0.904	0.889
	0.5	45	0.047	0.933	0.920	0.900	0.895
400	0	3	0.907	0.935	0.918	0.913	0.900
	0.1	5	0.897	0.945	0.933	0.929	0.927
	0.2	10	0.850	0.949	0.938	0.934	0.925
	0.3	18	0.740	0.949	0.925	0.916	0.909
	0.4	33	0.476	0.942	0.934	0.918	0.918
	0.5	60	0.088	0.946	0.942	0.928	0.928
625	0	3	0.935	0.951	0.938	0.929	0.928
	0.1	6	0.918	0.950	0.945	0.939	0.932
	0.2	11	0.885	0.946	0.929	0.928	0.918
	0.3	21	0.812	0.942	0.933	0.938	0.932
	0.4	39	0.590	0.950	0.940	0.926	0.929
	0.5	75	0.130	0.947	0.945	0.930	0.927
900	0	3	0.931	0.947	0.940	0.935	0.932
	0.1	6	0.922	0.951	0.945	0.943	0.935
	0.2	12	0.902	0.950	0.930	0.930	0.932
	0.3	23	0.849	0.946	0.942	0.942	0.945
	0.4	46	0.625	0.935	0.938	0.938	0.929
	0.5	90	0.172	0.947	0.946	0.940	0.933

Table 5: Coverage probabilities of PCELs (s = 1, 2, 3, 4) confidence regions with $\epsilon_i \sim N(0,1)$

n	*index*	p	PCEL1	PCEL2	PCEL3	PCEL4
100	0.6	48	0.936	0.925	0.907	0.905
	0.7	75	0.944	0.942	0.934	0.922
	0.8	119	0.944	0.937	0.918	0.898
	0.9	189	0.945	0.933	0.920	0.905
	1	300	0.941	0.925	0.925	0.906
	1.1	475	0.948	0.924	0.921	0.912
225	0.6	77	0.936	0.940	0.936	0.929
	0.7	133	0.948	0.946	0.944	0.941
	0.8	228	0.944	0.942	0.945	0.935
	0.9	393	0.939	0.938	0.935	0.930
	1	675	0.941	0.939	0.941	0.931
	1.1	1160	0.946	0.942	0.945	0.934
400	0.6	109	0.956	0.950	0.946	0.951
	0.7	199	0.946	0.954	0.947	0.950
	0.8	362	0.954	0.946	0.944	0.937
	0.9	659	0.945	0.961	0.952	0.950
	1	1200	0.953	0.951	0.955	0.955
	1.1	2185	0.951	0.953	0.946	0.944
625	0.6	143	0.947	0.952	0.947	0.947
	0.7	272	0.954	0.949	0.940	0.939
	0.8	517	0.945	0.948	0.938	0.943
	0.9	985	0.952	0.952	0.958	0.946
	1	1875	0.948	0.952	0.957	0.950
	1.1	3569	0.951	0.942	0.949	0.939
900	0.6	178	0.947	0.951	0.946	0.953
	0.7	351	0.957	0.953	0.954	0.950
	0.8	693	0.956	0.949	0.941	0.941
	0.9	1368	0.956	0.952	0.955	0.953
	1	2700	0.945	0.953	0.952	0.953
	1.1	5331	0.948	0.943	0.947	0.945

Proof. In the beginning, by (17), we get that

$$\tilde{Z}_n = \max_{1\le i\le n} \|\Gamma_s \omega_i(\theta)\| \le \sum_{j=1}^{s} \max_{1\le i\le n} |\xi_j^\tau \omega_i(\theta)| \le \sum_{j=1}^{s} \max_{1\le i\le n} \|\xi_j\|\|\omega_i(\theta)\| = sZ_n = o_p(n^{1/2}).$$

Thus (21) is proved. And then, by (18), we get that

$$E(\sum_{i=1}^{n} \tilde{\omega}_i) = 0, \quad Cov(\sum_{i=1}^{n} \tilde{\omega}_i) = \Gamma_s Cov(\sum_{i=1}^{n} \omega_i)\Gamma_s^\tau = \Gamma_s \Sigma_{p+3}\Gamma_s^\tau = \Lambda_s,$$

which implies

$$\sum_{i=1}^{n} \tilde{\omega}_i \xrightarrow{d} N(0, \Lambda_s).$$

The proof of (22) is complete. By (19), we get that

$$n^{-1}\sum_{i=1}^{n} \tilde{\omega}_i(\theta)\tilde{\omega}_i^\tau(\theta) = \Gamma_s \left(n^{-1}\sum_{i=1}^{n} \omega_i(\theta)\omega_i^\tau(\theta) \right) \Gamma_s^\tau = n^{-1}\Lambda_s + o_p(1),$$

which shows (23) is proved. Fianlly, we will prove (24). Note that

$$\sum_{i=1}^{n} E\|\tilde{\omega}_i(\theta)\|^3 \le \sum_{j=1}^{s}\sum_{i=1}^{n} E|\xi_j^\tau \omega_i|^3 \le \sum_{j=1}^{s}\sum_{i=1}^{n} E\|\xi_j\|^3\|\omega_i\|^3 \le C\sum_{i=1}^{n} E\|\omega_i\|^3,$$

by (20) and Markov inequality, we obtain $\sum_{i=1}^{n} \|\tilde{\omega}_i\|^3 = O_p(n)$.

Proof of Theorem 1 The proof is analogous to the proof of Theorem 1 in Qin (2021).

Acknowledgements This work was partially supported by the National Natural Science Foundation of China (12061017, 12161009), the Innovation Project of Guangxi Graduate Education (YJSCXP202104), and the Research Fund of Guangxi Key Lab of Multi-source Information Mining & Security (22-A-01-01).

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*. Springer Verlag, 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.

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.

Chen, S. X., Peng, L., Qin, Y. L. (2009). Effects of data dimension on empirical likelihood. *Biometrika, 96*, 711-722.

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.

Hjort, N. L., McKeague, I., Van Keilegom, I. (2009). Extending the scope of empirical likelihood. *Ann. Statist., 37*, 1079-1111.

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.

Leng, C. L., Tang, C. Y. (2012). Penalized empirical likelihood and growing dimensional general estimating equations. *Biometrika, 99*, 703-716.

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.

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.

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.

Shi, J. (2007). Empirical likelihood for high dimensional linear models (in Chinese). *J. Syst. Sci. Complex, 27*, 124-133.

Tang, C. Y., Leng, C. L. (2010). Penalized high dimensional empirical likelihood. *Biometrika, 97*, 905-920.

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.

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