Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 11, Number 4 (2015), pp. 2279-2284

http://www.ripublication.com

Spatial Regression Analisys of National Standar Education with National Examination at National Islamic Senior High School Java 2012-2013

Aman Abadi[1], Muhammad Nur Aidi[2], Anik Djuraidah[3]

Departement of Statistics
Bogor Agricultural University
Jalan Pajajaran, Kampus IPB Baranangsiang, Bogor 1615, Indonesia
1.abadi.aman0703@gmail.com, 2.nuraidi@yahoo.com, 3.anikdjuraidah@gmail.com

Abstract

Education is the most fundamental requirement for human. Regression model is used to know the factors that influence National Standards Education (NSE) against National Examination (NE). The level of education in a region affects education in other areas. Therefore, this study used a spatial regression that includes Spatial Autoregressive (SAR), Spatial Error Model (SEM) and General Spatial Model (GSM). Weighting matrix are weighted queen contiguity and weighted based on distance. The model selection criteria is the lowest of Akaike Information Criterion (AIC) value. The results shows that the SAR with a weighted queen contiguity is better than SEM with a weighted queen contiguity. The factors that influence the mean of the National Examination districts/citiesvalue are the mean of the processvalue, graduation, teachers and education personnel, infrastructures and management. The other are the meanof the SchoolsExamination districts/citiesvalue, the difference National Examination number of participants, the difference distance to the capital of the province and the provincial differences.

Keywords: Education, National Examination, SAR, SEM, GSM.

1. Introduction

Regression linear model is used to know the factors that influence National Standar Education (NSE) for National Examination (NE). The level of education in a region affects education in other areas, it causes the autocorrelation between locations and inefficient estimators. These problem can be resolved by spatial regression model that includes Spatial Autoregressive (SAR), Spatial Error Model (SEM) and General

Spatial Model (GSM). Weighting matrix are weighted queen contiguity and weighted based on the distance.

Some relevant researches are spatial regression model to determine the factors that influence the poverty in East Java [2]. The relationship between the National Examination with accreditation rated at districts/cities in West Java [7]. The aim of this researches are build a model between the NSE with the NE at National Islamic Senior High School (NISHS) on Java 2012-2013. Then determine the factors that influence the NSE for the NE in Java.

2. Methodolgy

3.1. Data

3.2. The data used from the Ministry of National Education in 2013. This data includes the value of the National Examination (NE), School Examination (SE), number of National Examination participants and National Standar Education (NSE) 112 districts/cities in National Islamic Senior High School (NISHS) Java 2012-2013 [4]. Dependent variable is mean the NE districs/cities value. Independent variable is the mean of NSE districs/cities value that include are the mean of thecontentvalue(X_1), process (X_2), graduation (X_3), teachers and education personnel (X_4), infrastructures (X_5), management (X_6), financing (X_7) and assessment (X_8)[3].The other independent variables is mean the SchoolsExamination districts/cities value, dummy provincial differences, dummy difference distance to the capital of the province and dummy difference National Examination number of participants.

3.3.

3.4. Methodologhy

The step of this research are:

1. Exploration of data, to determine ratio of the dependent and independent variable value that include minimum value, maximum and mean.
2. Index moran test used to detected correlation spatial about dependent variable with the waighting matrix. Weighting matrix are weighted queen contiguity and weighted based on the distance
3. The estimation parameter that includes SAR, SEM and GSM with the maximum likelihood estimation method.
4. Lagrange Multiple test to identify the spatial dependence [5].
5. Testing spatial regression model assumption of normality and homogenity variance with Breusch Pagan test [1]. If not normal and variety is not homogeneous then transformed and then back to step 3,4.
6. Selection of best model because difference value **W**usedAkaike Information Criterion(AIC) the smallest.

$$\text{AIC} = -2\log(\text{maksimum likelihood}) + 2\text{p}[6]$$

3. Materials and Methods

3.5. *Spatial Autoregressive Model*(SAR)

Spatial Autoregressive (SAR) is modeling the spatial dependence of the response variable. SAR model is expressed in equation form as follows [1]:

$$y = \rho Wy + X\beta + \varepsilon \quad (1)$$

$$\boldsymbol{\varepsilon} \sim N(0, \sigma^2 I)$$

withρ is an coefficient of spatial lag, **W**is an $(n \times n)$spatial weight matrix, **y**is an $(n \times 1)$ vectors of dependent variable,**X**is an $(n \times k)$ matrix of independent variable.

3.6. *Spatial Error Model*(SEM)

Spatial Error Model (SEM) is a linear regression model on the error variable spatial correlation. SEM model is expressed in equation form [1]:

$$y = X\beta + u \quad (2)$$

$$u = \lambda Wu + \varepsilon$$

$$\varepsilon \sim N(0, \sigma^2 I)$$

with ε_i is an error in the i location are assumed to normal spreads with zero mean and constant variance σ^2.

3.7. *General Spatial Model* (GSM)

General spatial model (GSM) is a linear regression model where the independent variable x to-i correlated with the dependent variable y to-j likewise error dependent to-i and to-j. GSM model is expressed in the following form [1]:

$$y = \rho Wy + X\beta + u \quad (3)$$

$$u = \lambda Wu + \varepsilon$$

$$\varepsilon \sim N(0, \sigma^2 I)$$

4. Results and Discussion

4.1. Exploration of National Examination on Java

Based on Table 1, shows that the achievement value of NE districts/citiesin Java is the range of 5.07 to 8.62 and the mean value of 6.85.DIY is the lowest value of 5.89 while the higest mean value of 7.60 in East Java. The mean its value of five provinces in Java (Banten, Yogyakarta, Jakarta, West Java, Central Java) or 83 % from provinces in Java smaller than the mean its value in Java amounted to 6.85. Therefore,

conducted modeling to determine the variables that affect the value of NE to improving the achievmentof the its value.

Table 1Descriftive of the National Examination

Province	Minimum	Maximum	Mean	Java Mean
Banten	5.07	7.24	6.19	6.85
Yogyakrta	5.63	6.44	5.89	
Jakarta	6.16	6.98	6.62	
West Java	6.03	8.04	6.64	
Central Java	5.52	7.91	6.61	
East Java	6.00	8.62	7.60	

4.2. Spatial Correlation

4.2.1. Indeks Moran Test

Indexs Moran test used to detect spatial correlation, the results showed that used queen contiguty weighted with moran index values of0.45 and a p-value of 0.00 which is a smaller than the α of 0.05, so H_0 is reject. This fact means that there is a spatial correlation at 5% significance level. The next Indexs Moran tes usedthe based on the distance weighted with moran index value of0.33 and a p-value of 0.00which is a smaller α of 0.05, so H_0 is reject. This fact means that there is a spatial correlation at 5% significance level.

4.3. Spatial Effect Test

Based on Table 2, the results LM test of the SAR model with queen contiguty wighted have p-values of 0.00 which is a smaller than the α of 0.05, so H_0 is reject. This fact means that there is a spatial lag effect at 5% significance level. Than results LM test of the SEM model with queen contiguty wighted have p-values of 0.08 which is a smaller than the α of 0.1, so H_0 is reject. This fact means that there is a spatial galat effect at10% significance level.

The other resultsLM test of the SAR and SEM modelsused based on distance wighted witha p-values of 0.16 and of 0.72 which are a larger the α of 0.05, so H_1arerejected. This factsmeans that there are not spatial lag and galat effect at 5% significance level. Therefore, the modelsare developed SAR and SEM with queen contiguity weighted.

Table 2.The LM Test

Type of the Test	LM Statistics	P-value	Decission
LM_ρqueen contiguity	11.59	0.00**	Tolak H_0
LM_λqueen contiguity	3.03	0.08*	Tolak H_0
LM_ρbased on distance	1.98	0.16	Tolak H_1
LM_λbased on distance	0.12	0.72	Tolak H_1

4.4. DevelopSpatial Autoregressive Model With Queen Contiguity Weighted
4.5. The SAR model build is as follows:
4.6. $\hat{y} = 10.86 + 0.20\ Wy - 0.03\ X_2 - 0.02\ X_3 + 0.04\ X_4 - 0.03\ X_5 + 0.03\ X_6 - 0.80\ X_9 - 1.63\ D_1 - 1.63\ D_2 - 1.04\ D_3 - 1.23\ D_4 - 1.01\ D_5 + 0.27\ D_8 + 0.41\ D_{10}$.
4.7. with AIC value is 168.91.

These coefficients of the model can be interpreted as follows. The estimate of parameter ρ is 0.20 and significant, this meanof the NE districts/cities value to-i contributed 0.20 of the mean NE value neighboring district to-i. If teachers and education personnel (X_4) increases 1% then the mean of NE districts/cities value increases about 0.04 if remaind the other variable. The coefficient $D_1$of-1.63 and significant, this means that the Banten Province lower 1.63 NE its value compared other provinces (Jakarta, Yogyakarta, West Java, Central Java and East Java). The coefficient D_8of0.27 and significant, this means that districts/cities with middle distance to the larger provincial capitals of 0.27 NE its value compared to the districts/cities within near and far. The coefficient $D_{10}$is 0.41 and significant, this means that districts/citieswith a much larger number of participants of 0.41 NE its value compared to the districts/citieswith the number of small participants.

4.8. Develop Spatial Error Model With Queen Contiguity Weighted
4.9. The SEM model build is as follows:
4.10. $\hat{y} = 11.28 + 0.40\ Wu + 0.04\ X_4 - 0.03\ X_5 + 0.02\ X_7 - 0.64\ X_9 - 1.75\ D_1 - 1.80\ D_2 - 1.31\ D_3 - 1.38\ D_4 - 1.17\ D_5 + 0.32\ D_{10}$.
4.11. with AIC value is 175.11.

These coefficient each model can be interpreted as follows. The estimate of parameter λ is 0.40 and significant, this means that the mean of the NE district value to-i contributed to the location 0.40 by the mean ressidual that location.If teachers and education personnel (X_4) increases of one then the means of NE districts/cities value increases about 0.04 if remaind the other variable. The coefficient on $D_1$is -1.75 and significant, this means that the Banten province lower 1.75 NE its value compared other provinces(Jakarta, Yogyakarta, West Java, Central Java and East Java).The coefficient $D_{10}$is 0.32 and significant, this means that districts/citieswith a much larger number of participants 0.32NE its value compared to the districts/citieswith the number of small participants.

4.12. Selection of Model

SAR model with weighted of the queen contiguity is selected modelbecause Akaike Information Criterion (AIC) lowest value of 168.91. The factors that influences the NE are mean of the process value (X_2), graduation standards (X_3), teachers and education (X_4) facilities and infrastructure (X_5), management (X_6) and the mean of the SchoolsExamination districts/citiesvalue (X_9). The other factors are the spatial lag effect, provincial dummy, distance dummy and number dummy of the NE participants.

4.13. Assumptions Model Selected Test

Kolmogorv Smirnov (KS) test to identify normality on ressidual of the SAR model. The result shows a p-valueof0.07 which is a larger theα of0.05, so H_1 is reject. This fact means that residual of the SARmodel is normal distribution. Then Breusch Pagan (BP) test to identify variance homogenity. The result shows a p-values of0.07 which is a larger the α of0.05, so H_1 is reject. This fact means that varietyof the SARmodel is homogeneous at 5% significance level.

5. Conclusion and Remarks

SAR models with queen contiguity is better than the SEM models with queen contiguityto determinemean of the NSE districs/cities value which influencemean of the NE districts/cities value at National Islamic Senior High School on Java. The mean of NSE districts/cities value which influence are the mean of the processvalue, graduation, teachers and education, infrastructures and management. The other is the spatial lag effect, mean of the SchoolsExamination districts/cities value, the difference NE number of participants, the difference distance to the capital of the province and the provincial differences.

References

[1] Anselin, L(1988), "Spatial Econometrics: Methods and Models," Netherlands, Kluwer Academic Publishers.

[2] Arisanti, R(2011),"Model Regresi Spasial untuk Deteksi Faktor-Faktor Kemiskinan di Provinsi Jawa Timur[tesis]" Bogor (ID): Institut Pertanian Bogor.

[3] [Depdiknas]. Departemen Pendidikan Nasional (2011),"Analisis Sistem Akreditasi Sekolah/Madrasah", Jakarta: Pusat Informasi dan Humas Kemendiknas.

[4] ----------. (2013),"Laporan Hasil Ujian Nasional dan Akreditasi", Jakarta: Balitbang Direktorat Jenderal Mandikdasmen Depdiknas.

[5] Fischer, M.M., Wang, J. (2011),"Spatial Data Analysis: Models, Methods and Techniques", Springer Heidelberg Dordrecht London New York.

[6] Fotheringham, A.S., Brunsdon, C., Chartlon, M.,(2002),"Geographically Weighted Regression, the analysis of spatially varying relationships", John Wiley and Sons, LTD.

[7] Mongi, C.,(2014),"KeterkaitanAntaraNilaiUNdenganPeringkatAkreditasipadaKabupaten/Kota di ProvinsiJawa Barat [tesis]", Bogor: Statistika Terapan. FMIPA IPB.