Appl. Math. J. Chinese Univ.
2024, 39(1): 125-139

Empirical likelihood for spatial cross-sectional data models with matrix exponential spatial specification

LIU Yan[1,2] RONG Jian-rong[2] QIN Yong-song[2,*]

Abstract. In this paper, we study spatial cross-sectional data models in the form of matrix exponential spatial specification (MESS), where MESS appears in both dependent and error terms. The empirical likelihood (EL) ratio statistics are established for the parameters of the MESS model. It is shown that the limiting distributions of EL ratio statistics follow chi-square distributions, which are used to construct the confidence regions of model parameters. Simulation experiments are conducted to compare the performances of confidence regions based on EL method and normal approximation method.

§1 Introduction

After several decades of development, the theories and applications of spatial econometric models have become relatively mature. Among them, the computational problems related to maximum likelihood estimation have always been a research hot spot. However, even with the rapid development and improvement in the speed and capacity of today's computers, the huge volume of spatial data is still an unbearable burden. Therefore, it is an effective way to find a method to speed up the calculation process. In spatial econometrics, there exist model specifications that cannot be nested with each other, such as spatial autoregressive (SAR) models and matrix exponential spatial specification (MESS) models proposed by LeSage and Pace (2007). LeSage and Pace (2007) propose the use of MESS to describe spatial correlation, which not only greatly simplifies the log-likelihood function, but also has estimates and inferences similar to the conventional spatial autoregressive classes of specific targets. The matrix exponential can replace the spatial autoregressive process. In essence, the exponential decay in space replaces the geometric decay in the spatial autoregressive process, which has the advantages of calculation and theory (see Chiu et al. 1996).

In recent years, there has been a lot of research progress about MESS. Han and Lee (2013) propose to use the J test method to study the selection of SAR and MESS models, and the

Received: 2021-05-03. Revised: 2021-07-07.
MR Subject Classification: 62G05, 62E20.
Keywords: MESS, empirical likelihood, confidence region.
Digital Object Identifier(DOI): https://doi.org/10.1007/s11766-024-4446-4.
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).
*Corresponding author.

Monte Carlo experiment shows that the J test statistic can distinguish SAR and MESS models well. Piribauer and Fischer (2014) study the uncertainty of the MESS model on the basis of LeSage and Pace (2007), and at the same time propose the spatial Durbin model in the form of matrix exponential. Figueiredo and Silva (2015) extend the MESS model to the panel form and compare the estimated results of the model with the autoregressive model of the spatial panel. Debarsy et al. (2015) study the property of large samples of the MESS model and make a comparative analysis of the MESS and SAR models by using the quasi-maximum likelihood (QML) method. The study shows that in the case of heteroskedasticity, the QML estimation results of the MESS model are consistent, while the QML estimation results for SAR can not be satisfied. In the case of unknown heteroskedasticity, the QML parameter estimation results of MESS model also satisfy the consistency. Meanwhile, GMM is asymptotically as efficient as QML when the disturbances are i.i.d. with normal distribution (not just under homoskedasticity), and GMM can be asymptotically more efficient (not effective) than QML when the disturbances are not normally distributed (when the moments are properly selected). Liu (2017) studies the model selection problem of SARAR and MESS, uses the adaptive LASSO to select variables and estimate parameters at the same time, develops the Vuong type test and adaptive LASSO program, and supplements the existing spatial model selection method. Zhang et al. (2019) study QML estimation of spatial panel data model in the case of fixed effects and heteroskedasticity, in which spatial effects of dependent variables and disturbances take the form of MESS. The asymptotic properties of QML estimation with large sample and finite or large T are established. It is shown that the QML estimator (QMLE) can be consistent and asymptotically normal under unknown heteroscedasticity when the spatial weight matrices in two MESS processes are commutative. It provides a consistent estimator for the standard deviation of QMLE under regular conditions and can be used for inference.

In the MESS literature above, there are two main methods of estimation. One is QML method, and the other is a method with higher computational efficiency, GMM (see Debarsy et al., 2015; Zhang et al., 2019). When the confidence intervals/regions in the model parameters is constructed using the above method, the normal approximation (NA) method is usually applied, and a consistent estimator of the asymptotic covariance of the estimator must be obtained, which may reduce the accuracy of the interval estimator. In this article, we propose to construct the confidence region of the parameters in the MESS model using the empirical likelihood (EL) method introduced by Owen(1988, 1990, 1991). There are many references about EL method, such as Qin and Lawlees (1994), Chen and Keilegom (2009) and Owen (2001). The idea of using the EL method in MESS models is to introduce martingale sequences, which convert the linear-quadratic form of the MESS model's estimation equations into linear form. It is worth noting that the estimation equations of other spatial data models also have the form of linear-quadratic form. By exploring the inherent martingale structures, Qin (2021) and Jin and Lee (2019) successfully construct the confidence intervals/regions of the spatial autoregressive model by using EL method. Li et al. (2020) further apply the above methods to the spatial panel data model. However, we have not seen any published research results on the the EL method for MESS models. The EL method has the advantage that the shape of the confidence interval is determined by the sample itself and does not require the estimation of covariance, which is the main motivation of our current research. The major difference of

Qin (2021) and this article is the model specification and the advantage of the models studied in this article is mentioned in the first paragraph.

In this paper, we use EL method to construct the MESS model with weight matrix in exponential form in both dependent variable and error terms, and compare the performance of confidence intervals based on NA and EL methods through simulation. The results show that the confidence interval based on EL method can achieve the same effect as that of NA method, and it is better than NA method in some cases. We find that the execution speed of the EL method is much faster than the NA method.

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. Some concluding remarks are given in Section 5.

§2 Main Results

In this article, we consider the following MESS model, where MESS appears in the dependent variable and the error terms, denoted as MESS(1,1). The model is as follows:

$$e^{\iota W_n} y_n = X_n\beta + u_n, \quad e^{\tau M_n} u_n = \epsilon_n, \tag{1}$$

where $\epsilon_n = (\epsilon_{n1}, \cdots, \epsilon_{nn})'$ is an n-dimensional vector of i.i.d. random variables with mean zero and finite variance σ^2, y_n is an n-dimensional column vector of observed dependent variables, and $X_n = (x_1, x_2, \cdots, x_n)'$ is the non-random $n \times k$ matrix of exogenous variables with the corresponding regression coefficient vector β. W_n and M_n are $n \times n$ spatial weight matrices, which have zero diagonals and may be the same or different. The quasi log likelihood function for the MESS(1,1) in model (1) is

$$\begin{aligned} L_n(\psi) \quad = \quad & -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) + \ln|e^{\iota W_n}| + \ln|e^{\tau M_n}| \\ & -\frac{1}{2\sigma^2}(e^{\iota W_n} y_n - X_n\beta)' e^{\tau M_n'} e^{\tau M_n}(e^{\iota W_n} y_n - X_n\beta), \end{aligned} \tag{2}$$

where $\psi = (\theta, \sigma^2)'$ with $\theta = (\iota, \tau, \beta')'$. Since $|e^{\iota W_n}| = e^{\iota \cdot \mathrm{tr}(W_n)}$ and $|e^{\tau M_n}| = e^{\tau \cdot \mathrm{tr}(M_n)}$ (e.g., Chiu et al., 1996), it follows that $\ln|e^{\iota W_n}| = \ln|e^{\tau M_n}| = 0$, given that W_n and M_n have zero diagonals, i.e., $w_{ii} = m_{ii} = 0$ for all i. Therefore, the quasi log likelihood function in (2) can be rewritten as:

$$L_n(\psi) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}(e^{\iota W_n} y_n - X_n\beta)' e^{\tau M_n'} e^{\tau M_n}(e^{\iota W_n} y_n - X_n\beta). \tag{3}$$

From (3), the QMLE of θ is obviously the minimization of the function

$$T_n(\theta) = (e^{\iota W_n} y_n - X_n\beta)' e^{\tau M_n'} e^{\tau M_n}(e^{\iota W_n} y_n - X_n\beta).$$

In our work, the parameters we are interested in do not include variance parameters. For given σ^2, our aim is to derive the EL statistic of θ in MESS (1,1) model. It is noted that the QMLE of θ is also the solutions of the estimating equations (4)-(6) below. The first order derivatives of $T_n(\theta)$ with respect to ι, τ, and β are, respectively,

$$\begin{aligned} \frac{\partial T_n(\theta)}{\partial \iota} &= 2(e^{\tau M_n} W_n X_n \beta)' \epsilon_n + 2\epsilon_n' e^{-\tau M_n'} W_n' e^{\tau M_n'} \epsilon_n, \\ \frac{\partial T_n(\theta)}{\partial \tau} &= 2\epsilon_n' M_n \epsilon_n, \quad \frac{\partial T_n(\theta)}{\partial \beta} = -2X_n' e^{\tau M_n'} \epsilon_n. \end{aligned}$$

Setting the above derivatives to zero, we obtain the following estimating equations:

$$X_n' e^{\tau M_n'} \epsilon_n = 0, \tag{4}$$

$$(e^{\tau M_n} W_n X_n \beta)' \epsilon_n + \epsilon_n' e^{-\tau M_n'} W_n' e^{\tau M_n'} \epsilon_n = 0, \tag{5}$$

$$\epsilon_n' M_n \epsilon_n = 0. \tag{6}$$

We note that when the disturbances ϵ_{ni}'s are i.i.d. with mean 0 and variance σ^2,

$$E(\epsilon_n' e^{-\tau M_n'} W_n' e^{\tau M_n'} \epsilon_n) = \sigma^2 \mathrm{tr}(W_n' e^{\tau M_n'} e^{-\tau M_n'}) = \sigma^2 \mathrm{tr}(W_n') = 0$$

and $E(\epsilon_n' M_n \epsilon_n) = \sigma^2 \mathrm{tr}(M_n) = 0$.

As the equations (5) and (6) include the linear-quadratic forms of ϵ_n, to use the EL method, we need to change them into the linear forms of martingale difference arrays. Firstly, we let $H_n = e^{\tau M_n} W_n e^{-\tau M_n}$, $\tilde{H}_n = \frac{1}{2}(H_n + H_n')$, and $\tilde{M}_n = \frac{1}{2}(M_n + M_n')$. Use $\tilde{h}_{ij}$, $\tilde{m}_{ij}$, a_i, b_i to denote the (i, j) element of matrix $\tilde{H}_n$, the (i, j) element of matrix $\tilde{M}_n$, the i-th column of the matrix $X_n' e^{\tau M_n'}$, the i-th component of the vector $(e^{\tau M_n} W_n X_n \beta)'$, respectively, and adopt the convention that any sum with an upper index of less than one is zero. Secondly, we introduce two martingale difference arrays as follows. Define the σ-fields: $\mathcal{F}_0 = \{\phi, \Omega\}, \mathcal{F}_i = \sigma(\epsilon_1, \epsilon_2, \cdots, \epsilon_i), 1 \le i \le n$. Let

$$\tilde{Y}_{in} = \tilde{h}_{ii}\left(\epsilon_i^2 - \sigma^2\right) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{h}_{ij}\epsilon_j, \quad \tilde{G}_{in} = \tilde{m}_{ii}\left(\epsilon_i^2 - \sigma^2\right) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{m}_{ij}\epsilon_j. \tag{7}$$

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

$$\epsilon_n' \tilde{H}_n \epsilon_n - \sigma^2 \mathrm{tr}(\tilde{H}_n) = \sum_{i=1}^n \tilde{Y}_{in}, \quad \epsilon_n' \tilde{M}_n \epsilon_n - \sigma^2 \mathrm{tr}(\tilde{M}_n) = \sum_{i=1}^n \tilde{G}_{in}. \tag{8}$$

Based on (4)-(8), we propose the following EL ration statistic for $\theta \in R^{k+2}$

$$L(\theta) = \sup_{p_i, 1 \le i \le n} \prod_{i=1}^n (np_i),$$

where $\{p_i\}$ satisfy

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

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

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

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

Let

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

where ϵ_i is the i-th component of $e^{\tau M_n}(e^{\iota W_n}y_n - X_n\beta)$. By Owen (1990), one can show that

$$\ell(\theta) \hat{=} -2\log L(\theta) = 2\sum_{i=1}^{n}\log\left\{1+\lambda'(\theta)\omega_i(\theta)\right\}, \tag{9}$$

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

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\omega_i(\theta)}{1+\lambda'(\theta)\omega_i(\theta)} = 0. \tag{10}$$

Let $\mu_j = E(\epsilon_1^j), j = 3, 4$. Use $Vec_D(A)$ to denote the vector formed by the diagonal elements of a matrix A and $\|a\|$ to denote the L_2-norm of a vector a. To obtain the asymptotic distribution of $\ell(\theta)$, we need following assumptions:

A1. $\{\epsilon_i, 1 \le i \le n\}$ are i.i.d. with mean 0, variance $\sigma^2 > 0$ and $E|\epsilon_1|^{4+\eta} < \infty$ for some $\eta > 0$.

A2. Let W_n, M_n and X_n be as described above. They satisfy the following conditions: (i) The row and column sums of W_n and M_n are uniformly bounded in absolute value. The diagonal elements of W_n and M_n are zero. (ii) The elements of X_n are uniformly bounded.

A3. There is a constant $\delta > 0$, such that $|\iota| \le \delta, |\tau| \le \delta$ and the true value (ι_0, τ_0) of (ι, τ) is in the interior of the parameter space $[-\delta, \delta] \times [-\delta, \delta]$.

A4. There are constants $c_j > 0, j = 1, 2$, such that

$$0 < c_1 \le \lambda_{min}\left(n^{-1}\Sigma_{k+2}\right) \le \lambda_{max}\left(n^{-1}\Sigma_{k+2}\right) \le c_2 < \infty,$$

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

$$\Sigma_{k+2} = \Sigma'_{k+2} = \text{Cov}\left\{\sum_{i=1}^{n}\omega_i(\theta)\right\} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ * & \Sigma_{22} & \Sigma_{23} \\ * & * & \Sigma_{33} \end{pmatrix}_{(k+2)\times(k+2)} \tag{11}$$

where

$$\begin{aligned}
\Sigma_{11} &= \sigma^2 X_n' e^{\tau M_n'} e^{\tau M_n} X_n, \Sigma_{12} = \sigma^2 X_n' e^{\tau M_n'} e^{\tau M_n} W_n X_n \beta + \mu_3 X_n' e^{\tau M_n'} Vec_D(\tilde{H}_n), \\
\Sigma_{13} &= \mu_3 X_n' e^{\tau M_n'} Vec_D(\tilde{M}_n), \Sigma_{22} = 2\sigma^4 \text{tr}(\tilde{H}_n^2) + (\mu_4 - 3\sigma^4)\|Vec_D(\tilde{H}_n)\|^2 \\
&\quad + \sigma^2 (e^{\tau M_n} W_n X_n \beta)' e^{\tau M_n} W_n X_n \beta + 2\mu_3 (e^{\tau M_n} W_n X_n \beta)' Vec_D(\tilde{H}_n), \\
\Sigma_{23} &= (\mu_4 - 3\sigma^4) Vec_D'(\tilde{H}_n) Vec_D(\tilde{M}_n) + 2\sigma^4 \text{tr}(\tilde{H}_n \tilde{M}_n) + \mu_3 (e^{\tau M_n} W_n X_n \beta)' Vec_D(\tilde{M}_n), \\
\Sigma_{33} &= (\mu_4 - 3\sigma^4)\|Vec_D(\tilde{M}_n)\|^2 + 2\sigma^4 \text{tr}(\tilde{M}_n^2).
\end{aligned}$$

A5. $\lim_{n\to\infty} \frac{1}{n} X_n' e^{\tau M_n'} e^{\tau M_n} X_n$ exists and is nonsingular for any $\tau \in [-\delta, \delta]$, and the sequence of the smallest eigenvalues of $e^{\tau M_n'} e^{\tau M_n}$ is bounded away from zero uniformly in $\tau \in [-\delta, \delta]$.

Remark 1. Conditions A1 to A3 are common assumptions for spatial models. For example, A1 and A2 , which are used in assumptions 4 and 5 in Lee (2004), A3 and A5 are used in Assumptions 3 and Assumptions 4 in Debarsy et al. (2015), where A3 is to guarantee that uniform convergence is possible on a compact parameter space. The analog of $0 < c_1 \le \lambda_{min}\left(n^{-1}\Sigma_{k+2}\right)$ is employed in the assumption of Theorem 1 in Kelejian and Prucha (2001). From Conditions A1 and A2, it can be seen that $\lambda_{max}\left(n^{-1}\Sigma_{k+2}\right) \le c_2 < \infty$. For the sake of discussion, we put this consequence of A1 and A2 as a part of A4.

We now present the main results.

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

$$\ell(\theta) \xrightarrow{d} \chi^2_{k+2},$$

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

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

§3 Simulations

Let $\theta = (\iota, \tau, \beta')'$. Denote $\mathbb{W}_n = e^{\tau M_n} W_n e^{-\tau M_n}$, and $A^s = A + A'$ for any square matrix A. It can be shown, e.g. Debarsy et al. (2015), that $\frac{1}{n}\frac{\partial^2 T_n(\theta)}{\partial\theta\partial\theta'} = C_n + o_p(1)$, where $C_n = E(\frac{1}{n}\frac{\partial^2 T_n(\theta)}{\partial\theta\partial\theta'})$ is positive semi-definite, which has the elements:

$$C_{n,\iota\iota} = E(\frac{1}{n}\frac{\partial^2 T_n(\theta)}{\partial\iota\partial\iota}) = \frac{1}{n}\{\sigma^2 \mathrm{tr}(\mathbb{W}_n^s \mathbb{W}_n^s) + 2(\mathbb{W}_n e^{\tau M_n} X_n \beta)'(\mathbb{W}_n e^{\tau M_n} X_n \beta)\},$$

$$C_{n,\iota\tau} = E(\frac{1}{n}\frac{\partial^2 T_n(\theta)}{\partial\iota\partial\tau}) = \frac{1}{n}\{\sigma^2 \mathrm{tr}(\mathbb{W}_n^s M_n^s)\},$$

$$C_{n,\iota\beta'} = E(\frac{1}{n}\frac{\partial^2 T_n(\theta)}{\partial\iota\partial\beta'}) = \frac{1}{n}\{-2(e^{\tau M_n} X_n)'\mathbb{W}_n e^{\tau M_n} X_n \beta\},$$

$$C_{n,\tau\tau} = E(\frac{1}{n}\frac{\partial^2 T_n(\theta)}{\partial\tau\partial\tau}) = \frac{1}{n}\sigma^2 \mathrm{tr}(M_n^s M_n^s), C_{n,\tau\beta'} = E(\frac{1}{n}\frac{\partial^2 T_n(\theta)}{\partial\tau\partial\beta'}) = 0,$$

$$C_{n,\beta\beta'} = E(\frac{1}{n}\frac{\partial^2 T_n(\theta)}{\partial\beta\partial\beta'}) = \frac{2}{n}(e^{\tau M_n} X_n)' e^{\tau M_n} X_n.$$

According to Debarsy et al. (2015), if $\epsilon_n \sim N(0, \sigma^2 I_n)$; $\tau = 0$; or W_n and M_n are commutative, the QMLE $\hat\theta$ of θ satisfies: $\sqrt{n}(\hat\theta - \theta) \xrightarrow{d} N(0, \Sigma)$, where $\Sigma = 2\sigma^2 \lim_{n\to\infty} C_n^{-1}$.

Based on the above asymptotic result, we can obtain the NA based confidence region for θ. However, we note that the NA method depends on the availability of a consistent estimator of the asymptotic covariance matrix in practical applications, while the EL method does not.

Using R software, we conducted a small simulation study to compare the finite sample performances of the confidence regions based on EL and NA methods with confidence level $\gamma = 0.95$, and reported the proportion of $\ell(\theta) \le z_{0.95}(k+2)$ and $(\hat\theta - \theta)'(\Sigma/n)^{-1}(\hat\theta - \theta) \le z_{0.95}(k+2)$ respectively in 1,000 replications.

In the simulations, we used the following MESS(1,1) model to generate data:

$$e^{\iota M_n} y_n = X_n \beta + u_n, \quad e^{\tau W_n} u_n = \epsilon_n,$$

where y_n, X_n were all $n \times 1$ vectors. X_n were generated independently from $U(1,5)$. We selected $\beta = 3.5$, and (ι, τ) were taken as $(-2,-1), (-2,1), (2,-1), (2,1)$, respectively, and $\epsilon_n' s$ were i.i.d. 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 were neighbor by queen contiguity rule, $W_{ij} = 0$ otherwise (Anselin, 1988, P.18). In addition, we took $M_n = W_n$. We considered five ideal cases of spatial units: $n = m \times m$ regular grid with $m = 7, 10, 13, 16, 20$, denoting W_n as $grid_{49}, grid_{100}$, $grid_{169}, grid_{256}$, and $grid_{400}$, respectively. A transformation is often used in applications to normalize W_n to have row sums equal to one. We also reported the mean, standard deviation (SD) and root mean square errors (RMSE) of the $1,000$ EL and NA estimators $\hat\theta$ to show the parameter estimates were close. The results of simulations are reported in Tables 1-4.

As shown in Table 1, when the error term follows a normal distribution and the number of spatial units is large enough, the confidence region coverage probabilities based on NA and EL are very close to the nominal level of 0.95. In general, the two methods have good coverage when the sample size is large. Tables 2 and 3 reflect the same trend. It is worth noting that when the error term follows t or chi-square distribution, the coverage probabilities of the confidence regions based on NA method are far inferior to that of the confidence regions based on EL method. In addition, the execution speed of the EL method is much faster than the NA method. Table 4 reports the simulation results of mean, SD and RMSE for MESS model ($\iota_0 = -2, \tau_0 = -1, \beta_0 = 3.5$). From the results, we can see that both the EL and NA estimators perform well and both estimators are close with a larger sample size $n = 400$.

Although MESS eliminates logarithmic determinant in NA method, it still brings time trouble to the simulation with covariance in the form of matrix exponent. Because EL method does not need to calculate covariance and its computational efficiency is better than NA method, thus EL method is recommended in constructing confidence regions of parameters.

Table 1. Coverage probabilities of the NA and EL confidence regions with $\epsilon_i \sim N(0,1)$.

(ι,τ)	$W_n = M_n$	NA	EL	(ι,τ)	$W_n = M_n$	NA	EL
$(-2,-1)$	$grid_{49}$	0.910	0.912	$(-2,1)$	$grid_{49}$	0.852	0.873
	$grid_{100}$	0.914	0.924		$grid_{100}$	0.914	0.928
	$grid_{169}$	0.940	0.936		$grid_{169}$	0.922	0.932
	$grid_{256}$	0.938	0.932		$grid_{256}$	0.960	0.956
	$grid_{400}$	0.943	0.952		$grid_{400}$	0.957	0.938
$(2,-1)$	$grid_{49}$	0.900	0.876	$(2,1)$	$grid_{49}$	0.880	0.858
	$grid_{100}$	0.920	0.934		$grid_{100}$	0.912	0.916
	$grid_{169}$	0.944	0.952		$grid_{169}$	0.906	0.932
	$grid_{256}$	0.952	0.934		$grid_{256}$	0.952	0.944
	$grid_{400}$	0.954	0.956		$grid_{400}$	0.960	0.948

Table 2. Coverage probabilities of the NA and EL confidence regions with $\epsilon_i \sim t(5)$.

(ι,τ)	$W_n = M_n$	NA	EL	(ι,τ)	$W_n = M_n$	NA	EL
$(-2,-1)$	$grid_{49}$	0.888	0.822	$(-2,1)$	$grid_{49}$	0.860	0.886
	$grid_{100}$	0.890	0.904		$grid_{100}$	0.904	0.912
	$grid_{169}$	0.932	0.936		$grid_{169}$	0.916	0.924
	$grid_{256}$	0.940	0.940		$grid_{256}$	0.956	0.960
	$grid_{400}$	0.933	0.956		$grid_{400}$	0.930	0.948
$(2,-1)$	$grid_{49}$	0.892	0.838	$(2,1)$	$grid_{49}$	0.870	0.886
	$grid_{100}$	0.904	0.906		$grid_{100}$	0.922	0.926
	$grid_{169}$	0.930	0.936		$grid_{169}$	0.936	0.942
	$grid_{256}$	0.928	0.943		$grid_{256}$	0.938	0.950
	$grid_{400}$	0.946	0.954		$grid_{400}$	0.942	0.952

Table 3. Coverage probabilities of the NA and EL confidence regions with $\epsilon_i + 4 \sim \chi^2(4)$.

(ι, τ)	$W_n = M_n$	NA	EL	(ι, τ)	$W_n = M_n$	NA	EL
$(-2, -1)$	$grid_{49}$	0.894	0.862	$(-2, 1)$	$grid_{49}$	0.846	0.864
	$grid_{100}$	0.896	0.916		$grid_{100}$	0.908	0.920
	$grid_{169}$	0.928	0.936		$grid_{169}$	0.914	0.948
	$grid_{256}$	0.930	0.940		$grid_{256}$	0.924	0.955
	$grid_{400}$	0.942	0.956		$grid_{400}$	0.922	0.936
$(2, -1)$	$grid_{49}$	0.844	0.866	$(2, 1)$	$grid_{49}$	0.838	0.888
	$grid_{100}$	0.886	0.914		$grid_{100}$	0.888	0.894
	$grid_{169}$	0.912	0.924		$grid_{169}$	0.922	0.926
	$grid_{256}$	0.925	0.932		$grid_{256}$	0.932	0.940
	$grid_{400}$	0.944	0.950		$grid_{400}$	0.930	0.956

§4 Proofs

Lemma 1. *Suppose that Assumptions A1-A5 are satisfied. Then as $n \to \infty$,*

$$Z_n = \max_{1 \le i \le n} \|\omega_i(\theta)\| = o_p(n^{1/2}) \ a.s., \tag{12}$$

$$\Sigma_{k+2}^{-1/2} \sum_{i=1}^{n} \omega_i(\theta) \xrightarrow{d} N(0, I_{k+2}), \tag{13}$$

$$n^{-1} \sum_{i=1}^{n} \omega_i(\theta) \omega_i'(\theta) = n^{-1} \Sigma_{k+2} + o_p(1), \tag{14}$$

$$\sum_{i=1}^{n} \|\omega_i(\theta)\|^3 = O_p(n), \tag{15}$$

where Σ_{k+2} is given in (11).

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

Note that

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

By Conditions A1 and A2, we have

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

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

$$\max_{1 \le i \le n} |\tilde{h}_{ii} \left(\epsilon_i^2 - \sigma^2\right)| = \max_{1 \le i \le n} |\tilde{h}_{ii}| o_p(n^{1/2}) = o_p(n^{1/2}),$$

$$\max_{1 \le i \le n} |\epsilon_i \sum_{j=1}^{i-1} \tilde{h}_{ij} \epsilon_j| \le (\max_{1 \le i \le n} |\epsilon_i|)^2 \cdot \max_{1 \le i \le n} (\sum_{j=1}^{i-1} |\tilde{h}_{ij}|) = o_p(n^{1/2}).$$

Similarly,

$$\max_{1 \le i \le n} |\tilde{m}_{ii} \left(\epsilon_i^2 - \sigma^2\right)| = o_p(n^{1/2}), \quad \max_{1 \le i \le n} |\epsilon_i \sum_{j=1}^{i-1} \tilde{m}_{ij} \epsilon_j| = o_p(n^{1/2}).$$

Table 4. The mean, (SD) and [RMSE] for MESS model based on EL and NA estimations.

	$W_n = M_n$	$\iota_0 = -2$	$\tau_0 = -1$	$\beta_0 = 3.5$
$\epsilon_i \sim N(0,1)$	$n = 49$			
	EL	-1.993(0.034)[0.034]	-1.035(0.344)[0.340]	3.493(0.065)[0.064]
	NA	-2.001(0.040)[0.040]	-0.991(0.298)[0.317]	3.497(0.085)[0.085]
	$n = 100$			
	EL	-1.996(0.031)[0.031]	-0.963(0.206)[0.209]	3.500(0.070)[0.070]
	NA	-2.000(0.027)[0.027]	-0.945(0.201)[0.208]	3.499(0.066)[0.066]
	$n = 400$			
	EL	-1.999(0.014)[0.013]	-1.010(0.051)[0.046]	3.498(0.024)[0.023]
	NA	-2.000(0.012)[0.011]	-0.993(0.100)[0.093]	3.499(0.029)[0.027]
$\epsilon_i \sim t(5)$	$n = 49$			
	EL	-2.003(0.123)[0.122]	-1.010(0.350)[0.348]	3.478(0.168)[0.168]
	NA	-2.000(0.043)[0.043]	-0.926(0.299)[0.307]	3.499(0.092)[0.092]
	$n = 100$			
	EL	-1.999(0.050)[0.048]	-0.942(0.217)[0.223]	3.500(0.107)[0.106]
	NA	-1.999(0.035)[0.035]	-0.968(0.186)[0.188]	3.498(0.089)[0.089]
	$n = 400$			
	EL	-2.007(0.016)[0.016]	-1.049(0.079)[0.084]	3.483(0.022)[0.025]
	NA	-1.999(0.018)[0.018]	-0.993(0.102)[0.102]	3.499(0.038)[0.039]
$\epsilon_i + 4 \sim \chi^2(4)$	$n = 49$			
	EL	-1.994(0.251)[0.251]	-1.031(0.384)[0.384]	3.465(0.283)[0.284]
	NA	-2.001(0.108)[0.108]	-0.914(0.303)[0.314]	3.487(0.224)[0.224]
	$n = 100$			
	EL	-1.997(0.160)[0.161]	-0.958(0.279)[0.280]	3.490(0.184)[0.186]
	NA	-2.002(0.074)[0.074]	-0.954(0.201)[0.202]	3.493(0.176)[0.176]
	$n = 400$			
	EL	-1.999(0.011)[0.018]	-1.000(0.126)[0.121]	3.501(0.093)[0.091]
	NA	-1.999(0.037)[0.037]	-0.990(0.106)[0.106]	3.497(0.103)[0.103]

Thus $Z_n = o_p(n^{1/2})$, and (12) is proved.

We now prove (13). For any given $l = (l_1', l_2, l_3)' \in R^{k+2}$ with $\|l\| = 1$, where $l_1 \in R^k, l_2, l_3 \in R$. Then

$$\begin{aligned} l'\omega_i(\theta) =& l_1' a_i \epsilon_i + l_2\{b_i\epsilon_i + \tilde{h}_{ii}\left(\epsilon_i^2 - \sigma^2\right) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{h}_{ij,1}\epsilon_j\} \\ &+ l_3\{\tilde{m}_{ii}\left(\epsilon_i^2 - \sigma^2\right) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{m}_{ij}\epsilon_j\} \\ =& (l_1' a_i + l_2 b_i)\epsilon_i + (l_2\tilde{h}_{ii} + l_3\tilde{m}_{ii})(\epsilon_i^2 - \sigma^2) + 2\epsilon_i \sum_{j=1}^{i-1} (l_2\tilde{h}_{ij} + l_3\tilde{m}_{ij})\epsilon_j. \end{aligned}$$

Thus

$$\begin{aligned} \sum_{i=1}^{n} l'\omega_i(\theta) =& \sum_{i=1}^{n} (l_1' a_i + l_2 b_i)\epsilon_i + \sum_{i=1}^{n} (l_2\tilde{h}_{ii} + l_3\tilde{m}_{ii})(\epsilon_i^2 - \sigma^2) \\ &+ 2\sum_{i=1}^{n}\sum_{j=1}^{i-1} (l_2\tilde{h}_{ij} + l_3\tilde{m}_{ij})\epsilon_i\epsilon_j. \end{aligned}$$

Denote

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

where

$$u_{ii} = l_2\tilde{h}_{ii} + l_3\tilde{m}_{ii}, \quad u_{ij} = l_2\tilde{h}_{ij} + l_3\tilde{m}_{ij}(i \neq j), \quad \nu_i = l_1'a_i + l_2b_i.$$

Then

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

We firstly try to obtain the variance of Q_n. It can be shown that

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{n} u_{ij}^2 = &\sum_{i=1}^{n}\{(l_2\tilde{h}_{ii} + l_3\tilde{m}_{ii})^2 + \sum_{i\neq j}(l_2\tilde{h}_{ij} + l_3\tilde{m}_{ij})^2\}\\
&= l_2^2\mathrm{tr}(\tilde{H}_n^2) + l_3^2\mathrm{tr}(\tilde{M}_n^2) + 2l_2l_3\mathrm{tr}(\tilde{H}_n\tilde{M}_n),
\end{aligned}$$

$$\begin{aligned}
\sum_{i=1}^{n} u_{ii}^2 = &\sum_{i=1}^{n}(l_2\tilde{h}_{ii} + l_3\tilde{m}_{ii})^2\\
&= l_2^2\|Vec_D(\tilde{H}_n)\|^2 + l_3^2\|Vec_D(\tilde{M}_n)\|^2\\
&\quad + 2l_2l_3Vec_D'(\tilde{H}_n)Vec_D(\tilde{M}_n),
\end{aligned}$$

$$\begin{aligned}
\sum_{i=1}^{n}\nu_i^2 = &\sum_{i=1}^{n}(l_1'a_i + l_2b_i)^2\\
&= l_1'(\sum_{i=1}^{n} a_ia_i')l_1 + l_2^2\sum_{i=1}^{n} b_ib_i' + 2l_1'(\sum_{i=1}^{n} a_ib_i')l_2\\
&= l_1'\{e^{\tau M_n}X_n\}'e^{\tau M_n}X_nl_1 + l_2^2\{e^{\tau M_n}W_nX_n\beta\}'e^{\tau M_n}W_nX_n\beta\\
&\quad + 2l_1'l_2\{e^{\tau M_n}X_n\}'e^{\tau M_n}W_nX_n\beta,
\end{aligned}$$

and

$$\begin{aligned}
\sum_{i=1}^{n} u_{ii}\nu_i = &\sum_{i=1}^{n}(l_2\tilde{h}_{ii} + l_3\tilde{m}_{ii})(l_1'a_i + l_2b_i)\\
&= l_1'l_2\sum_{i=1}^{n}\tilde{h}_{ii}a_i + l_2^2\sum_{i=1}^{n}\tilde{h}_{ii}b_i + l_1'l_3\sum_{i=1}^{n}\tilde{m}_{ii}a_i + l_2l_3\sum_{i=1}^{n}\tilde{m}_{ii}b_i\\
&= l_1'l_2\{e^{\tau M_n}X_n\}'Vec_D(\tilde{H}_n) + l_2^2\{e^{\tau M_n}W_nX_n\beta\}'Vec_D(\tilde{H}_n)\\
&\quad + l_1'l_3\{e^{\tau M_n}X_n\}'Vec_D(\tilde{M}_n) + l_2l_3\{e^{\tau M_n}W_nX_n\beta\}'Vec_D(\tilde{M}_n).
\end{aligned}$$

Thus the variance of Q_n is

$$\sigma_{Q_n}^2 = 2\sum_{i=1}^{n}\sum_{j=1}^{n} u_{ij}^2\sigma^4 + \sum_{i=1}^{n}\nu_i^2\sigma^2 + \sum_{i=1}^{n}\{u_{ii}^2(\mu_4 - 3\sigma^4) + 2u_{ii}\nu_i\mu_3\} = l'\Sigma_{k+2}l,$$

where Σ_{k+2} is given in (11). From Condition A4, one can see that $n^{-1}\sigma_{Q_n} \geq c_1 > 0$. By Lemma 1 in Qin (2021) or Theorem 1 in Kelejian and Prucha (2001), we have

$$\frac{Q_n - E(Q_n)}{\sigma_{Q_n}} \xrightarrow{d} N(0,1).$$

Noting that $E(Q_n) = 0$, we thus have (13).

Next we will prove (14). i.e.,

$$n^{-1}\sum_{i=1}^{n}\{l'\omega_i(\theta)\}^2 = n^{-1}\sigma_{Q_n}^2 + o_p(1). \tag{16}$$

Let

$$G_{in} = l'\omega_i(\theta) = u_{ii}(\epsilon_i^2 - \sigma^2) + 2\sum_{j=1}^{i-1} u_{ij}\epsilon_i\epsilon_j + \nu_i\epsilon_i = u_{ii}(\epsilon_i^2 - \sigma^2) + R_i\epsilon_i, \tag{17}$$

where $R_i = 2\sum_{j=1}^{i-1} u_{ij}\epsilon_j + \nu_i$. Let $\mathcal{F}_0 = \{\phi, \Omega\}, \mathcal{F}_i = \sigma(\epsilon_1, \epsilon_2, \cdots, \epsilon_i), 1 \le i \le n$. Then $\{G_{in}, \mathcal{F}_i, 1 \le i \le n\}$ form a martingale difference array. Note that

$$\begin{aligned}
& n^{-1}\sum_{i=1}^{n}\{l'\omega_i(\theta)\}^2 - n^{-1}\sigma_{Q_n}^2 = n^{-1}\sum_{i=1}^{n}\left(G_{in}^2 - EG_{in}^2\right) \\
= \; & n^{-1}\sum_{i=1}^{n}\left\{G_{in}^2 - E\left(G_{in}^2 \mid \mathcal{F}_{i-1}\right) + E\left(G_{in}^2 \mid \mathcal{F}_{i-1}\right) - EG_{in}^2\right\} = n^{-1}S_{n1} + n^{-1}S_{n2},
\end{aligned} \tag{18}$$

where

$$S_{n1} = \sum_{i=1}^{n}\left\{G_{in}^2 - E(G_{in}^2 \mid \mathcal{F}_{i-1})\right\}, S_{n2} = \sum_{i=1}^{n}\left\{E\left(G_{in}^2 \mid \mathcal{F}_{i-1}\right) - EG_{in}^2\right\}.$$

And then we need to prove that

$$n^{-1}S_{n1} = o_p(1), \tag{19}$$

and

$$n^{-1}S_{n2} = o_p(1). \tag{20}$$

Obviously, it is sufficient to prove $n^{-2}E\left(S_{n1}^2\right) \to 0$ and $n^{-2}E\left(S_{n2}^2\right) \to 0$ separately. Since

$$G_{in}^2 = u_{ii}^2\left(\epsilon_i^2 - \sigma^2\right)^2 + R_i^2\epsilon_i^2 + 2u_{ii}R_i\left(\epsilon_i^2 - \sigma^2\right)\epsilon_i,$$

then

$$E\left(G_{in}^2 \mid \mathcal{F}_{i-1}\right) = u_{ii}^2 E\left(\epsilon_i^2 - \sigma^2\right)^2 + R_i^2\sigma^2 + 2u_{ii}R_i\mu_3.$$

It follows that

$$\begin{aligned}
& n^{-2}E\left(S_{n1}^2\right) = n^{-2}\sum_{i=1}^{n} E\Big\{G_{in}^2 - E\left(G_{in}^2 \mid \mathcal{F}_{i-1}\right)\Big\}^2 \\
= \; & n^{-2}\sum_{i=1}^{n} E\Big[u_{ii}^2\Big\{\left(\epsilon_i^2 - \sigma^2\right)^2 - E\left(\epsilon_i^2 - \sigma^2\right)^2\Big\} \\
& + R_i^2(\epsilon_i^2 - \sigma^2) + 2u_{ii}R_i\left(\epsilon_i^3 - \sigma^2\epsilon_i - \mu_3\right)\Big]^2 \\
\le \; & Cn^{-2}\sum_{i=1}^{n} E\Big[u_{ii}^4\left\{\left(\epsilon_i^2 - \sigma^2\right)^2 - E\left(\epsilon_i^2 - \sigma^2\right)^2\right\}^2\Big] \\
& + Cn^{-2}\sum_{i=1}^{n} E\left\{R_i^4\left(\epsilon_i^2 - \sigma^2\right)^2\right\} + Cn^{-2}\sum_{i=1}^{n} E\left\{u_{ii}^2R_i^2\left(\epsilon_i^3 - \sigma^2\epsilon_i - \mu_3\right)^2\right\}.
\end{aligned} \tag{21}$$

By Conditions A1, we have the following

$$\begin{aligned}
& n^{-2}\sum_{i=1}^{n} E\Big[u_{ii}^4\left\{\left(\epsilon_i^2 - \sigma^2\right)^2 - E\left(\epsilon_i^2 - \sigma^2\right)^2\right\}^2\Big] \\
\le \; & Cn^{-2}\sum_{i=1}^{n} u_{ii}^4 \le Cn^{-2}\sum_{i=1}^{n}\Big|l_2\tilde{h}_{ii} + l_3\tilde{m}_{ii}\Big|^4 \le Cn^{-1} \to 0,
\end{aligned} \tag{22}$$

and

$$
\begin{aligned}
& n^{-2}\sum_{i=1}^{n}E\left\{R_i^4\left(\epsilon_i^2-\sigma^2\right)^2\right\}\leq\ Cn^{-2}\sum_{i=1}^{n}E\Big(\sum_{j=1}^{i-1}u_{ij}\epsilon_j+\nu_i\Big)^4\\
\leq\ & Cn^{-2}\sum_{i=1}^{n}E\Big(\sum_{j=1}^{i-1}u_{ij}\epsilon_j\Big)^4+Cn^{-2}\sum_{i=1}^{n}\nu_i^4\leq Cn^{-2}\sum_{i=1}^{n}\sum_{j=1}^{i-1}u_{ij}^4\mu_4+Cn^{-2}\sum_{i=1}^{n}\Big(\sum_{j=1}^{i-1}u_{ij}^2\sigma^2\Big)^2\\
& +Cn^{-2}\sum_{i=1}^{n}\left(l_1'a_i+l_2b_i\right)^4\leq Cn^{-1}\to 0. \qquad (23)
\end{aligned}
$$

Similarly, we can prove that

$$
n^{-2}\sum_{i=1}^{n}E\left\{u_{ii}^2R_i^2\left(\epsilon_i^3-\sigma^2\epsilon_i-\mu_3\right)^2\right\}\to 0. \qquad (24)
$$

From (21)-(24), we have $n^{-2}E\left(S_{n1}^2\right)\to 0$. Furthermore,

$$
\begin{aligned}
EG_{in}^2&=E\left\{E\left(G_{in}^2\mid\mathcal{F}_{i-1}\right)\right\}=u_{ii}^2E\left(\epsilon_i^2-\sigma^2\right)^2+\sigma^2E\left(R_i^2\right)+2u_{ii}\mu_3E\left(R_i\right)\\
&=u_{ii}^2E\left(\epsilon_i^2-\sigma^2\right)^2+\sigma^2\Big(4\sum_{j=1}^{i-1}u_{ij}^2\sigma^2+\nu_i^2\Big)+2u_{ii}\mu_3\nu_i.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
& n^{-2}E\left(S_{n2}^2\right)=n^{-2}E\Big[\sum_{i=1}^{n}\left\{E\left(G_{in}^2\mid\mathcal{F}_{i-1}\right)-EG_{in}^2\right\}\Big]^2\\
=\ & n^{-2}E\Big[\sum_{i=1}^{n}\Big\{R_i^2\sigma^2-\sigma^2\Big(4\sum_{j=1}^{i-1}u_{ij}^2\sigma^2+\nu_i^2\Big)+2u_{ii}\mu_3\left(R_i-\nu_i\right)\Big\}\Big]^2\\
=\ & n^{-2}\sum_{i=1}^{n}E\Big[\sigma^2\Big\{\Big(2\sum_{j=1}^{i-1}u_{ij}\epsilon_j\Big)^2-4\sum_{j=1}^{i-1}u_{ij}^2\sigma^2\Big\}+4\Big(\sum_{j=1}^{i-1}u_{ij}\epsilon_j\Big)\nu_i\sigma^2+2u_{ii}\mu_3\Big(2\sum_{j=1}^{i-1}u_{ij}\epsilon_j\Big)\Big]^2\\
\leq\ & Cn^{-2}\sum_{i=1}^{n}E\Big\{\sigma^2\Big(\sum_{j=1}^{i-1}u_{ij}\epsilon_j\Big)^2-\sum_{j=1}^{i-1}u_{ij}^2\sigma^2\Big\}^2+Cn^{-2}\sum_{i=1}^{n}E\Big\{\Big(\sum_{j=1}^{i-1}u_{ij}\epsilon_j\Big)\nu_i\sigma^2\Big\}^2\\
& +Cn^{-2}\sum_{i=1}^{n}E\Big\{2u_{ii}\mu_3\Big(\sum_{j=1}^{i-1}u_{ij}\epsilon_j\Big)\Big\}^2. \qquad (25)
\end{aligned}
$$

Note that

$$
\begin{aligned}
& n^{-2}\sum_{i=1}^{n}E\Big[\sigma^2\Big\{\Big(\sum_{j=1}^{i-1}u_{ij}\epsilon_j\Big)^2-\sum_{j=1}^{i-1}u_{ij}^2\sigma^2\Big\}\Big]^2\leq n^{-2}\sigma^4\sum_{i=1}^{n}E\Big(\sum_{j=1}^{i-1}u_{ij}\epsilon_j\Big)^4\\
\leq\ & Cn^{-2}\sum_{i=1}^{n}\sum_{j=1}^{i-1}u_{ij}^4\mu_4+Cn^{-2}\sum_{i=1}^{n}\Big(\sum_{j=1}^{i-1}u_{ij}^2\sigma^2\Big)^2\leq Cn^{-1}\to 0, \qquad (26)
\end{aligned}
$$

$$
n^{-2}\sum_{i=1}^{n}E\Big\{\Big(\sum_{j=1}^{i-1}u_{ij}\epsilon_j\Big)\nu_i\sigma^2\Big\}^2=n^{-2}\sigma^6\sum_{i=1}^{n}\nu_i^2\sum_{j=1}^{i-1}u_{ij}^2\leq Cn^{-2}\to 0, \qquad (27)
$$

and

$$
n^{-2}\sum_{i=1}^{n}E\Big\{2u_{ii}\mu_3\Big(\sum_{j=1}^{i-1}u_{ij}\epsilon_j\Big)\Big\}^2=4\mu_3^2\sigma^2n^{-2}\sum_{i=1}^{n}u_{ii}^2\sum_{j=1}^{i-1}u_{ij}^2\leq Cn^{-1}\to 0, \qquad (28)
$$

where we have used Conditions A1 and A2. Form (25)-(28), we have $n^{-2}E(S_{n2})^2 \to 0$. The proof of (14) is thus completed.

Finally, we prove (15). Note that

$$\begin{aligned}\sum_{i=1}^{n} E\|\omega_i(\theta)\|^3 &\le \sum_{i=1}^{n} E\|a_i\epsilon_i\|^3 + \sum_{i=1}^{n} E\Big|b_i\epsilon_i + \tilde{h}_{ii}(\epsilon_i^2-\sigma^2) + 2\epsilon_i\sum_{j=1}^{i-1}\tilde{h}_{ij}\epsilon_j\Big|^3 \\ &\quad + \sum_{i=1}^{n} E\Big|\tilde{m}_{ii}(\epsilon_i^2-\sigma^2) + 2\epsilon_i\sum_{j=1}^{i-1}\tilde{m}_{ij}\epsilon_j\Big|^3. \end{aligned} \tag{29}$$

Combining Conditions A1 and A2, we have

$$\sum_{i=1}^{n} E\|a_i\epsilon_i\|^3 \le Cn\Big(\max_{1\le i\le n}\|a_i\|\Big)^3 E\left|\epsilon_i\right|^3 = O\left(n\right), \tag{30}$$

$$\begin{aligned} &\sum_{i=1}^{n} E\Big|b_i\epsilon_i + \tilde{h}_{ii}(\epsilon_i^2-\sigma^2) + 2\epsilon_i\sum_{j=1}^{i-1}\tilde{h}_{ij}\epsilon_j\Big|^3 \\ \le\ & C\sum_{i=1}^{n} E\left|b_i\epsilon_i\right|^3 + C\sum_{i=1}^{n} E\left|\tilde{h}_{ii}(\epsilon_i^2-\sigma^2)\right|^3 + C\sum_{i=1}^{n} E\Big|2\epsilon_i\sum_{j=1}^{i-1}\tilde{h}_{ij}\epsilon_j\Big|^3 \\ \le\ & C\sum_{i=1}^{n} E\left|b_i\epsilon_i\right|^3 + C\sum_{i=1}^{n} E\left|\tilde{h}_{ii}(\epsilon_i^2-\sigma^2)\right|^3 \\ & +C\sum_{i=1}^{n} E\left|\epsilon_i\right|^3\sum_{j=1}^{i-1} E\left|\tilde{h}_{ij}\epsilon_j\right|^3 + C\sum_{i=1}^{n} E\left|\epsilon_i\right|^3\Big\{\sum_{j=1}^{i-1} E\left(\tilde{h}_{ij}\epsilon_j\right)^2\Big\}^{3/2} \\ =\ & O\left(n\right). \end{aligned} \tag{31}$$

Similarly,

$$\sum_{i=1}^{n} E\Big|\tilde{m}_{ii}(\epsilon_i^2-\sigma^2) + 2\epsilon_i\sum_{j=1}^{i-1}\tilde{m}_{ij}\epsilon_j\Big|^3 = O\left(n\right). \tag{32}$$

From (29)-(32), one can prove that

$$\sum_{i=1}^{n} E\left\|\omega_i(\theta)\right\|^3 = O\left(n\right). \tag{33}$$

Using (33) and Markov inequality, we have $\sum_{i=1}^{n}\left\|\omega_i(\theta)\right\|^3 = O_p\left(n\right)$. Thus (15) is proved. □

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

§5 Conclusions

In this paper, we study EL method for spatial cross-sectional data models in the form of MESS. By using the QML method and then changing the quadratic forms of the error terms into the linear forms of martingale difference arrays to derive the estimating functions for the EL method, the EL ratio statistics are established for the parameters of the MESS model. The limiting distributions of EL ratio statistics are then obtained and we use this result to construct the confidence regions of model parameters. Simulation results show that the computational efficiency of EL method is better than NA method and thus EL method is recommended in constructing confidence regions of parameters. We will study EL method for spatial panel data models in the form of MESS in our future work.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

[1] L Anselin. *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, The Netherlansd, 1988.

[2] S X Chen, I V Keilegom. *A review on empirical likelihood methods for regression (with discussions)*, Test, 2009, 18: 415-447.

[3] T Y M Chiu, T Leonard, K W Tsui. *The matrix-logarithmic covariance model*, J Amer Statist Assoc, 1996, 91: 198-210.

[4] N Debarsy, F Jin, L F Lee. *Large sample properties of the matrix exponential spatial specification with an application to FDI*, J Econometrics, 2015, 188: 1-21.

[5] C Figueiredo, A R D Silva. *A matrix exponential spatial specification approach to panel data models*, Empir Econ, 2015, 49: 115-129.

[6] X Han, L F Lee. *Model selection using J-test for the spatial autoregressive model vs. the matrix exponential spatial model*, Regional Science and Urban Economics, 2013 43: 250-271.

[7] F Jin, L F Lee. *GEL estimation and tests of spatial autoregressive models*, Journal of Econometrics, 2019, 208: 585-612.

[8] H H Kelejian, I R Prucha. *On the asymptotic distribution of the Moran I test statistic with applications*, J Econometrics, 2001, 104: 219-257.

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

[10] J P LeSage, R K Pace. *A matrix exponential spatial specification*, J Econometrics, 2007, 140: 190-214.

[11] T Liu. *Model selection and adaptive lasso estimation of spatial models (Ph.D)*, Ohio State University, 2017.

[12] Y Li, Y Li, Y Qin. *Empirical likelihood for panel data models with spatial errors*, Communications in Statistics-Theory and Methods, 2022, 51(9): 2838-2857.

[13] A B Owen. *Empirical likelihood ratio confidence intervals for a single functional*, Biomertika, 1988, 75: 237-249.

[14] A B Owen. *Empirical likelihood ratio confidence regions*, Ann Statist, 1990, 18(1): 90-120.

[15] A B Owen. *Empirical likelihood for linear models*, Ann Statist, 1991, 19(4): 1725-1747.

[16] A B Owen. *Empirical likelihood*, Chapman & Hall, London, 2001.

[17] P Piribauer, M M Fischer. *Model Uncertainty in Matrix Exponential Spatial Growth Regression Models*, Geographical Analysis, 2014, 47: 240-261.

[18] J Qin, J Lawless. *Empirical likelihood and general estimating equations*, Ann Statist, 1994, 22: 300-325.

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

[20] Y Zhang, S Feng, F Jin. *QML estimation of the matrix exponential spatial specification panel data model with fixed effects and heteroskedasticity*, Economics Letters, 2019, 180: 1-5.

[1]Nanning University, Nanning 530200, China.
Email: liuyan@unn.edu.cn
[2]Center for Applied Mathematics of Guangxi and College of Mathematics and Statistics, Guangxi Normal University, Guilin 541004, China.
Email: ysqin@gxnu.edu.cn