Application of spatial error model using GMM estimation in impact of education on poverty alleviation in Java, Indonesia

Ryan Willmanda Januardi[*], Agung Priyo Utomo

Department of Social Statistics, Sekolah Tinggi Ilmu Statistik
Jl. Otto Iskandardinata 64C, East Jakarta, 13330. Indonesia

Abstract

Java Island is the center of development in Indonesia, and yet poverty remains its major problem. The pockets of poverty in Java are often located in urban and rural areas, dominated by productive age group population with low education. Taking into account spatial factors in determining policy, policy efficiency in poverty alleviation can be improved. This paper presents a Spatial Error Model (SEM) approach to determine the impact of education on poverty alleviation in Java. It not only focuses on the specification of empirical models but also in the selection of parameter estimation methods. Most studies use Maximum Likelihood Estimator (MLE) as a parameter estimation method, but in the presence of normality disturbances, MLE is generally biased. The assumption test on the poverty data of Java showed that the model error was not normally distributed and there was spatial autocorrelation on the error terms. In this study we used SEM using Generalized Methods of Moment (GMM) estimation to overcome the biases associated with MLE. Our results indicate that GMM is as efficient as MLE in determining the impact of education on poverty alleviation in Java and robust to non-normality. Education indicators that have significant impact on poverty alleviation are literacy rate, average length of school year, and percentage of high schools and university graduates.

Keywords: poverty ratio; education; spatial error model; robustness; MLE; GMM.

1. Introduction

Java Island is the center of development in Indonesia, even though poverty remains one of its major problems. This can be seen from the contribution of GDP of Java Island from 2008 to 2013, which has been consistent in the 56 percent rate [1]. Still, in 2013 the majority of people who lived in poverty were concentrated in Java. As shown in Fig. 1, 54.45 percent of the total poor population in Indonesia live in Java [2]. So, even though the Island has a significant contribution to the national economy, it still cannot get out of the poverty problem, with more than half of Indonesia's poor population living in the Island.

The pockets of poverty in Java are located in urban and rural areas, dominated by productive age group population with low education and productivity, which put this group at a disadvantage in the labor market. The increasing poverty rate in both urban and rural areas is due to the low quality of human resources, of which education is an important indicator [3], [4]. Additionally, poverty is also associated with inter-regional spatial interactions associated with population mobility and spatial impoverishment [5]. Crandall and Weber [6] explained that poverty has a spatial interaction. A region with high poverty rate would affect and be affected by other regions around it.
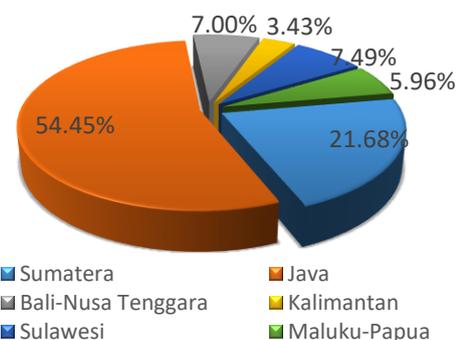

Fig. 1. Percentage of people living in poverty in Indonesia 2013 [2]

Taking into account this spatial interaction, a policy on poverty alleviation can benefit from and made more efficient by using spatial analysis. Numerous researches have incorporated spatial statistics when examining poverty. However, many of such research have focused on the specification of spatial statistical models rather than the selection of parameter estimation methods. Although the use of parameter estimation methods might not make as significant difference as model specification, the existence of assumption requirements and data conditions requires appropriate parameter estimation methods [7]. To be best of our knowledge, existing spatial analyses do not take into account these assumption requirements explicitly –perhaps inadvertently. Spatial regression parameters estimation [8] can

* Corresponding author. Tel.: +62 852 40421858.
Email: rw.januardi@gmail.com

be obtained through several estimation methods such as Maximum Likelihood Estimation (MLE) and Generalized Method of Moments (GMM). These two parameters estimation methods are used for different assumption requirements. We select GMM as a parameter estimation method to overcome the bias associated with MLE.

2. Theoretical Framework, Materials, and Methods

2.1. Theoretical Framework

Poverty is defined as a lack of means necessary to meet basic needs such as food and non-food as measured by expenditure [2]. People who live in poverty are the ones with an average consumption per capita per month below this established poverty line. Poverty can be measured in two dimensions: the monetary dimension that covers insufficient income or consumption; and non-monetary dimensions that covers insufficient outcomes with respect to human capital such as education, health, and nutrition [9]. One of the efforts to reduce and cut the vicious cycle of poverty is to improve the education of the population [10]. The low education level of the poor will lead to a vicious cycle of poverty in the next generation. People with low education will have low productivity; and low productivity will lead to low income, resulting in poverty. Poor households will find it difficult to finance their children's schooling so that it will produce the next generation with similar low education thus creating the unfortunate cycle of poverty.

Throughout the world, it has been found that the probability of finding employment rises with higher levels of education, and that earnings are higher for people with higher level of education [11]. This connection between education and poverty works through three mechanisms. Firstly, more educated people earn more. Secondly, more (and especially better quality) education improves economic growth and thereby economic opportunities and income. Thirdly, education brings wider social benefits, such as economic development, which will have a positive ripple effect on the poor regions. The theoretical framework compiled by Janjua and Kamal [12] states that education has direct and indirect effects in poverty alleviation. From this theoretical framework, we consider education and skills of the individual as the direct effects in poverty alleviation. Fig. 2 depicts the impact of education on poverty alleviation.

2.2. Materials

The analysis of this study included all regencies/cities in Java Island. This study used secondary data from the 2013 Indonesian Central Bureau of Statistics (BPS) and Indonesian Ministry of Education and Culture for its 118 regencies/cities. Data used include: literacy rate (X_1), average length of school year (X_2), percentage of high schools and university graduates (X_3), ratio of junior high school availability (X_4), ratio of senior high school availability (X_5), and poverty ratio (Y). Poverty ratio is the proportion of the poor population with total population in a region (regency/city). It indicated the incidence of poverty in a region, but ignored the differences in well-being between different poor households.

The data used in this study was aggregate data in every regency/city. This study was conducted at the spatial poverty level, not on the individual or household poverty. Meanwhile, the spatial data used in this study was derived from BPS mapping.
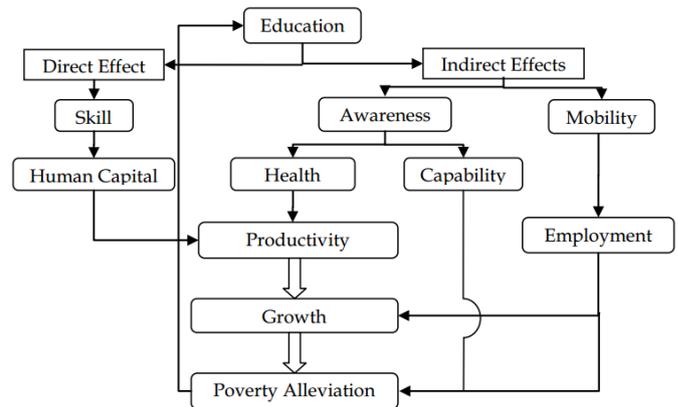

Fig. 2. Impact of education on poverty alleviation [12]

2.3. Methods

The methods used in this research was Exploratory Spatial Data Analysis (ESDA) and inference analysis using spatial regression. The details of each analysis are explained below.

2.3.1. Exploratory Spatial Data Analysis

Exploratory Spatial Data Analysis (ESDA) was used to describe spatial patterns of poverty in Java Island. ESDA [13] is a collection of techniques to describe and visualize spatial distributions; identify spatial outliers; discover patterns of spatial association, clusters or hot-spots; and suggest spatial regimes or other forms of spatial heterogeneity. ESDA was applied based on Global Moran's I, Moran Scatterplot, and LISA statistics. To identify spatial patterns, spatial clustering association patterns, and outlier data, we used the following statistical techniques for exploratory spatial data analysis.

1. Constructing Spatial Weighted Matrices

The basic form of spatial weighted matrices is a square symmetric weighted matrices (denoted **W**) $n \times n$ (row standardized) matrices that define which areas are neighbors of a given area. Spatial weighted matrices is a weight denoting the strength of the connection between areas *i* and *j*.

In this study, we used contiguity-based relations based on modified queen contiguity. Contiguity-based relations are mostly used in the presence of irregular polygons with varying shape and surface, since contiguity ignores distance and focuses instead on the location of an area. It was appropriate for areas in Java that had irregular polygons. Queen contiguity defined a neighbor when at least one point on the boundary of one polygon is shared with at least one point of its neighbor (common border or corner). Queen contiguity weighted matrices (denoted $\mathbf{W_Q}$) is denoted as follows:

$$\mathbf{W_Q} = \begin{bmatrix} w_{11} & \cdots & w_{1j} \\ \vdots & \ddots & \vdots \\ w_{i1} & \cdots & w_{ij} \end{bmatrix} \qquad w_{ij} = \begin{cases} 1, & \text{if i neighbor j} \\ 0, & \text{otherwise} \end{cases} \qquad (1)$$

In addition, the queen contiguity weighted matrices used in this study was manually modified to maintain connectivity

between areas. Thus, separate cross-islands areas like Bangkalan and Surabaya or Pulau Seribu and Jakarta Utara still have access to interact as neighboring areas. Connectivity between areas based on modified queen contiguity weighted matrices was illustrated in Fig. 3.
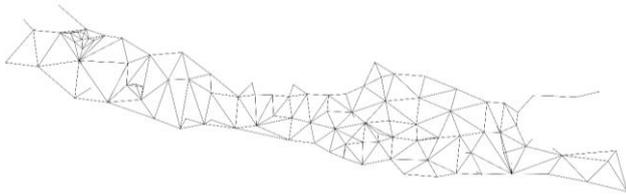

Fig. 3. Connectivity between areas based on modified queen contiguity weighted matrices

The obtained queen contiguity weighted matrices were transformed into the normality matrices (row-standardized), which is the spatial weighted value (denoted Wij) for each neighbor which forms the spatial weighted matrices **W**, according to following equation:

$$W_{ij} = \frac{w_{ij}}{w_j} \tag{2}$$

$$\mathbf{W} = \begin{bmatrix} W_{11} & \cdots & W_{1j} \\ \vdots & \ddots & \vdots \\ W_{i1} & \cdots & W_{ij} \end{bmatrix} \tag{3}$$

2. Analyzing Global Indicators of Spatial Autocorrelation by Using Global Moran's I Statistics

Moran's I [14] was used in this study to determine whether the value of neighboring areas were more similar than would be expected under the null hypothesis. Mathematically, Global Moran's I statistics for *n* observation and *i*-th observation at the *j*-th location can be formulated in following equation:

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}} \times \frac{\sum_i \sum_j W_{ij}(y_i-\bar{y})(y_j-\bar{y})}{\sum_i (y_i-\bar{y})^2}, i \neq j \tag{4}$$

$$E(1) = -\frac{1}{n-1} \tag{5}$$

$$\mathrm{var}_R(I) =$$
$$\frac{n[(n^2-3n+3)S_1-b[(n^2-n)S_1-2nS_2+6S_0^2]}{(n-1)(n-2)(n-3)S_0^2} - E(I)^2 \tag{6}$$

$$b = \frac{n \sum_{i=1}^{n} (y_i-\bar{y})^4}{\left(\sum_{i=1}^{n}(y_i-\bar{y})^2\right)^2} \tag{7}$$

$$S_0 = \sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij} \tag{8}$$

$$S_1 = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} (W_{ij}+W_{ji})^2 \tag{9}$$

$$S_2 = \sum_{i=1}^{n} (W_{i.}+W_{.i})^2 \tag{10}$$

$$W_{i.} = \sum_{j=1}^{n} W_{ij} \tag{11}$$

$$W_{.i} = \sum_{j=1}^{n} W_{ji} \tag{12}$$

Under the randomization assumption (denoted R), the rates were random samples from a population whose distribution was unknown. Assumption R is less restrictive since their theoretical distribution is often unknown.

The value of Global Moran's I ranges between -1 and 1. If $I > E(I)$, then the spatial pattern is clustered indicating a positive spatial autocorrelation. If $I = E(I)$, the pattern spread unevenly (no spatial autocorrelation), and if $I < E(I)$, the pattern is diffused indicating negative spatial autocorrelation [15].

3. Analyzing Local Indicators of Spatial Autocorrelation by Using LISA Cluster Map

In contrast to the previously described Global Moran's I, which is a global indicators of spatial autocorrelation, LISA indicated local autocorrelation. In this case, LISA identified the relationship between a location of observation to another location of observation. Furthermore, the clustering of areas belonged to four types of spatial associations and visualized through the LISA cluster map [16]. The four possible scenarios were as follows:

a. Hot Spots, high-value location would be surrounded by high-value neighbors (high-high)
b. Cold Spots, the low-value location would be surrounded by low-value neighbors (low-low)
c. Outliers, high-value locations would be surrounded by low-value neighbors (high-low)
d. Outliers, low-value locations would be surrounded by high-value neighbors (low-high)

If the LISA cluster map showed 'not significant' results, it meant that the proximity of the area was not closely related to the events studied.

2.3.2. Spatial Regression

Spatial regression is closely related to the autoregressive process, indicated by the dependence relationship between a set of observations or locations. The relations could also be expressed with the location value depending on another neighboring location value. Spatial regression was used to analyze the impact of education, with predictor variables as follows: literacy rate (X_1), average length of school year (X_2), percentage of high schools and university graduates (X_3), ratio of junior high school availability (X_4), and ratio of senior high school availability (X_5); and poverty ratio in Java in 2013 as response variable, taking into account spatial factors.

There are two common types of spatial regression:

1. Spatial Lag Model (SLM)

SLM [17] is a model that combines a classic regression model with spatial lag in response variables using cross-

sectional data so often called spatial lag model. The SLM is formed when $\rho \neq 0$ and $\lambda = 0$. This model assumed that the autoregressive process only occurs in the response variable. The spatial lag model that was possible to be formed in this research is as follows:

$$Y_i = \rho \sum_{j=1,i\neq j}^{n} W_{ij}\, Y_j + \beta_0 - \beta_1 X1_i - \beta_2 X2_i - \beta_3 X3_i - \beta_4 X4_i - \beta_5 X5_i + \varepsilon_i \quad (13)$$

with ρ is spatial lag coefficient parameters on the response variable.

2. Spatial Error Model (SEM)

SEM [17] is a model in which the model-error has spatial correlation. The SEM is formed when $\rho = 0$ and $\lambda \neq 0$. This model assumed that the autoregressive process only occurred in the model error. The spatial error model that was formed in this research was as follows:

$$Y_i = \beta_0 - \beta_1 X_1 - \beta_2 X_2 - \beta_3 X_3 - \beta_4 X_4 - \beta_5 X_5 + \lambda \sum_{j=1,i\neq j}^{n} W_{ij}\, u_j + \varepsilon_i \quad (14)$$

with λ as spatial error coefficient parameters on error u, and u as spatial error vector ($n \times 1$).

To specify the appropriate model, we follow the steps illustrated in Fig. 4. In particular the focus was on detecting model specification due to spatial dependence (in the form of an omitted spatially lagged dependent variable and spatial residual autocorrelation). Four tests were performed to assess the spatial dependence of the model. The statistics were the simple LM diagnostics for a missing spatially lagged dependent variable (Lagrange Multiplier (lag)), the simple LM diagnostics for error dependence (Lagrange Multiplier (error)), variants of these robust to the presence of the other (Robust LM (lag) and Robust LM (error) which diagnoses for error dependence in the possible presence of a missing lagged dependent variable, Robust LM (lag) is the other way round). All modelling process were calculated by R programming using 'spdep' packages and GeoDaSpace, two softwares for advanced spatial econometrics.

2.3.3. Parameter Estimation Methods

1. Maximum Likelihood Estimator (MLE)

The underlying assumption of this estimator is the normal distribution of model errors, i.e., $N(0. \sigma^2 I)$. The log-likelihood function and estimator of SLM was:

$$\ln L = \ln|I - \rho\mathbf{W}| - \frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{(\mathbf{Y} - \rho\mathbf{WY} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \rho\mathbf{WY} - \mathbf{X}\boldsymbol{\beta})}{2\sigma^2} \quad (15)$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{I} - \rho\mathbf{W})\mathbf{Y} \quad (16)$$

The log-likelihood function and estimator of SEM was:

$$\ln L = \ln|I - \lambda\mathbf{W}| - \frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{I} - \lambda\mathbf{W})'(\mathbf{I} - \lambda\mathbf{W})(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})}{2\sigma^2} \quad (17)$$

$$\hat{\boldsymbol{\beta}} = ((\mathbf{X} - \lambda\mathbf{WX})'(\mathbf{X} - \lambda\mathbf{WX}))^{-1}(\mathbf{X} - \lambda\mathbf{WY})'(\mathbf{Y} - \lambda\mathbf{WY}) \quad (18)$$

2. Generalized Methods of Moment (GMM)

The basic principle of GMM is to estimate β so that the moment of condition in the sample will be equal to the moment of condition in the population by minimizing objective function of the moment of sample condition. Where $g(\hat{\beta}) = \frac{1}{n}\sum_{i=1}^{n} z_i(Y_i - X_i\hat{\beta})$ is the moment of sample condition, z_i is the instrument variable, and $E(g(\beta)) = 0$ is the moment of the population condition.
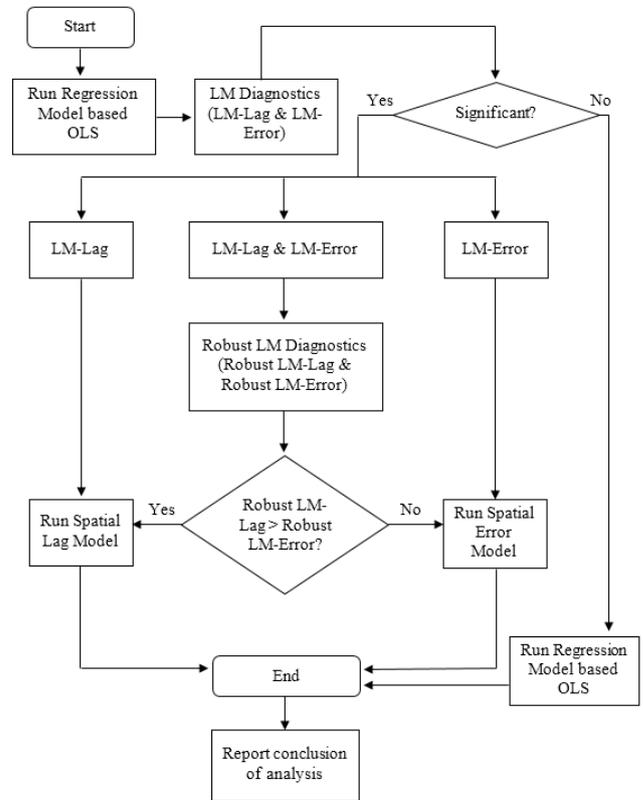

Fig. 4. Schematic process spatial regression model specification

Kelejian dan Prucha (in Anselin) [8] argue that GMM is as efficient as MLE. In addition, GMM is parameter estimator that did not require the normal distribution assumption for model errors as required by MLE. This results in the GMM estimator for SEM. The GMM estimator can be produced in three steps [18]:

a. Build objective function of the moment of sample condition, that is the quadratic function of moment of sample condition based on specified spatial weighting matrices:

$$Q(\hat{\boldsymbol{\beta}}) = \mathbf{g}(\hat{\boldsymbol{\beta}})'\mathbf{W}\mathbf{g}(\hat{\boldsymbol{\beta}}) \quad (19)$$

b. Obtain a consistent but inefficient estimate of β by minimizing the objective function of the moment of sample condition as follows:

$$\hat{\beta}_{[1]} = \arg\min_{\beta}\, \mathbf{g}(\hat{\boldsymbol{\beta}})'\mathbf{W}\mathbf{g}(\hat{\boldsymbol{\beta}}) \quad (20)$$

c. Obtain a consistent and efficient estimate of β by minimizing the objective function of the moment of sample condition with an optimized spatial weighting matrices based on $\hat{\beta}_{[1]}$, as follows:

$$\hat{\beta}_{[2]} = \arg\min_{\beta}\, \mathbf{g}(\hat{\boldsymbol{\beta}})'\mathbf{W}_{\mathbf{opt}}\mathbf{g}(\hat{\boldsymbol{\beta}}) \quad (21)$$

3. Results

Java had a wide spread poverty among its regions. Through ESDA, we could describe the spatial distribution pattern which included patterns of spatial association and identification of outlier data from poverty ratio. ESDA was applied based on: (1) Global Moran's I statistics score which described the effects of global spatial autocorrelation, and (2) LISA cluster map which described the effects of local spatial autocorrelation through spatial weighted matrices based on modified queen contiguity.

Table 1. Global Moran's I statistics calculation

Variable	I	E(I)	*pseudo p-value*
Y	0.5354	-0.0085	0.001

The calculation results in Table 1 showed that there were statistically spatial autocorrelation effects in poverty among regions in Java in 2013. This was indicated by pseudo *p-value* less than 5%. The table also indicated that that $I > E(I)$, which meant that the spatial pattern of poverty among regions in Java Island was clustered.

To determine which region has a significant effect on spatial association in general, we used LISA cluster map. As shown in Fig. 5, there were three significant spatial distribution patterns of poverty based on the LISA calculation results.
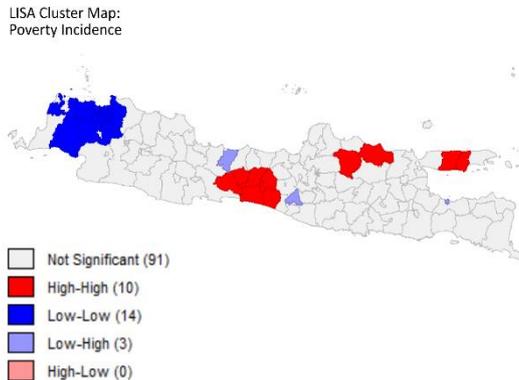

Fig. 5. Poverty cluster in Java Island based on LISA cluster map

Table 2. Significant spatial distribution patterns of poverty based on LISA cluster map

Patterns	Regions
High-High	Banyumas Regency, Purbalingga Regency, Banjarnegara Regency, Kebumen Regency, Purworejo Regency, Wonosobo Regency, Blora Regency, Tuban Regency, Sampang Regency, and Pamekasan Regency
Low-Low	Lebak Regency, Serang Regency, Tangerang Regency, Tangerang City, Tangerang Selatan City, Jakarta Barat City, Jakarta Selatan City, Jakarta Utara City, Jakarta Pusat City, Jakarta Timur City, Bekasi Regency, Bogor Regency, Bekasi City, and Depok City
Low-High	Tegal Regency, Sleman Regency, and Probolinggo City

To see the spatial effects on poverty in Java Island, we used spatial regression model. Before determining the appropriate model, multicollinearity diagnostics using Variance Inflation Factor (VIF) was conducted to see if there were correlations among the predictor variables. The presence of multicollinearity in certain predictor variables would cause greater of standard error and thus interfere the results of the analysis. If the VIF value is less than 10 then it can be concluded that there is no multicollinearity. If a multicolinearity is found in the model, one solution is to remove one of the variables from the model. The goal is to extract information that is already represented by the other predictor variables. The results of the multicollinearity diagnostics are displayed on Table 3.

Table 3. Multicollinearity diagnostics results

Variable	VIF value	
	Step 1	Step 2
X_1	2,477	1,465
X_2	10.987	Removed
X_3	7,447	1,895
X_4	1,965	1,96
X_5	3,023	2,986

From the above table, appropriate predictor variables that could be used in this study were obtained. The average length of school year variable was removed because it had VIF > 10. There was a possibility that average length of school year (X_2) correlated with percentage of high schools and university graduates variable (X_3). After all the predictor variables were free from multicollinearity, we built the spatial regression model. Before determining the appropriate model, we did model specification between SLM and SEM. The results of model specification are displayed on Table 4. Model specification by LM and Robust LM diagnostics showed that Spatial Error Model (SEM) was better suited for this study. Additionally, the Robust LM (lag) value was smaller than Robust LM (error) value, and the Robust LM (error) was more significant than Robust LM (lag).

Table 4. LM and Robust LM diagnostics results for spatial regression model specification

Diagnostics	Value	*p-value*
Lagrange Multiplier (lag)	18.156	0.0000*
Robust LM (lag)	0.4512	0.5018
Lagrange Multiplier (error)	25.463	0.0000*
Robust LM (error)	7.7582	0.0053*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Meanwhile, based on normality test results, data error in this study was not normally distributed. The result of normality test by Jarque-Bera test is displayed on Table 5. The *p-value* of Jarque-Bera is 0.0013 (i.e. reject null hypothesis where null hypothesis is error normally distributed).

Table 5. Normality test result

Test	Value	*p-value*
Jarque-Bera	13.246	0.0013*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Spatial regression modelling by SEM was estimated using Generalized Method of Moments (GMM) estimator. The GMM was used because the normal distribution assumption for model errors was not met. The following table is a summary of the model.

Table 6. Model parameter estimates, estimates of standard error of the parameters, and pseudo R-Squared of SEM by MLE and GMM

Parameter	SEM MLE		SEM GMM	
	Estimates (*p-value*)	Std. Error	Estimates (*p-value*)	Std. Error
(Intercept)	52.7044 (0.0000)*	7.5309	54.0939 (0.0000)*	7.3672
β_1	-0.3705 (0.0000)*	0.0314	-0.3853 (0.0000)*	0.0312
β_3	-0.1298 (0.0000)*	0.0858	-0.1282 (0.0000)*	0.0841
β_4	-0.0260 (0.3122)	0.0565	-0.0452 (0.2148)	0.0572
β_5	-0.0430 (0.2234)	0.0531	-0.0296 (0.2901)	0.0535
λ	0.5579 (0.0000)*	0.0866	0.5114 (0.0000)*	0.0837
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
Pseudo R-Squared	0.5987		0.6002	

Table 7. Variance of estimators in SEM by MLE and GMM

Parameter	SEM MLE	SEM GMM
β_1	0.0107	0.0106
β_3	0.0800	0.0768
β_4	0.0347	0.0355
β_5	0.0306	0.0311
λ	0.0815	0.0761

In terms of the estimates of standard errors of parameters, MLE produced slightly larger standard errors for the significant parameter estimates and slightly smaller standard errors for the non-significant parameter estimates than GMM. The MLE and GMM also produced slightly different results in the parameter estimates (i.e. $\hat{\lambda}, \widehat{\beta_1}, \widehat{\beta_3}, \widehat{\beta_4}, \widehat{\beta_5}$). GMM, which was free of distributional assumption for the model errors, had *pseudo R-squared* comparable to that of MLE. The *pseudo R-squared* of SEM GMM was 0.6002. The results indicated that GMM was better than MLE in terms of *pseudo R-squared.* *Pseudo R-Squared* is used to describe how close the data to the fitted regression line. In this study, these results meant that the SEM GMM with variables of literacy rate (X_1), percentage of high schools and university graduates (X_3), ratio of junior high school availability (X_4), and ratio of senior high school availability (X_5) could explain 60.02% the variability of the poverty ratio in Java.

The variance of each parameter estimator for the parameter $\widehat{\beta_1}, \widehat{\beta_3}, \widehat{\beta_4}, \widehat{\beta_5}$, and $\hat{\lambda}$ was also computed for MLE and SEM GMM (Table 7). Theoretically, MLE is most efficient (producing lowest variance) if the normality assumption is met. But in this study, MLE produced much larger variance than GMM for the significant parameter. Under the non-normality, GMM was better in terms of the variance than MLE. This results indicated that GMM was as efficient as MLE and robust to non-normality.

Based on the results obtained in Table 6, the SEM equation formed is as follows:

$$\widehat{Y}_i = 54{,}0939 - 0.3853X_1 - 0.1282X_3 - 0.0452X_4 - 0.0296X_5 + u_i \quad (22)$$

$$u_i = 0.5114 \sum_{j=1, i \neq j}^{n} W_{ij}\, u_j \quad (23)$$

Significance test of model parameter estimates in Table 6 showed that the variable of literacy rate (X_1), percentage of high schools and university graduates (X_3), and spatial error (λ) had significant effect to poverty ratio in Java in 2013. The significant coefficient λ indicated that autoregressive process on model error significantly influenced the regions' poverty ratio in Java in 2013. If the variable of literacy rate (X_1), percentage of high schools and university graduates (X_3), ratio of junior high school availability (X_4), and ratio of senior high school availability (X_5) was ignored or equal to zero, the poverty ratio in Java was estimated at 54.09%.

Assuming the condition of other variables is constant, the increase in literacy rate in a region by 1% can reduce the poverty ratio of a region by 0.3853%. Similarly, if the percentage of high schools and university graduates rose 1% in a region then the poverty ratio in the region will reduce by 0.1282%. The variables of ratio of junior high school availability (X_4) and ratio of senior high school availability (X_5) also had a negative relationship with poverty ratio. That is, the higher the ratio of school availability to both junior and senior high school, the lower the poverty ratio. However, the variable of ratio of junior high school availability and ratio of high school availability had no significant effect on poverty ratio.

4. Discussion

Based on the results in Table 2, regions with high poverty tended to be surrounded by regions with high poverty as well, and vice versa, regions with low poverty tended to be surrounded by similar low poverty. This phenomenon has been described by Crandall and Weber [6] in which they argued that poverty has a spatial interaction. However, we also found the outliers in this phenomenon. There were three regions which belonged to low-high clusters, low poverty regions surrounded by high poverty neighbors. The regions were:

- Tegal Regency (10,75%) as low-poverty region surrounded by Brebes Regency (21,12%), Banyumas Regency (19,44%), Purbalingga Regency (21,19%), and Pemalang Regency (19,27%) as high-poverty neighbors;
- Sleman Regency (10,44%) as low-poverty region surrounded by Kulon Progo Regency (23,31%) and Gunung Kidul Regency (22,71%) as high-poverty neighbors; and
- Probolinggo City (10,92%) as low-poverty region surrounded by Probolinggo Regency (22,22%) as high-poverty neighbors.

This condition might lead to two possible scenarios: the low-poverty region affects or be affected by the high-poverty neighbors. Which of these two scenarios will likely to happen depends on many factors.

One of the contributing factors is literacy rate. According to this model, a significant increase in literacy rate could reduce poverty ratio. As Murray and Shillington [19] describe, a person with low literacy skill tends to be unsuitable for a job compared to those with higher literacy skill. Regions with higher literacy rate have a population with a higher chance of entering the labor market and earning income so as to avoid poverty. In aggregate, it can reduce the poverty ratio in a region. In addition, an increase in the percentage of high schools and university graduates also has a significant impact

on poverty alleviation. Silva [20] documents that poverty declines with increasing years of education. Increasing one year of education will increase the human capital. The increased human capital will contribute negatively to the possibility of being in poverty.

The variables of ratio of junior high school availability and ratio of senior high school availability also had a negative relationship with poverty ratio, but they had no significant effect on poverty ratio. This might be due to the difference in calculation approach of poverty ratio and the ratio of school availability in both junior and senior high schools. The poverty ratio calculation used a household approach, while the school availability ratio of both junior and senior high schools used an individual approach of school age (ages 13 to 15 years for junior high school and 16 to 18 years old for high school). Increasing the ratio of school availability to both junior and senior high schools would only affect the increased chance of a certain school-age population to have a certain education. When there is a decrease in one poor household in a region, then its poverty ratio will also decrease. Meanwhile, when a household in a region in which a household member has the opportunity to go to school and receive a certain education due to an increase in the number of schools, the poverty level in the region does not necessarily decrease, but the school availability ratio will.

Based on the resulting spatial error equation in Equation 23, the poverty of a region would increase by a multiple of 0.5114% of the spatial weighting of each region, if the average error of the neighboring region rises by 1%. For example, Sampang Regency had a spatial weighting with its neighboring area of 0.50 (provided in Appendix A) and the spatial error equation of Sampang was:

$$u_{Sampang}=0.5114 \sum_{j=1,i \neq j}^{3} 0.50 \, u_j$$

$$u_{Sampang}=0.2557 u_{Bangkalan}+0.2557 u_{Pamekasan}$$

Sampang Regency had two neighboring regions: Bangkalan Regency and Pamekasan Regency. If one or all of regions variable of error (u_j) was increased so that the average of all neighboring regions were increased by 1%, then Sampang Regency would get the effect of increased poverty rate of 0.2557%. These results meant that there were influences of the predictor variables other than the ones used in this study from the neighboring areas.

These results were consistent with those presented by Henninger and Snel [21], in which they argued that spatial variations in poverty level are often caused by factors with spatial dimension of the surrounding areas. In this study, Sampang regency was the region with the highest poverty rate and lowest literacy rate and percentage of high schools and university graduates. Furthermore, Sampang regency was also surrounded by high-poverty regions. Meanwhile Tangerang Selatan city was the region with the lowest poverty rate and high literacy rate and percentage of high schools and university graduates, so as Yogyakarta city and Cimahi city. They were also surrounded by low-poverty regions.

Based on these results, the government can improve the efficiency and effectiveness education and poverty alleviation

policy by paying more attention to the cluster of high poverty and low education region.

5. Conclusion

We compare MLE and GMM parameter estimation methods for spatial error model. Our results indicate that SEM using GMM estimation significantly better than MLE, in terms of *pseudo R-squared* and variance under the non-normality. However, in terms of model specification, the results do not make a significant difference between GMM and MLE. The estimates of parameters and its standard error have a slight difference. These results indicate that GMM is as efficient as MLE and robust to non-normality. Therefore, the selection of the parameter estimation methods may depend on the distribution of data and variables, as well as the purpose of the specific research.

Based on these results, education indicators that have significance impact to poverty alleviation in Java are literacy rate, average length of school year, and percentage of high schools and university graduates. By ESDA, there were positive spatial autocorrelation effects in poverty among regions in Java in 2013 so as to form the clusters of poverty regions.

This study showcases one alternative to spatial statistics and parameter estimation methods besides the commonly used MLE, and compare it with other methods like GMM. Future studies can also test other alternatives such as Quasi Maximum Likelihood.

References

1. Badan Pusat Statistik. *Penghitungan dan Analisis Makro Kemiskinan Indonesia 2014*. Jakarta, Indonesia: BPS, 2014.
2. Badan Pusat Statistik. *Data dan Informasi Kemiskinan Kabupaten/Kota 2013*. Jakarta, Indonesia: BPS, 2013.
3. H. Widodo. *Potret Pendidikan di Indonesia dan Kesiapannya Dalam Menghadapi Masyarakat Ekonomi Asia (MEA)*. Jurnal Cendekia. 2 (13) (2015) 209–307.
4. C. Neilson, D. Contreras, R. Cooper, and J. Hermann. *The Dynamic Poverty in Chile*. Journal of American Latin Studies. 40 (2008) 251-273.
5. M. D. Partridge and D. S. Rickman. *The Geography of American Poverty: Is There a Need for Place-Based Policies?*. Oklahoma, USA: Oklahoma State University, 2006.
6. Crandall, S. M. & Bruce A. W. *Local Social and Economic Conditions, Spatial Concentrations of Poverty, and Poverty Dynamics*. USA: Oregon State University, 2004.
7. J. Lu and L. Zhang. *Evaluation of Parameter Estimation Methods for Fitting Spatial Regression Models*. Forest Science. 56 (5) (2010) 505-515.
8. L. Anselin. *Spatial Econometrics*. A Companion of Theoretical Econometrics. 14 (2003) 313–330.
9. A. Couduel, S. H. Jesko, and T. W. Quentin. *Poverty Measurement and Analysis*. Washington DC, USA: World Bank, 2001.
10. C. M. Eireann. *Poverty and Educational Disadvantage: Breaking the Cycle*. Dublin, Ireland: INTO Publication, 1994.
11. S. van der Berg. *Poverty and Education*. Paris, France: International Institute for Educational Planning and International Academy of Education, 2008.
12. P. Z. Janjua & U. A. Kamal. *The Role of Education and Income in Poverty Alleviation: Cross Country Analysis*. The Lahore Journal of Economics. 16 (1) (2011) 143–172.
13. L. Anselin. *Exploratory Spatial Data Analysis in a Geocomputational Environment*. New York, USA: Wiley and Sons, 1998.

14. P. A. P. Moran. *Notes on Continuous Stochastic Phenomena*. Biometrika. 37 (1950) 17–23.
15. J. Lee & D. W. S. Wong. *Statistical Analysis with Arcview GIS*. New York, USA: John Wiley & Sons, 2001.
16. L. Anselin. *Local indicators of spatial association–LISA*. Geographical Analysis. 27 (1995) 93–115.
17. L. Anselin. *Spatial Econometrics: Methods and Models*. Netherlands: Kluwer Academic Publishers, 1988.
18. H. B. Nielsen. *Generalized Method of Moments (GMM) Estimation [Online]*. Available: http://www.econ.ku.dk/metrics. [Accessed: 22-Jan-2017]
19. S. Murray & R. Shillington. *From Poverty to Prosperity: Literacy's Impact on Canada's Economic Success*. Ottawa, Canada: Canadian Literacy and Learning Network, 2011.
20. I. de Silva. *Micro-Level Determinants of Poverty in Sri Lanka: a Multivariate Approach*. International Journal of Social Economics. 35(3) (2008) 140-158.
21. N. Henninger and M. Snel. *Where are the Poor? Experiences with the Development and Use of Poverty Maps*. Washington DC, USA: World Resource Institute, 2002

Appendix A

Neighborhood and spatial weighting value of regions in Java Island in 2013

Region Code	Region	Number of Neighbor	Regions Neighbor Code	Spatial Weighting Value
1	Kab. Kepulauan Seribu	1	6	1.00
2	Kota Jakarta Selatan	6	3, 4, 5, 29, 118, 115	0.17
3	Kota Jakarta Timur	6	2, 4, 6, 22, 28, 29	0.17
4	Kota Jakarta Pusat	4	2, 3, 5, 6	0.25
5	Kota Jakarta Barat	5	2, 4, 6, 113, 115	0.20
6	Kota Jakarta Utara	6	1, 3, 4, 5, 22, 113	0.17
7	Kab. Bogor	11	8, 9, 20, 21, 22, 24, 28, 29, 112, 113, 118	0.09
8	Kab. Sukabumi	4	7, 9, 25, 112	0.25
9	Kab. Cianjur	6	7, 8, 10, 11, 20. 23	0.17
10	Kab. Bandung	7	9, 11, 17, 19, 23, 26, 30	0.14
11	Kab. Garut	4	9, 10, 12, 17	0.25
12	Kab. Tasikmalaya	5	11, 13, 16, 17, 31	0.20
13	Kab. Ciamis	6	12, 14, 16, 31, 32, 33	0.17
14	Kab. Kuningan	5	13, 15, 16, 33, 61	0.20
15	Kab. Cirebon	5	14, 16, 18, 27, 61	0.20
16	Kab. Majalengka	6	12, 13, 14, 15, 17, 18	0.17
17	Kab. Sumedang	6	10, 11, 12, 16, 18, 19	0.17
18	Kab. Indramayu	4	15, 16, 17, 19	0.25
19	Kab. Subang	6	10, 17, 18, 20, 21, 23	0.17
20	Kab. Purwakarta	5	7, 9, 19, 21, 23	0.20
21	Kab. Karawang	4	7, 19, 20, 22	0.25
22	Kab. Bekasi	5	3, 6, 7, 21, 28	0.20
23	Kab. Bandung Barat	6	9, 10, 19, 20. 26, 30	0.17
24	Kota Bogor	1	7	1,00
25	Kota Sukabumi	1	8	1,00
26	Kota Bandung	3	10, 23, 30	0.33
27	Kota Cirebon	1	15	1,00
28	Kota Bekasi	4	3, 7, 22, 29	0.25
29	Kota Depok	5	2, 3, 7, 28, 118	0.20
30	Kota Cimahi	3	10, 23, 26	0.33
31	Kota Tasikmalaya	2	12, 33	0.50
32	Kota Banjar	2	13, 33	0.50
33	Kab. Cilacap	6	13, 4, 32, 34, 37, 61	0.17
34	Kab. Banyumas	7	33, 35, 36, 37, 59, 60, 61	0.14
35	Kab. Purbalingga	4	34, 36, 58, 59	0.25
36	Kab. Banjarnegara	6	34, 35, 37, 39, 57, 58	0.17
37	Kab. Kebumen	5	33, 34, 36, 38, 39	0.20
38	Kab. Purworejo	4	37, 39, 40, 68	0.25
39	Kab. Wonosobo	7	36, 37, 38, 40, 55, 56, 57	0.14
40	Kab. Magelang	8	38, 39, 41, 54, 55, 62, 68, 71	0.13
41	Kab. Boyolali	9	40, 42, 43, 45, 46, 47, 54, 63, 71	0.11
42	Kab. Klaten	4	41, 43, 70, 71	0.25

Region Code	Region	Number of Neighbor	Regions Neighbor Code	Spatial Weighting Value
43	Kab. Sukoharjo	6	41, 42, 44, 45, 63, 70	0.17
44	Kab. Wonogiri	6	43, 45, 70, 73, 74, 92	0.17
45	Kab. Karanganyar	7	41, 43, 44, 46, 63, 92, 93	0.14
46	Kab. Sragen	4	41, 45, 47, 93	0.25
47	Kab. Grobogan	8	41, 46, 48, 50, 51, 53, 54, 93	0.13
48	Kab. Blora	6	47, 49, 50, 93, 94, 95	0.17
49	Kab. Rembang	3	48, 50, 95	0.33
50	Kab. Pati	5	47, 48, 49, 51, 52	0.20
51	Kab. Kudus	4	47, 50, 52, 53	0.25
52	Kab. Jepara	3	50, 51, 53	0.33
53	Kab. Demak	5	47, 51, 52, 54, 65	0.20
54	Kab. Semarang	8	40, 41, 47, 53, 55, 56, 64, 65	0.13
55	Kab. Temanggung	4	39, 40, 54, 56	0.25
56	Kab. Kendal	5	39, 54, 55, 57, 65	0.20
57	Kab. Batang	5	36, 39, 56, 58, 66	0.20
58	Kab. Pekalongan	5	35, 36, 57, 59, 66	0.20
59	Kab. Pemalang	4	34, 35, 58, 60	0.25
60	Kab. Tegal	4	34, 59, 61, 67	0.25
61	Kab. Brebes	6	14, 15, 33, 34, 60, 67	0.17
62	Kota Magelang	1	40	1,00
63	Kota Surakarta	3	41, 43, 45	0.33
64	Kota Salatiga	1	54	1,00
65	Kota Semarang	3	53, 54, 56	0.33
66	Kota Pekalongan	2	57, 58	0.50
67	Kota Tegal	2	60, 61	0.50
68	Kab. Kulon Progo	4	38, 40, 69, 71	0.25
69	Kab. Bantul	4	68, 70, 71, 72	0.25
70	Kab. Gunung Kidul	5	42, 43, 44, 69, 71	0.20
71	Kab. Sleman	7	40. 41, 42, 68, 69, 70, 72	0.14
73	Kab. Pacitan	3	44, 74, 75	0.33
74	Kab. Ponorogo	7	44, 73, 75, 76, 90, 92	0.14
75	Kab. Trenggalek	3	73, 74, 76	0.33
76	Kab. Tulungagung	5	74, 75, 77, 78, 90	0.20
77	Kab. Blitar	4	76, 78, 79, 103	0.25
78	Kab. Kediri	6	76, 77, 79, 89, 90. 102	0.17
79	Kab. Malang	9	77, 78, 80, 85, 86, 88, 89, 104, 110	0.11
80	Kab. Lumajang	3	79, 81, 85	0.33
81	Kab. Jember	4	80, 82, 83, 85	0.25
82	Kab. Banyuwangi	3	81, 83, 84	0.33
83	Kab. Bondowoso	4	81, 82, 84, 85	0.25
84	Kab. Situbondo	3	82, 83, 85	0.33
85	Kab. Probolinggo	7	79, 80, 81, 83, 84, 86, 105	0.14
86	Kab. Pasuruan	6	79, 85, 87, 88, 106, 110	0.17
87	Kab. Sidoarjo	4	86, 88, 97, 109	0.25
88	Kab. Mojokerto	8	79, 86, 87, 89, 96, 97, 107, 110	0.13

Region Code	Region	Number of Neighbor	Regions Neighbor Code	Spatial Weighting Value
89	Kab. Jombang	6	78, 79, 88, 90, 94, 96	0.17
90	Kab. Nganjuk	6	74, 76, 78, 89, 91, 94	0.17
91	Kab. Madiun	6	74, 90, 92, 93, 94, 108	0.17
92	Kab. Magetan	6	44, 45, 74, 91, 93, 108	0.17
93	Kab. Ngawi	7	45, 46, 47, 48, 91, 92, 94	0.14
94	Kab. Bojonegoro	7	48, 89, 90, 91, 93, 95, 96	0.14
95	Kab. Tuban	4	48, 49, 94, 96	0.25
96	Kab. Lamongan	5	88, 89, 94, 95, 97	0.20
97	Kab. Gresik	4	87, 88, 96, 109	0.25
98	Kab. Bangkalan	2	99, 109	0.50
99	Kab. Sampang	2	98, 100	0.50
100	Kab. Pamekasan	2	99, 101	0.50
101	Kab. Sumenep	1	100	1,00
102	Kota Kediri	1	78	1,00
103	Kota Blitar	1	77	1,00
104	Kota Malang	1	79	1,00
105	Kota Probolinggo	1	85	1,00
106	Kota Pasuruan	1	86	1,00
107	Kota Mojokerto	1	88	1,00
108	Kota Madiun	2	91, 92	0.50
109	Kota Surabaya	3	87, 97, 98	0.33
110	Kota Batu	3	79, 86, 88	0.33
111	Kab. Pandeglang	2	112, 114	0.50
112	Kab. Lebak	5	7, 8, 111, 113, 114	0.20
113	Kab. Tangerang	7	5, 6, 7, 112, 114, 115, 118	0.14
114	Kab. Serang	5	111, 112, 113, 116, 117	0.20
115	Kota Tangerang	4	2, 5, 113, 118	0.25
116	Kota Cilegon	1	114	1.00
117	Kota Serang	1	114	1.00
118	Kota Tangerang Selatan	5	2, 7, 29, 113, 115	0.20