

Lagrange Multiplier Test for Spatial Autoregressive Model with Latent Variables

Anik Anekawati ^{1,2} , Bambang Widjanarko Otok ^{1,*}, Purhadi Purhadi ¹ and Sutikno Sutikno ¹

¹ Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia; anik@wiraraja.ac.id (A.A.); purhadi@statistika.its.ac.id (P.P.); sutikno@statistika.its.ac.id (S.S.)

² Faculty of Teacher Training and Education, Universitas Wiraraja, Sumenep 69451, Indonesia

* Correspondence: bambang_wo@statistika.its.ac.id

Received: 15 July 2020; Accepted: 17 August 2020; Published: 18 August 2020



Abstract: The focus of this research is to develop a Lagrange multiplier (LM) test of spatial dependence for the spatial autoregressive model (SAR) with latent variables (LVs). It was arranged by the standard SAR, where the independent variables were replaced by factor scores of the exogenous latent variables from a measurement model (in structural equation modeling) as well as their dependent variables. As a result, an error distribution of the SAR-LVs should have a different distribution from the standard SAR. Therefore, this LM test for the SAR-LVs is based on the new distribution. The estimation of the latent variables used a weighted least squares (WLS) method. The estimation of the SAR-LVs parameter used a two-stage least squares (2SLS) method. The SAR-LVs model was applied to the model with a positive and negative spatial autoregressive coefficient to illustrate how it was interpreted.

Keywords: Lagrange multiplier; spatial autoregressive; latent variable

1. Introduction

Researchers have often faced models involving latent variables and analyzed the relationships between two or more of those latent variables simultaneously. Latent variables are unobserved or unmeasured variables [1] and measured by connecting to the observed variables because they cannot be directly measured [2]. The statistical methodology that is able to accommodate these two objectives is structural equation modeling (SEM). SEM is a statistical method used to test the relationships between latent variables (path models) and between its observed variables (confirmatory factor models) [2]. In general, SEM has two submodels, namely the measurement model and the structural model. The structural model describes the relationship between latent variables, while the measurement model is the relationship between indicators and the latent variables that construct it.

In social research, analyses involving latent variables and at the same time having a spatial effect have often been found. Spatial dependence may be caused by different kinds of spatial spillover effects. There are two frameworks that involve spatial data in the SEM model, namely at the level of the measurement model or the structural model. The involvement of spatial data at the level of the measurement model is commonly used to analyze multivariate spatial data [3], i.e., when several variables are measured at the same locations over a spatial area, and they are often correlated with each other. Each of them might also be correlated across the locations because of geographic similarities of the different locations.

Christensen and Amemiya [4] suggested a model with a latent variable that is distribution-free to analyze multivariate spatial data. However, it was limited by the assumption of a linear relationship between observed and latent variables that might not apply to Poisson and binomial data. The parameters of the model were estimated by means of a moment method. Wang and Wall [3]

proposed the generalized common spatial factor, which was an extension of the traditional factor analysis model. In this model, it was assumed that the common factors were spatially correlated and extended to handle more types of observed data from exponential families, especially Poisson and binomial data.

Hogan and Tchernis [5] proposed the method of a Bayesian hierarchical model for analysis factors of spatially correlated multivariate data. At the first level, the distribution of a vector of manifest variables was conditional on an underlying latent factor in each location, whereas at the second level, the area-specific latent factors had a joint distribution that combined spatial correlation.

In contrast to previous researchers that only analyzed multivariate spatial data in measurement models in the SEM model, Liu et al. [6] developed a generalized spatial structural equation model (GSSEM). They joined the generalized common spatial factor model proposed by Wang and Wall in [3] and SEM that calculated spatial correlations. The GSSEM can also be extended to spatial correlations that can be added to the measurement model. Congdon [7] used the factor analysis on the measurement model. In this model, the construct was observed through indicators. Indicators allowed both spatial correlation and correlation with one another. The relation among constructs that are nonlinear was approached using a spline regression.

Oud and Folmer [8] proposed a SEM approach to the spatial dependence model. They combined the standard spatial model in [9] with the multiple indicators multiple causes (MIMIC) model in [10]. In this approach, the spatial weight that described the spatial spillover effects was located in the structural model. This approach was more flexible and informative compared to modeling that gave the spatial weight to the measurement model. The parameters of this model were estimated using Full Information Maximum Likelihood (FIML) and resulted in an estimator to control the bias of endogeneity due to the interaction between the dependent variable and its lag.

Anekawati et al. [11] conducted modeling of education quality in the senior high school level using the spatial SEM approach. Although they allocated the spatial weight on the structural model, their work had a different perspective on the model in [8]. They developed the spatial SEM model from the standard spatial model by Anselin in [9] but replaced the dependent variables by endogenous latent variables, as well as independent variables. The latent variables were estimated as the factor scores using the partial least squares (PLS) method through iterative estimation developed by Trujillo in [12]. The factor scores were modeled by involving the spatial effects, and the spatial dependence of this model was tested using the Lagrange multiplier (LM) test. The results of the spatial dependence test led to the spatial autoregressive (SAR) model. Furthermore, this model was called the spatial autoregressive model with latent variables (SAR-LVs). The parameters of the SAR-LVs model were estimated using maximum likelihood estimation (MLE).

Anekawati et al. [13] estimated the parameters of the SAR-LVs model from Anekawati et al. in [11] using the generalized method of moment (GMM), which was developed by Kelejian and Prucha in [14,15]. The SAR-LVs model in [13] indicated a better fit for the model than the MLE method since it produced a higher R-square. Additionally, the GMM was computationally easier than MLE.

This idea can be used as alternative modeling involving latent variables and spatial data simultaneously, as the research limitation in [16]. The research purposes in [16] were to identify the relationship between vulnerability factors related to social, economic, and environmental aspects, and economic losses from natural disasters in 230 local communities in South Korea. The social, economic, and environmental aspects were latent variables measured by connecting to the observed variables. The social aspect was constructed by two indicators, namely the percentage of the population over age 15 without elementary school completion and a minority percentage of foreigners. Additionally, the economic and environmental aspects were latent variables. However, the relationship was modeled based on indicators, which were not latent variables. It was less precise to identify the research purpose. That study revealed the limitations of the study, which was an indicator-based approach for the identification of vulnerability factors. Therefore, the method in this study provides an alternative solution for the spatial model that involves latent variables, especially focusing on the spatial dependency test using the LM test.

Oud and Folmer [8] did not perform a diagnostic test of spatial dependence, so there was no direction in determining whether the model led to the spatial autoregressive model or the spatial error model. Anekawati's research works in [11,13] tested the spatial dependence using the Lagrange multiplier (LM) to diagnose spatial dependence.

One of the constructions of the test of parametric hypotheses based on asymptotic theory is the LM test [17]. Anselin [18] developed the diagnostics for spatial dependence using the LM test. The LM approach seems reasonable and relatively easy based on estimation under the null hypotheses [17], namely, in its most simple form. Yang [19] introduced a residual-based bootstrap method for asymptotically refined approximations to the finite sample critical values of the LM statistics.

The LM test of spatial dependence [18,19] did not involve latent variables for the standard SAR model. Anekawati's work [11,13] used the LM test of spatial dependence for the SAR-LVs model, but an error distribution of the model was assumed the same as the standard SAR model in [9,18]. The LM test for the standard SAR model in [9,18] had an assumption that error was normally distributed $\varepsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$. Meanwhile, the SAR-LVs model was modeled based on the standard SAR model, where factor scores replaced the independent and dependent variables. The factor scores were the estimation result of the latent variables in the submodel in SEM, namely, the measurement model. The measurement model had the assumption that the error was normally distributed, namely, $\delta \sim N(\mathbf{0}, \Theta_\delta)$ for exogenous and $\varepsilon^* \sim N(\mathbf{0}, \Theta_{\varepsilon^*})$ for endogenous, while the error distribution in the standard SAR model was $\varepsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$. As a result, the error distribution of the SAR-LVs model should have a different distribution from the error of the standard SAR model. In this paper, an attempt is made to fill this gap. The focus of this study is to develop a Lagrange multiplier test of spatial dependence for the spatial autoregressive model (SAR) with latent variables (LVs).

To complete this paper, the estimation of latent variables into factor scores uses the weighted least squares (WLS) method, so that the error distribution of the SAR-LVs model is constructed from the result of this estimation. The estimation of parameters of the SAR-LVs model uses the two-stage least squares (2SLS) method. In the last section, the SAR-LVs model is applied for cases of the positive and negative spatial autoregressive coefficient to provide an interpretation of the spatial autoregressive coefficient.

2. Materials and Methods

2.1. Spatial Autoregressive Model with Latent Variable (SAR-LVs Model)

SEM consists of two basic components—the structural model and measurement model in [2]. The measurement model represents the relationship between the manifest variable and exogenous latent variables (1) or endogenous latent variables (2), while the structural model describes the relationship among the latent variables (3). Bollen in [1] wrote the measurement and structural model as Equations (1)–(3).

$$\mathbf{x} = \Lambda_x \boldsymbol{\xi} + \boldsymbol{\delta} \text{ with } \boldsymbol{\delta} \sim N_A(\mathbf{0}, \Theta_\delta) \quad (1)$$

$$\mathbf{y} = \Lambda_y \boldsymbol{\eta} + \boldsymbol{\varepsilon}^* \text{ with } \boldsymbol{\varepsilon}^* \sim N_B(\mathbf{0}, \Theta_{\varepsilon^*}) \quad (2)$$

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma} \boldsymbol{\xi} + \boldsymbol{\zeta} \quad (3)$$

where $\boldsymbol{\eta}$ is the $(q \times 1)$ endogenous random vector, q is the number of the endogenous variables, $\boldsymbol{\xi}$ is the $(p \times 1)$ exogenous random vector, p is the number of the exogenous variables, $\mathbf{B}$ is the $(q \times q)$ coefficient matrix that shows the effect of the relationship of an endogenous latent variable to another endogenous variable, $\boldsymbol{\Gamma}$ is the $(q \times p)$ coefficient matrix which shows the effect of $\boldsymbol{\xi}$ relationship to $\boldsymbol{\eta}$, $\boldsymbol{\zeta}$ is the $(q \times 1)$ random error vector, $\mathbf{y}$ is the $(B \times 1)$ observed vector of the endogenous variables, B is a total number of indicators of the endogenous variables, $\mathbf{x}$ is the $(A \times 1)$ observed vector of the exogenous, A is a total number of indicators of the exogenous variables, Λ_y is the $(B \times q)$ coefficient matrix which shows the relationship of $\mathbf{y}$ to $\boldsymbol{\eta}$, Λ_x is the $(A \times p)$ coefficient matrix which shows the relationship of $\mathbf{x}$ to $\boldsymbol{\xi}$, $\boldsymbol{\varepsilon}^*$ is the $(B \times 1)$ measurement error vector of $\mathbf{y}$, and $\boldsymbol{\delta}$ is the $(A \times 1)$ measurement error vector of $\mathbf{x}$.

Assumptions that must be fulfilled are $E(\eta) = \mathbf{0}$, $E(\xi) = \mathbf{0}$, $E(\zeta) = \mathbf{0}$, $E(\xi\zeta') = \mathbf{0}$, $E(\varepsilon^*) = \mathbf{0}$, $E(\delta) = \mathbf{0}$, $(\mathbf{I} - \mathbf{B})$ is nonsingular, and the element of error vectors of measurements, namely δ_i and ε_i^* , are homoscedastic and nonautocorrelated across observations (see [1]).

Anselin wrote the standard SAR model in [9]:

$$\mathbf{y}^* = \lambda \mathbf{W}\mathbf{y}^* + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \text{ with } \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}), \quad (4)$$

where $\mathbf{y}^*$ is the $(T \times 1)$ spatially lagged dependent vector, T is the number of the observed units, $\mathbf{X}$ is the $(T \times (p + 1))$ exogenous matrix, λ is coefficient of $\mathbf{y}^*$, $\boldsymbol{\beta}$ is the $((p + 1) \times 1)$ parameters vector of exogenous, $\mathbf{W}$ is the $(T \times T)$ spatial weight matrix with the main diagonal elements being zero, $\boldsymbol{\varepsilon}$ is the $(T \times 1)$ disturbance vector, where it is the classic homoscedastic situation. The queen contiguity method was used for spatial weighting. In queen contiguity, $\mathbf{W}_{ij}$ is defined as 1 for the entity where the common side or the common vertex meets the region of concern, and $\mathbf{W}_{ij}$ is defined as 0 for other regions [20].

Oud and Folmer in [8] wrote the SAR-LVs model in the form of the MIMIC. The proposed model was $\hat{\eta} = \lambda \mathbf{W}\hat{\eta} + \tilde{\mathbf{X}}\boldsymbol{\gamma} + \tilde{\boldsymbol{\zeta}}$, where λ was the spatial lag coefficient of the endogenous variable, $\mathbf{W}$ was the contiguity matrix, and $\tilde{\mathbf{X}}$ was the observation matrix of the explanatory variable.

The SAR-LVs model is a model that involves latent variables, and the unit of observation is location. The SAR-LVs model is a standard SAR model in which the independent and dependent variables are latent variables. In the SEM model, there are latent variables that cannot be measured directly as a sample unit. Therefore, in this work, to represent the latent variable in the standard SAR in Equation (4) is changed by the factor score. The latent variable is replaced by the factor score from the measurement model in Equations (2) and (3) as a measured and random unit sample. As seen in Equation (4), the spatially lagged dependent variable ($\mathbf{y}^*$) is changed by the endogenous latent variable (η), and the exogenous variable ($\mathbf{X}$) is changed by the exogenous latent variable (ξ), which is previously estimated using the WLS method. The result of estimation of the latent variable is denoted by $\hat{\eta} = \mathbf{I}$ and $\hat{\xi} = \mathbf{K}$. This SAR-LVs model does not use the MIMIC model, since there are no exogenous or endogenous variables that are observed variables, and the endogenous variable is limited to only one.

Thus, the SAR-LVs model in Equation (4) changes to:

$$\mathbf{I} = \mathbf{K}\boldsymbol{\beta} + \lambda \mathbf{W}\mathbf{I} + \boldsymbol{\varepsilon}, \quad (5)$$

where $\mathbf{I}$ is the $(T \times 1)$ endogenous factor score vector, $\mathbf{K}$ is the $(T \times (p + 1))$ exogenous factor score matrix, and $\boldsymbol{\beta}$ is the $((p + 1) \times 1)$ regression coefficient vector.

2.2. Estimation of Score of Latent Variable

The factor score is the estimation result of the latent variables, both the endogenous and exogenous variables in the measurement model. The method used is the WLS, which is by minimizing the sum squared errors that are weighted by the error variant matrix. In the estimation process, to obtain a factor score, it is assumed that the value of the loading factor and the error variant matrix are constant.

In the equation of the measurement model of the exogenous latent variable (1), where p is the number of the exogenous latent variable, a_i is the number of indicators of the i th exogenous latent variable, and $\sum_{i=1}^p a_i = A$. δ is an error, where $\delta \sim N_A(\mathbf{0}, \boldsymbol{\Theta}_\delta)$ with $\boldsymbol{\Theta}_\delta(A \times A)$ is the covariant–variant matrix of measurement error of observed variable $\mathbf{x}$, namely

$$\boldsymbol{\Theta}_\delta = \text{diag}(\sigma_{\delta_{(1)1}}^2, \sigma_{\delta_{(2)1}}^2, \dots, \sigma_{\delta_{(a1)1}}^2, \sigma_{\delta_{(a1)2}}^2, \sigma_{\delta_{(a2)2}}^2, \dots, \sigma_{\delta_{(a2)2}}^2, \dots, \sigma_{\delta_{(1)p}}^2, \sigma_{\delta_{(2)p}}^2, \dots, \sigma_{\delta_{(ap)p}}^2).$$

The distribution of $\mathbf{x}$ is obtained through the properties of the expected value and variance of a random variable. It is assumed that the value of $\boldsymbol{\Lambda}_x$ and $\boldsymbol{\Theta}_\delta$ are constant. $E(\mathbf{x}) = E(\boldsymbol{\Lambda}_x \boldsymbol{\xi} + \delta) = \boldsymbol{\Lambda}_x \boldsymbol{\xi}$ and $\text{var}(\mathbf{x}) = \text{var}(\boldsymbol{\Lambda}_x \boldsymbol{\xi} + \delta) = \text{var}(\delta) = \boldsymbol{\Theta}_\delta$. Therefore, if $\mathbf{x} = \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \delta$ and $\delta \sim N_A(\mathbf{0}, \boldsymbol{\Theta}_\delta)$ then

the distribution of $\mathbf{x}$ is $\mathbf{x} \sim N_A(\Lambda_x \xi, \Theta_\delta)$. Suppose that a random sample T is given from a random variable $\mathbf{x}$

$$(x_{(x)11}, x_{(2)12}, \dots, x_{(a_i)1T}, x_{(1)21}, x_{(2)22}, \dots, x_{(a_2)2T}, \dots, x_{(1)p1}, x_{(2)p2}, \dots, x_{(a_p)pT}, \xi_{11}, \xi_{12}, \dots, \xi_{1T}, \xi_{21}, \xi_{22}, \dots, \xi_{2T}, \dots, \xi_{p1}, \xi_{p2}, \dots, \xi_{pT})$$

with $t = 1, 2, \dots, T$ so $\mathbf{x}_t \sim N_A(\Lambda_x \xi_t, \Theta_\delta)$. The probability function of $\mathbf{x}_t$ is

$$f(\mathbf{x}_t) = (2\pi)^{-\frac{A}{2}} |\Theta_\delta|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{x}_t - \Lambda_x \xi_t)' \Theta_\delta^{-1} (\mathbf{x}_t - \Lambda_x \xi_t)\right\}$$

The likelihood function is

$$L(\xi, \Theta_\delta) = (2\pi)^{-\frac{AT}{2}} |\Theta_\delta|^{-\frac{T}{2}} \exp\left\{-\frac{1}{2} \mathbf{Q}_x\right\} \text{ with } \mathbf{Q}_x = \sum_{t=1}^T (\mathbf{x}_t - \Lambda_x \xi_t)' \Theta_\delta^{-1} (\mathbf{x}_t - \Lambda_x \xi_t)$$

The latent variable ξ is estimated using the WLS method with optimization $L(\xi, \Theta_\delta)$. Maximum $L(\xi, \Theta_\delta) \iff \min \mathbf{Q}_x$ by adding the weight of an error variant matrix Θ_δ is obtained as $\sum_{t=1}^T \hat{\xi}_t = (\Lambda_x' \Theta_\delta^{-1} \Lambda_x) (\Lambda_x' \Theta_\delta^{-1}) \sum_{t=1}^T \mathbf{x}_t$ or can be written in matrix form and contain each element as

$$\begin{pmatrix} \hat{\xi}_{11} & \hat{\xi}_{12} & \cdots & \hat{\xi}_{1T} \\ \hat{\xi}_{21} & \hat{\xi}_{22} & \cdots & \hat{\xi}_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\xi}_{p1} & \hat{\xi}_{p2} & \cdots & \hat{\xi}_{pT} \end{pmatrix} = (\Lambda_x' \Theta_\delta^{-1} \Lambda_x) (\Lambda_x' \Theta_\delta^{-1}) \begin{pmatrix} x_{(1)11} & x_{(1)12} & \cdots & x_{(1)1T} \\ x_{(2)11} & x_{(2)12} & \cdots & x_{(2)1T} \\ \vdots & \vdots & \ddots & \vdots \\ x_{(a_1)11} & x_{(a_1)12} & \cdots & x_{(a_1)1T} \\ x_{(1)21} & x_{(1)22} & \cdots & x_{(1)2T} \\ x_{(2)21} & x_{(2)22} & \cdots & x_{(2)2T} \\ \vdots & \vdots & \ddots & \vdots \\ x_{(a_2)21} & x_{(a_2)22} & \cdots & x_{(a_2)2T} \\ \vdots & \vdots & \ddots & \vdots \\ x_{(1)p1} & x_{(1)p2} & \cdots & x_{(1)pT} \\ x_{(2)p1} & x_{(2)p2} & \cdots & x_{(2)pT} \\ \vdots & \vdots & \ddots & \vdots \\ x_{(a_p)p1} & x_{(a_p)p2} & \cdots & x_{(a_p)pT} \end{pmatrix}$$

By assuming matrix $\begin{pmatrix} \hat{\xi}_{11} & \hat{\xi}_{12} & \cdots & \hat{\xi}_{1T} \\ \hat{\xi}_{21} & \hat{\xi}_{22} & \cdots & \hat{\xi}_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\xi}_{p1} & \hat{\xi}_{p2} & \cdots & \hat{\xi}_{pT} \end{pmatrix} = \mathbf{K}'$ and $\mathbf{X} = \begin{pmatrix} x_{(1)11} & x_{(1)12} & \cdots & x_{(1)1T} \\ x_{(2)11} & x_{(2)12} & \cdots & x_{(2)1T} \\ \vdots & \vdots & \ddots & \vdots \\ x_{(a_1)11} & x_{(a_1)12} & \cdots & x_{(a_1)1T} \\ x_{(1)21} & x_{(1)22} & \cdots & x_{(1)2T} \\ x_{(2)21} & x_{(2)22} & \cdots & x_{(2)2T} \\ \vdots & \vdots & \ddots & \vdots \\ x_{(a_2)21} & x_{(a_2)22} & \cdots & x_{(a_2)2T} \\ \vdots & \vdots & \ddots & \vdots \\ x_{(1)p1} & x_{(1)p2} & \cdots & x_{(1)pT} \\ x_{(2)p1} & x_{(2)p2} & \cdots & x_{(2)pT} \\ \vdots & \vdots & \ddots & \vdots \\ x_{(a_p)p1} & x_{(a_p)p2} & \cdots & x_{(a_p)pT} \end{pmatrix}$, so

$$\mathbf{K}' = (\Lambda_x' \Theta_\delta^{-1} \Lambda_x) (\Lambda_x' \Theta_\delta^{-1}) \mathbf{X}. \quad (6)$$

$\mathbf{X}$ in Equation (6) is a random observation matrix. Based on Definition A.1 in [21] (Appendix A), $\mathbf{X}$ has normal distribution $\mathbf{X} \sim N_{A,T}(\Lambda_x \xi_t \mathbf{e}', \Theta_\delta \otimes \mathbf{I}_T)$, where $\mathbf{e}_{T \times 1} = (1, \dots, 1)$.

The definition of the characteristic function of the random matrix $\mathbf{X}$ is $\phi_x(\mathbf{Z}) = E[\text{etr}(\iota \mathbf{X} \mathbf{Z}')] = E[\text{etr}(\iota \sqrt{-1} \mathbf{X} \mathbf{Z}')] = E[\text{etr}(\iota \mathbf{P} \mathbf{X} \mathbf{Z}')] = E[\text{etr}(\iota \mathbf{P} (\mathbf{X} \mathbf{Z}'))]$. If part of Equation (6) is assumed $(\Lambda_x' \Theta_\delta^{-1} \Lambda_x)(\Lambda_x' \Theta_\delta^{-1}) = \mathbf{P}$ then Equation (6) can be simplified into $\mathbf{K}' = \mathbf{X} \mathbf{P}$. The characteristic function of $\mathbf{K}'$ is $\phi_{\mathbf{K}'}(\mathbf{Z}) = E[\text{etr}(\iota (\mathbf{P} \mathbf{X}) \mathbf{Z}')] = E[\text{etr}(\iota \mathbf{P} (\mathbf{X} \mathbf{Z}'))]$. Based on Theorem B.1, the characteristic function of $\mathbf{K}'$ can be changed into $\phi_{\mathbf{K}'}(\mathbf{Z}) = E[\text{etr}(\iota \mathbf{X} (\mathbf{Z}' \mathbf{P}))]$. If $\mathbf{Z}'_1 = \mathbf{Z}' \mathbf{P}$ then the characteristic function of $\mathbf{K}'$ can be changed into $\phi_{\mathbf{K}'}(\mathbf{Z}) = E[\text{etr}(\iota \mathbf{X} \mathbf{Z}'_1)]$.

The distribution of $\mathbf{X}$ is $\mathbf{X} \sim N_{A,T}(\Lambda_x \xi_t \mathbf{e}', \Theta_\delta \otimes \mathbf{I}_T)$ and $\mathbf{K}' = \mathbf{X} \mathbf{P}$. Based on Theorem A.2, the characteristic function of $\mathbf{K}'$ is $\phi_{\mathbf{K}'}(\mathbf{Z}) = \text{etr}(\iota \mathbf{Z}'_1 \Lambda_x \xi_t \mathbf{e}' - 1/2 \mathbf{Z}'_1 \Theta_\delta \mathbf{Z}_1 \mathbf{I}_T)$. If $\mathbf{Z}'_1$ is changed by $\mathbf{Z}'_1 = \mathbf{Z}' \mathbf{P}$ then the characteristic function of $\mathbf{K}'$ is $\phi_{\mathbf{K}'}(\mathbf{Z}) = \text{etr}(\iota (\mathbf{Z}' \mathbf{P}) \Lambda_x \xi_t \mathbf{e}' - 1/2 (\mathbf{Z}' \mathbf{P})' \Theta_\delta (\mathbf{Z}' \mathbf{P}) \mathbf{I}_T)$. Furthermore, $\mathbf{P}$ is changed by $\mathbf{P} = (\Lambda_x' \Theta_\delta^{-1} \Lambda_x)(\Lambda_x' \Theta_\delta^{-1})$, so the characteristic function of $\mathbf{K}'$ is

$$\phi_{\mathbf{K}'}(\mathbf{Z}) = \text{etr}(\iota \mathbf{Z}' \xi_t \mathbf{e}' - 1/2 \mathbf{Z}' (\Lambda_x' \Theta_\delta^{-1} \Lambda_x)^{-1} \mathbf{Z} \mathbf{I}_T) \quad (7)$$

Based on Theorem A.2 and Equation (7), $\mathbf{K}'$ is the matrix variate normal distribution with mean $\xi_t \mathbf{e}'$ and covariate matrix $(\Lambda_x' \Theta_\delta^{-1} \Lambda_x)^{-1} \otimes \mathbf{I}_T$ and is notated by $\mathbf{K}' \sim N_{T,p}(\xi_t \mathbf{e}', (\Lambda_x' \Theta_\delta^{-1} \Lambda_x)^{-1} \otimes \mathbf{I}_T)$. Based on Theorem A.1, $\mathbf{K}$ is the matrix variate normal distribution that is notated by:

$$\mathbf{K} \sim N_{T,p}(\mathbf{e} \xi_t, \mathbf{I}_T \otimes (\Lambda_x' \Theta_\delta^{-1} \Lambda_x)^{-1}). \quad (8)$$

In this paper, the number of the endogenous variables is limited only to one. The covariance matrix of the measurement error of the observed variable for $\mathbf{y}$ is Θ_{ε^*} , namely $\Theta_{\varepsilon^*} = \text{diag}(\sigma_{\varepsilon_1^*}^2, \sigma_{\varepsilon_2^*}^2, \dots, \sigma_{\varepsilon_B^*}^2)$, where B is the number of indicators of the endogenous latent variable. In the same way as the previous estimation with the exogenous latent variable and by assuming that vector $(\hat{\eta}_1 \hat{\eta}_2 \dots \hat{\eta}_t) = \mathbf{l}'$ and

$$\mathbf{Y} = \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1T} \\ y_{21} & y_{22} & \cdots & y_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ y_{B1} & y_{B2} & \cdots & y_{BT} \end{pmatrix}, \text{ it is obtained}$$

$$\mathbf{l}' = (\Lambda_y' \Theta_{\varepsilon^*}^{-1} \Lambda_y)^{-1} \Lambda_y' \Theta_{\varepsilon^*}^{-1} \mathbf{Y}, \quad (9)$$

and its distribution is

$$\mathbf{l} \sim N_T(\mathbf{e} \eta_t, (\Lambda_y' \Theta_{\varepsilon^*}^{-1} \Lambda_y)^{-1} \mathbf{I}_T). \quad (10)$$

Theorem A.1, A.2, and B.1 are provided in Appendix A.

2.3. The Error Distribution of The SAR-LVs Model

The equation of the SAR-LVs model from the estimated factor score can be arranged as Equation (6) and adding β_0 , it can be written as

$$\mathbf{l} = \mathbf{K} \beta + \lambda \mathbf{W} \mathbf{l} + \varepsilon, \quad (11)$$

or can be written in the matrix form

$$\begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_T \end{pmatrix} = \begin{pmatrix} 1 & k_{11} & k_{12} & \cdots & k_{1p} \\ 1 & k_{21} & k_{22} & \cdots & k_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & k_{T1} & k_{T2} & \cdots & k_{Tp} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{pmatrix}$$

where $\mathbf{K} \sim N_{T,p}(\mathbf{e}\xi_t, \mathbf{I}_T \otimes (\Lambda'_x \Theta_\delta^{-1} \Lambda_x)^{-1})$ is Equation (6) and $\mathbf{l} \sim N_T(\mathbf{e}\eta_t, (\Lambda'_y \Theta_{\varepsilon^*}^{-1} \Lambda_y)^{-1} \mathbf{I}_T)$ is Equation (10).

The SAR-LVs model (11) is a spatial regression model by considering $\mathbf{l}$ as a response variable and $\mathbf{K}$ as a predictor variable, both of which are random. Thus, the function of $\mathbf{l}$ is $f(\mathbf{l}|\mathbf{K})$, so that the variable $\mathbf{K}$ is no longer random but fixed. As a result, the error in Equation (11) is $\varepsilon = (\mathbf{I} - \lambda \mathbf{W})\mathbf{l} - \mathbf{K}\beta$ where $\mathbf{l}$ is a random variable, $\mathbf{K}$ is fixed and is assumed not to correlate with ε , and $cov(\mathbf{l}, \mathbf{K}) \neq \mathbf{0}$.

The expectation value of ε is $E(\varepsilon) = (\mathbf{I} - \lambda \mathbf{W})\mathbf{e}\eta_t - \mathbf{K}\beta$, and the variance of ε is $var(\varepsilon) = (\mathbf{I} - \lambda \mathbf{W})(\Lambda'_y \Theta_{\varepsilon^*}^{-1} \Lambda_y)^{-1}(\mathbf{I} - \lambda \mathbf{W})'$, so the error distribution of ε is

$$\varepsilon \sim N_T((\mathbf{I} - \lambda \mathbf{W})\mathbf{e}\eta_t - \mathbf{K}\beta, \Theta), \quad (12)$$

where

$$\Theta = (\mathbf{I} - \lambda \mathbf{W})(\Lambda'_y \Theta_{\varepsilon^*}^{-1} \Lambda_y)^{-1}(\mathbf{I} - \lambda \mathbf{W})'. \quad (13)$$

The error distribution of the standard SAR model is $\varepsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$, while the error distribution of the SAR-LVs model is as Equation (12). This error distribution is used to construct the LM test.

2.4. Test of Dependency Spatial

The Lagrange multiplier (LM) test of the SAR-LVs model, as shown in Equation (11), is a test based on estimation under the null hypothesis. The likelihood function l in the SAR-LVs model is obtained by replacing ε and multiplying by the Jacobian in the Gaussian function so that the likelihood function for the SAR-LVs model is obtained: $L(\lambda, \beta, \Theta; \mathbf{l}) = \pi^{-T/2} |\Theta|^{-1/2} |\mathbf{C}| \exp(-\frac{1}{2} \varepsilon' \Theta^{-1} \varepsilon)$, where $\mathbf{C} = (\mathbf{I} - \lambda \mathbf{W})$.

The log-likelihood function for the SAR-LVs model is $\mathcal{L}(\lambda, \beta, \Theta; \mathbf{l}) = -\frac{T}{2} \ln |\pi| - \frac{1}{2} \ln |\Theta| + \ln |\mathbf{C}| - \frac{1}{2} (\varepsilon' \Theta^{-1} \varepsilon)$, where the value of Θ is as in Equation (13), and $\varepsilon = (\mathbf{I} - \lambda \mathbf{W})\mathbf{l} - \mathbf{K}\beta$.

Breusch and Pagan in [17] defined LM test statistics as follows: $LM_\lambda = \hat{\mathbf{D}}'_\lambda \hat{\Psi}_{\lambda\lambda}^{-1} \hat{\mathbf{D}}_\lambda$, where $\hat{\Psi}_{\lambda\lambda}^{-1}$ is an element of an information matrix measuring $k \times k$ whose elements are the second derivative of each parameter estimated as $\hat{\Psi}_\theta = E\left[-\frac{\partial^2 \mathcal{L}(\theta)}{\partial \theta \partial \theta'}\right]$. The test was under the null hypothesis, so $\hat{\mathbf{D}}_\lambda$ is the first derivative of the log-likelihood function of λ where $\lambda = \mathbf{0}$.

The value of $\hat{\Psi}_{\lambda\lambda}^{-1}$ and $\hat{\mathbf{D}}_\lambda$ were decomposed in Appendix A, which obtained $\hat{\Psi}_{\lambda\lambda}^{-1} = -p^{-1}((\mathbf{e}\eta_t - \mathbf{K}\beta)' \mathbf{W} \mathbf{W}' (\mathbf{e}\eta_t - \mathbf{K}\beta))^{-1}$ so $\hat{\mathbf{D}}_\lambda = p(\mathbf{W} \mathbf{K} \beta)' (\mathbf{l} - \mathbf{K} \beta)$ where $(\mathbf{l} - \mathbf{K} \beta)$ is an error of the OLS regression model, $\mathbf{l} - \mathbf{K} \beta = \tilde{\varepsilon}$, so $\hat{\mathbf{D}}_\lambda = p(\mathbf{W} \mathbf{K} \beta)' \tilde{\varepsilon}$. The LM statistic test is

$$LM_\lambda = \frac{-(p(\mathbf{W} \mathbf{K} \beta)' \tilde{\varepsilon})^2}{p(\mathbf{e}\eta_t - \mathbf{K} \beta)' \mathbf{W} \mathbf{W}' (\mathbf{e}\eta_t - \mathbf{K} \beta)}$$

where $p = (\Lambda'_y \Theta_{\varepsilon^*}^{-1} \Lambda_y)$ and $d = (\mathbf{e}\eta_t - \mathbf{K} \beta)' \mathbf{W} \mathbf{W}' (\mathbf{e}\eta_t - \mathbf{K} \beta)$, so the value of test statistic LM becomes

$$LM_\lambda = \frac{-(p(\mathbf{W} \mathbf{K} \beta)' \tilde{\varepsilon})^2}{pd} \quad (14)$$

The LM statistics LM_λ follows the asymptotic distribution of $\chi^2_{(1)}$.

2.5. Estimation of Parameter of SAR-LVs Model

If the parameter of the SAR-LVs model in Equation (11) and its error distribution as Equation (12) are estimated by the OLS, then the estimator is biased and inconsistent, since there is a case where the regression variable $(\mathbf{W} \mathbf{l})$ correlates with the error ε or $cov(\mathbf{W} \mathbf{l}, \varepsilon) \neq \mathbf{0}$. If the model is estimated

by the moment method, the overidentified condition is obtained. The following explains the overidentified condition.

The Equation (11) can be simplified as follows:

$$l = Z\alpha + \varepsilon, \quad (15)$$

where Z is $Z = (K|Wl)$ and has the $(T \times (p+2))$ -sized, α is $\alpha = (\beta'|\lambda)'$ and has the $((p+2) \times 1)$ -sized.

In this work, δ as Equation (15) was estimated by the 2SLS method as performed by [14,15], namely two steps of the ordinary least squares (OLS) method as follows: (i). The 2SLS method requires an H instrument variable, which is a joint of the K matrix and the WK matrix or written as $H = (K|WK)$. The instrument variable H is valid because it does not correlate with ε and correlates with regressor Wl ; (ii). Regress Wl on instrument variable H to obtain $\hat{Wl} = H(H'H)^{-1}H'(Wl)$; (iii). Regress l on $\hat{Z}$ to obtain

$$\hat{\alpha} = (\hat{Z}'\hat{Z})^{-1}\hat{Z}'l, \quad (16)$$

where $\hat{Z} = (K|\hat{Wl})$ and $\hat{\alpha}$ which contains $\hat{\beta}$ and $\hat{\lambda}$.

3. Results and Discussion

In the discussion, this study examined two cases or models developed with the results of positive and negative spatial autoregressive coefficient to provide an interpretation of the spatial autoregressive coefficient. The first case was the education quality model developed by [13], with updated data in 2018 and showing a negative spatial autoregressive coefficient. Meanwhile, the second case related to a poverty model conducted by [22] producing a positive spatial autoregressive coefficient.

The education quality model for senior high schools in Sumenep Regency involved 27 observation units, one endogenous latent variable, and two exogenous latent variables. The endogenous latent variable was the education quality with three indicators. Indicators of education quality were the ratio of the gross enrolled number of senior high school students to the number of children aged between 15 and 18 years in each district (Y11)—the ratio of the number of accredited senior high schools with at least B levels to the total number of senior high schools in each district (Y12) and the average of national exam scores of senior high school students in each district (Y13).

Exogenous latent variables were school infrastructure and socioeconomic conditions. Indicators of school infrastructure were the proportion of the number of schools with a minimum classroom space according to the regulations of the national education ministry (X11), the proportion of the numbers of schools with laboratories according to the regulations of the national education ministry (X12), and the proportion of the number of schools with libraries according to the regulations of the national education ministry (X13).

Indicators of socioeconomic conditions were the ratio of the number of households running a home industry or with a shop at home to the total number of households in each district (X21) and the ratio of the number of households using clean water to the total number of households in each district (X22). The model is shown in Figure 1.

Latent variables were estimated as in the Equations (6) and (9), then were modeled as in the Equation (11). The estimation of the model parameters as the Equation (16) used the Matlab software, and the results are shown in Table 1. The LM test based on the Equation (14) used Matlab software, and the obtained value LM_{λ} was -2.3272 . The value of LM_{λ} was compared to χ^2 with degrees of freedom of one and $\alpha = 5\%$. Then, the result was significant towards the SAR-LVs model. Additionally, the spatial autoregressive coefficient was negative.

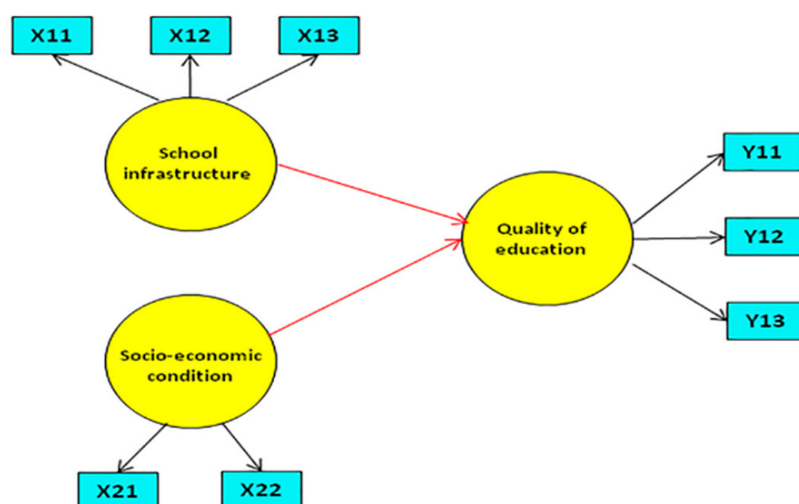


Figure 1. The education quality model.

Table 1. The estimation result of parameter and spatial autoregressive coefficient for the education model.

Variable	Coefficient
School Infrastructure (b1)	2.3121
Socioeconomic condition (b2)	0.1286
Constant (b0)	9.6604
Spatial Autoregressive Coefficient (λ)	−0.002

In general, the SAR-LVs model for the education quality of the senior high school is: $l_i = 9.6604 - 0.002 \sum_{j=1, i \neq j}^{27} W_{ij} l_j + 2.3121 k_1 + 0.1286 k_2$, where l_i is the education quality in the i -th district, k_1 is the infrastructure, and k_2 is socioeconomic condition.

The spatial dependence test result was significant. This means there was a correlation between the education quality of the senior high schools in one district and the one in other contiguous districts. The negative spatial autoregressive coefficient interpreted the opposite of the common spillover effect, i.e., a district was supported by or gained a spillover effect of the neighboring districts' education quality.

The spillover effect of the neighboring districts' education quality was generally due to the migration of students to find high-quality schools in the neighboring districts. As a result, the districts that had high-quality schools, gained the spillover effect of quality education through high-achieving students from the neighboring districts. In general, high-quality schools in Sumenep Regency are public schools. Figure 2 draws the distribution of districts with and without public schools. To illustrate the student migration, an example is provided in the Gapura district (see Figure 2). Table 2 shows the number of junior high school graduates and new senior high school students for seven districts in the same year, 2018, based on [23]. This table provides an overview of the migration data of students entering senior high school among districts. The Gapura District was used as an example to illustrate student migration (see Figure 2 and Table 2). The Gapura District has one public school and six neighboring neighbors (queen contiguity). Based on Table 2, the Gapura District has 485 junior high school graduates and 622 new high school students. This means that there was a migration of students to the Gapura District.

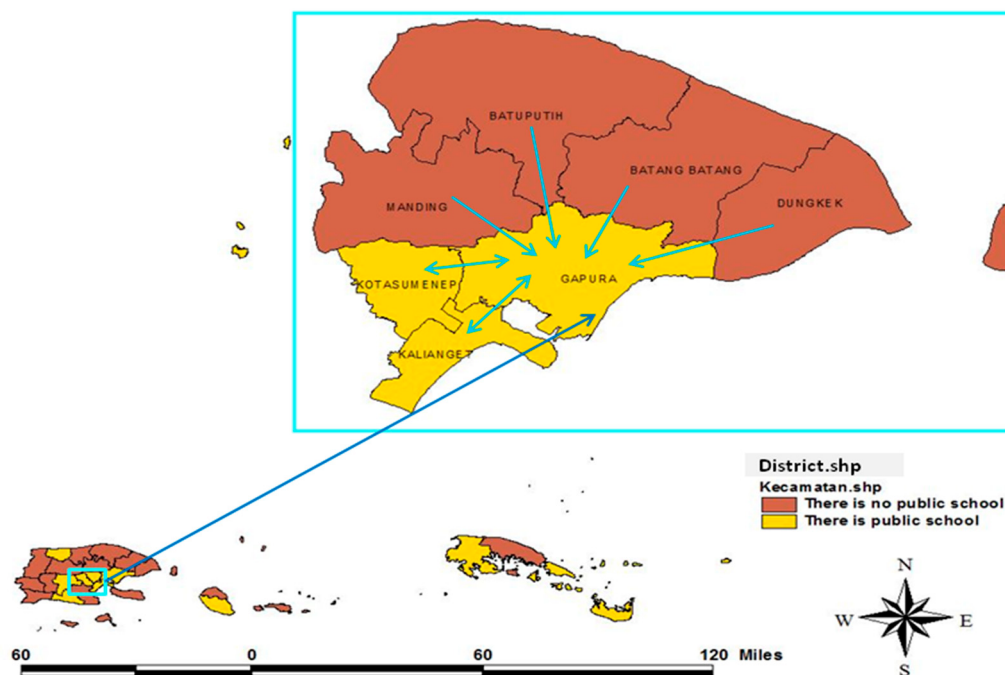


Figure 2. Distribution of public schools in Sumenep Regency.

Table 2. Number of students and public schools for the neighboring of Gapura district.

District	Number of Public Senior High Schools	Number of Junior High School Graduate Students	Number of New Senior High School Students
Kalianget	1	488	744
Kota Sumenep	3	1582	2041
Manding	0	201	93
Batuputih	0	226	143
Gapura	1	485	622
Batang Batang	0	552	360
Dungkek	0	409	156

The Gapura District received the spillover effect of quality education from neighboring districts, especially those with no public schools. The quality spillover was due to schools in the Gapura district having the opportunity to select the best students from the district itself or the neighboring districts.

The second case was the poverty model in East Java province, with an observation unit of 38 regencies. This model had one endogenous latent variable, namely poverty, and three exogenous latent variables, namely Economy, Human Resource, and Health. Poverty indicators were the percentage of the poor population (Y1), the index of poverty depth (Y2), and the index of poverty severity (Y3). Indicators of economics were the percentage of poor people around 15 years old or more who were unemployed (X1), the percentage of poor people aged 15 years old or more who were working in agriculture (X2), and the percentage of households gaining Raskin (X3). Raskin is an Indonesian subsidy program to provide rice for people who live under the poverty line. Indicators of Human Resources were the percentage of poor people aged 15 years old and over who did not complete elementary education (X4), the literacy rate of the poor aged from 15 to 55 years (X5), and the participation rates in schools for the poor aged from 13 to 15 years (X6). Health indicators were the percentage of women using KB (Family planning program) devices in poor households (X7), the percentage of children under five in poorly immunized households (X8), the percentage of poor households using drinking water (X9), and the percentage of poor households using private/together latrines (X10). The model is shown in Figure 3.

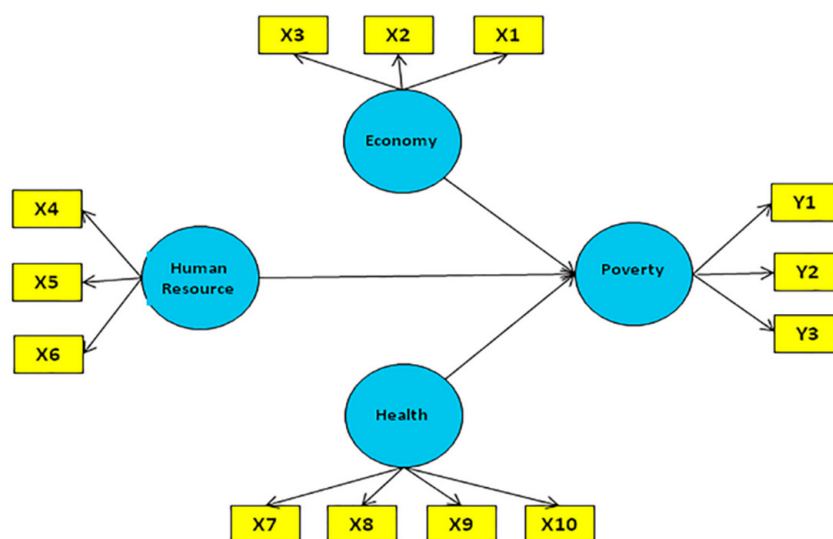


Figure 3. Model of poverty.

Latent variables were estimated as in the Equations (6) and (9), then were modeled as in the Equation (11). The estimation of the model parameters as in the Equation (16) used the Matlab software, and the results are shown in Table 3. The LM test based on Equation (14) used the Matlab software, and the obtained value LM_{λ} was -4965 . The value of LM_{λ} was compared to χ^2 with degrees of freedom of one and $\alpha = 5\%$. Then, the result was significant towards the SAR-LVs model. Meanwhile, the spatial autoregressive coefficient was positive.

Table 3. The estimation result of parameter and spatial autoregressive coefficient for the poverty model.

Variable	Coefficient
Economy (b1)	0.0742
Human Resource (b2)	-0.0722
Health (b3)	0.0155
Constant (b0)	7.0881
Spatial Autoregressive Coefficient (λ)	0.2345

In general, the SAR-LVs model for the poverty is $l_i = 7.0881 + 0.2345 \sum_{j=1, i \neq j}^{38} W_{ij} l_j + 0.0742 k_1 - 0.0722 k_2 + 0.0155 k_3$, where l_i is poverty in the i -th regency, k_1 is Economy, k_2 is Human Resource, and k_3 is Health.

The spatial dependence test result was significant. This means there was a correlation between poverty in one regency and the one in other contiguous regencies. The positive spatial autoregressive coefficient interpreted the common spillover effect, i.e., a regency gives a poverty spillover effect to the neighboring regencies. Figure 4 describes the poor people distribution in the East Java province measured in percent. The percentage of poor people was based on the poverty data in [24] and was clustered into four quartiles. The regency with the high percentage of poor people was categorized as the first quartile (red zone), with 20.71–13.01%. On the other hand, the fourth quartile (green zone) includes the regency with 7.13–3.80% of the poor group. The first quartile was the regency group with the highest percentage of poor people, and so on until the fourth quartile was the regency group with the lowest percentage of poor people. According to Figure 4, the location of regencies in the first quartile was always close to the regencies in the first and the second quartile. This visualization reinforces the results of the multiplier Lagrange test that there is a spatial effect where one regency gives the influence of poverty on its neighboring regencies.

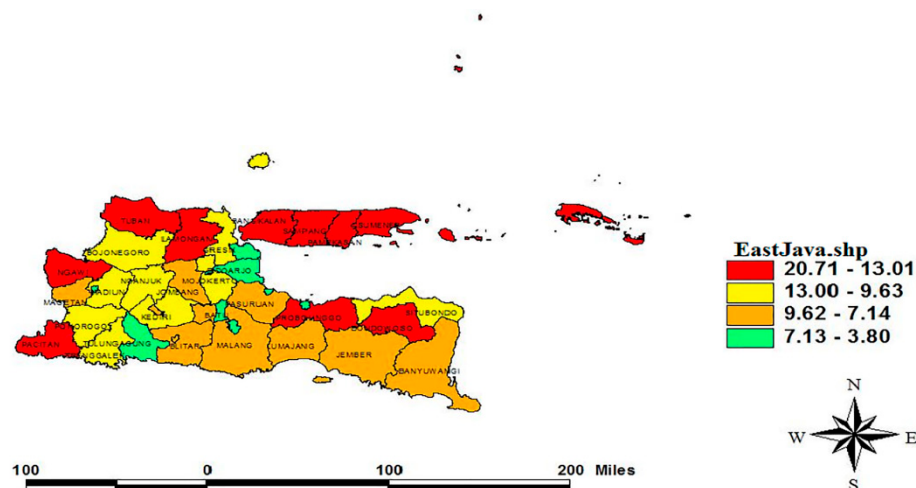


Figure 4. A map of the distribution of poor people in the East Java province (in percent).

The finding of this method provides value for policymakers relating to existing problems. The first case was the negative spatial autoregressive coefficient for the education quality model. It interpreted that a district was supported by or gained a spillover effect of the neighboring districts' education quality through students' migration. This case needs the policy to strive for quality standardization for all schools. The second case was the positive spatial autoregressive coefficient for the poverty model. It interpreted that a regency gives a poverty spillover effect to the neighboring regencies. Policymakers need this information to assist at the locus of poverty appropriately.

4. Conclusions

The SAR-LVs model is a standard SAR model in which the independent and dependent variables are latent variables. The standard SAR model is $\mathbf{y}^* = \lambda \mathbf{W} \mathbf{y}^* + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}$. The variable of $\mathbf{y}^*$ is changed by the endogenous factor score ($\hat{\eta}$), and $\mathbf{X}$ is changed by the exogenous factor score ($\hat{\xi}$). The estimation of the latent variables uses the WLS method and assumes that the value of Λ_x and Θ_δ are constant. Therefore, the SAR-LVs model can be modeled as $\mathbf{l} = \mathbf{K} \boldsymbol{\beta} + \lambda \mathbf{W} \mathbf{l} + \boldsymbol{\varepsilon}$, where $\hat{\eta} = \mathbf{l}$ and $\hat{\xi} = \mathbf{K}$. The distribution of $\mathbf{K}$ and $\mathbf{l}$ are $\mathbf{K} \sim N_{T,p}(\mathbf{e} \xi_t, \mathbf{I}_T \otimes (\Lambda'_x \Theta_\delta^{-1} \Lambda_x)^{-1})$ and $\mathbf{l} \sim N_T(\mathbf{e} \eta_t, (\Lambda'_y \Theta_{\varepsilon^*}^{-1} \Lambda_y)^{-1} \mathbf{I}_T)$. The variables of $\mathbf{l}$ and $\mathbf{K}$ are random. Thus, the function of $\mathbf{l}$ is $f(\mathbf{l} | \mathbf{K})$, so that the variable $\mathbf{K}$ is no longer random but fixed. The error distribution is obtained through the properties of the expected value and the variance of the error in the SAR-LVs model, namely $\boldsymbol{\varepsilon} \sim N_T((\mathbf{I} - \lambda \mathbf{W}) \mathbf{e} \eta_t - \mathbf{K} \boldsymbol{\beta}, \boldsymbol{\Theta})$, where $\boldsymbol{\Theta} = (\mathbf{I} - \lambda \mathbf{W}) (\Lambda'_y \Theta_{\varepsilon^*}^{-1} \Lambda_y)^{-1} (\mathbf{I} - \lambda \mathbf{W})'$. Based on its error distribution model, so under the null hypothesis, the LM statistic is $LM_\lambda = \frac{-(\mathbf{p}(\mathbf{W} \mathbf{K} \boldsymbol{\beta})' \hat{\boldsymbol{\varepsilon}})^2}{\text{pd}}$ and follows the asymptotic distribution of $\chi^2_{(1)}$.

Some significant limitations of this study need to be considered. Firstly, the number of endogenous latent variables is one. Future studies can be developed for the higher number of endogenous latent variables. Secondly, the LM test developed was merely for SAR-LVs. Future studies can be developed for a spatial error model with latent variables (SEM-LVs).

Author Contributions: Conceptualization, A.A., B.W.O., P.P., and S.S.; data curation, B.W.O.; formal analysis, A.A. and P.P.; funding acquisition, A.A.; methodology, A.A., B.W.O., P.P., and S.S.; project administration, A.A.; resources, S.S.; software, A.A.; supervision, B.W.O.; validation, P.P.; visualization, S.S.; writing—original draft, A.A.; writing—review and editing, B.W.O. All authors have read and agreed to the published version of the manuscript.

Funding: The authors would like to thank the Ministry of Research, Technology, and Higher Education of Indonesia (Kemenristekdikti) for providing a research grant to carry out preliminary research with the contract number 046/SP2H/K7/KM/2017. The first author would also like to thank Universitas Wiraraja for its general financial support.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Appendix A.1. Matrix Variate Normal Distribution

Definition A.1 based on [21] page 56.

The matrix variate normal distribution arises when sampling from the multivariate normal population. Let $x_1, \dots, x_N$ be a random sample of size N from $N_p(\mu, \Sigma)$. Define the observation random matrix as

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1N} \\ x_{21} & x_{22} & \cdots & x_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pN} \end{pmatrix} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \begin{pmatrix} \mathbf{x}'_1 \\ \mathbf{x}'_2 \\ \vdots \\ \mathbf{x}'_p \end{pmatrix} \text{ then } \mathbf{X}' \sim N_{N,p}(\mathbf{e}\mu', \mathbf{I}_N \otimes \Sigma)$$

where $\mathbf{e}_{N \times 1} = (1, \dots, 1)'$

Theorem A.1 based on [21] page 56.

If $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \Sigma \otimes \Psi)$, then $\mathbf{X}' \sim N_{p,n}(\mathbf{M}', \Sigma \otimes \Psi)$

Teorema A.2 based on [21] page 56.

If $\mathbf{X} \sim N_{p,n}(\mathbf{M}, \Sigma \otimes \Psi)$, then the characteristic function of $\mathbf{X}$ is $\phi_{\mathbf{X}}(\mathbf{Z}) = \text{etr}(\iota \mathbf{Z}' \mathbf{M} - 1/2 \mathbf{Z}' \Sigma \mathbf{Z} \Psi)$

where $\iota = \sqrt{-1}$

Appendix A.2. Properties of Matrix and Derivative of Matrix/Vector

Theorem B.1 based on [21] page 56.

$\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$ with $\mathbf{A}_{p \times q}$ and $\mathbf{B}_{q \times p}$

Properties of derivative matrix/vector

B.2. $\partial \ln|\mathbf{X}| = \text{Tr}(\mathbf{X}^{-1} \partial \mathbf{X})$

B.3. $\partial(\mathbf{X}^{-1}) = -\mathbf{X}^{-1}(\partial \mathbf{X})\mathbf{X}^{-1}$

B.4. $\frac{\partial \mathbf{A}_{(x)}^T \mathbf{A}_{(x)}}{\partial x} = 2 \left(\frac{\partial \mathbf{A}_{(x)}}{\partial x} \right)^T \mathbf{A}_{(x)}$

B.5. $\frac{\partial}{\partial \mathbf{s}} = (\mathbf{x} - \mathbf{As})^T (\mathbf{x} - \mathbf{As}) = -2\mathbf{A}^T (\mathbf{x} - \mathbf{As})$

B.6. $\frac{\partial \mathbf{A}_{(x)}^T \mathbf{B}_{(x)}}{\partial x} = \left(\frac{\partial \mathbf{A}_{(x)}}{\partial x} \right)^T \mathbf{B}_{(x)} + \left(\frac{\partial \mathbf{B}_{(x)}}{\partial x} \right)^T \mathbf{A}_{(x)}$

Appendix A.3. Derivative of the Element of the Information Matrix for the SAR-LVs Model

- The first partial derivative of the log-likelihood function $L(\lambda, \beta, \Theta; \mathbf{I})$ to λ based on the error distribution in Equation (12) where $\Theta = \mathbf{Ap}^{-1}\mathbf{A}' = p^{-1}\mathbf{AA}'$; $p = (\Lambda_y' \Theta_{\varepsilon^*}^{-1} \Lambda_y)$; $\mathbf{A} = (\mathbf{I} - \lambda \mathbf{W})$ and $\varepsilon = \mathbf{Al} - \mathbf{K}\beta$.

- The first partial derivative of Θ to λ

$$\frac{\partial \lambda}{\partial \Theta} = p^{-1} \frac{\partial ((\mathbf{I} - \lambda \mathbf{W})(\mathbf{I} - \lambda \mathbf{W})')}{\partial \lambda} = -p^{-1}(\mathbf{W}(\mathbf{I} - \lambda \mathbf{W})' + (\mathbf{I} - \lambda \mathbf{W})\mathbf{W}')$$

$$\frac{\partial \lambda}{\partial \Theta} = -p^{-1}(\mathbf{WA}' + \mathbf{AW}')$$

- The first partial derivative of $\ln|\Theta|$ to λ Based on B.2

$$\frac{\partial \ln|\Theta|}{\partial \lambda} = \text{Tr} \left((p^{-1}\mathbf{AA}')^{-1} \frac{\partial \Theta}{\partial \lambda} \right) = \text{Tr} \left(p(p^{-1}\mathbf{AA}')(-p^{-1}(\mathbf{WA}' + \mathbf{AW}')) \right) = -2\text{Tr}(\mathbf{A}^{-1}\mathbf{W})$$

- c. The first partial derivative of $\ln|A|$ to λ Based on B.2:

$$\frac{\partial \ln|A|}{\partial \lambda} = \text{Tr} \left((I - \lambda W)^{-1} \frac{\partial (I - \lambda W)}{\partial \lambda} \right) = \text{Tr}(A^{-1}W)$$

- d. The first partial derivative of A^{-1} to λ Based on B.3:

$$\frac{\partial \ln A^{-1}}{\partial \lambda} = -(I - \lambda W)^{-1}(-W)(I - \lambda W)^{-1} = A^{-1}WA^{-1}$$

- e. The first partial derivative of $\varepsilon' \Theta^{-1} \varepsilon$ to λ

$$\varepsilon' \Theta^{-1} \varepsilon = [(A I - K \beta)' (p^{-1} A A')^{-1} (A I - K \beta)] \text{ or } \varepsilon' \Theta^{-1} \varepsilon = p(1 - A^{-1} K \beta)'(1 - A^{-1} K \beta)$$

Based on B.4:

$$\frac{\partial (\varepsilon' \Theta^{-1} \varepsilon)}{\partial \lambda} = -2p(A^{-1}WA^{-1}K\beta)'(I - A^{-1}K\beta)$$

The first partial derivative of the log-likelihood function $L(\lambda, \beta, \Theta; I)$ to λ based on the error distribution in Equation (12), point b, c, and d

$$\frac{\partial L(\lambda, \beta, \Theta; I)}{\partial \lambda} = -\frac{1}{2} (-2\text{Tr}(A^{-1}W)) - \text{Tr}(A^{-1}W) + p(A^{-1}WA^{-1}K\beta)'(I - A^{-1}K\beta)$$

$$\frac{\partial L(\lambda, \beta, \Theta; I)}{\partial \lambda} = p(A^{-1}WA^{-1}K\beta)'(I - A^{-1}K\beta)$$

2. The first partial derivative of the log-likelihood function $L(\lambda, \beta, \Theta; I)$ to β based on the error distribution in Equation (12) Based on B.5:

$$\frac{\partial L(\lambda, \beta, \Theta; I)}{\partial \beta} = p(A^{-1}K)'(I - A^{-1}K\beta)$$

3. The second partial derivative $\frac{\partial^2 L(\lambda, \beta, \Theta; I)}{\partial \lambda^2}$ Based on B.3:

$$\begin{aligned} \frac{\partial^2 L(\lambda, \beta, \Theta; I)}{\partial \lambda^2} &= p[2(K\beta)'((A^{-1}WA^{-1})A^{-1}(A^{-1}WA^{-1}))'(I - A^{-1}K\beta) - \\ &\quad (A^{-1}WA^{-1}K\beta)'(A^{-1}WA^{-1})(K\beta)] \end{aligned}$$

4. The second partial derivative $\frac{\partial^2 L(\lambda, \beta, \Theta; I)}{\partial \beta \partial \beta'}$

$$\frac{\partial^2 L(\lambda, \beta, \Theta; I)}{\partial \beta \partial \beta'} = -p(A^{-1}K)'(A^{-1}K)$$

5. The second partial derivative $\frac{\partial^2 L(\lambda, \beta, \Theta; I)}{\partial \beta \partial \lambda}$

$$\frac{\partial^2 L(\lambda, \beta, \Theta; I)}{\partial \beta \partial \lambda} = p[(I - A^{-1}K\beta)'(A^{-1}WA^{-1}K) - (A^{-1}WA^{-1}K\beta)'(A^{-1}K)]$$

6. The element of the information matrix

- a. The Element (1,1), namely $\tilde{\Psi}_{\lambda\lambda}$

$$\tilde{\Psi}_{\lambda\lambda} = E\left(-\frac{\partial^2 L(\lambda, \beta, \Theta; I)}{\partial \lambda^2}\right)$$

$$\tilde{\Psi}_{\lambda\lambda} = p\left[\left(\mathbf{A}^{-1}\mathbf{W}\mathbf{A}^{-1}\mathbf{K}\beta\right)'(\mathbf{A}^{-1}\mathbf{W}\mathbf{A}^{-1})(\mathbf{K}\beta) - 2(\mathbf{K}\beta)'((\mathbf{A}^{-1}\mathbf{W}\mathbf{A}^{-1})\mathbf{A}^{-1}(\mathbf{A}^{-1}\mathbf{W}\mathbf{A}^{-1}))'(\mathbf{e}\eta_t - \mathbf{A}^{-1}\mathbf{K}\beta)\right]$$

if $\lambda = 0$ then

$$\tilde{\Psi}_{\lambda\lambda} = p\left[(\mathbf{W}\mathbf{K}\beta)'(\mathbf{W}\mathbf{K}\beta) - 2(\mathbf{W}\mathbf{K}\beta)'\mathbf{W}(\mathbf{e}\eta_t - \mathbf{A}^{-1}\mathbf{K}\beta)\right]$$

- b. The Element (2,2), namely $\tilde{\Psi}_{\beta\beta}$

$$\tilde{\Psi}_{\beta\beta} = E\left(-\frac{\partial^2 L(\lambda, \beta, \Theta; I)}{\partial \beta \partial \beta'}\right) = E\left(-(-p(\mathbf{A}^{-1}\mathbf{K})'(\mathbf{A}^{-1}\mathbf{K}))\right) = p(\mathbf{A}^{-1}\mathbf{K})'(\mathbf{A}^{-1}\mathbf{K})$$

if $\lambda = 0$ then $\tilde{\Psi}_{\beta\beta} = p(\mathbf{K}'\mathbf{K})$

- c. The Element (1,2), namely $\tilde{\Psi}_{\lambda\beta}$

$$\tilde{\Psi}_{\lambda\beta} = E\left(-\frac{\partial^2 L(\lambda, \beta, \Theta; I)}{\partial \lambda \partial \beta'}\right)$$

$$\tilde{\Psi}_{\lambda\beta} = p\left[(\mathbf{A}^{-1}\mathbf{W}\mathbf{A}^{-1}\mathbf{K}\beta)'(\mathbf{A}^{-1}\mathbf{K}) - (\mathbf{e}\eta_t - \mathbf{A}^{-1}\mathbf{K}\beta)'(\mathbf{A}^{-1}\mathbf{W}\mathbf{A}^{-1}\mathbf{K})\right]$$

if $\lambda = 0$ then $\tilde{\Psi}_{\lambda\beta} = p\left[(\mathbf{W}\mathbf{K}\beta)'\mathbf{K} - (\mathbf{e}\eta_t - \mathbf{K}\beta)'(\mathbf{W}\mathbf{K})\right]$

- d. The Element (2,1), Namely $\tilde{\Psi}_{\beta\lambda}$

$$\tilde{\Psi}_{\beta\lambda} = E\left(-\frac{\partial^2 L(\lambda, \beta, \Theta; I)}{\partial \beta \partial \lambda}\right)$$

$$\tilde{\Psi}_{\beta\lambda} = p\left[(\mathbf{A}^{-1}\mathbf{K})'(\mathbf{A}^{-1}\mathbf{W}\mathbf{A}^{-1}\mathbf{K}\beta) - (\mathbf{A}^{-1}\mathbf{W}\mathbf{A}^{-1}\mathbf{K})'(\mathbf{e}\eta_t - \mathbf{A}^{-1}\mathbf{K}\beta)\right]$$

if $\lambda = 0$ then

$$\tilde{\Psi}_{\beta\lambda} = p\left[\mathbf{K}'(\mathbf{W}\mathbf{K}\beta) - (\mathbf{W}\mathbf{K})'(\mathbf{e}\eta_t - \mathbf{K}\beta)\right]$$

7. The information matrix if $\lambda = 0$ then the information matrix is $\tilde{\Psi}_{\theta} = \begin{bmatrix} \tilde{\Psi}_{\lambda\lambda} & \tilde{\Psi}_{\lambda\beta} \\ \tilde{\Psi}_{\beta\lambda} & \tilde{\Psi}_{\beta\beta} \end{bmatrix}$
8. Invers of the information matrix when $\lambda = 0$ If the partition matrix is $\mathbf{C} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \\ \mathbf{C}_3 & \mathbf{C}_4 \end{bmatrix}$ then the invers matrix is $\mathbf{C}^{-1} = \begin{bmatrix} \mathbf{C}_{E1} & \mathbf{C}_{E2} \\ \mathbf{C}_{E3} & \mathbf{C}_{E4} \end{bmatrix}$ where $\mathbf{C}_{E1} = (\mathbf{C}_1 - \mathbf{C}_2\mathbf{C}_4^{-1}\mathbf{C}_3)^{-1}$; $\mathbf{C}_{E2} = (-\mathbf{C}_{E1}\mathbf{C}_2\mathbf{C}_4^{-1})$; $\mathbf{C}_{E3} = -\mathbf{C}_4^{-1}\mathbf{C}_3\mathbf{C}_{E1}$; and $\mathbf{C}_{E4} = (\mathbf{C}_4^{-1} - \mathbf{C}_4^{-1}\mathbf{C}_3\mathbf{C}_{E2})$ The element (1,1) of the information matrix invers is $\tilde{\Psi}_{\lambda\lambda}^{-1} = (\tilde{\Psi}_{\lambda\lambda} - \tilde{\Psi}_{\lambda\beta}(\tilde{\Psi}_{\beta\beta})^{-1}\tilde{\Psi}_{\beta\lambda})^{-1}$

$$\tilde{\Psi}_{\lambda\lambda}^{-1} = (p(\mathbf{W}\mathbf{K}\beta)'(\mathbf{W}\mathbf{K}\beta) - 2(\mathbf{W}\mathbf{K}\beta)'\mathbf{W}'(\mathbf{e}\eta_t - \mathbf{K}\beta) - p((\mathbf{W}\mathbf{K}\beta)'(\mathbf{W}\mathbf{K}\beta) - (\mathbf{e}\eta_t - \mathbf{K}\beta)'$$

$$\mathbf{W}(\mathbf{W}\mathbf{K}\beta) - (\mathbf{W}\mathbf{K}\beta)\mathbf{W}'(\mathbf{e}\eta_t - \mathbf{K}\beta) + (\mathbf{e}\eta_t - \mathbf{K}\beta)'\mathbf{W}\mathbf{W}'(\mathbf{e}\eta_t - \mathbf{K}\beta)))^{-1}$$

$$\tilde{\Psi}_{\lambda\lambda}^{-1} = p^{-1}(-(\mathbf{e}\eta_t - \mathbf{K}\beta)'\mathbf{W}\mathbf{W}'(\mathbf{e}\eta_t - \mathbf{K}\beta))^{-1}$$

$$\Psi_{\lambda\lambda}^{-1} = -p^{-1} \left((\mathbf{e}\boldsymbol{\eta}_t - \mathbf{K}\boldsymbol{\beta})' \mathbf{W}\mathbf{W}' (\mathbf{e}\boldsymbol{\eta}_t - \mathbf{K}\boldsymbol{\beta}) \right)^{-1}$$

References

1. Bollen, K.A. *Structural Equations with Latent Variables*; John Wiley & Sons: New York, NY, USA, 1989.
2. Civelek, M.E. *Essentials of Structural Equation Modeling*; The University of Nebraska: Lincoln, NE, USA, 2018; ISBN 978-1-60962-129-2.
3. Wang, F.; Wall, M.M. Generalized common spatial factor model. *Biostatistics* **2003**, *4*, 569–582. [[CrossRef](#)] [[PubMed](#)]
4. Christensen, W.F.; Amemiya, Y. Latent variable analysis of multivariate spatial data. *J. Am. Stat. Assoc.* **2002**, *97*, 302–317. [[CrossRef](#)]
5. Hogan, J.W.; Tchernis, R. Bayesian Factor Analysis for Spatially Correlated Data, With Application to Summarizing Area-Level Material Deprivation from Census Data. *J. Am. Stat. Assoc.* **2004**, *99*, 314–324. [[CrossRef](#)]
6. Liu, X.; Wall, M.M.; Hodges, J.S. Generalized spatial structural equation models. *Biostatistics* **2005**, *6*, 539–557. [[CrossRef](#)] [[PubMed](#)]
7. Congdon, P. A spatial structural equation model for health outcomes. *J. Stat. Plan. Inference* **2008**, *138*, 2090–2105. [[CrossRef](#)]
8. Oud, J.H.L.; Folmer, H. A Structural Equation Approach to Models with Spatial Dependence. *Geogr. Anal.* **2008**, *40*, 152–166. [[CrossRef](#)]
9. Anselin, L. *Spatial Econometrics: Methods and Models*; Kluwer Academic Publisher: Dordrecht, The Netherlands, 1988; Volume 4, ISBN 978-90-481-8311-1.
10. Joreskog, K.G.; Sorbom, D. *Lisrel 8: User's Reference Guide*; Scientific Software: Chicago, IL, USA, 1997.
11. Anekawati, A.; Widjanarko Otok, B. Modelling of the education quality of a high schools in Sumenep Regency using spatial structural equation modelling. *J. Phys. Conf. Ser.* **2017**, *890*, 012094. [[CrossRef](#)]
12. Trujillo, G.S. *Pathmox Approach: Segmentation Trees in Partial Least Squares Path Modeling*; Universitat Politècnica de Catalunya: Barcelona, Spain, 2009.
13. Anekawati, A.; Otok, B.W.; Sutikno, P. Generalized method of moments approach to spatial structural equation modeling. *FJMS* **2018**, *103*, 1057–1076. [[CrossRef](#)]
14. Kelejian, H.H.; Prucha, I.R. A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. *J. Real Estate Financ. Econ.* **1998**, *17*, 99–121. [[CrossRef](#)]
15. Kelejian, H.H.; Prucha, I.R. A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. *Int. Econ. Rev.* **1999**, *40*, 509–533. [[CrossRef](#)]
16. Jeong, S.; Yoon, D. Examining Vulnerability Factors to Natural Disasters with a Spatial Autoregressive Model: The Case of South Korea. *Sustainability* **2018**, *10*, 1651. [[CrossRef](#)]
17. Breusch, T.S.; Pagan, A.R. The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics. *Rev. Econ. Stud.* **1980**, *47*, 239–253. [[CrossRef](#)]
18. Anselin, L. Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity. *Geogr. Anal.* **1988**, *20*, 1–17. [[CrossRef](#)]
19. Yang, Z. LM tests of spatial dependence based on bootstrap critical values. *J. Econom.* **2015**, *185*, 33–59. [[CrossRef](#)]
20. LeSage, J.P. *The Theory and Practice of Spatial Econometrics*, 1st ed.; The University of Toledo: Toledo, OH, USA, 1999.
21. Gupta, A.K.; Nagar, D.K. *Matrix Variate Distributions*; Monographs and Surveys in Pure and Applied Mathematics; Chapman and Hall/CRC: New York, NY, USA, 2000; ISBN 1-58488-046-5.
22. Otok, B.W.; Standsyah, R.E.; Suharsono, A.; Purhadi. Development of Model Poverty in Java Using Meta-Analysis Structural Equation Modeling (MASEM). In Proceedings of the 2nd International Conference on Science, Mathematics, Environment, and Education, Surakarta, Indonesia, 26–28 July 2019; AIP Conference Proceedings Volume 2194, p. 020078. [[CrossRef](#)]

23. BPS. *Sumenep in Figures 2018*; BPS-Statistics of Sumenep Regency: Sumenep, Indonesia, 2018; ISBN 0215.2193.
24. BPS. *Data dan Informasi Kemiskinan Kabupaten/Kota tahun 2019*; BPS-Statistics Indonesia: Jakarta, Indonesia, 2019.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).