Journal of Physics: Conference Series

PAPER • **OPEN ACCESS**

Spatial Econometric Model for Mapping Poverty Area in West Sulawesi

To cite this article: Asrirawan Asrirawan *et al* 2021 *J. Phys.: Conf. Ser.* **1752** 012048

Spatial Econometric Model for Mapping Poverty Area in West Sulawesi

Asrirawan[1*], Rahmawati[1], Hikmah[1], and M Abdy[2]

[1] Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Sulawesi Barat, Indonesia.
[2] Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Makassar, Indonesia

*e-mail: asrirawan@unsulbar.ac.id

Abstract. This paper depicts the frameworks of spatial regression methods which cover the spatial interaction effects and apply it to analyze the spatial distribution of poverty area for five districts in West Sulawesi. These methods consider queen contiguity-based spatial weights to create spatially lagged variables and apply to Spatial Lag Model, Spatial Error Model (SEM), Spatial Durbin Model and Ordinary Least Square (OLS) Model. The percentage of poverty area represents the observed dependent variable in this article while several independent variables are recorded by the previous research and correlated to poverty. The analysis results showed that the spatial effects should be involved to the econometric model because an estimated spatial parameter is statistically significant. To select model criteria, AIC and Likelihood ratio test are used and this criteria highlight Spatial Lag Model is better than the other model.

Keyword: Spatial econometric, mapping poverty area, west sulawesi

1. Introduction

Poverty alleviation is one of the nation's most important challenges and need some strategies to reduce this root causes. In 2017, the national poverty line reached its lowest point for almost two decades, which was 10.12 percent. Data released by the Badan Pusat Statistik (BPS), Indonesia, in the year 2017 shows that the percentage of Indonesia's poor is reduced by 0.58 percentage points (year-on-year). However, based on the latest statistical data and socioeconomic development, Indonesia is continuing its efforts to eradicate extreme poverty both in rural and urban areas today. From the dynamics of the poverty rate in 2009-2017, the percentage of rural poverty rate is higher than urban poverty. It was 13.47 percent accounted for roughly 16.31 million people, while urban areas was 7.26 percent at approximate 10.27 million. In the period 2010-2014, the level of poverty reduction was at rural areas faster than urban areas [1].

The number of poor population who have monthly per capita expenditure (MPCE) below the Poverty Line in West Sulawesi reached 151.78 thousand people (11.25 percent) for the year 2018. It experienced an increase of 2.3 thousand people compared to the number of indigent population for the previous year, who accounted for 149.47 thousand people (11.18 percent). 9.50 percent of the urban population was poor, compared to 11.70 percent of the rural population in 2017. In 2018, the percentage increase in the number of indigent population of urban and rural be 9.64% and 11.75% respectively. The role of food commodities towards the poverty line is greater than the role of non-food commodities such as housing, clothing, education, and health. Donations Food Poverty Line to Poverty Line in March 2018 was recorded at 78.27 percent. This condition is not much different from the condition in September 2017 in the amount of 78.99 percent [2].

To reduce poverty, it is important to know what factors are actually related to the level of poverty in Indonesia so that an effective public policy can be formulated to diminish it in the future and not only the number but also qualitative scale [3]. Modeling poverty using spatial econometrics has been done by [3] Zelinsky [4], Paraguas and Kamil [5], and Zewdie, Aidi and Sartono [6]. Moreover, a spatial modelling framework which explicitly calculate the correlation from a region to neighbouring regions may be more suitable. Particular locations contribute to increase the number of poverty incidence, suggesting the significance of neighbourhood influence [7].

2. Method

2.1. Case Data

The data were collected from BPS Majene and BPS West Sulawesi on Measurement of Macroeconomic and Social Work in West Sulawesi. The response variable used is the percentage of poverty in five districts (Mamuju, Majene, Polewali Mandar, Mamasa and North Mamuju) with involving 240 observations. On the other side, the predictor variables that are assumed to have spatial correlation with the poverty are the percentage of literacy rate (X_1), the percentage of regional budget (X_2), the percentage of labour force participation rate (X_3), the percentage of open unemployment rate (X_4), percentage of income (X_5) and the percentage of drinking-water sources (X_6). Then, framework of R software is provided to analyze the model and assumptions.

2.2. Panel Data Regression

Panel data is a combination of cross-section data and time series data. In the panel data, the same cross-section units are surveyed over several time periods. So, panel data has dimensions of space and time. If each cross-section unit has the same number of time series observations, then the panel data is called balanced panel data, whereas if the number of time series observations is different in each unit, it is called unbalanced panel data data). The panel regression model can generally be stated in the following form:

$$y_{it} = \alpha_{it} + \boldsymbol{\beta}\mathbf{X_{it}} + u_{it};\ i = 1,2,3,\ldots,N;\ t = 1,2,3,\ldots,T \tag{1}$$

where y_{it} is i-th cross section unit for the t-time period, $\mathbf{X}$ is Observation vector on the independent variable, $\boldsymbol{\beta}$ is constant vector, α_{it} is intercept i-th object t-time, u_{it} is regression error for t-i group the t-time, and $u_{it} \sim IIDN\left(0,\sigma^2\right)$ [8] [9] [10].

Journal of Physics: Conference Series **1752** (2021) 012048 doi:10.1088/1742-6596/1752/1/012048

Common effect Model (CEM) approach assumes that the intercept and slope values of each variable are the same for all cross section units and time series

$$y_{it} = \alpha + \boldsymbol{\beta}\mathbf{X_{it}} + u_{it};\ i = 1,2,3,\ldots,N;\ t = 1,2,3,\ldots,T \tag{2}$$

Fixed effect Model (FEM) approach is assumed that the slope value of each variable is fixed but the intercept value is different for each unit cross section and fixed for each unit time series

$$y_{it} = \alpha_i + \boldsymbol{\beta}\mathbf{X_{it}} + u_{it};\ i = 1,2,3,\ldots,N;\ t = 1,2,3,\ldots,T \tag{3}$$

2.3. *Spatial Weight Matrix*

Spatial weighting matrix is a matrix that states the relationship of the observation area that is sized $n \times n$ and symbolized by $\mathbf{W}$. The general form of the spatial weighting matrix ($\mathbf{W}$) is:

$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix} \tag{4}$$

The elements $\mathbf{W}$ from above are w_{ij} by i are rows in elements $\mathbf{W}$ and j are columns in elements $\mathbf{W}$ and are the area around the observation i site. The above element $\mathbf{W}$ can have two values namely zero and one. Where the value $w_{ij} = 1$ for the area adjacent to the observation location, while the value $w_{ij} = 0$ for the area that is not close to the observation location. In general, there are four types of interactions or border crossings.

2.3.1. Rook Contiguity. Rook Contiguity is the contact of one side of the territory with the other side of the neighboring territory. The value of each element is that if the locations of i and j are side to side then $w_{ij} = 1$. However, if the locations of i and j do not touch the sides then $w_{ij} = 0$.

2.3.2. Bishop Contiguity. Bishop contiguity is the percentage of a point of one region with another neighboring region. The value of each element is that if the locations of i and j are in contact with the angles then $w_{ij} = 1$. However, if the locations of i and j do not touch the angles then $w_{ij} = 0$.

2.3.3. Queen Contiguity. Queen Contiguity, namely contact between the sides and vertices of one region with another region, which is a combination of rook contiguity and bishop contiguity. The value of each element is if the locations of i and j touch the sides or vertices then $w_{ij} = 1$. However, if the locations of i and j do not touch the sides or vertices then $w_{ij} = 0$.

2.3.4. Customized Contiguity. Customized contiguity, this method defines $w_{ij} = 1$ for regions that coexist or regions with the same characteristics as the region of interest, and $w_{ij} = 0$ for other regions.

2.4. Spatial Autoregressive Model

The spatial autocorrelation model is a combination of spatial lag effect model and spatial error model which calls most of the time Simultaneous autoregressive model or general spatial model according to (Lesage, 2009)

$$y_{it} = \rho \sum_{j=1}^{N} Wy_{jt} + \mathbf{\beta X_{it}} + u_{it};\ i, j = 1,2,3,\ldots,N;\ t = 1,2,3,\ldots,T \tag{5}$$

$$u_{it} = \lambda \sum_{j=1}^{N} Wu_{jt} + \varepsilon_{it} \tag{6}$$

The positive and negative spatial autocorrelation can be determined with using value cluster in a map. If similar value cluster together, it tend to be a positive spatial correlation while if dissimilar value cluster together, it tend to be a negative spatial correlation [7].

2.5. Spatial Lag Model

When λ is equal 0 in Spatial Autocorrelation Model, it can derive other models SAR. Meaning, a spatial lag model or following SAR model can be formulated which is comparable to the time-series lagged dependent variable

$$y_{it} = \rho \sum_{j=1}^{N} W_{ij} y_{jt} + \mathbf{\beta X_{it}} + \mu_i + \varepsilon_{it};\ i = j = 1,2,3,\ldots,N;\ t = 1,2,3,\ldots,T \tag{7}$$

2.6. Spatial Error Model

When $\rho = 0$ in Spatial Autocorrelation Model, the error term of a spatial error model (SEM) with spatial effect can be derived the form

$$\begin{gathered} y_{it} = \mathbf{\beta X_{it}} + \mu_i + \phi_{it};\ i = j = 1,2,3,\ldots,N;\ t = 1,2,3,\ldots,T \\ \phi_{it} = \rho \sum_{j=1}^{N} W_{ij} \phi_{it} + \varepsilon_{it} \end{gathered} \tag{8}$$

2.7. Spatial Durbin Model

Spatial Durbin Model [11] can be expressed by formulation below

$$y_{it} = \rho \sum_{j=1}^{N} W_{ij} y_{jt} + \mathbf{\beta X_{it}} + W\mathbf{X_{it}}\theta + \varepsilon_{it};\ i = j = 1,2,3,\ldots,N;\ t = 1,2,3,\ldots,T \tag{9}$$

3. Results and Discussions

3.1. Descriptive Analysis of Poverty Rate

Figure 1 displays the poverty rate in five districts over 8 years old (2010-2017). Overall, the poverty rate for all locations decrease significantly from 2010 to 2017. Polewali Mandar was the highest percentage of poverty rate in 2010 and 2017, whilst the other categories such as Majene, Mamasa, Mamuju and North Mamuju each represented about 18%, 16%, 8% and 6% respectively. However, the percentage of poverty rate in North Mamuju was the lowest for all over year and the second was Mamuju. These locations had characteristically the same poverty rate for all years.

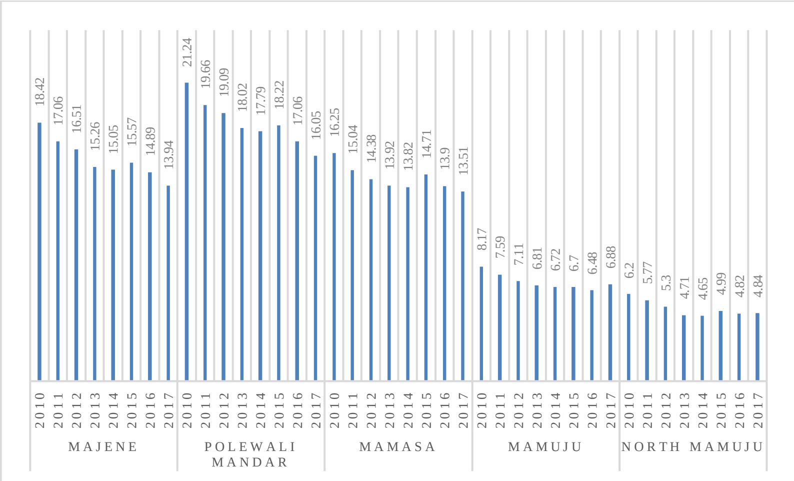

Figure 1. Poverty Rate in West Sulawesi for year 2010-2017

3.2. Spatial Autocorrelation Test

After creating weighted matrix, the spatial autocorrelation test and spatial dependency test need to be analyzed to allow for spatial interpolation and Morans'I is commonly used statistic to measure spatial autocorrelation which is identified by a correlation in neighboring locations in space. Spatial regression model will be the main priority model to analyze this case when spatial auto-correlation test is significant based on Moran's I statistics.

Table 1. Spatial Autocorrelation and Dependency Test

Moran I statistic standard deviate = 6.932, p-value < 2.2e-16		
Moran I statistic	Expectation	Variance
0.218276163	-0.00739271	0.00237284
Lagrange multiplier diagnostics for spatial dependence		
LMerr = 16.21, d lf = 1, p-value = 1.9234e-04		
LMlag = 28.36, df = 1, p-value = 1.5832e-06		
Robust Lagrange multiplier diagnostics for spatial dependence		
RLMerr = 0,56121290, df = 1,p-value = 0,0672		
RLMlag = 7.9210,df = 1, p-value = 0,000561		

In general, table 1 shows data that the value of Moran I test statistic is 0.2183. This value defines that a positive autocorrelation occurs in this poverty issues. Furthermore, it can be seen that p-value or calculated probability of Morans' I is 2.2e-16 which indicates that α value (0.05) is greater than p-value so we tend to reject the null hypothesis. Consequently, there is a positive spatial autocorrelation which means that one location is similar to other nearby locations. In a different way Moran's I test statistics can be explained as the correlation between poverty, variables, and the spatial lag of poverty rate. To determine it, we need to calculate the average of poverty rate value for the neighboring polygons. After identifying the spatial correlation, the next step is how to interpret the spatial dependence. Based on Table 1, it highlights that the traditional model using OLS is not appropriate because Lagrange multiplier shows all spatial lag and spatial error dependence significantly exist. Also, Robust LM test implies that the best model for poverty rate may be SAR model due to its p-value criteria. P-value for Robust LM lag is small than another one. Next, we have to consider AC and LR test as comparison.

3.3. Parameter Estimation of OLS, SAR, SEM and SDM

Table 2. Parameter Estimation by Using OLS, SAR, SEM and SDM

Significant Variable	OLS	SAR	SEM	SDM
Intercept	12.0671 (0.0000321)	5.9211 (0.000561)	8.9004 (0.00199)	9.7882 (0.000612)
X_1	0.03271 (0.005021)			
X_2	-0.7981 (0.02257)	-0.2542 (0.001243)	-0.3367 (0.00545)	-0.4912 (0.006913)
X_3				
X_4	0.0783 (0.000062)	0.00239 (0.0239)	0.00716 (0.00942)	0.00862 (0.00967)
X_5	-0.2153 (0.0034)	-0.0178 (0.000129)	-0.1341 (0.000412)	-0.1782 (0.0000937)
X_6	0.3412 (0.04018)			
Lagged log y(ρ)		0.3217 (0.00002221)		0.4918 (0.0007215)
Lagged error y(λ)			0.7206 (0.00000882)	
Lag X_1				-0.048(0.021)
Lag X_2				0.0721(0.039)
Lag X_3				-0.0635(0.0426)
AIC	78.34	56.31	59.34	62.05
LR Test				42.90 (0.0336)
N	5	5	5	5
RLM			7.921 (0.000561)	0.5612 (0.0672)

Overview, table 2 describes that the best model for poverty rate in West Sulawesi is spatial lag model (SAR) which has the smallest AIC (56.31) compared to OLS, SEM and SDM which have AIC 78.34, 59.34, and 62.05 respectively. Moreover, OLS and spatial lag model will be the best choice model because of the likelihood ratio test from the above table. It represents value to reduce the spatial error model because this model differs from the spatial durbin model. Nonetheless, based on the previous analysis, OLS is not appropriate based on LM test so we can finally conclude that the best model for poverty rate is SAR model. On the other side, the percentage of regional budget (X2), the percentage of open unemployment rate (X4) and the percentage of income (X5) have significantly influences to the percentage of the poverty rate with looking at the p-value (0.001243, 0.0239, and 0,000129). On contrary, the percentage of literacy rate, labor force participation rate and drinking-water sources is not significant, suggesting that we these factors will not have an impact on the poverty rate. Regional budget and income affect negatively to the poverty rate while open unemployment rate has a positive impact to this issue. It means that increased regional budget and income in neighboring locations have a negative impact on poverty rate. In another interpretation, we can see that the lag model on the partial lag effect is significant, which is 0,3217 with p-value 0.00002221. It means that average 100 percent increased in poverty rate in a location will result 32.17 percentage to increase in poverty rate in a neighboring location. The spatial lag model for poverty rate can be formulated as follow:

$$y_{it} = 5.9211 - 0,2542X_1 + 0,00239X_2 - 0.0178X_3 + 3217 \sum_{j=1}^{5} W_{ij} y_{jt} + \mu_i + \varepsilon_{it}; \quad i, j = 1,2,3,4,5; \; t = 1,2,3,\ldots,8 \tag{10}$$

3.4. Assumption Check

The next step is to meet the requirement of assumptions which are normality and homoscedasticity test. The errors are normally distributed on the large sample and the same errors variance is homoscedastic.

Tabel 3. Normality and Homoscedasticity Test

KS= 0.07802, p-value = 0.8562 alternative hypothesis: two-sided for SAR model
Studentized Breusch-Pagan test BP =6.087, df = 5, p-value =0.3288 for SAR model
Durbin-Watson statistic=2.25382617 for SAR (dl= 1.562, du=1.978)
For comparison DW = 1.893, p-value = 0.03762 for OLS (dl= 1.587 , du =1.922)

Table 3 displays that the formed model has fulfilled the normality assumption based on the p-value of Kolmogorov-Smirnov (KS), 0.8562. This value leads to accept the null hypothesis (p-value> α) so the error term follow normal distribution while Studentized Breusch-Pagan Test gives information that the error variance is constant. It can be seen from the BP p-value which is greater than α value (0.05). this value also indicates to accept the null hypothesis.

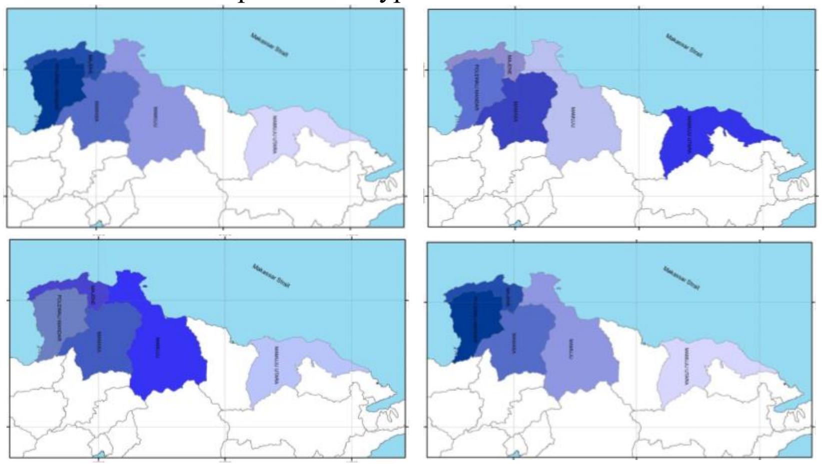

Figure 2. The Spread of Poverty Rate in West Sulawesi

4. Conclusions

Modelling the poverty rate by using several explanatory variables compared to traditional and spatial regression model (OLS, SAR, SEM and SDM) can be concluded that SAR model is the best model to see the influence of these variables to dependent variable. The percentage of the poverty rate is significantly influenced by the percentage of regional budget (X2), open unemployment rate (X4) and income (X5).

Acknowledgment

The authors acknowledge the BPS in Majene and West Sulawesi for making surveillance data available and Ministry of Research, Technology and Higher Education of the Republic of Indonesia through DIPA contributed to give the cost for researcher.

References

[1] Badan Pusat Statistik Indonesia 2018 *Information and Data of Poverty in Districts/City in 2018*. Catalog BPS No. 3205014, Jakarta: CV Nario Sari
[2] Badan Pusat Statistik Sulawesi Barat 2018 *Poverty Profile on March 2018 in West Sulawesi*, West Sulawesi: Majene
[3] Setiawati AK and Setiawan 2012 *Modeling Percentage of Poor Population in East Java with the Spatial Panel Econometrics Approach. Art and Sciences Journal ITS* **1(1)**
[4] Zelinsky T 2014 *Regional Poverty Levels in the European Union: A Spatial Econometric Approach*, Economy of region, Centre for Economic Security, Institute of Economics of Ural Branch of Russian Academy of Sciences, **1(2)** 62-69
[5] Paraguas FJ and Kamil AA 2005 *Spatial Econometrics Modeling of Poverty*, Proceedings of the 8th WSEAS International Conference on Applied Mathematics, Tenerife, Spain, December 16-18, 2005 159-164
[6] Zewdie MA, Aidi MN and Sartono B 2015 *Spatial Econometric Model of Poverty in Java Island*, American Journal of Theoretical and Applied Statistics **4(6)** 420-425
[7] Anselin L 1988 *Spatial Econometrics: Methods and Models*, Netherlands: Kluwer Academic Publisher
[8] Gujarati DN 2005 *Basic Econometric 5th Edition*, New York: Mc Graw Hill Companies
[9] Hsiao C 2003 *Analysis of Panel Data*, New York: Cambridge University Press
[10] Setiawan and Dwi EK 2010 *Ekonometrics*, Yogyakarta: Andi Offset
[11] Lesage JP 1999 *The Theory and Practice of Spatial Econometrics*, Department of Economics University of Toledo.