Evaluation Of MLE And TSLS Methods For Fitting Spatial Lag Model And It’s Application

by Budi Pratikno

Far East Journal of Mathematical Sciences (FJMS)

http://www.pphmj.com
http://dx.doi.org/10.17654/MS107010037
Volume 107, Number 1, 2018, Pages 37-52 ISSN: 0972-0871

EVALUATION OF MLE AND TSLS METHODS FOR FITTING SPATIAL LAG MODEL AND ITS APPLICATION

Jajang, B. Pratikno and Mashuri

Department of Mathematics
Jenderal Soedirman University (UNSOED)
Purwokerto 53122, Indonesia

Abstract

In this paper, we study two parameter estimation methods, maximum likelihood estimation (MLE) and two-stage least squares (TSLS) for fitting spatial lag model (SLM). We use Monte Carlo simulation to generate the data used and evaluate the methods by the root of mean squared error (RMSE) criterion. Statistical analysis of real data sets is presented to demonstrate the conclusion of the results. In this data, we use poverty data and factors affecting poverty which consist of the number of people who graduated from junior high school, shares of industry GDP, agricultural GDP, trading GDP and services GDP. The result shows that the best choice in fitting SLM is MLE. Analysis results based on MLE method conclude that the increase in share of agriculture GDP causes significant increase in the number of poor people. Conversely, the increase in share of services GDP causes significant decrease in the number of poor people. The increase in the number of people who graduated from junior high school, share of industry GDP, share of trading GDP causes decrease in the number of poor people, but they are not significant.

Received: December 12, 2017; Revised: July, 26 2018; Accepted: July 30, 2018
2010 Mathematics Subject Classification: 62H11, 65C05.
Keywords and phrases: MLE, TSLS, SLM, poverty.

1. Introduction

In recent years, spatial regression model has been developed to take spatial dependence. The models that involve statistical dependence are often more realistic [7, 8]. A fundamental concern of spatial analysts is to find patterns in spatial data that lead to the identification of spatial autocorrelation or association [15]. Taking spatial dependences into account when dealing with spatial data is very important, and neglecting them can cause problems. For example, ignoring spatial lag structures causes ordinary least squares (OLS) estimators to become bias and inconsistent. The spatial weights matrix is one of the most convenient ways to summarize spatial relationship in the data. Spatial weights matrix is a nonnegative matrix that specifies the neighborhood set for each observation. Here, the data are collected from different spatial locations. Spatial weights matrix characterizes cross-section dependence in useful ways. Their measurement has an important effect on the estimation of a spatial dependence model [1, 9, 14]. The prediction result becomes accurate if we found a representative spatial weight matrix and parameter estimation method. There are many to create spatial weights matrix [10]. However, the most commonly used spatial weights matrix is a binary matrix based on geographic distance and contiguity.

In the spatial model, we found endogenous problem in the model. Therefore, classic method such as ordinary least squares (OLS) is not relevant to solve the problem. The OLS estimator will be biased as well as inconsistent for the parameters of the spatial model [2]. The inappropriateness of the least squares estimator for models that incorporate spatial dependence has focused attention on the maximum likelihood estimation (MLE), generalized method of moment (GMM) and two-stage least squares (TSLS) methods approach as an alternative [11, 12]. In this paper, we characterize and compare MLE and TSLS methods to estimate the parameters of SLM model.

The rest of the paper is organized as follows: Section 2 presents spatial lag model and parameter estimation methods. Monte Carlo simulation is given in Section 3. Statistical analysis of real case sets is presented in Section 4. Concluding remarks are provided in Section 5.

2. Spatial Lag Model and Parameter Estimation Method

2.1. Spatial lag model

Spatial lag dependence or spatial lag model in a regression model is similar to the inclusion of serially autoregressive term for dependent variable in a time-series context. Spatial lag model (SLM) is specified as [2, 3]

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{1}$$

where $\mathbf{y}$ is the $n \times 1$ of the response variable, $\mathbf{X}$ is the $n \times 1$ matrix of the non-stochastic explanatory variables, $\mathbf{W}$ is the $n \times 1$ non-stochastic weights matrix, ρ is a spatial autoregressive parameter, $\boldsymbol{\beta}$ is a parameter vector, and $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_i, \dots, \varepsilon_n)$ is an $n \times 1$ vector of innovations.

2.2. Maximum likelihood estimation

Maximum likelihood estimation of the SLM models described involves maximizing the log-likelihood function with respect to the parameters. The method of maximum likelihood selects the set of values-values of the model parameters that maximize the likelihood function. The model (1) represents an equilibrium, so $(I - \rho W)$ is assumed to be invertible. The equilibrium vector $\mathbf{y}$ is given by $\mathbf{y} = (\mathbf{I} - \rho \mathbf{W})^{-1} (\mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon})$. It follows that $\mathbf{W}\mathbf{y} = \mathbf{W}(\mathbf{I} - \rho \mathbf{W})^{-1} (\mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon})$. We assume that the errors are normally distributed $(\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{I}\sigma^2)) \sup_i | \omega_i | < 1$ and the matrices $(\mathbf{I} - \rho \mathbf{W})$ are nonsingular. For errors are normally distributed, it can be expressed by

$$f(\boldsymbol{\varepsilon}) = \frac{1}{(2\pi)^{n/2} \sigma^n} \exp\left[-\frac{\boldsymbol{\varepsilon}' \boldsymbol{\varepsilon}}{2\sigma^2} \right]. \tag{2}$$

Based on nonsingularity condition of the matrix $(\mathbf{I} - \rho \mathbf{W})$, we can rewrite equation in (1) as $\mathbf{y} = (\mathbf{I} - \rho \mathbf{W})^{-1} (\mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon})$ or $\boldsymbol{\varepsilon} = (\mathbf{I} - \rho \mathbf{W}) \mathbf{y} - \mathbf{X} \boldsymbol{\beta}$. Furthermore, from (2), we can find pdf of $\mathbf{y}$ by Jacobi transformation method

[4, 6]. Let $J = \left|\frac{d\varepsilon}{dy}\right| = |\mathbf{I} - \rho\mathbf{W}|$ be the Jacobi transformation from ε to $\mathbf{y}$. By using the Jacobi method, we can denote pdf $\mathbf{y}$ as

$$f(\mathbf{y}) = f(\boldsymbol{\varepsilon})|\mathbf{J}|. \tag{3}$$

Hereinafter, substituting $\mathbf{J} = |\mathbf{I} - \rho\mathbf{W}|$ and $\boldsymbol{\varepsilon} = (\mathbf{I} - \rho\mathbf{W})\mathbf{y} - \mathbf{X}\beta$ into $f(\boldsymbol{\varepsilon})$ in equation (2) results in

$$f(\mathbf{y}) = \frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left[\frac{((\mathbf{I} - \rho\mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'((\mathbf{I} - \rho\mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta})}{2\sigma^2}\right]\cdot|\mathbf{I} - \varrho\mathbf{W}|. \tag{4}$$

We can see that equation (4) has ρ, σ^2 and $\boldsymbol{\beta}$. Furthermore, we use $f(\mathbf{y}; \rho, \sigma^2, \beta)$ instead of $f(\mathbf{y})$ for expression of the pdf. Thus, with reference to (4), the likelihood function is

$$\mathcal{L}(\rho, \sigma^2, \boldsymbol{\beta}\mathbf{y})$$
$$= \frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left[\frac{((\mathbf{I} - \rho\mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'((\mathbf{I} - \rho\mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta})}{2\sigma^2}\right]\cdot|\mathbf{I} - \varrho\mathbf{W}|. \tag{5}$$

The expression in (5) is actually quite a pain to differentiate, so it is almost always simplified by taking the natural logarithm of the expression. This is absolutely fine because the natural logarithm is a monotonically increasing function. This is important because it ensures that the maximum value of the log of the probability occurs at the same point as the original probability function. Therefore, we can work with the simpler log-likelihood instead of the original likelihood. The logarithm of likelihood function of (5) can be denoted as

$$\ln(\mathcal{L}(\rho, \sigma^2, \boldsymbol{\beta}\mathbf{y})) = \ln|\mathbf{I} - \rho\mathbf{W}| - \frac{n}{2}\ln(2\pi)$$
$$- \frac{n}{2}\ln\sigma^2 - \frac{((\mathbf{I} - \rho\mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'((\mathbf{I} - \rho\mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta})}{2\sigma^2}. \tag{6}$$

There are the following requirements: the existence of the log-likelihood function for the parameter values under consideration, continuous differentiability of the log-likelihood, boundedness of various partial derivatives; the existence, nonsingularity of covariance matrices; and the finiteness of various quadratic forms. Here, there are the conditions to ensure that these assumptions hold. These conditions are all diagonal elements of $\mathbf{W}$ are zero, $\sup_i |\omega_i| < 1$, the matrices $(\mathbf{I} - \rho\mathbf{W})$ are nonsingular for $-1 < \omega_i < 1$, $i = 1, 2, \ldots n$. The innovations ε_i are independent identically distribution, $E(\varepsilon_i) = 0$, $E(\varepsilon_i^2) = \sigma^2 > 0$, and $E(|\varepsilon_i|^{4+\eta}) < \infty$, for some η. To avoid calculating $\ln|\mathbf{I} - \rho\mathbf{W}|$ in (6), Ward and Kristiani [17] proposed that $\ln|\mathbf{I} - \rho\mathbf{W}| = \sum_i \ln(1 - \rho\omega_i)$,

$$\ln(\mathcal{L}(\rho, \sigma^2, \boldsymbol{\beta}; \mathbf{y})) = \sum_i \ln(\mathbf{I} - \rho\omega_i) - \frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{((\mathbf{I} - \rho\mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'((\mathbf{I} - \rho\mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta})}{2\sigma^2}, \quad (7)$$

where $\omega_i, i = 1, 2, \ldots, n$ are eigenvalues of the matrix $\mathbf{W}$. The resulting vector of first partial derivatives of equation (7) is a set equal to zero and needs to be solved for the parameter values.

Let $\theta = (\rho, \sigma^2, \boldsymbol{\beta})$ and $\hat{\theta} = (\hat{\rho}, \hat{\sigma}^2, \hat{\boldsymbol{\beta}})$. The values $\hat{\rho}, \hat{\sigma}^2$ and $\hat{\boldsymbol{\beta}}$ are estimators for ρ, σ^2, respectively. The maximum value of $\ln(\mathcal{L}(\rho, \sigma^2, \boldsymbol{\beta}\mathbf{y}))$ is obtained when we consider $\hat{\rho}, \hat{\sigma}^2$ and $\hat{\boldsymbol{\beta}}$.

To obtain maximum of the function in (7), we have to find critical points by partial differential [16], $\frac{\partial}{\partial\theta}\ln(\mathcal{L}(\theta; \mathbf{y}))|_{\theta=\hat{\theta}} = 0$,

$$\frac{\partial}{\partial\boldsymbol{\beta}}\ln(\mathcal{L}(\rho, \sigma^2, \boldsymbol{\beta}; \mathbf{y})) = 0 \Leftrightarrow ((\mathbf{I} - \rho\mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{X} = 0, \quad (8)$$

$$\frac{\partial}{\partial\rho}\ln(\mathcal{L}(\rho, \sigma^2, \boldsymbol{\beta}; \mathbf{y}))$$

$$= \mathbf{0} \Leftrightarrow -tr(\mathbf{I}-\rho\mathbf{W})^{-1}\mathbf{W} + ((\mathbf{I}-\rho\mathbf{W})\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{W}\mathbf{y} = \mathbf{0}. \tag{9}$$

Under the usual regularity conditions, the maximal likelihood estimates that are found as solutions to the system (8)-(9) will be asymptotically efficient [2]. Clearly, this system of highly nonlinear equations does not have an analytic solution and needs to be solved by numerical methods. Part of the first order conditions has solution which can be used to construct a concentrated likelihood function. There are some methods/algorithm can be used to solve this nonlinear equation, such as Newton-Raphson, gradient descent and Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. In this paper, we use BFGS algorithm in optimization problem in equation (7).

Let H and ∇ be Hessian matrices and gradients operator, respectively. A starting point $\theta^{(0)}$ and estimate of $\nabla^2 \ln(\theta^{(0)})$ must be given. The iteration is then $k = 1, 2, 3, \ldots$

(1) $\theta^{(k+1)} = \theta^{(k)} - H_k^{-1}\nabla\ln(\theta^{(k)})$

(2) $s^{(k)} = \theta^{(k+1)} - \theta^{(k)}$

(3) $y^{(k)} = \nabla\ln(\theta^{(k+1)}) - \nabla\ln(\theta^{(k)})$

(4) $H_{k+1} = H_k - \dfrac{H_k s^{(k)}(s^{(k)})' H_k}{s^{(k)} H_k s^{(k)}} + \dfrac{y^{(k)}(y^{(k)})'}{y^{(k)} s^{(k)}}.$

2.3. Two-stage least squares

Two-stage least squares (TSLS) method is one of the methods that can be used to solve the endogenous problem. In the TSLS method is used the instrument variable. The instrument variable is a new variable correlated to the response variable, but uncorrelated with residual. From equation (1), we can express $\mathbf{y} = (\mathbf{I}-\rho\mathbf{W})^{-1}(\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon})$, so we can show that $E(\mathbf{W}\mathbf{y}\boldsymbol{\varepsilon}') =$

$E(\mathbf{W}(\mathbf{I} - \rho\mathbf{W})^{-1}\boldsymbol{\varepsilon\varepsilon}') \neq 0$. Therefore, we cannot use OLS method to estimate the parameters of the SLM model. As an alternative, we can solve this problem by using TSLS method. Let $Z = (\mathbf{WyX})$ and $\boldsymbol{\theta} = \begin{pmatrix} \rho \\ \boldsymbol{\beta} \end{pmatrix}$. Then the model in equation (1) can be written as

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\theta} + \boldsymbol{\varepsilon}. \tag{10}$$

Due to $E(\mathbf{Wy}\boldsymbol{\varepsilon}') \neq 0$, Kelejian and Prucha [11] suggested a TSLS based on instruments $\mathbf{H} = (\mathbf{X}, \mathbf{WX}, \mathbf{W^2X}, \mathbf{W^3X}, \ldots)$. Therefore, the estimator of $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}} = (\mathbf{Z'H}(\mathbf{H'H})^{-1}\mathbf{H'Z})^{-1}\mathbf{Z'H}(\mathbf{H'H})^{-1}\mathbf{H'y}. \tag{11}$$

The TSLS method can be produced in three steps [13]:

1. Obtain the consistent estimates of $\boldsymbol{\beta}$ by instrument variables, where $\mathbf{X}, \mathbf{WX}, \mathbf{W^2X}$ are instruments in SLM.

2. Estimate ρ and σ by GMM using samples constructed from the functions of model errors.

3. Use estimate of ρ and σ to perform a spatial Cochrane-Orcutt transformation of the data and obtain more efficient estimate of β.

3. Monte Carlo Simulation

In this study, we simulated data of the SLM model used by Monte Carlo simulation method. Here, we set intercept $\beta_0 = 2$, slope $\beta_1 = 2$ for different coefficients of spatial $\text{lag}(\rho)$ and sample size $(n) \ldots \rho = 0.3, 0.4, 0.5, 0.6, 0.7$ and 0.8. Furthermore, the process of data generation that is used to evaluate the estimation methods is conducted as follows:

(1) Given $\mathbf{W}$ as a contiguity matrix.

(2) Fix the parameters $\beta_0 = 2$, $\beta_1 = 2$ and $\rho = 0.3, 0.4, 0.5, 0.6, 0.7$ and 0.8.

(3) Generate X explanatory variables: uniform, $U(20, 50)$.

(4) Generate $\varepsilon_N(t) \sim iid\ N(0, I)$.

(5) Generate $y_N(t)$ from equation (1).

(6) Estimate the parameters β_0, β_1 and ρ by MLE and TSLS.

(7) Repeat stage (1) until (6) B time $(B = 500)$.

(8) Determine average of RMSE of MLE's and RMSE of TSLS's.

The RMSE's MLE and RMSE's TSLS yielded from Monte Carlo simulation for various sample sizes (n) and coefficients of spatial lag(ρ) are listed in Table 1.

Table 1. The RMSE's MLE and RMSE's TSLS for variations n and ρ

	$\rho = 0.3$		$\rho = 0.4$		$\rho = 0.5$		$\rho = 0.6$		$\rho = 0.7$		$\rho = 0.8$	
n	MLE	TSLS	MLE	TSLS	MLE	TSLS	MLE	TSLS	MLE	TSLS	MLE	TSLS
20	1.024	1.111	1.175	1.274	1.436	1.558	1.899	2.058	2.728	2.950	4.412	4.728
40	0.984	1.023	1.025	1.066	1.112	1.157	1.278	1.329	1.621	1.685	2.445	2.536
60	1.008	1.035	1.071	1.098	1.186	1.216	1.403	1.439	1.833	1.880	2.853	2.920
80	1.047	1.068	1.143	1.165	1.315	1.340	1.649	1.681	2.297	2.340	3.684	3.747
100	1.095	1.112	1.240	1.259	1.500	1.523	1.965	1.995	2.824	2.866	4.674	4.736
200	1.202	1.211	1.455	1.466	1.876	1.890	2.583	2.602	3.837	3.864	6.414	6.455
300	1.005	1.010	1.023	1.028	1.058	1.064	1.134	1.139	1.310	1.317	1.772	1.780
400	1.012	1.016	1.045	1.049	1.099	1.103	1.220	1.224	1.482	1.488	2.135	2.142
500	1.030	1.033	1.071	1.074	1.158	1.161	1.335	1.339	1.703	1.708	2.555	2.562

Based on Table 1, we can see that RMSE's MLE methods are smaller than RMSE's TSLS, e.g., for $\rho = 0.3$ and $n = 100$, the RMSE's MLE = 1.095 is smaller than the RMSE's TSLS $= 1.112$. Further, all of RMSE's

MLE are smaller than RMSE's TSLS for various n and ρ. However, the differences between RMSE's MLE and RMSE's TSLS are small, especially for n large. These results are consistent with the properties of MLE estimation method. In this data simulation, the errors are normally distributed, so we can say that in that case, MLE is most efficient.

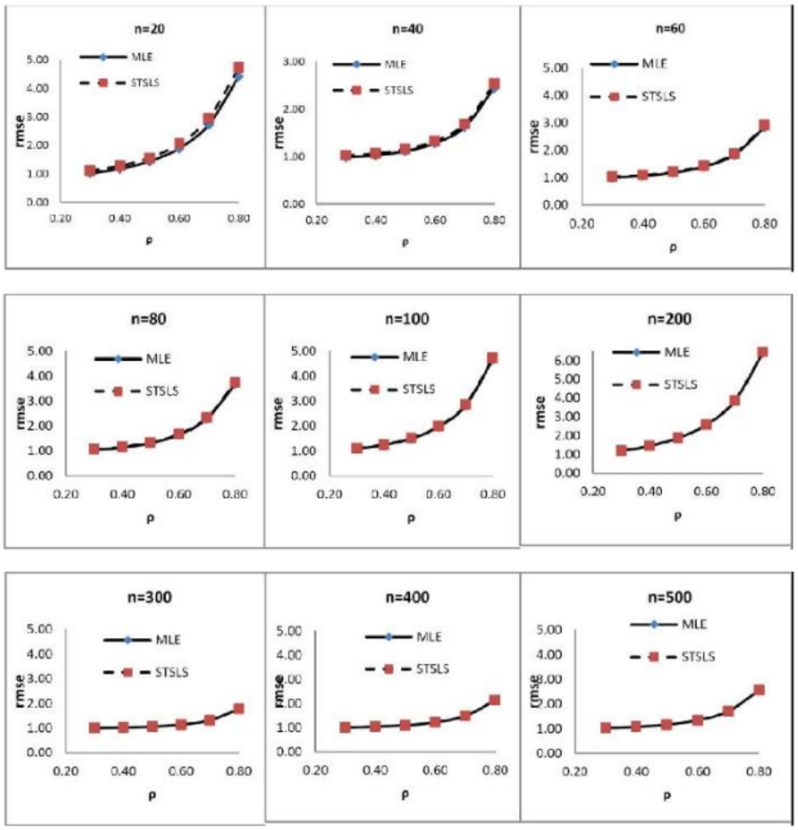

Figure 1. RMSE's MLE and RMSE's TSLS of the SLM model from various sample sizes (n) and coefficients of spatial lag(ρ).

Figure 1 shows the RMSE from different parameter estimation methods for $n = 20$ to 500 and $\rho = 0.2$ to 0.8. For $n = 20$, we see that RMSE's

MLE is smaller than RMSE's TSLS, whereas for $n > 20$, the differences of RMSE's MLE and RMSE's TSLS are very small. Further, all of the plots RMSE vs coefficients of spatial $\text{lag}(\rho)$ have trend positive, so we can say that the increase in coefficients of spatial lag causes increase in RMSEs.

4. Implementation of Estimation Methods

4.1. Data

The data used in this study were taken from BPS statistics of Central Java province [5]. The estimation method of spatial lag model is implemented to poverty data in Central Java province. The poverty data consisted of one response variable and five predictor variables. The response variable is the number of poor people, whereas five predictor variables are percentage of the number of people who graduated from junior high school (EDU), share of industry GDP (industry), share of agricultural GDP (AGRI), share of trading GDP (trading), and share of services GDP (services). Firstly, we describe distribution of the number of poor people in Central Java province using quantile map analysis. Here, we use GeoDa software version 1.8 to create quantile map. Figure 2 shows the distribution of the number of poor people for 35 districts. First class is described by the white color which represents the number of poor less than 9.87%. The second class is described by light orange color which represents the number of poor people between 10.9%-12.6%. The third class is described by the orange color which represents the number of poor people between 13%-14.9%. Finally, the fourth class is described by the dark orange color which represents the number of poor more than 14.9 percent (Figure 2).

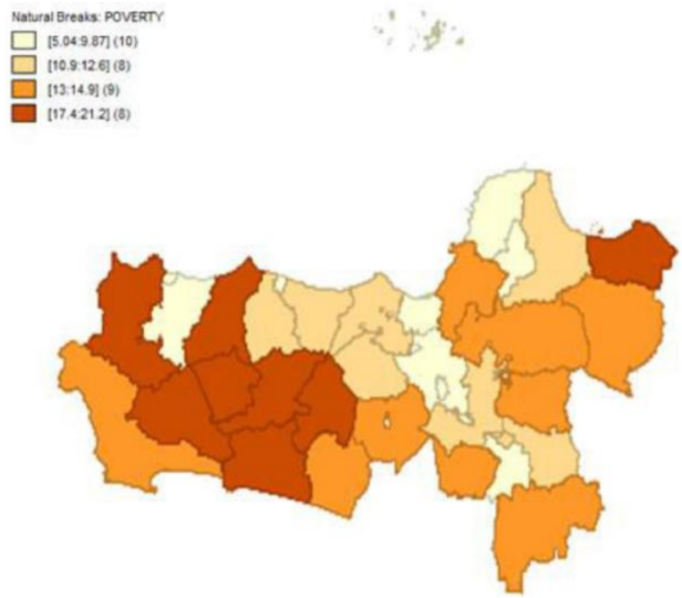

Figure 2. Quantile map of poverty data in Central Java province.

4.2. Model fitting

After a descriptive analysis, the next analysis is to find a relevant model for poverty and its factors. Here, we specified spatial lag model (SLM) and spatial error model (SEM). The SLM and SEM models for this poverty data are expressed in equations (12) and (13):

$$y_i = \rho \sum_{j \neq i}^{n} w_{ij} y_j + \beta_0 + \beta_1 EDU_i + \beta_2 IND_i + \beta_3 TRD_i + \beta_4 AGR_i + \beta_5 SER_i + \varepsilon_i, \quad i = 1, 2, \ldots, n \tag{12}$$

and

$$y_i = \beta_0 + \beta_1 EDU_i + \beta_2 IND_i + \beta_3 TRD_i + \beta_4 AGR_i + \beta_5 SER_i + \varepsilon_i,$$

$$\varepsilon_i = \lambda \sum_{j \neq i}^{n} w_{ij} \varepsilon_j + \nu, \quad i = 1, 2, \ldots, n. \tag{13}$$

To choose the best model between SLM (12) and SEM (13), we use the statistics LM_{lag} and LM_{error}. The statistics LM_{lag} and LM_{error} are expressed in equations (14) and (15):

$$\text{LM}_{\text{lag}} = \frac{[\varepsilon' W\varepsilon/(\varepsilon'\varepsilon)/n]^2}{[(WX\beta)'(I - X(X'X)^{-1}X')WX\beta] + tr(W^2 + W'W)} \tag{14}$$

and

$$\text{LM}_{\text{error}} = \frac{[\varepsilon' W\varepsilon/(\varepsilon'\varepsilon)/n]^2}{tr(W^2 + W'W)}. \tag{15}$$

The distributions of LM_{lag} and LM_{error} are Chi-square distribution with degree of freedom 1. We use minimum p value to choose the model. The values of statistic, parameter and p value of LM_{lag} and LM_{error} are listed in Table 2.

Table 2. Statistic test of SLM and SEM models

Statistic	Parameter	p value
LM_{lag}	4.0388	0.045
LM_{error}	6.1162	0.013

We can see that the p value of LM_{lag} is less than LM_{error}. Therefore, we use SLM for modeling poverty data. After choosing SLM model, we estimate parameter of SLM. Here, we again use MLE and TSLS parameter estimation methods. Figure 3 shows RMSE's MLE and RMSE's TSLS of the SLM model. There are two graphs in the chart. The orange graph shows the RMSE of MLE method and the blue graph shows RMSE of TSLS method. We can see that RMSE of MLE is less than RMSE of TSLS method. Therefore, we can say that the MLE is better than TSLS for modeling the poverty data.

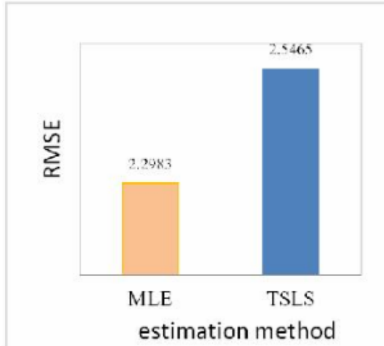

Figure 3. RMSE of MLE and TSLS methods of poverty modeling in Central Java province.

Based on Figure 3, we then use ANOVA's MLE for modeling and analysis. Table 3 shows ANOVA's MLE of SLM model.

Table 3. Analysis of variance of poverty model

Variable	Coefficient	Std. error	*z*-value	Probability
W. poverty	0.4591	0.1560	2.9429	0.00160
Constant	24.23953	10.3837	2.3344	0.01958
EDU	−0.25014	0.13428	−1.8628	0.06248
IND	−0.04021	0.04915	−0.8181	0.41329
AGR	0.30412	0.07407	4.1056	0.00004
TRD	−0.02310	0.13740	−0.1681	0.86651
SER	−8.64677	4.05475	−2.1325	0.03297

Based on Table 3, the coefficient of W. poverty is significant, so we can say that adjacent districts influence each other. Further, from Table 3, we see that there are two significant predictors, namely, shares of GDP agriculture and GDP services. The coefficient of share of agriculture GDP is positive. This means that when share of the agriculture GDP increases, the number of poor people increases. The coefficient of services GDP is negative. This means that when the share of services GDP value increases, the number of poor people decreases. The relationship between education, share of GDP industry and GDP trading is negative. This means that when all of them decrease, the number of poor people decreases, but they are not significant.

5. Conclusion

Estimation methods are influenced by the error distribution characteristic. In the SLM model, we simulate normal distributed errors for sample size (n) from 20 to 500. Based on data simulation, all of RMSE's MLE are smaller than RMSE's TSLS, so we can conclude that the MLE's method is more efficient than TSLS's method when the errors are normally distributed.

Analysis of the estimation methods MLE and TSLS on SLM for modeling real data shows that the best choice in fitting SLM is MLE. Based on ANOVA's MLE, the coefficients of spatial lag and two predictors are significant, whereas the others are not significant. The increase in share of agriculture GDP causes significant increase in the number of poor people. The increase in share of services GDP causes significant decrease in the number of poor people. The increase in the number of people who graduated from junior high school, share of industry GDP and share of trading GDP causes decrease in the number of poor people, but they are not significant.

Acknowledgments

This research was supported by Jenderal Soedirman University (UNSOED). The authors thank the financial support from UNSOED through research fund of research competence with contract number: DIPA 042.01.2.400901/2017.

The authors also thank the anonymous referees for valuable suggestions and comments leading to the improvement of the manuscript.

References

[1] J. Aldstadt and A. Getis, Constructing the spatial weights matrix using local statistic, Geographical Analysis 36 (2004), 90-104.

[2] L. Anselin, Spatial Econometrics: Methods and Models, Kluwer Academic Publishers, Dordrecht, Netherlands, 1988.

[3] L. Anselin, Local indicators of spatial association-LISA, Geographical Analysis 27 (1995), 93-115.

[4] L. J. Bain and M. Engelhardt, Introduction to Probability and Mathematical Statistics, Brooks/Cole, 1992.

[5] BPS - Statistics of Central Java Province, Central Java Province in Figures, 2016.

[6] G. Casella and R. Berger, Statistical Inference, Wadsworth, Belmont, California, 1990.

[7] N. A. C. Cressie, Statistics for Spatial Data, John Wiley and Sons, New York, 1993.

[8] C. Gaetan and X. Guyon, Spatial Statistics and Modelling, John Wiley and Sons, New York, 2010.

[9] A. Getis and J. K. Ord, The analysis of spatial association by use of distance statistics, Geographical Analysis 24 (1992), 189-206.

[10] Jajang, A. Saefuddin, I. W. Mangku and H. Siregar, Comparing performances of WG, WG_{new} and WC on dynamic spatial panel model by Monte Carlo simulation, Far East J. Math. Sci. (FJMS) 90(1) (2014), 17-34.

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

[12] H. H. Kelejian and I. R. Prucha, A generalized method of moments estimator for autoregressive parameter in a spatial model, Int. Econ. 40 (1999), 509-533.

[13] J. Lu and L. Zhang, Evaluation of parameter estimation methods for fitting spatial regression models, Forest Science 56(5) (2010), 505-514.

[14] K. Ord, Estimation methods for models of spatial interaction, J. Amer. Statist. Assoc. 70 (1975), 120-126.

[15] J. K. Ord and A. Getis, Testing for local spatial autocorrelation in the presence of global autocorrelation, Journal of Regional Science 41(3) (2001), 411-432.

[16] E. J. Purcell and D. Varberg, Calculus, 9th ed., John Wiley and Sons, New York, 2007.

[17] M. D. Ward and S. G. Kristiani, Spatial Regression Models Series: Quantitative Application in the Social Science, SAGE Publications Inc., California, 2008.

Jajang: rzjajang@yahoo.com

Evaluation Of MLE And TSLS Methods For Fitting Spatial Lag Model And It's Application

ORIGINALITY REPORT

9%	%	%	9%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Georgia Institute of Technology Main Campus Student Paper	1%
2	Submitted to University of Florida Student Paper	1%
3	Submitted to Delhi University Student Paper	1%
4	Submitted to University of New South Wales Student Paper	1%
5	Submitted to Auckland University of Technology Student Paper	1%
6	Submitted to Osmaniye Korkut Ata University (deneme) Student Paper	1%
7	Submitted to Nanyang Technological University, Singapore Student Paper	1%

8	Submitted to Universitas Negeri Surabaya The State University of Surabaya Student Paper	1 %
9	Submitted to University of Leicester Student Paper	1 %
10	Submitted to University of Sheffield Student Paper	1 %
11	Submitted to Universitas Jember Student Paper	<1 %
12	Submitted to University of Central Florida Student Paper	<1 %
13	Submitted to University of Newcastle upon Tyne Student Paper	<1 %
14	Submitted to CSU, Pomona Student Paper	<1 %
15	Submitted to Cardiff University Student Paper	<1 %

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		