Empirical likelihood for linear models with spatial errors

Yongsong Qin

Department of Statistics, Guangxi Normal University
Guilin, Guangxi 541004, China

Abstract. For linear models with spatial errors, the empirical likelihood ratio statistics are constructed for the parameters of the models. It is shown that the limiting distributions of the empirical likelihood ratio statistics are chi-squared distributions, which are used to construct confidence regions for the parameters of the models.

Keywords: linear model; spatial error; empirical likelihood; confidence region

AMS 2000 subject classification: Primary 62G05 secondary 62E20

Short title: Empirical likelihood for linear models with spatial errors

1Corresponding author. Tel: +86-773-5851556
Email address: ysqin@gxnu.edu.cn (Y. Qin)
1. Introduction

The linear regression models are the most important statistical models for explaining the relationship between response and explanatory variables. Whenever the variables in a linear regression model refer to attributes of a particular location (height of a plant, population of a country, position in a social network, etc.), one often allows for correlation among the errors (disturbances) by assuming that the errors follow a spatial autoregressive correlation (e.g. Dow et al., 1982; Ord, 1975; Krämer and Donninger, 1987). Then we have the following linear regression model with spatial autoregressive errors:

\[ Y_n = X_n \beta + u(n), \quad u(n) = \rho W_n u(n) + \epsilon(n), \quad (1) \]

where \( n \) is the number of spatial units, \( \beta \) is the \( k \times 1 \) vector of regression parameters, \( X_n = (x_1, x_2, \cdots, x_n)^T \) is the non-random \( n \times k \) matrix of observations on the independent variable, \( Y_n = (y_1, y_2, \cdots, y_n)^T \) is an \( n \times 1 \) vector of observations on the dependent variable, \( u(n) \) is an \( n \times 1 \) vector of errors (disturbances), \( \rho \) is the scalar autoregressive parameter with \( |\rho| < 1 \), \( W_n \) is an \( n \times n \) spatial weighting matrix of constants, \( \epsilon(n) \) is an \( n \times 1 \) vector of innovations which satisfies

\[ E\epsilon(n) = 0, \quad Var(\epsilon(n)) = \sigma^2 I_n. \]

Model (1) is also called spatial error model (SEM). The development in testing and estimation of SEM models has been summarized in Anselin (1988), Cliff and Ord (1973), Ord (1975), Krämer and Donninger (1987) and Helejian and Prucha (1999), among others.

There are two competing estimation approaches for the corresponding parameters. One is the maximum likelihood (ML) method (e.g. Anselin, 1988). The other is the computationally more efficient method, the generalized method of moment (GMM) approach by Kelejian and Prucha (1999). The asymptotic properties of the maximum likelihood estimator (MLE) and the GMM estimator for the SEM model are investigated by Anselin
(1988) and Kelejian and Prucha (1999), respectively. However, it may not be easy to use these normal approximation results to construct confidence region for the parameters in the SEM 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. In this article, we propose to use the empirical likelihood (EL) method introduced by Owen (1988, 1990) to construct confidence region for the parameters in the SEM model. The shape and orientation of the EL confidence region are determined by data and the confidence region is obtained without covariance estimation. These features of the EL confidence region are the major motivations for our current proposal. Owen (1991) has used the EL method to construct confidence regions for the vector of regression parameters in a linear model with independent errors. A comprehensive review on EL for regressions can be found in Chen and Keilegom (2009). More references on EL methods can be found in Owen (2001), Qin and Lawless (1994), Chen and Qin (1993), Zhong and Rao (2000) and Wu (2004), among others.

The idea in using the EL method for the SEM is to introduce a martingale sequence to transform the linear-quadratic form of the estimating equations (e.g. (2)-(4)) for the SEM into a linear form. It is interesting to note that the estimation equations for other spatial models may have the linear-quadratic forms. Therefore this approach of transformation also opens a way to use EL methods to more general spatial models.

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

2. Main Results

We continue with model (1). Let $A_n(\rho) = I_n - \rho W_n$ and suppose that $A_n(\rho)$ is nonsingular. Then

$$Y_n = X_n \beta + A_n^{-1}(\rho) \varepsilon_n.$$
At this moment, suppose that $\epsilon(n)$ is normally distributed, which is used to derive the EL statistic only and not employed in our main results. Then the log-likelihood function based on the response vector $Y_n$ is

$$L = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 + \log |A_n(\rho)| - \frac{1}{2\sigma^2} \epsilon^\tau(n) \epsilon(n),$$

where $\epsilon(n) = A_n(\rho)(Y_n - X_n\beta)$. Let $G_n = W_nA_n^{-1}(\rho)$ and $\tilde{G}_n = \frac{1}{2}(G_n + G_n^\tau)$. It can be shown that (e.g. Anselin, 1988, pp. 74-75)

$$\frac{\partial L}{\partial \beta} = \frac{1}{\sigma^2} X^\tau_n A_n(\rho) \epsilon(n),$$

$$\frac{\partial L}{\partial \rho} = \frac{1}{\sigma^2} \{\epsilon^\tau(n)W_nA_n^{-1}(\rho)\epsilon(n) - \sigma^2 \text{tr}(W_nA_n^{-1}(\rho))\} = \frac{1}{\sigma^2} \{\epsilon^\tau(n)\tilde{G}_n\epsilon(n) - \sigma^2 \text{tr}(\tilde{G}_n)\},$$

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

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

$$X^\tau_n A_n(\rho) \epsilon(n) = 0, \quad (2)$$

$$\epsilon^\tau(n)\tilde{G}_n\epsilon(n) - \sigma^2 \text{tr}(\tilde{G}_n) = 0, \quad (3)$$

$$\epsilon^\tau(n)\epsilon(n) - n\sigma^2 = 0. \quad (4)$$

We use $g_{ij}$, $\tilde{g}_{ij}$ and $b_i$ to denote the $(i, j)$ element of the matrix $G_n$, the $(i, j)$ element of the matrix $\tilde{G}_n$ and the $i$-th column of the matrix $X^\tau_n A_n(\rho)$, 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 (3), we follow Kelejian and Prucha (2001) to introduce a martingale difference array. Define the $\sigma$-fields: $\mathcal{F}_0 = \{\emptyset, \Omega\}, \mathcal{F}_i = \sigma(\epsilon_1, \epsilon_2, \ldots, \epsilon_i), 1 \leq i \leq n$. Let

$$\tilde{Y}_{in} = \tilde{g}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{g}_{ij} \epsilon_j. \quad (5)$$
Then $\mathcal{F}_{i-1} \subseteq \mathcal{F}_i, \tilde{Y}_m$ is $\mathcal{F}_i$-measurable and $E(\tilde{Y}_m|\mathcal{F}_{i-1}) = 0$. Thus $\{\tilde{Y}_m, \mathcal{F}_i, 1 \leq i \leq n\}$ form a martingale difference array and

$$
\epsilon_{(n)} \hat{G}_n \epsilon_{(n)} - \sigma^2 tr(\hat{G}_n) = \sum_{i=1}^{n} \tilde{Y}_m. \tag{6}
$$

Based on (2) to (6), we propose the following EL ratio statistic for $\theta = (\beta^\tau, \rho, \sigma^2)^\tau \in R^{k+2}$:

$$
L_n(\theta) = \sup_{p_i, 1 \leq i \leq n} \prod_{i=1}^{n} (n p_i),
$$

where $\{p_i\}$ satisfy

$$
p_i \geq 0, 1 \leq i \leq n, \sum_{i=1}^{n} p_i = 1,
$$

$$
\sum_{i=1}^{n} p_i b_i \epsilon_i = 0,
$$

$$
\sum_{i=1}^{n} p_i \{\tilde{g}_{ii}(\epsilon_i^2 - \sigma^2) + 2 \epsilon_i \sum_{j=1}^{i-1} \tilde{g}_{ij} \epsilon_j\} = 0,
$$

$$
\sum_{i=1}^{n} p_i (\epsilon_i^2 - \sigma^2) = 0.
$$

Let

$$
\omega_i(\theta) = \left(\begin{array}{c}
\frac{b_i \epsilon_i}{\tilde{g}_{ii}(\epsilon_i^2 - \sigma^2) + 2 \epsilon_i \sum_{j=1}^{i-1} \tilde{g}_{ij} \epsilon_j} \\
\epsilon_i^2 - \sigma^2
\end{array}\right)_{(k+2) \times 1},
$$

where $\epsilon_i$ is the $i$-th component of $\epsilon_{(n)} = A_n(\rho)(Y_n - X_n\beta)$. Following Owen (1990), one can show that

$$
\ell_n(\theta) \triangleq -2 \log L_n(\theta) = 2 \sum_{i=1}^{n} \log \{1 + \lambda^\tau(\theta) \omega_i(\theta)\}, \tag{7}
$$

where $\lambda(\theta) \in R^{k+2}$ is the solution of the following equation:

$$
\frac{1}{n} \sum_{i=1}^{n} \frac{\omega_i(\theta)}{1 + \lambda^\tau(\theta) \omega_i(\theta)} = 0. \tag{8}
$$
Let $\mu_j = E\varepsilon^j, j = 3, 4$. Use $Vec(diagA)$ to denote the vector formed by the diagonal elements of a matrix $A$ and $||a||$ to denote the $L_2$-norm of a vector $a$. Furthermore, Let $1_n$ present the $n$-dimensional (column) vector with 1 as its components. To obtain the asymptotical distribution of $\ell_n(\theta)$, we need following assumptions.

A1. \{\varepsilon_i, 1 \leq i \leq n\} are independent and identically distributed random variables with mean 0, variance $\sigma^2 > 0$ and $E|\varepsilon_1|^{4 + \eta_1} < \infty$ for some $\eta_1 > 0$.

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

(i) The row and column sums of $W_n$ and $A_n^{-1}(\rho)$ are uniformly bounded in absolute value;  
(ii) $\{x_i\}$ are uniformly bounded.

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

$$\Sigma_{k+2} = \Sigma_{k+2}^r = Cov \left\{ \sum_{i=1}^{n} \omega_i(\theta) \right\} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{pmatrix}, \quad (9)$$

$$\Sigma_{11} = \sigma^2 X_n^r A_n(\rho) A_n^r(\rho) X_n, \Sigma_{12} = \mu_3 X_n^r A_n(\rho) Vec(diag\hat{G}_n),$$

$$\Sigma_{13} = \mu_3 X_n^r A_n(\rho) 1_n, \Sigma_{22} = 2\sigma^4 tr(\hat{G}_n^2) + (\mu_4 - 3\sigma^4)||Vec(diag\hat{G}_n)||^2,$$

$$\Sigma_{23} = (\mu_4 - \sigma^4) tr(\hat{G}_n), \Sigma_{33} = n(\mu_4 - \sigma^4).$$

**Remark 1.** Conditions A1 to A3 are common assumptions for SAR models. For example, A1 and A2 are used in Assumptions 1, 4, 5 and 6 in Lee (2004), the analog of $0 < c_1 \leq \lambda_{\min}(n^{-1}\Sigma_{k+2})$ (e.g. $n^{-1}\sigma^2_Q \geq c$ for some constant $c > 0$ in Lemma 1 in this article) is employed in the assumption of Theorem 1 in Kelejian and Prucha (2001). From Conditions A1 and A2, one can see that $\lambda_{\max}(n^{-1}\Sigma_{k+2}) \leq c_2 < \infty$. For the sake of argument, we list this consequence of A1 and A2 as a condition here.

We now state the main results.
Theorem 1 Suppose that Assumptions (A1) to (A3) are satisfied. Then under model (1), as \( n \to \infty \),

\[
\ell_n(\theta) \xrightarrow{d} \chi^2_{k+2},
\]

where \( \chi^2_{k+2} \) is a chi-squared distributed random variable with \( k+2 \) degrees of freedom.

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

\[
\{ \theta : \ell_n(\theta) \leq z_\alpha(k+2) \}.
\]

3. Simulations

According to Anselin (1988), when the error term \( \epsilon_{(n)} \) is normal distributed, the likelihood ratio (LR) \( LR(\theta_0) = 2(L(\hat{\theta}) - L(\theta_0)) \) is asymptotically distributed as \( \chi^2_{k+2} \) under the null hypothesis: \( \theta = \theta_0 \), where \( L \) is the corresponding log-likelihood and \( \hat{\theta} \) is the maximum likelihood estimator. It follows that the LR based confidence region for \( \theta \) with asymptotically correct coverage probability \( \alpha \) can be constructed as

\[
\{ \theta : LR(\theta) \leq z_\alpha(k+2) \}.
\]

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 the LR method theoretically when the population distribution is not normal. Our following simulation results do confirm this conclusion.

We conducted a small simulation study to compare the finite sample performances of the confidence regions based on EL and LR methods with confidence level \( \alpha = 0.95 \), and report the proportion of \( LR(\theta_0) \leq z_{0.95}(k+2) \) and \( \ell_n(\theta_0) \leq z_{0.95}(k+2) \) respectively in our 2,000 simulations, where \( \theta_0 \) is the true value of \( \theta \). The results of simulations are reported in tables 1 to 3.
In the simulations, we used the model: \( Y_n = X_n \beta + u(n), \quad u(n) = \rho W_n u(n) + \epsilon(n) \) with \( X_n = (x_1, x_2, \cdots, x_n)^\tau, \quad x_i = \frac{i}{n+1}, \quad 1 \leq i \leq n, \quad \beta = 3.5, \quad \rho \) were taken as \(-0.85, -0.15, 0.15\) and \(0.85\), respectively, and \(\epsilon_i\)s were taken from \(N(0, 1), \quad t(5)\) and \(\chi^2 - 4\), respectively.

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 first considered three ideal cases of spatial units: \(n = m \times m\) regular grid with \(m = 7, 10, 13\), denoting \(W_n\) as \(grid_{49}, grid_{100}\) and \(grid_{169}\), respectively. Secondly, we used the weight matrix \(W_{49}\) related to 49 contiguous planning neighborhoods in Columbus, Ohio, U.S., which appeared in Anselin(1988, P. 187). Thirdly, \(W_n = I_5 \otimes W_{49}\) was considered, where \(\otimes\) is kronecker product. This corresponds to the pooling of five separate districts with similar neighboring structures in each district. Finally, weight matrix \(W_{345}\) was included in the simulations, which is related to 345 major cities in China.

A transformation is often used in applications to convert the matrix \(W_n\) to the unity of row-sums. We used the standardized version of \(W_n\) in our simulations, namely \(W_{ij}\) was replaced by \(W_{ij}/\sum_{j=1}^{n} W_{ij}\).

Simulation results show that the confidence regions based on LR behave well with coverage probabilities very close to the nominal level 0.95 when the error term \(\epsilon_i\) is normally distributed, but not well in other cases. The coverage probabilities of the confidence regions based on LR fall to the range \([0.8045, 0.8560]\) for \(t\) distribution and \([0.8295, 0.8615]\) for \(\chi^2\) distribution, which are far from the nominal level 0.95.

We can see, from tables 1 to 3, 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 \(\epsilon_i\) is normally distributed or not. Our simulation results recommend EL method when we can not confirm the normal distribution of the error term.

Tables 1-3 are about here.
4. Proofs

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

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

where \( \epsilon_{ni} \) are real valued random variables, and the \( a_{nij} \) and \( b_{ni} \) denote the real valued coefficients of the linear-quadratic form. We need the following assumptions in Lemma 1.

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

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

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

\[
\mu_{\tilde{Q}} = \sum_{i=1}^{n} a_{nii} \sigma_{ni}^2, \\
\sigma_{\tilde{Q}}^2 = 2 \sum_{i=1}^{n} \sum_{j=1}^{n} a_{nij}^2 \sigma_{ni}^2 \sigma_{nj}^2 + \sum_{i=1}^{n} b_{ni}^2 \sigma_{ni}^2 + \sum_{i=1}^{n} \{ a_{nii}^2 (\mu_{ni}^{(4)} - 3 \sigma_{ni}^4) + 2 b_{ni} a_{nii} \mu_{ni}^{(3)} \}, \tag{10}
\]

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

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

\[ \frac{\tilde{Q}_n - \mu_{\tilde{Q}}}{\sigma_{\tilde{Q}}} \overset{d}{\rightarrow} N(0,1). \]

**Proof.** See Theorem 1 and Remark 12 in Kelejian and Prucha (2001).
Lemma 2 Let $\eta_1, \eta_2, \ldots, \eta_n$ be a sequence of stationary random variables, with $E|\eta_i|^s < \infty$ for some constants $s > 0$ and $C > 0$. Then

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

Proof. It is straightforward.  

Lemma 3 Suppose that Assumptions (A1) to (A3) are satisfied. Then as $n \to \infty$,

$$Z_n = \max_{1 \leq i \leq n} ||\omega_i(\theta)|| = o_p(n^{1/2}) \; a.s., \quad (11)$$

$$\Sigma_{k+2}^{-1/2} \sum_{i=1}^{n} \omega_i(\theta) \xrightarrow{d} N(0, I_{k+2}), \quad (12)$$

$$n^{-1} \sum_{i=1}^{n} \omega_i(\theta)\omega_i^T(\theta) = n^{-1}\Sigma_{k+2} + o_p(1), \quad (13)$$

$$\sum_{i=1}^{n} ||\omega_i(\theta)||^3 = O_p(n), \quad (14)$$

where $\Sigma_{k+2}$ is given in (9).

Proof. Note that

$$Z_n \leq \max_{1 \leq i \leq n} ||b_i\epsilon_i|| + \max_{1 \leq i \leq n} \left| \tilde{g}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{g}_{ij}\epsilon_j \right|$$

$$+ \max_{1 \leq i \leq n} |\epsilon_i^2 - \sigma^2|$$

$$\leq \max_{1 \leq i \leq n} ||b_i\epsilon_i|| + \max_{1 \leq i \leq n} |\tilde{g}_{ii}(\epsilon_i^2 - \sigma^2)| + \max_{1 \leq i \leq n} \left| 2\epsilon_i \sum_{j=1}^{i-1} \tilde{g}_{ij}\epsilon_j \right|$$

$$+ \max_{1 \leq i \leq n} |\epsilon_i^2 - \sigma^2|. $$

By Conditions A1 and A2 and Lemma 2, we have

$$\max_{1 \leq i \leq n} ||b_i\epsilon_i|| = \max_{1 \leq i \leq n} ||b_i||o_p(n^{1/4}) = o_p(n^{1/4}),$$

10
Thus $Z_n = o_p(n^{1/2})$. (11) is proved.

For any given $l = (l_1^r, l_2, l_3) \in R^{k+2}$ with $||l|| = 1$, where $l_1 \in R^k, l_2, l_3 \in R$. Then

$$l^r \omega_i(\theta) = l_1^r b_i \epsilon_i + l_2 \{ \tilde{g}_{ii}(\epsilon_i^2 - \sigma^2) + 2 \epsilon_i \sum_{j=1}^{i-1} \tilde{g}_{ij} \epsilon_j \} + l_3 (\epsilon_i^2 - \sigma^2)$$

$$= (l_2 \tilde{g}_{ii} + l_3)(\epsilon_i^2 - \sigma^2) + 2 \epsilon_i \sum_{j=1}^{i-1} l_2 \tilde{g}_{ij} \epsilon_j + l_1^r b_i \epsilon_i.$$

Thus

$$\sum_{i=1}^{n} l^r \omega_i(\theta) = \sum_{i=1}^{n} (l_2 \tilde{g}_{ii} + l_3)(\epsilon_i^2 - \sigma^2) + 2 \sum_{i=1}^{n} \sum_{j=1}^{i-1} l_2 \tilde{g}_{ij} \epsilon_j \epsilon_j + \sum_{i=1}^{n} l_1^r b_i \epsilon_i.$$

Let

$$Q_n = \sum_{i=1}^{n} \sum_{j=1}^{n} u_{ij} \epsilon_i \epsilon_j + \sum_{i=1}^{n} v_i \epsilon_i,$$

where

$$u_{ii} = l_2 \tilde{g}_{ii} + l_3, u_{ij} = l_2 \tilde{g}_{ij}(i \neq j), v_i = l_1^r b_i.$$

Then

$$Q_n = \sum_{i=1}^{n} l^r \omega_i(\theta) = \sum_{i=1}^{n} \{ u_{ii}(\epsilon_i^2 - \sigma^2) + \sum_{j=1}^{i-1} u_{ij} \epsilon_j \epsilon_j + v_i \epsilon_i \}.$$

To obtain the asymptotic distribution of $Q_n$, we need to check Condition C2. From Condition A2(i), it can be shown that

$$\sum_{i=1}^{n} |u_{ij}| \leq |l_2| \sum_{i=1}^{n} |\tilde{g}_{ij}| + |l_3| \leq C. \quad (15)$$

Further,

$$n^{-1} \sum_{i=1}^{n} |v_i|^3 = n^{-1} \sum_{i=1}^{n} |l_1^r b_i|^3 \leq C \max_{1 \leq i \leq n} ||x_i||^3 \max_{1 \leq i \leq n} \left( \sum_{k=1}^{n} |a_{ik}| \right)^3 \leq C, \quad (16)$$
where \(a_{ik}\) is the \((i, k)\)-element of \(A_n(\rho)\). From (15) and (16), it follows that \(n^{-1} \sum_{i=1}^{n} |v_i|^3 \leq C\). Therefore, Condition C2 is satisfied.

We now derive the variance of \(Q_n\). Let \(e_i\) be the unit vector in the \(i\)-th coordinate direction. It can be shown that

\[
\sum_{i=1}^{n} \sum_{j=1}^{n} u_{ij}^2 = \sum_{i=1}^{n} \left\{ (l_2 \tilde{g}_{ii} + l_3)^2 + \sum_{i \neq j} (l_2 \tilde{g}_{ij})^2 \right\}
\]

\[
= \sum_{i=1}^{n} \left\{ (l_2 \tilde{g}_{ii})^2 + 2l_2 l_3 \tilde{g}_{ii} + l_3^2 + \sum_{i \neq j} (l_2 \tilde{g}_{ij})^2 \right\}
\]

\[
= 2l_2 l_3 \sum_{i=1}^{n} \tilde{g}_{ii} + nl_3^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} (l_2 \tilde{g}_{ij})^2
\]

\[
= 2l_2 l_3 \text{tr}(\tilde{G}_n) + nl_3^2 + l_2^2 \text{tr}(\tilde{G}_n^2),
\]

\[
\sum_{i=1}^{n} u_{ii}^2 = \sum_{i=1}^{n} (l_2 \tilde{g}_{ii} + l_3)^2
\]

\[
= l_2^2 \sum_{i=1}^{n} \tilde{g}_{ii}^2 + 2l_2 l_3 \text{tr}(\tilde{G}_n) + nl_3^2
\]

\[
= l_2^2 ||\text{Vec}(\text{diag}\tilde{G}_n)||^2 + 2l_2 l_3 \text{tr}(\tilde{G}_n) + nl_3^2;
\]

\[
\sum_{i=1}^{n} v_i^2 = \sum_{i=1}^{n} (l_1^T b_i)^2 = l_1^T \left( \sum_{i=1}^{n} b_i b_i^T \right) l_1
\]

\[
= l_1^T \left( \sum_{i=1}^{n} X_n^T A_n(\rho) e_i e_i^T A_n^T(\rho) X_n \right) l_1
\]

\[
= l_1^T X_n^T A_n(\rho) \left( \sum_{i=1}^{n} e_i e_i^T \right) A_n^T(\rho) X_n l_1
\]

\[
= l_1^T X_n^T A_n(\rho) A_n^T(\rho) X_n l_1,
\]

and that

\[
\sum_{i=1}^{n} u_{ii} v_i = \sum_{i=1}^{n} (l_2 \tilde{g}_{ii} + l_3) l_1^T b_i
\]

\[
= l_1^T X_n^T A_n(\rho) \text{Vec}(\text{diag}\tilde{G}_n) l_2 + l_1^T X_n^T A_n(\rho) l_1 l_3,
\]

where \(1_n\) is the \(n\)-dimensional vector with 1 as its components. It follows from (10) that the variance of \(Q_n\) is

\[
\sigma_Q^2 = 2 \sum_{i=1}^{n} \sum_{j=1}^{n} u_{ij}^2 \sigma^4 + \sum_{i=1}^{n} v_i^2 \sigma^2 + \sum_{i=1}^{n} u_{ii} (\mu_4 - 3\sigma^4) + 2u_{ii} v_i \mu_3
\]

12
\[= 2\sigma^4\{l_2^2 tr(\tilde{G}_n^2) + 2l_2l_3 tr(\tilde{G}_n) + nl_3^2\} + \sigma^2 l_1^2 X_n^T A_n(\rho) A_n^T(\rho) X_n l_1 \\
+ (\mu_4 - 3\sigma^4)\{l_2^2||Vec(diag\tilde{G}_n)||^2 + 2l_2l_3 tr(\tilde{G}_n) + nl_3^2\} + 2\mu_3\{l_1^2 X_n^T A_n(\rho) Vec(diag\tilde{G}_n)l_2 + l_1^2 X_n^T A_n(\rho) l_1 l_3\} \]

where \(\Sigma_{k+2}\) is given in (9). From Condition A3, one can see that \(n^{-1}\sigma_Q^2 \geq c_1 > 0\). From Lemma 1, we have

\[\frac{Q_n - E(Q_n)}{\sigma_Q} \xrightarrow{d} N(0, 1).\]

Noting that \(E(Q) = 0\), we thus have (12).

Next we will prove (13), i.e.

\[n^{-1}\sum_{i=1}^{n}(l^T \omega_i(\theta))^2 = n^{-1}\sigma^2_Q + o_p(1). \quad (17)\]

Let

\[Y_{in} = l^T \omega_i(\theta) = u_{ii}(\epsilon_i^2 - \sigma^2) + 2\sum_{j=1}^{i-1} u_{ij}\epsilon_i\epsilon_j + v_i\epsilon_i \]

\[= u_{ii}(\epsilon_i^2 - \sigma^2) + B_i\epsilon_i, \quad (18)\]

where \(B_i = 2\sum_{j=1}^{i-1} u_{ij}\epsilon_j + v_i\). Let \(\mathcal{F}_0 = \{\emptyset, \Omega\}, \mathcal{F}_i = \sigma(\epsilon_1, \epsilon_2, \ldots, \epsilon_i), 1 \leq i \leq n\). Then \(\{Y_{in}, \mathcal{F}_i, 1 \leq i \leq n\}\) form a martingale difference array. Note that

\[n^{-1}\sum_{i=1}^{n}(l^T \omega_i(\theta))^2 - n^{-1}\sigma^2_Q = n^{-1}\sum_{i=1}^{n}(Y_{in}^2 - EY_{in}^2) \]

\[= n^{-1}\sum_{i=1}^{n}\{Y_{in}^2 - E(Y_{in}^2|\mathcal{F}_{i-1}) + E(Y_{in}^2|\mathcal{F}_{i-1}) - EY_{in}^2\} \]

\[= n^{-1}S_{n1} + n^{-1}S_{n2}, \quad (19)\]

where \(S_{n1} = \sum_{i=1}^{n}\{Y_{in}^2 - E(Y_{in}^2|\mathcal{F}_{i-1})\}, S_{n2} = \sum_{i=1}^{n}\{E(Y_{in}^2|\mathcal{F}_{i-1}) - EY_{in}^2\}\).

Next we will prove

\[n^{-1}S_{n1} = o_p(1), \quad (20)\]
and

\[ n^{-1}S_{n2} = o_p(1). \]  \hspace{1cm} (21) 

It suffices to prove \( n^{-2}E(S_{n1}^2) \to 0 \) and \( n^{-2}E(S_{n2}^2) \to 0 \) respectively. Obviously,

\[ Y_i^2 = u_{ni}^2(\epsilon_i^2 - \sigma^2)^2 + B_i^2(\epsilon_i^2 - \sigma^2) \epsilon_i. \]

Thus

\[ E(Y_i^2 \mid F_{i-1}) = u_{ni}^2E(\epsilon_i^2 - \sigma^2)^2 + B_i^2\sigma^2 + 2u_{ni}B_i \mu_3. \]

It follows that

\[
\begin{align*}
n^{-2}E(S_{n1}^2) &= n^{-2} \sum_{i=1}^{n} E\{Y_i^2 - E(Y_i^2 \mid F_{i-1})\}^2 \\
&= n^{-2} \sum_{i=1}^{n} E[u_{ni}^2((\epsilon_i^2 - \sigma^2)^2 - E(\epsilon_i^2 - \sigma^2)^2) + B_i^2(\epsilon_i^2 - \sigma^2) \\
&\quad + 2u_{ni}B_i(\epsilon_i^3 - \sigma^2 \epsilon_i - \mu_3)]^2 \\
&\leq Cn^{-2} \sum_{i=1}^{n} E[u_{ni}^4((\epsilon_i^2 - \sigma^2)^2 - E(\epsilon_i^2 - \sigma^2)^2)^2] + Cn^{-2} \sum_{i=1}^{n} E[B_i^4(\epsilon_i^2 - \sigma^2)^2] \\
&\quad + Cn^{-2} \sum_{i=1}^{n} E\{u_{ni}^2B_i^2(\epsilon_i^2 - \sigma^2 \epsilon_i - \mu_3)^2\}. \hspace{1cm} (22)
\end{align*}
\]

By Condition A1, we have

\[
n^{-2} \sum_{i=1}^{n} E[u_{ni}^4((\epsilon_i^2 - \sigma^2)^2 - E(\epsilon_i^2 - \sigma^2)^2)^2] \leq Cn^{-1} \to 0, \hspace{1cm} (23)
\]

and

\[
\begin{align*}
n^{-2} \sum_{i=1}^{n} E\{B_i^4(\epsilon_i^2 - \sigma^2)^2\} &\leq Cn^{-2} \sum_{i=1}^{n} E\{(\sum_{j=1}^{i-1} u_{ij} \epsilon_j + \nu_i)^4 \\
&\leq Cn^{-2} \sum_{i=1}^{n} E(\sum_{j=1}^{i-1} u_{ij} \epsilon_j)^4 + Cn^{-2} \sum_{i=1}^{n} \nu_i^4 \\
&\leq Cn^{-2} \sum_{i=1}^{n} \sum_{j=1}^{i-1} u_{ij}^4 \mu_4 + Cn^{-2} \sum_{i=1}^{n} (\sum_{j=1}^{i-1} u_{ij}^2 \sigma^2)^2 + Cn^{-2} \sum_{i=1}^{n} (l_i^* b_i)^4 \\
&\leq Cn^{-1} \to 0. \hspace{1cm} (24)
\end{align*}
\]
Similarly, one can show that
\[ n^{-2} \sum_{i=1}^{n} E\{u_{ii}^{2}B_{i}^{2}(\epsilon_{i}^{2} - \sigma^{2}\epsilon_{i} - \mu_{3})^{2}\} \to 0. \]  
(25)

From (22)-(25), we have \( n^{-2}E(S_{n1}^{2}) \to 0 \). Furthermore,
\[
E(Y_{in}^{2}) = E\{E(Y_{in}^{2}|F_{i-1})\} = u_{ii}^{2}E(\epsilon_{i}^{2} - \sigma^{2})^2 + \sigma^{2}E(B_{i}^{2}) + 2u_{ii}\mu_{3}E(B_{i})
\]
\[
= u_{ii}^{2}E(\epsilon_{i}^{2} - \sigma^{2})^2 + \sigma^{2}(4\sum_{j=1}^{i-1}u_{ij}\sigma^{2} + v_{i}^{2}) + 2u_{ii}\mu_{3}v_{i}.
\]

Thus,
\[
n^{-2}E(S_{n2}^{2}) = n^{-2}E[\sum_{i=1}^{n}\{E(Y_{in}^{2}|F_{i-1}) - EY_{in}^{2}\}]^{2}
\]
\[
= n^{-2}E[\sum_{i=1}^{n}\{B_{i}^{2}\sigma^{2} - \sigma^{2}(4\sum_{j=1}^{i-1}u_{ij}\sigma^{2} + v_{i}^{2}) + 2u_{ii}\mu_{3}(B_{i} - v_{i})\}]^{2}
\]
\[
= n^{-2}\sum_{i=1}^{n}E[\sigma^{2}\{(2\sum_{j=1}^{i-1}u_{ij}\epsilon_{j})^{2} - 4\sum_{j=1}^{i-1}u_{ij}\sigma^{2}\} + 4(\sum_{j=1}^{i-1}u_{ij}\epsilon_{j})v_{i}\sigma^{2}
\]
\[
+2u_{ii}\mu_{3}(2\sum_{j=1}^{i-1}u_{ij}\epsilon_{j})]^{2}
\]
\[
\leq Cn^{-2}\sum_{i=1}^{n}E\{\sigma^{2}\{(\sum_{j=1}^{i-1}u_{ij}\epsilon_{j})^{2} - (\sum_{j=1}^{i-1}u_{ij}\sigma^{2})^{2}\} + Cn^{-2}\sum_{i=1}^{n}E\{(\sum_{j=1}^{i-1}u_{ij}\epsilon_{j})v_{i}\sigma^{2}\}^{2}
\]
\[
+Cn^{-2}\sum_{i=1}^{n}E\{2u_{ii}\mu_{3}(\sum_{j=1}^{i-1}u_{ij}\epsilon_{j})\}^{2}.
\]  
(26)

Note that
\[
n^{-2}\sum_{i=1}^{n}E[\sigma^{2}\{(\sum_{j=1}^{i-1}u_{ij}\epsilon_{j})^{2} - (\sum_{j=1}^{i-1}u_{ij}\sigma^{2})^{2}\}^{2} \leq n^{-2}\sigma^{4}\sum_{i=1}^{n}E(\sum_{j=1}^{i-1}u_{ij}\epsilon_{j})^{4}
\]
\[
\leq Cn^{-2}\sum_{i=1}^{n}\sum_{j=1}^{i-1}u_{ij}\mu_{4} + Cn^{-2}\sum_{i=1}^{n}(\sum_{j=1}^{i-1}u_{ij}\sigma^{2})^{2} \leq Cn^{-1} \to 0,
\]  
(27)

\[
n^{-2}\sum_{i=1}^{n}(\sum_{j=1}^{i-1}u_{ij}\epsilon_{j})v_{i}\sigma^{2}\}^{2} = n^{-2}\sigma^{6}\sum_{i=1}^{n}v_{i}^{2}\sum_{j=1}^{i-1}u_{ij}^{2} \leq Cn^{-2} \to 0,
\]  
(28)
and
\[
 n^{-2} \sum_{i=1}^{n} E\{2u_{ii}\mu_3(\sum_{j=1}^{i-1} u_{ij}\epsilon_j)\} = 4\mu_3^2\sigma^2 n^{-2} \sum_{i=1}^{n} u_{ii}^2 \sum_{j=1}^{i-1} u_{ij}^2 \leq Cn^{-1} \to 0, \quad (29)
\]
where we have used Conditions A1 and A2. From (26)-(29), we have
\[
n^{-2}ES_n^2 \to 0. \quad \text{The proof of (17) is thus complete.}
\]
Finally, we will prove (14). Note that
\[
\sum_{i=1}^{n} E||\psi_i(\theta)||^3 \leq \sum_{i=1}^{n} E||b_i\epsilon_i||^3 + \sum_{i=1}^{n} E|\tilde{g}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{g}_{ij}\epsilon_j|^3 + \sum_{i=1}^{n} E|\epsilon_i^2 - \sigma^2|^3. \quad (30)
\]
By Conditions A1 and A2,
\[
\sum_{i=1}^{n} E||b_i\epsilon_i||^3 \leq Cn(\max_{1\leq i \leq n} ||x_i||)^3 E|\epsilon_i|^3 = O(n), \quad (31)
\]
\[
\sum_{i=1}^{n} E|\tilde{g}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{g}_{ij}\epsilon_j|^3 \leq C \sum_{i=1}^{n} E|\tilde{g}_{ii}(\epsilon_i^2 - \sigma^2)|^3 + C \sum_{i=1}^{n} E|2\epsilon_i \sum_{j=1}^{i-1} \tilde{g}_{ij}\epsilon_j|^3 \\
\leq C \sum_{i=1}^{n} E|\tilde{g}_{ii}(\epsilon_i^2 - \sigma^2)|^3 + C \sum_{i=1}^{n} E|\epsilon_i|^3 \sum_{j=1}^{i-1} E|\tilde{g}_{ij}\epsilon_j|^3 + C \sum_{i=1}^{n} E|\epsilon_i|^3 \left( \sum_{j=1}^{i-1} E(\tilde{g}_{ij}\epsilon_j)^2 \right)^{3/2} = O(n), \quad (32)
\]
\[
\sum_{i=1}^{n} E|\epsilon_i^2 - \sigma^2|^3 = O(n). \quad (33)
\]
From (30)-(33), we have
\[
\sum_{i=1}^{n} E||\psi_i(\theta)||^3 = O(n). \quad (34)
\]
Further, using (34) and Markov inequality, we obtain \( \sum_{i=1}^{n} ||\omega_i(\theta)||^3 = O_p(n) \). Thus (14) is proved.

We now in the position to prove the main results in this article.

**Proof of Theorem 1.** Let \( \lambda = \lambda(\theta), \rho_0 = ||\lambda||, \lambda = \rho_0 \eta_0 \). From (8), we have
\[
\frac{\eta_{0}^{\tau}}{n} \sum_{j=1}^{n} \omega_{j}(\theta) - \frac{\rho_0}{n} \sum_{j=1}^{n} \frac{(\eta_{0}^{\tau} \omega_{j}(\theta))^2}{1 + \lambda^\tau \omega_{j}(\theta)} = 0.
\]
It follows that
\[
|\eta_{0}^{\tau} \bar{\omega}| \geq \frac{\rho_0}{1 + \rho_0 Z_n} \lambda_{\text{min}}(S_0),
\]
where \( Z_n \) is defined in (11), \( \bar{\omega} = n^{-1} \sum_{i=1}^{n} \omega_i(\theta), S_0 = n^{-1} \sum_{i=1}^{n} \omega_i(\theta) \omega_i^\tau(\theta) \).

That is
\[
|\eta_{0}^{\tau} \Sigma_{k+2}^{1/2} \Sigma_{k+2}^{-1/2} \bar{\omega}| \geq \frac{\rho_0}{1 + \rho_0 Z_n} \lambda_{\text{min}}(S_0),
\]
i. e.
\[
\lambda_{\text{max}}(\Sigma_{k+2}^{1/2}) ||\eta_0|| \cdot ||\Sigma_{k+2}^{-1/2} \bar{\omega}|| \geq \frac{\rho_0}{1 + \rho_0 Z_n} \lambda_{\text{min}}(S_0).
\]
Combining with Lemma 3 and Condition A3, we have
\[
\frac{\rho_0}{1 + \rho_0 Z_n} = O_p(n^{-1/2}).
\]
Therefore, from Lemma 3,
\[
\rho_0 = O_p(n^{-1/2}).
\]
Let \( \gamma_i = \lambda^\tau \omega_i(\theta) \). Then
\[
\max_{1 \leq i \leq n} |\gamma_i| = o_p(1).
\]
(35)
Using (8) again, we have
\[
0 = \frac{1}{n} \sum_{j=1}^{n} \frac{\omega_{j}(\theta)}{1 + \lambda^\tau \omega_{j}(\theta)}
\]
\[
= \frac{1}{n} \sum_{j=1}^{n} \omega_{j}(\theta) - \frac{1}{n} \sum_{j=1}^{n} \frac{\omega_{j}(\theta) \{\lambda^\tau \omega_{j}(\theta)\}}{1 + \lambda^\tau \omega_{j}(\theta)}
\]
\[
= \frac{1}{n} \sum_{j=1}^{n} \omega_{j}(\theta) - \{\frac{1}{n} \sum_{j=1}^{n} \omega_{j}(\theta) \omega_{j}(\theta)^\tau\} \lambda + \frac{1}{n} \sum_{j=1}^{n} \frac{\omega_{j}(\theta) \{\lambda^\tau \omega_{j}(\theta)\}^2}{1 + \lambda^\tau \omega_{j}(\theta)}
\]
17
\[
\ell_n(\theta) = \frac{1}{n} \sum_{j=1}^{n} \omega_j(\theta) - \left\{ \frac{1}{n} \sum_{j=1}^{n} \omega_j(\theta) \omega_j(\theta)^T \right\} \lambda + \frac{1}{n} \sum_{j=1}^{n} \frac{\omega_j(\theta) \gamma_j^2}{1 + \gamma_j}
\]

\[
= \overline{\omega} - S_0 \lambda + \frac{1}{n} \sum_{j=1}^{n} \frac{\omega_j(\theta) \gamma_j^2}{1 + \gamma_j}.
\]

Combining with Lemma 3 and Condition A3, we may write

\[
\lambda = S_0^{-1} \overline{\omega} + \varsigma, \quad (36)
\]

where \(||\varsigma||\) is bounded by

\[
n^{-1} \sum_{j=1}^{n} ||\omega_j(\theta)||^2 ||\lambda||^2 = O_p(n^{-1}).
\]

By (35) we may expand \(\log(1 + \gamma_i) = \gamma_i - \gamma_i^2/2 + \nu_i\) where, for some finite \(B > 0\),

\[
P(|\nu_i| \leq B|\gamma_i|^3, 1 \leq i \leq n) \to 1, \text{ as } n \to \infty.
\]

Therefore, from (7), (36) and Taylor expansion, we have

\[
\ell_n(\theta) = 2 \sum_{j=1}^{n} \log(1 + \gamma_j) = 2 \sum_{j=1}^{n} \gamma_j - \sum_{j=1}^{n} \gamma_j^2 + 2 \sum_{j=1}^{n} \nu_j
\]

\[
= 2n \lambda^T \overline{\omega} - n \lambda^T S_0 \lambda + 2 \sum_{j=1}^{n} \nu_j
\]

\[
= 2n(\lambda^T S_0^{-1} \overline{\omega})^T \overline{\omega} - n \lambda^T S_0^{-1} \overline{\omega} - 2n \varsigma^T \overline{\omega} + 2 \sum_{j=1}^{n} \nu_j
\]

\[
= n \overline{\omega}^T S_0^{-1} \overline{\omega} - n \varsigma^T S_0 \varsigma + 2 \sum_{j=1}^{n} \nu_j
\]

\[
= \left\{ n \Sigma_{k+2}^{-1/2} \overline{\omega} \right\}^T \left\{ n \Sigma_{k+2}^{-1/2} S_0 \Sigma_{k+2}^{-1/2} \right\}^{-1} \left\{ n \Sigma_{k+2}^{-1/2} \overline{\omega} \right\} - n \varsigma^T S_0 \varsigma + 2 \sum_{j=1}^{n} \nu_j.
\]

From Lemma 3 and Condition A3, we have

\[
\left\{ n \Sigma_{k+2}^{-1/2} \overline{\omega} \right\}^T \left\{ n \Sigma_{k+2}^{-1/2} S_0 \Sigma_{k+2}^{-1/2} \right\}^{-1} \left\{ n \Sigma_{k+2}^{-1/2} \overline{\omega} \right\} \xrightarrow{d} \chi_{k+2}^2.
\]
On the other hand, using Lemma 3 and above derivations, we can see that
\[ n \zeta \tau s_0 s = O_p(n^{-1}) = o_p(1) \] and

\[
\left| \sum_{j=1}^{n} \nu_j \right| \leq B ||\lambda||^3 \sum_{j=1}^{n} ||\omega_j(\theta)||^3 = O_p(n^{-1/2}) = o_p(1).
\]

The proof of Theorem 1 is thus complete.

Acknowledgements This work was partially supported by the National Natural Science Foundation of China (11671102), the Natural Science Foundation of Guangxi (2016GXNSFAA3800163, 2017GXNSFAA198349) and the Program on the High Level Innovation Team and Outstanding Scholars in Universities of Guangxi Province.

References
Anselin, L., 1988, Spatial Econometrics: Methods and Models. The Netherlands: Kluwer Academic Publishers.
Anselin, L. and Bera, A. K., 1998, Spatial dependence in linear regression models with an introduction to spatial econometrics, Handbook of Applied Economics Statistics, ed. by Ullah A. and Giles, D. E. A., New York: Marcel Dekker.
Bell K. P. and Bockstael, N. E., 2000, Applying the generalized-moments estimation approach to spatial problems involving microlevel data, The Review of Economics and Statistics, 82, 72-82.
Besley, T. and Case, A., 1995, Incumbent behavior: vote-seeking, tax-setting, and yardstick competition, The American Economic Review, 85, 25-45.
Bertrandm, M, Luttmer, E. F. P. and Mullainathan, S., 2000, Network effects and welfare cultures, Quarterly Journal of Economics, 115, 1019-1055.
Brueckner, J. K., 1998, Testing for strategic interaction among local governments: the case of growth controls, Journal of Urban Economics, 44, 438-467.
Case, A. C., 1991, Spatial patterns in household demand, Econometrica, 59, 953-965.
Case, A. C., Rosen, H. S. and Hines, J. R., 1993, Budget Spillovers and fiscal policy interdependence: evidence from the States, Journal of Public Economics, 52, 285-307.

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. and Keilegom, I. V. 2009, A review on empirical likelihood for regressions (with discussions), Test, 3, 415-447.

Cliff, A. D. and Ord, J. K., 1973, Spatial Autocorrelation. London: Pion Ltd.

Cressie, N., 1993, Statistics for Spatial Data, New York: John Wiley & Sons.

Dow, M. M., Burton, M. L., and White, D. R., 1982, Network autocorrelation: a simulation study of a foundational problem in regression and survey research, Social Networks, 4, 169-200.

Kelejian, H. H. and Prucha, I. 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, Journal of Econometrics, 104, 219-257.

Krämer, W. and Donninger, C., 1987, Spatial autocorrelation among errors and the relative efficiency of OLS in the linear regression model, Publications of the American Statistical Association, 82, 577-579.

Lee, L. F., 2004, Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models, Econometrica, 72, 1899-1925.

Ord, K. 1979, Estimation methods for models of spatial interaction, Journal of the American Statistical Association, 70, 120-126.

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, Ann.
Owen, A. B., 1991, Empirical likelihood for linear models, Ann. Statist., 19, 1725-1747.
Owen, A. B., 2001, Empirical Likelihood, London: Chapman & Hall.
Qin, J. and Lawless, J., 1994, Empirical likelihood and general estimating equations, Ann. Statist., 22, 300-325.
Topa, G. 2001, Social interactions, local Spillovers and unemployment, Review of Economic Studies, 68, 261-295.
Wu, C. B., 2004, Weighted empirical likelihood inference, Statistics & Probability Letters, 66, 67-79.
Zhong, B., Rao, J.N.K., 2000, Empirical likelihood inference under stratified random sampling using auxiliary population information, Biometrika, 87, 929-938.
| $\rho$ | $W_n$ | LR   | EL   | $\rho$ | $W_n$ | LR   | EL   |
|-------|-------|------|------|--------|-------|------|------|
| -0.85 | grid$_{49}$ | 0.9715 | 0.8760 | -0.15  | grid$_{49}$ | 0.9435 | 0.8820 |
|       | grid$_{100}$ | 0.9655 | 0.9200 |         | grid$_{100}$ | 0.9450 | 0.9045 |
|       | grid$_{169}$ | 0.9595 | 0.9370 |         | grid$_{169}$ | 0.9455 | 0.9325 |
|       | $W_{49}$ | 0.9630 | 0.8840 |         | $W_{49}$ | 0.9405 | 0.8645 |
|       | $I_5 \otimes W_{49}$ | 0.9565 | 0.9370 |         | $I_5 \otimes W_{49}$ | 0.9455 | 0.9330 |
|       | $W_{345}$ | 0.9535 | 0.9260 |         | $W_{345}$ | 0.9460 | 0.9395 |
| 0.85  | grid$_{49}$ | 0.9285 | 0.8635 | 0.15   | grid$_{49}$ | 0.9290 | 0.8680 |
|       | grid$_{100}$ | 0.9320 | 0.9045 |         | grid$_{100}$ | 0.9435 | 0.9160 |
|       | grid$_{169}$ | 0.9435 | 0.9305 |         | grid$_{169}$ | 0.9470 | 0.9320 |
|       | $W_{49}$ | 0.9435 | 0.8680 |         | $W_{49}$ | 0.9450 | 0.8805 |
|       | $I_5 \otimes W_{49}$ | 0.9560 | 0.9500 |         | $I_5 \otimes W_{49}$ | 0.9525 | 0.9405 |
|       | $W_{345}$ | .9545  | 0.9445 |         | $W_{345}$ | 0.9485 | 0.9375 |

Table 1: Coverage probabilities of the LR and EL confidence regions with $\epsilon_i \sim N(0,1)$
Table 2: Coverage probabilities of the LR and EL confidence regions with $\epsilon_i \sim t(5)$

| $\rho$ | $W_n$ | LR   | EL   | $\rho$ | $W_n$ | LR   | EL   |
|--------|-------|------|------|--------|-------|------|------|
| -0.85  | grid$_{49}$ | 0.8640 | 0.8025 | -0.15  | grid$_{49}$ | 0.8695 | 0.8010 |
|        | grid$_{100}$ | 0.8575 | 0.8610 |         | grid$_{100}$ | 0.8310 | 0.8640 |
|        | grid$_{169}$ | 0.8400 | 0.8870 |         | grid$_{169}$ | 0.8160 | 0.8800 |
|        | $W_{49}$ | 0.8670 | 0.8065 |         | $W_{49}$ | 0.8355 | 0.7990 |
|        | $I_5 \otimes W_{49}$ | 0.8425 | 0.9155 |         | $I_5 \otimes W_{49}$ | 0.8175 | 0.8930 |
|        | $W_{345}$ | 0.8145 | 0.9010 |         | $W_{345}$ | 0.8290 | 0.9200 |
| 0.85   | grid$_{49}$ | 0.8180 | 0.7890 | 0.15   | grid$_{49}$ | 0.8520 | 0.8040 |
|        | grid$_{100}$ | 0.8160 | 0.8575 |         | grid$_{100}$ | 0.8440 | 0.8750 |
|        | grid$_{169}$ | 0.8115 | 0.9020 |         | grid$_{169}$ | 0.8210 | 0.8970 |
|        | $W_{49}$ | 0.8480 | 0.7855 |         | $W_{49}$ | 0.8495 | 0.7985 |
|        | $I_5 \otimes W_{49}$ | 0.8180 | 0.9010 |         | $I_5 \otimes W_{49}$ | 0.8090 | 0.8955 |
|        | $W_{345}$ | 0.8030 | 0.9110 |         | $W_{345}$ | 0.8065 | 0.9125 |
Table 3: Coverage probabilities of the LR and EL confidence regions with $\epsilon_i + 4 \sim \chi^2_i$

| $\rho$ | $W_n$ | LR   | EL   | $\rho$ | $W_n$ | LR   | EL   |
|-------|-------|------|------|-------|-------|------|------|
| -0.85 | grid$_{49}$ | 0.8670 | 0.8070 | -0.15 | grid$_{49}$ | 0.8560 | 0.8080 |
|       | grid$_{100}$ | 0.8530 | 0.8850 |       | grid$_{100}$ | 0.8370 | 0.8610 |
|       | grid$_{169}$ | 0.8570 | 0.8950 |       | grid$_{169}$ | 0.8450 | 0.8975 |
|       | $W_{49}$    | 0.8615 | 0.7985 |       | $W_{49}$    | 0.8490 | 0.8125 |
|       | $I_5 \otimes W_{49}$ | 0.8580 | 0.9185 |       | $I_5 \otimes W_{49}$ | 0.8385 | 0.9160 |
|       | $W_{345}$  | 0.8525 | 0.9270 |       | $W_{345}$  | 0.8275 | 0.9295 |
| 0.85  | grid$_{49}$ | 0.8365 | 0.7915 | 0.15  | grid$_{49}$ | 0.8505 | 0.7955 |
|       | grid$_{100}$ | 0.8320 | 0.8530 |       | grid$_{100}$ | 0.8430 | 0.8690 |
|       | grid$_{169}$ | 0.8395 | 0.8900 |       | grid$_{169}$ | 0.8320 | 0.9050 |
|       | $W_{49}$    | 0.8490 | 0.7820 |       | $W_{49}$    | 0.8445 | 0.7920 |
|       | $I_5 \otimes W_{49}$ | 0.8435 | 0.9050 |       | $I_5 \otimes W_{49}$ | 0.8385 | 0.9215 |
|       | $W_{345}$  | 0.8490 | 0.9325 |       | $W_{345}$  | 0.8430 | 0.9285 |