SPATIAL EDUCATION IN THAILAND: DOES SPATIAL DIMENSION AFFECT ENROLLMENT RATE AND LITERACY RATE OR NOT?

BY

MR. SUPHAWISIT BOONLA

PRESENT

ASSISTANT PROFESSOR
DR. NATTAPONG PUTTANAPONG

A REPORT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF ECONOMICS
(INTERNATIONAL PROGRAM)
FACULTY OF ECONOMICS
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2019

1. Introduction

The education system plays an important role in producing human capital for the economy. However, in recent years, from 2011 to 2013, American colleges' enrollment of students aged between 18 and 24 plummeted by nearly one million (Tomar, 2015). Furthermore, the knowledge level of students in each province is also different. Despite having the same educational system as mentioned, students in the urban are more successful than rural students because they have enough endowment to study and facilitate them. Therefore, this makes a significant difference in scores between urban and rural areas, including the biasing policy of the state that often subsidizes schools in the urban first. It affects more disparity between studies Therefore, we will foresee such problems. Therefore, is the origin of this study.

For this paper, we will analyze 2 steps, Firstly, we would like to investigate whether spatial clustering of educational variable and socio-economic variable can depict the distribution of these variables across the province in Thailand. Therefore, we describe and visualizes spatial distribution, identify spatial outliers, detects agglomerations and local spatial autocorrelations and highlights the types of spatial heterogeneities. Secondly, Spatial regression, with the dependent variable being the enrollment rate in secondary education and literacy rate. Hence, we have to review literature related to enrollment in education or demand for education and literacy rate in order to find more control variables for spatial regression.

2. Literature Review

2.1 The Empirical Study of Education and Human Capital

The number of skilled labor in the country will help to improve economic growth. There is a lot of evidence to mention that knowledge in humans, we call human capital, that produces output for the country and the main contribution in human capital comes from education. Many researchers have studied education and human capital as follows.

Gary S. Becker (1964) Education is considered the most important supporting factor to human capital, and human capital is the factor driving economic growth. In Solow's growth model, we decompose capital as physical and human capital in order to be part of production functions and Endogenous growth theory, it focuses on the importance of knowledge generation and spillover effects of education. Additionality, Çeçen *et al.* (2003) and Doğruel and Deniz (2008) which uses VAR model to analyze the relationship between growth and human capital. And they find that all levels of education expect for university-level contribute economic development in Turkey.

Several empirical researches support about the crucial importance of human capital stocks for growth and productivity. Glaeser and Saiz (2004) learning will help to adapt every changing scenario and there is still a positive relationship between human capital and economic growth. For the UK region, Monastiriotis (2002) finds a positive impact on human capital on production. There is evidence also of strong divergence across cities during the 1990s in levels of human capital, presumably mainly as a result of selective migration of the more skilled towards areas with an already better-educated population.

This is important that underscores the importance of education as it affects the economy that causes economic growth. In the next section, we will discuss the spatial analysis in education.

2.2 The Empirical Study of Spatial Analysis in Education

Sofia Ahmed (2009) From Tobler's first law of geography "proximate locations often share more similarities than locations far apart" and is incorporated in spatial modelling which typically aims to look for associations. So, spatial dimensions of economic growth begin to attract attention and we do not ignore the likely possibility of spatial interaction particularly within a country. Firstly, we usually compute Global and Local measurement[1] of spatial autocorrelation because spatial social sciences data usually has spatial dependence that refers to the lack of independence between observations. It can be considered as a functional relationship between what happens at one point in space and what happens in another.

In regression analysis, conventional regression is not enough for detailed spatial analysis because we do not consider the spatial effects and problems of spatial data analysis. Due to spatial autocorrelation, it causes measurement error that may exist for observations in contiguous spatial units. When we use conventional regression to analyze. It violates the standard assumption of independence among observations. Therefore, standard regression analysis that does not compensate for spatial dependency can yield possibly biased estimators and unreliable significance tests. As a remedy spatial autocorrelation statistics have been devised to detect, measure and analyze the degree of dependency among observations.

2.2.1 Predominant of Urban : Why the spatial dimension is important in Education ?

Xiaobo Zhang and Ravi Kanbur (2003) Biased policy from the China's central government, health care and school conditions for rural residents are much worse than their urban cohorts. They use educational indicators from sample survey data is publicly available

[1] Theil index and Moran's I index

at the national level such as the student-teacher ratio, enrollment rate and illiteracy rates for the years; 1964,1982,1990 and 1995. As a result, the illiteracy rate has declined steadily over the years, large rural-urban, and gender-gaps in 1995 and the rural illiteracy rate was 78% higher than the urban illiteracy rate. They display the spread in the illiteracy rate across rural and urban areas, with the Gini coefficient, Generalized Entropy and Theil index measure inequality. As a result, these value at national level declined from 1964 to 1981 and then increased from 1981 to 1995.So, it shows the differences between urban and rural communities in China.

Helen De la Fuente *et al.* (2013) According to data from the Chilean ministry of education for 2010, the best performance was concentrated in urban spaces, whereas less-advanced school, generally of public nature are found in spaces with high presence of lower-income economic classes. Furthermore, Kiatanantha (2013) Using the Thai PISA 2009, analyzes urban-rural student achievement differentials. It has been argued that such differentials exist because the school in the urban area enjoy more endowments than their rural counterparts; their students, therefore, are able to enjoy more benefits from these endowments. Therefore, it is important that we analyze the educational variables by using location as the determining factor.

2.2.2 Spatial Clustering Analysis in Education

Burhan *et al.* (2009) investigate inequality in Turkey using data covers the period 1997-2006. They use pupils to teacher ratio to measure educational inequality and use the Theil index to capture the regional inequalities. As a result, both between regional inequalities are evaluated at NUTS 1 level[2], between regional inequalities are slightly the same for primary and secondary education. For LISA analysis, university education no significant spatial relation is observed, whereas, for primary and secondary education, there exists a spatial dependency. The LISA cluster maps indicate that the southwestern and southeastern regions of Turkey have asymmetric cluster structures. Furthermore, Tselios (2008) using 94 European regions covering the year 1995-2000. He produces the descriptive spatial analysis of income and educational inequalities with Theil index and Moran’s I index together. Hence, the effects of geographical location are effective on inequalities. He finds that both geographical locations are influential in creating inequalities in income and education.

Sofia Ahmed (2009) uses socio-economic characteristics such as education, health, household income and expenditure data from Federal Bureau of Statistics of Pakistan since 2004 to analyze the relationship between the spatial clustering of income and education in the

[2] 12 geographic regions exists at NUTS 1 level in Turkey, www.tuik.gov.tr

districts of Pakistan and uses Global and Local measurement of spatial autocorrelation, we call Spatial exploratory data analysis (ESDA), are computed using the Moran's I index to obtain estimates of the existing spatial autocorrelation in income and education level across districts. The results show that do not have knowledge spillovers in term of reduction attainment rates across districts but for income reveal a clear spatial autocorrelation pattern whereby high-income district tend to be neighbors of other high-income districts.

Helen De la Fuente *et al.* (2013) intend to systematically compare the 493 schools of existing educational opportunities in the concepcion metropolitan area in Chile by using 4 dimensions; 1. Standard deviation ellipse, 2. Neighborhood index, 3. Sargent Florence Quotient, and 4. Supply-demand relation. The aim to evaluate the spatial disparity in order to improve a most inclusive urban planning. As a result, nearest units to the centre of the metropolitan area have better conditions than the farthest with respect to student-teacher ratio, teaching quality, and academic outcomes assessment. In addition, Bernardin Senadza (2011) would like to examine the nature and extent of gender and spatial inequalities in educational attainment in Ghana by using the education Gini coefficient, computed on the basis of years of schooling based on the distribution of the years of schooling of persons aged 15 years and above. As a result, the three northern regions have lower education attainment as well as higher education gini coefficients compared to the rest of the country and he also finds a positive correlation between poverty incidence and education inequality.

2.2.3 Spatial Regression Analysis in Education

David Brasington *et al.* (2014) investigate the determinants of inter-district open enrollment policies by school districts in Ohio that have 3 choice; 1. prohibit enrollment of students from any other school district in the state (no open enrollment), 2. permit enrollment of students from adjacent school districts, or 3. permit enrollment of students from any school district in the state, which is similar to the school quota system in Thailand. They employ a recently developed spatial autoregressive lag multinomial logit model to examine the determinants of adoption of each open enrollment alternative. Among the most influential factors are school districts demographic characteristics, financial factors, and competitive environment. More importantly, the results show evidence of strategic interaction, with a positive correlation in policy choice between neighboring school districts of about 0.4.

Therefore, from the review, the educational variables used are student-teacher ratio (pupils to teacher ratio), enrollment rate, illiteracy rates, teaching quality, academic outcomes assessment, and years of schooling of persons aged 15 years and above which are used to calculate the index. In addition; poverty, income, gender, and location are factors in determining educational variables. For the location variable, it affects the geographical disparity in education. Thus, we use nighttime light instead of location which indicates the urbanization density. Henderson *et al.* (2012), and Pinkovskiy and Sala-i-Martin (2016) intend to examine the association between nighttime light and income per capita, both published in the top journals in the field of economics. And Papaioannou (2013), Alesina *et al.* (2012), and Ebener *et al.* (2005) use nighttime light as a proxy for income per capita. So, the more densely packed areas of the nighttime light represent more urbanity than other areas.

For this study, we will analyze spatial regression with the dependent variable being enrollment rate in secondary education and the independent variables described above are poverty, income, and nighttime light. But these independent variables are not sufficient to describe the enrollment rate so that they do not endogeneity with the omitted variable. Thus, we have to review literature related to enrollment in education or demand for education in order to find more control variables for spatial regression in addition to the above factors.

2.3 The Empirical Study of Demand for Education

This section intends to examine the factors that affecting enrollment rate that is educational demand often use demand and investment theory to support the study analysis, with details of the result and methods of each study as follows.

To examine the factors that affect the demand for education. Correa (1962) lists the principal determinants of the demand for education; The total population and demographic groups are among the macroeconomics determinants while the preferences, income, and cost of study are microeconomics determinants. The preferences of the consumers for education are defined by intellectual capacity, vocation, parental influence. In addition, Correa utilizes the size of the school-age population, and per capita income to analyze the Harrod-Domar growth model that depicts the relationship among production, education of labor force, and education system. Next, Papi (1969) study suggests the factors affecting the demand for education are population, economic progress that is measured by country's income per capita increasing, and the supply of education such as number of educational institutions in the area.

Panitchpakdi (1974) suggests that among the factors that determined the demand for higher education are the demographic factors; these factors will affect the flows of a student from secondary schools who enrolled in the university. we include the attitude towards education; different social-economic groups foster different attitudes towards education; and economic factors such as cultural living, employment. Some of these factors had an impact on the parents' and students' attitudes towards education. Besides, Centra (1980) finds that the decline in population affects the enrollment in the U.S college in the 1980s by splitting the survey into three situations. Finally, Weiler (2018) applies the time series data and linear regression method to forecast the short-run enrollment in higher education. As a result, the impact of the change in school-age students on number of enrollment in each filed. The result exhibits that the number of high school students contributes a significant effect only on the enrollment of Liberal Art whereas; Technology, Agriculture, Forestry and Economics, demographic factors have not such a significant effect.

Oyer & Lazear (2003), Karanassou *et al.* (2006), Haraldsen *et al.* (2015) aim to identify the factors influencing the demand for education by using macroeconomics factors and labor demand. The study considers panel data and macroeconomic factors as the independent variable such as; Employment, Wage, Unemployment Rate, Real GDP, and Consumer Price Index. They find that the economic growth, wage, employment, and cost of study are the factors significant to determine the demand for education.

All of the above-mentioned studies, we find the control variable to determine enrollment rate. For example; demographic variable, real national income, the cost of education, unemployment, social-culture, and family factors also affect the demand for education. Therefore, the above-mentioned studies provide the basic support to this study.

2.4 Review Summary

From all the literature reviews mentioned, educational variables such as student-teacher ratio, enrollment rate, illiteracy rate, etc. are inequality. Therefore, we use the nighttime light variable to represent the urbanization density. From Sofia Ahmed (2009), Xiaobo Zhang and Ravi Kanbur (2003), and Helen De la Fuente *et al.* (2013) also show the relationship between gender and income that effect educational variables. For this study, we are interested in the demand for education that is represented by the enrollment rate.

Many researches, such as Correa (1962), and Panitchpakdi (1974) indicate that demographic variables will affect demand for education. In addition, population in each province, school-age population, and socio-economic models such as road network, student-teacher ratio (pupils to teacher ratio), illiteracy rates, and school building footprint that must be included in the model. From Oyer & Lazear (2003), Karanassou *et al.* (2006), and Haraldsen *et al.* (2015) include per capital income in each province which here will use the average monthly income, unemployment rate in each province, gross provincial product, and consumer price index in each province.

3. Database

Database from this table, we show the sources of database that we use as secondary data. And in our study, we select 2016 and 2017 to do the analysis because o-net (ordinary national educational test) is not available in 2018 and 2019 and this variable is relevant for our study. So, scope of study is 2016 and 2017.

Table 1

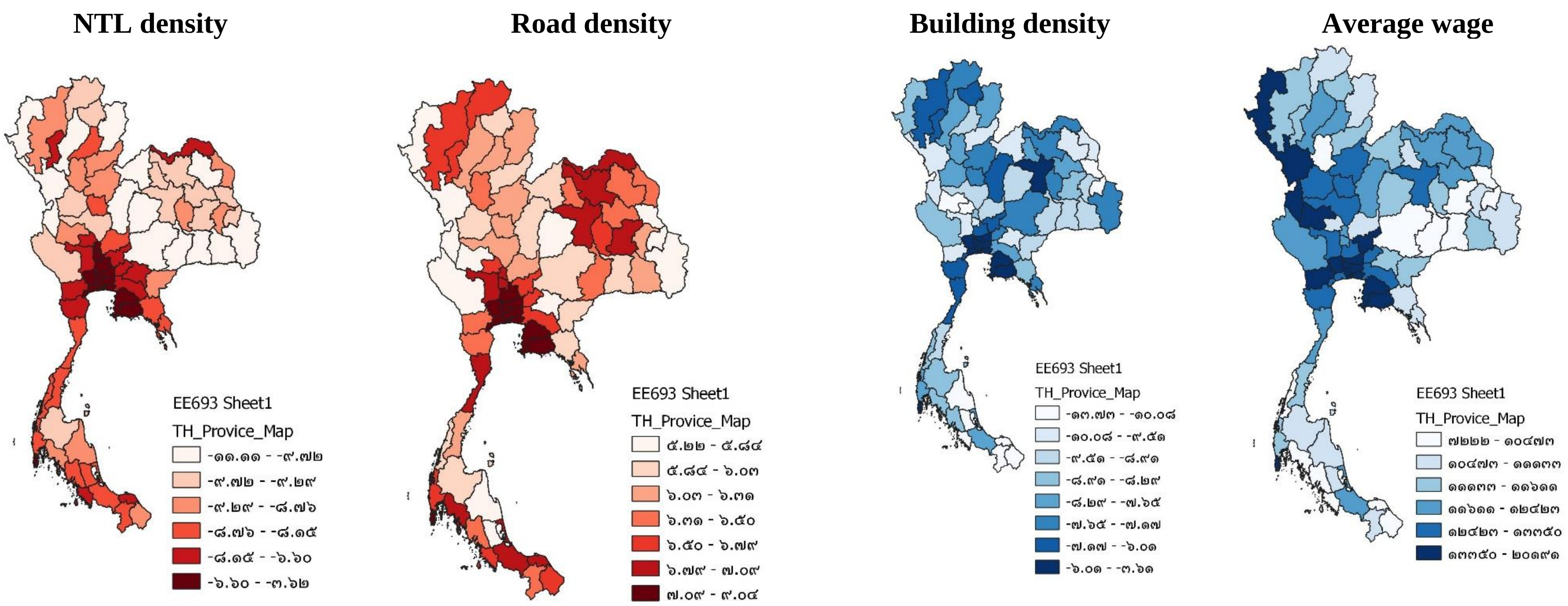
Sources of data

Variable	**Sources**
Nighttime light density	VIIRS-DNB
Road density and Building density	Open street map (http://download.geofabrik.de/)
Average wage for each province	Labor Force Survey
Gross provincial product : GPP	NSO
Consumer price index : CPI	NSO
Proportion of the poor as measured by expenditure	NSO
Poverty line	NSO
Population density	NSO
Enrollment rate	MOE-eduWH (http://www.eduwh.moe.go.th/)
O-NET score	NIETS (http://www.serviceapp.niets.or.th/onetmap/)

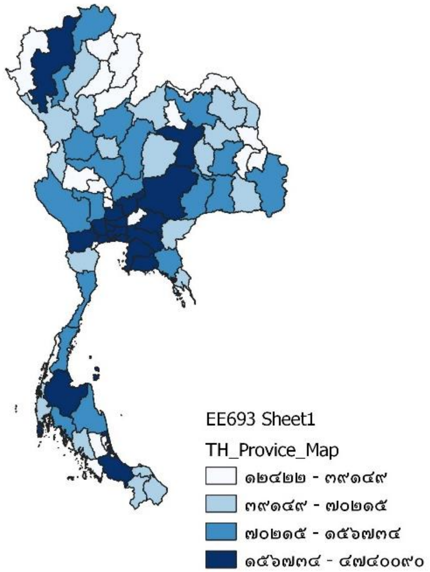
Source: Author’s summarizing

Figure 1

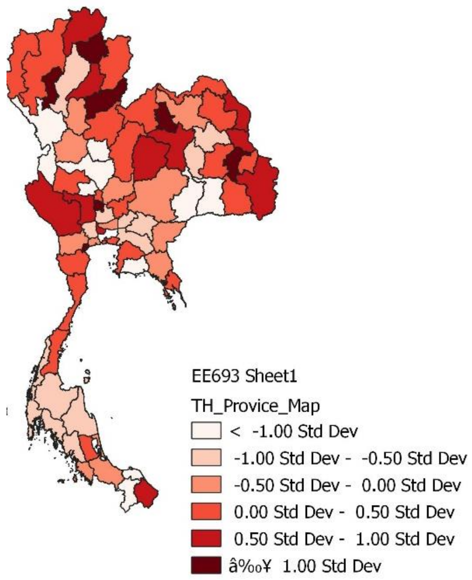
The distribution of each variable in 2016



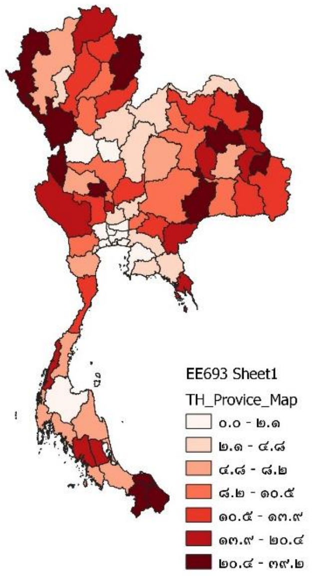
Gross provincial product



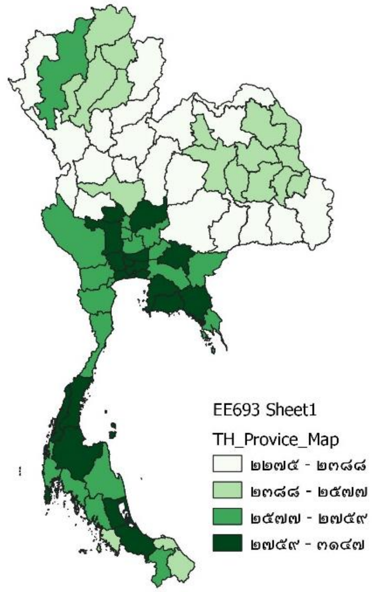
Consumer price index



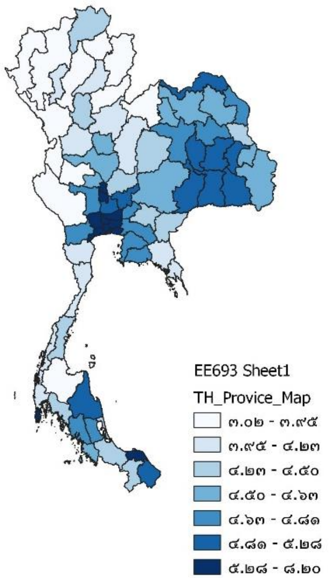
Proportion of the poor



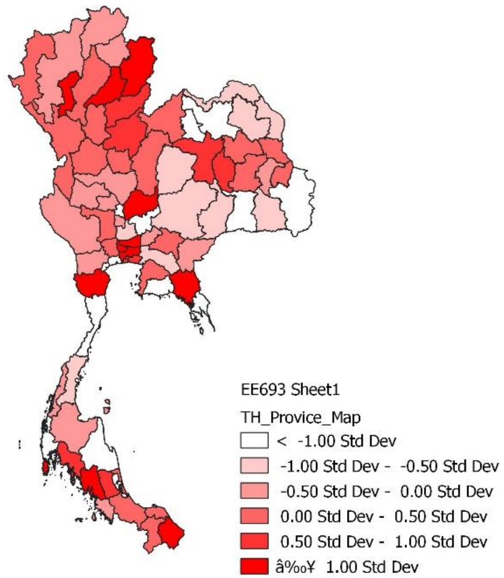
Poverty line



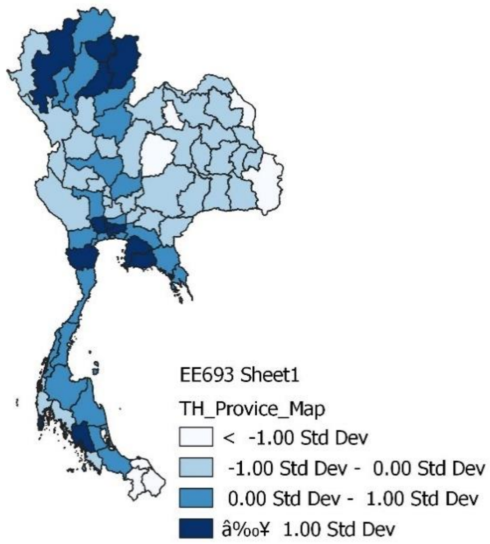
Population density



Enrollment rate



O-NET

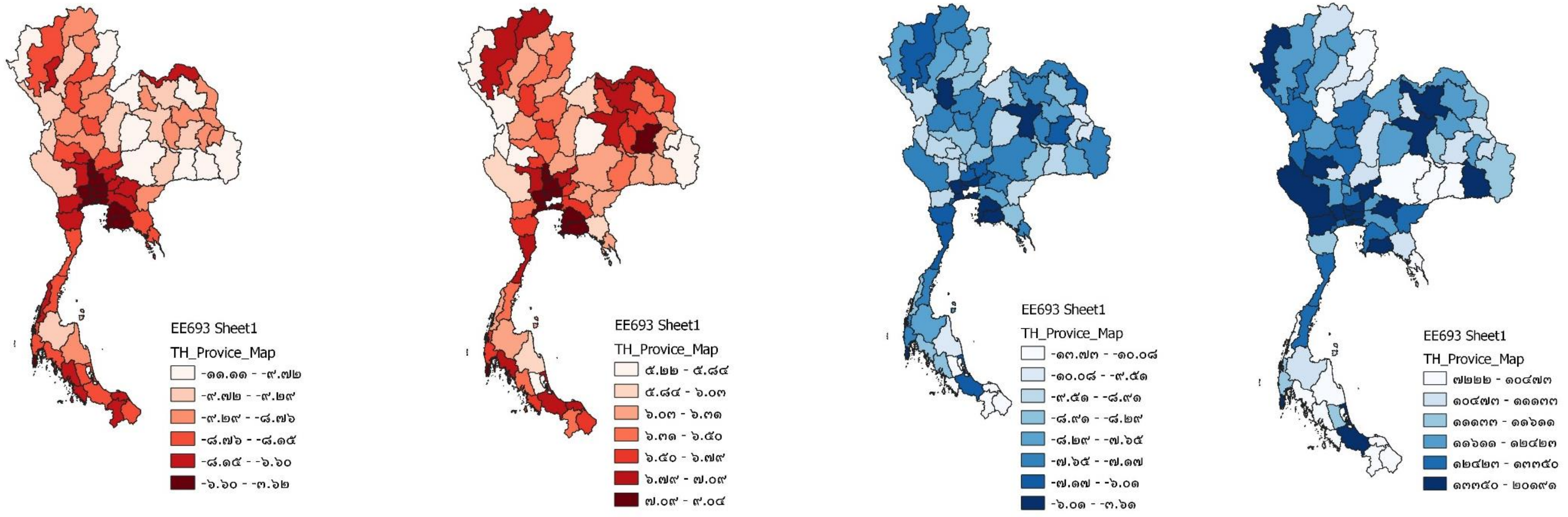


Source: Author’s computation

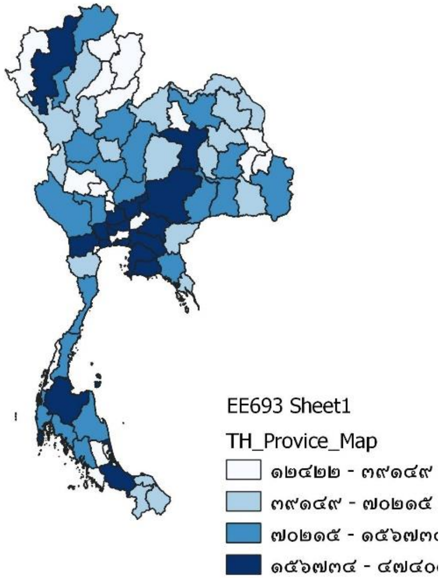
Figure 2

The distribution of each variable in 2017

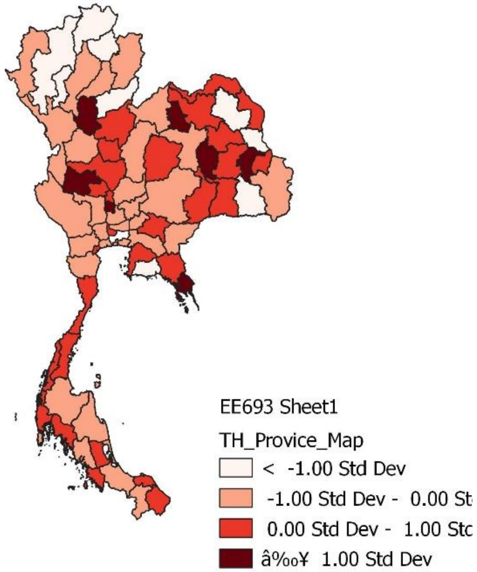
NTL density **Road density** **Building density** **Average wage**



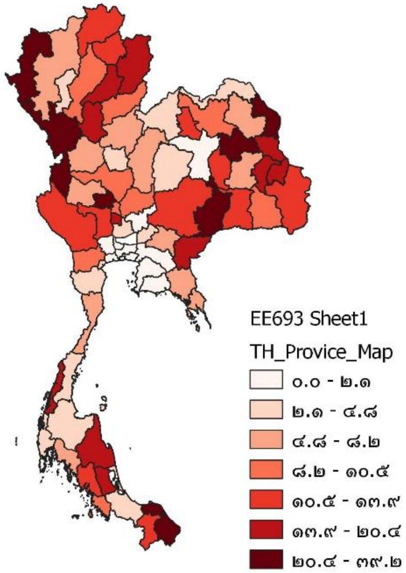
Gross provincial product



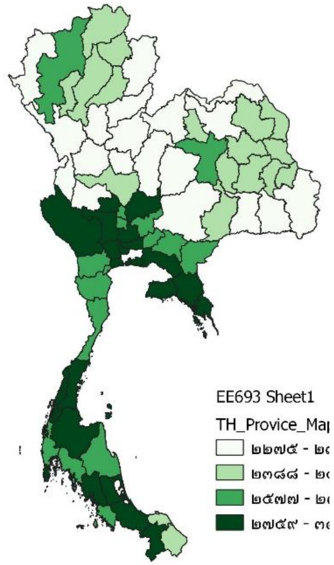
Consumer price index



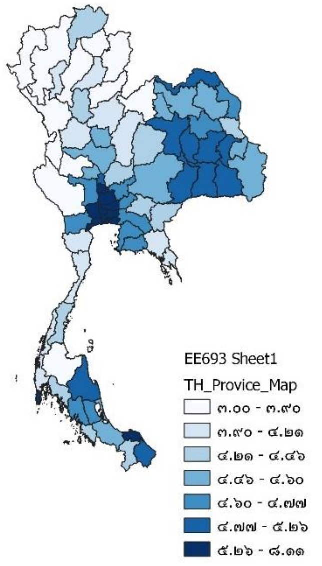
Proportion of the poor



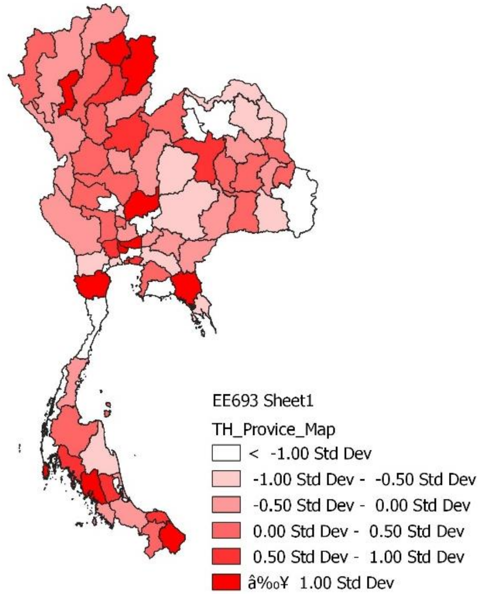
Poverty line



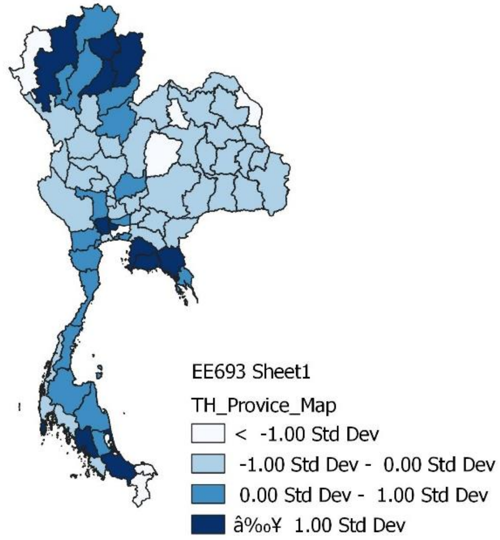
Population density



Enrollment rate



O-NET



Source: Author’s computation

4. Methodolody

4.1 Principle Component Analysis (PCA)

Since the O-NET scores are not able to be used to calculate the arithmetic mean for each province. Because each subject has a different weight or different factor loading, we use PCA to combine the scores of all 4 subjects into one variable.

PCA transforms data and generates the new indicator which is an aggregate index representing the main literacy rate of each province in Thailand.

Mathematically, PCA is the transformation of data generating the new set of uncorrelated components. Each is the linear weighted combination of original variable, we call the weight is factor loading. Relevant components must have eigenvalue that is greater than 1.

Table 2

Principle Component Analysis

Components	Linear combination	Eigenvalue
1	$ONET_1 = a_{11}THAI + a_{12}ENG + a_{13}MATH + a_{14}SCI$	λ_1
2	$ONET_2 = a_{21}THAI + a_{22}ENG + a_{23}MATH + a_{24}SCI$	λ_2
3	$ONET_3 = a_{31}THAI + a_{32}ENG + a_{33}MATH + a_{34}SCI$	λ_3
. . .	. . .	. . .
p	$ONET_p = a_{p1}THAI + a_{p2}ENG + a_{p3}MATH + a_{p4}SCI$	λ_p

Source: Author's summarizing

where $\lambda_1 > \lambda_2 > \lambda_3 > ... > \lambda_p > 1$ and a_{pp} identifies the factor-loading for the p-th principal component and the p-th variable. And to bring to new variable, we use the factor-loading from highest eigenvalue component.

4.2 Local Moran test

The local Moran test by Anselin (1995) detects local spatial autocorrelation in aggregated data by decomposing Moran's I Statistic into contributions for each area within a study region. Formula to calculation as follow

$$Moran's \quad I = \frac{n\sum_i \sum_j w_{ij}\left(X_i - \overline{X}\right)\left(X_j - \overline{X}\right)}{\sum_i \sum_j w_{ij}\left(X_i - \overline{X}\right)^2}$$

where X_i is the variable of interest, $\overline{X}$ is the mean of X_i, n is a number of spatial unit indexed by i and j (for this case is number of province) and w_{ij} is the spatial weight matrix.

For spatial weight matrix, we have to identify neighbors that can be defined based either on; adjacency and distance. But if we choose adjacency, Phuket will disappear in the analysis. Therefore, in this study we would like to use distance criteria to select neighbor. As a result, province in around 111.247 km will be the neighbors in our analysis.

4.3 Spatial Regression Models

Conventional regression method that we usually use to analyze the relationship amount variable is **Ordinary Least Squared Method** that set the conditional on error term.

4.3.1 Ordinary Least Squared Method (OLS)

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + ...\beta_j X_{ji} ... + \beta_k X_{ki} + u_i$$

$\mathrm{cov}(u_i, X_{2i}) = \mathrm{cov}(u_i, X_{3i}) = ... = 0$ or **Fixed X values or X values independent of the error term.**

$E(u_i) = 0$ or **the mean of error term is equal to 0.**

Homoscedasticity $Var(u_i | X_i) = \sigma_u^2$ **constant.**

No autocorrelation (serial correlation) between the disturbance $\mathrm{cov}(u_i, u_j) = 0$.

However, OLS estimation may not be appropriate if clustering patterns occur. Therefore, after estimating OLS, we will do post-estimation testing to select a suitable model such as Spatial Lagged Model (SLM) and Spatial error model (SEM).

4.3.2 Spatial Lagged Model (SLM)

$$Y_i = \rho W Y_j + \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + ...\beta_j X_{ji} ... + \beta_k X_{ki} + u_i$$

where WY_j on the right-hand side of the equation represents an additional spatially lagged dependent variable and ρ is a spatial autocorrelation coefficient.

4.3.3 Spatial Error Model (SEM)

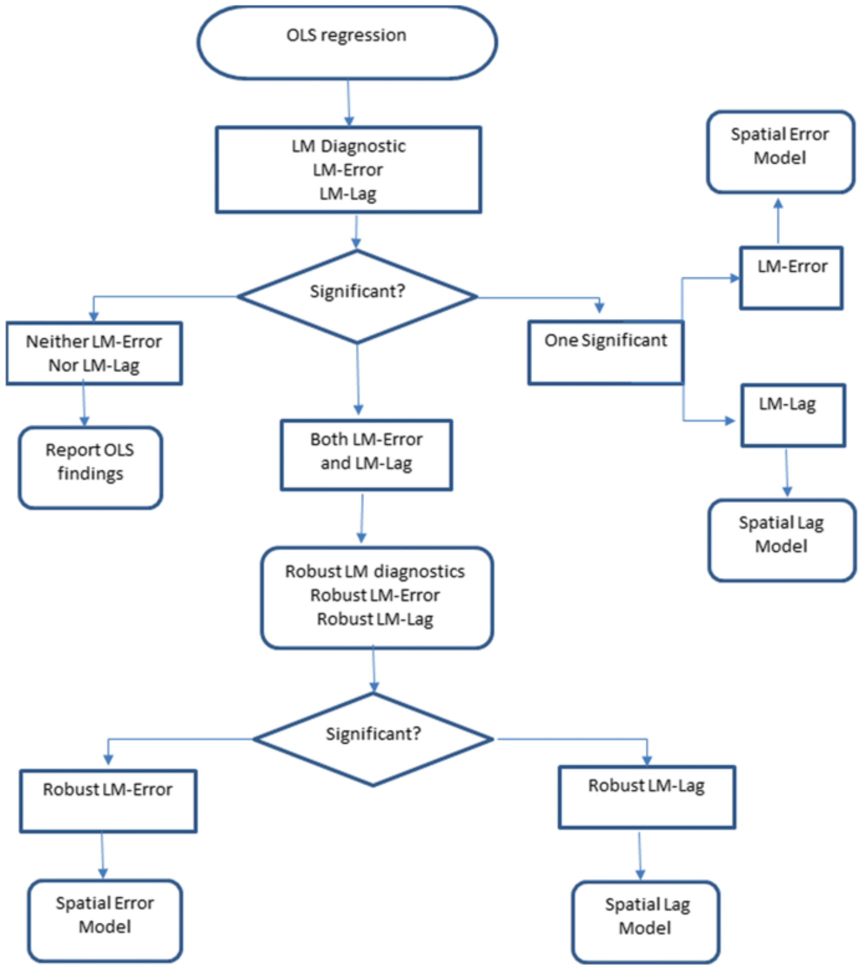
$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + ...\beta_j X_{ji} ... + \beta_k X_{ki} + u_i \quad ; u_i = \lambda W u_j + \varepsilon_i$$

where the stochastic disturbance is a function of the neighbor's disturbance and λ is the influences from neighbors.

To perform the test whether we should use OLS, SLM, or SEM after estimating OLS, we will test Moran's I test with residual and use LM test to select a suitable model.

Figure 3

Steps for finding the suitable model



Source: Nilima *et al.* (2018)

5. Result discussion

5.1 Literacy rate obtained from PCA

Table 3
Eigenvalue for each component for 2016 O-NET

Component	Eigenvalue	Difference	Proportion	Cumulative
Component1	3.66922	3.40401	0.9173	0.9173
Component2	0.26521	0.21612	0.0663	0.9836
Component3	0.04909	0.0326	0.0123	0.9959
Component4	0.01649	.	0.0041	1

Source: Author's computation

Table 4

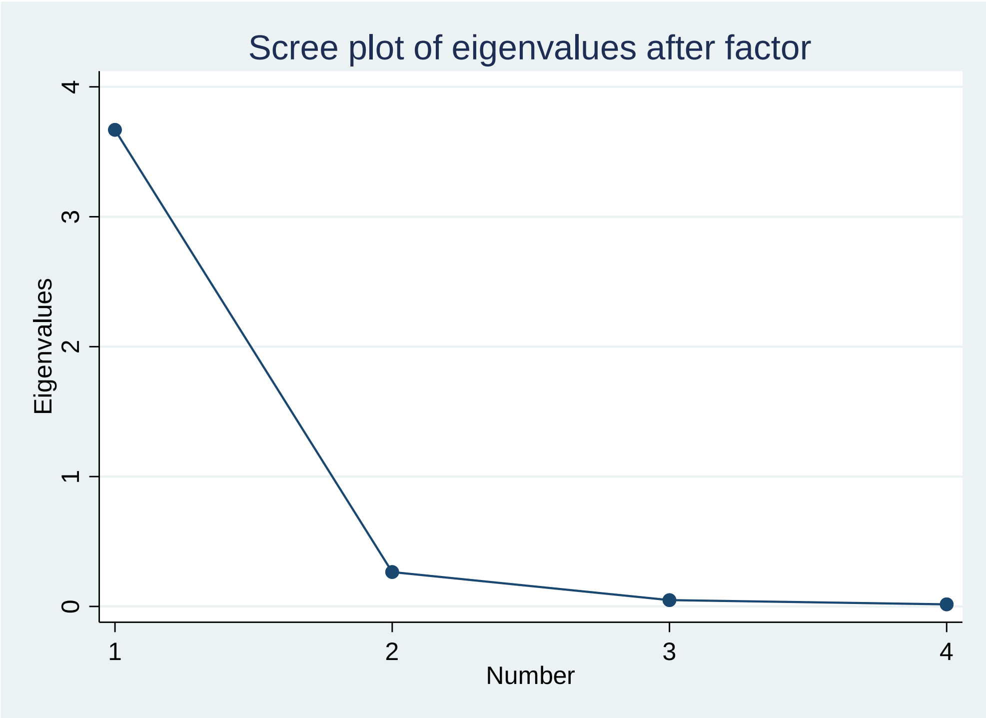
Factor loading for each component for 2016 O-NET

Variable	Factor loadings	Uniqueness
THAI_16	0.9472	0.1028
ENG_16	0.9059	0.1794
MATH_16	0.9896	0.0207
SCI_16	0.9859	0.028

Source: Author's computation

Figure 4

Screen plot for 2016 O-NET



Source: Author's computation

Table 5

Eigenvalue for each component for 2017 O-NET

Component	Eigenvalue	Difference	Proportion	Cumulative
Component1	3.70139	3.46405	0.9253	0.9253
Component2	0.23734	0.1906	0.0593	0.9847
Component3	0.04674	0.03221	0.0117	0.9964
Component4	0.01453	.	0.0036	1

Source: Author's computation

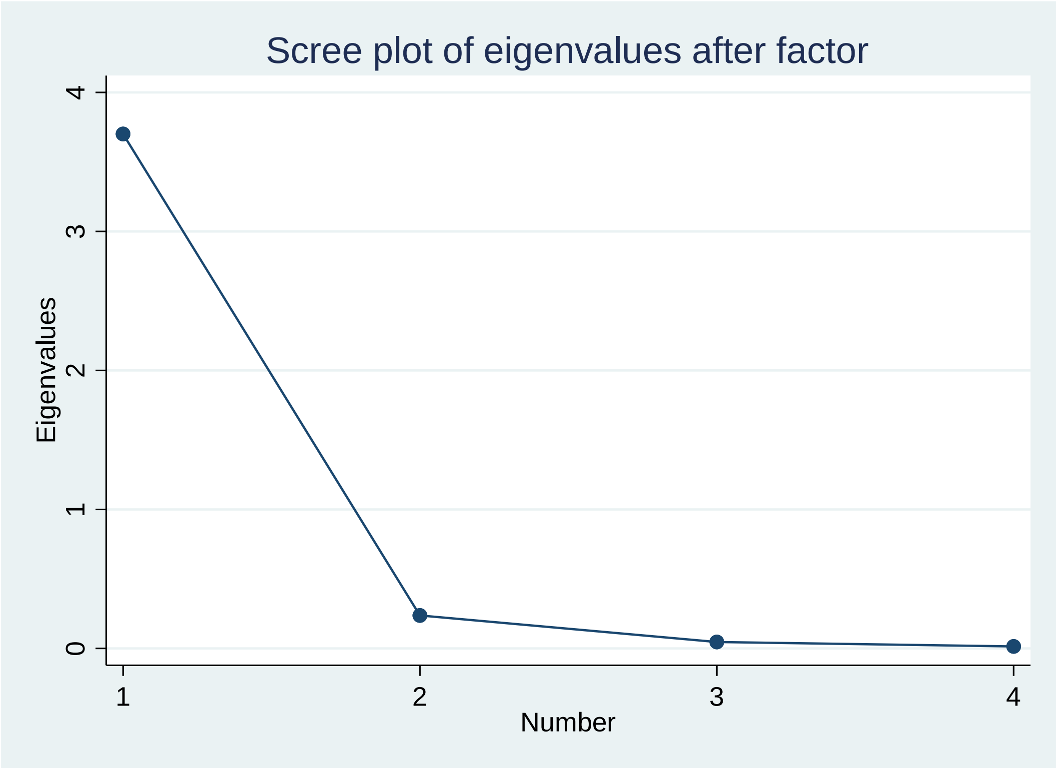
Table 6
Factor loading for each component for 2017 O-NET

Variable	Factor loadings	Uniqueness
THAI_17	0.9604	0.0776
ENG_17	0.9115	0.1693
MATH_17	0.991	0.0178
SCI_17	0.9829	0.0339

Source: Author's computation

Figure 5

Screen plot for 2017 O-NET



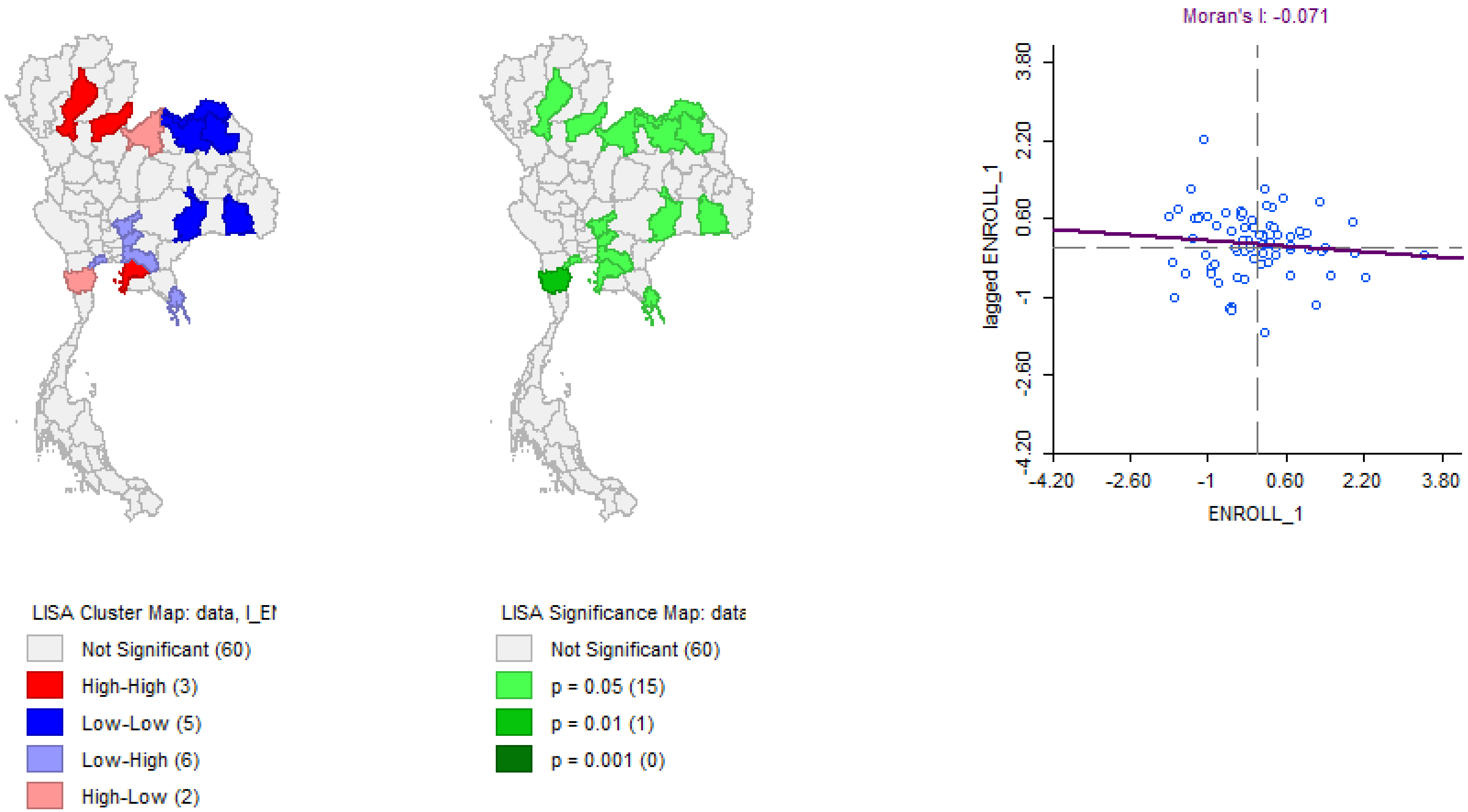
Source: Author's computation

It can be seen that from the PCA of o-net data in 2016 and 2017, will get only one important component. After we got factor-loading, then we will create a factor score that represents the o-net in all 4 subjects.

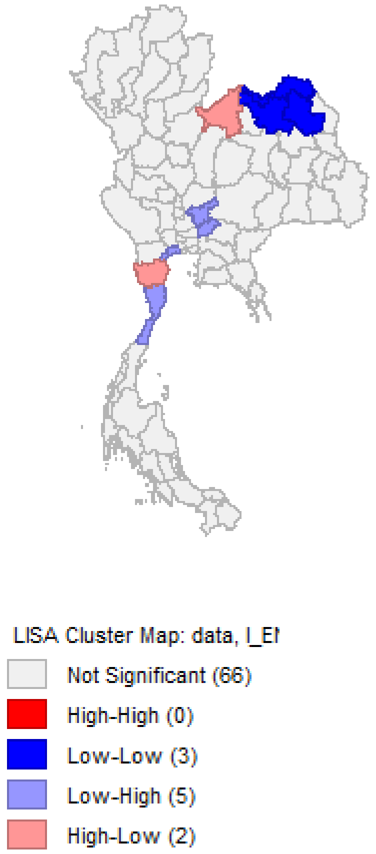
5.2 Spatial Cluster Analysis

Figure 6
Univariate analysis of Enrollment rate 2016 and 2017

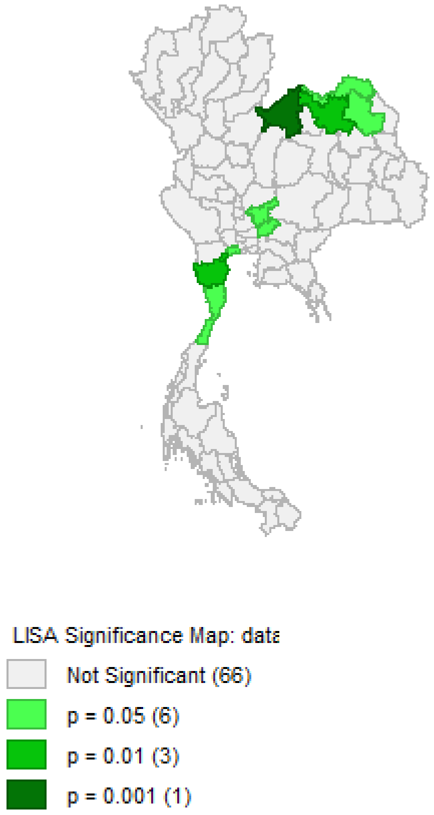
Cluster Map | Significance Map | Moran Scatter Plot



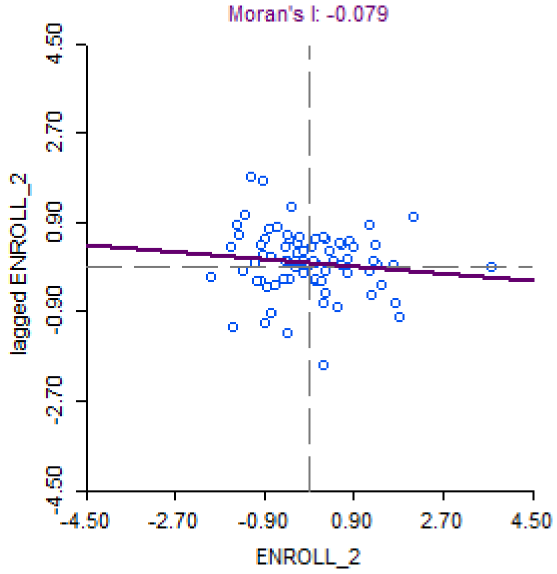
Cluster Map



Significance Map



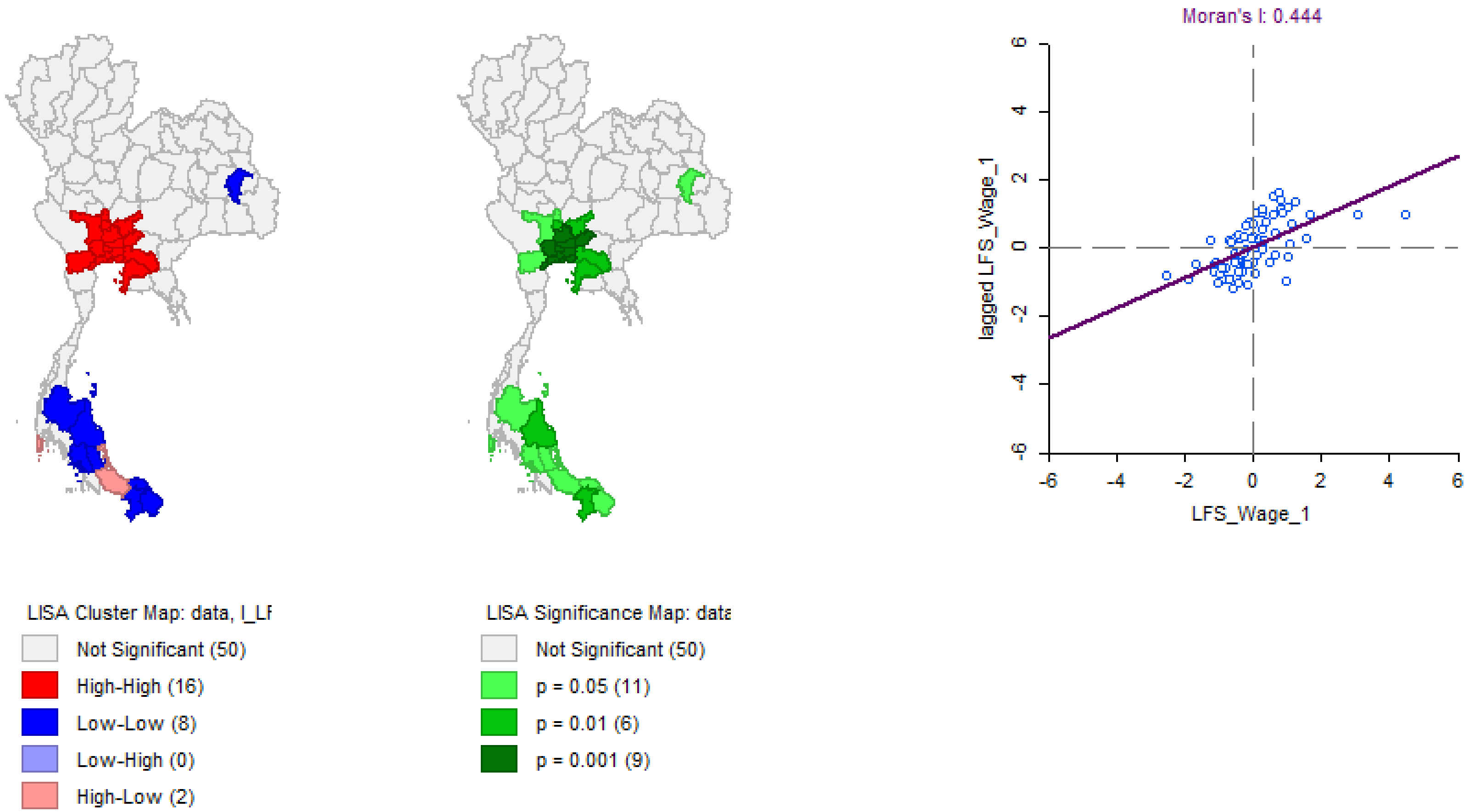
Moran Scatter Plot



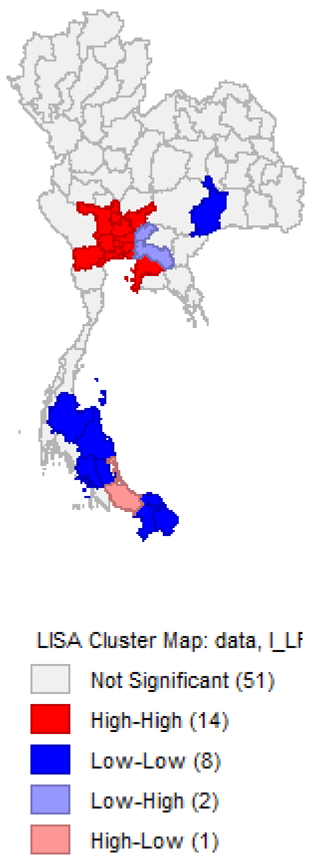
Source: Author’s computation

Figure 7
Univariate analysis of Average wage for each province 2016 and 2017

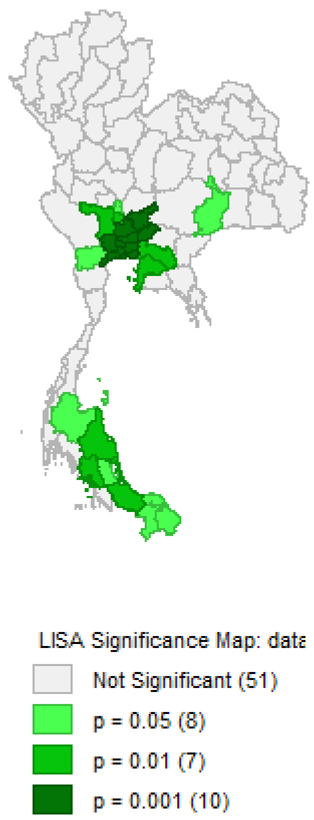
Cluster Map | Significance Map | Moran Scatter Plot



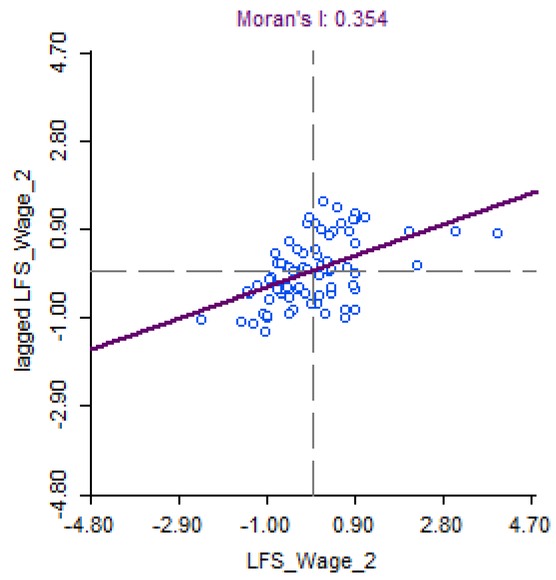
Cluster Map



Significance Map

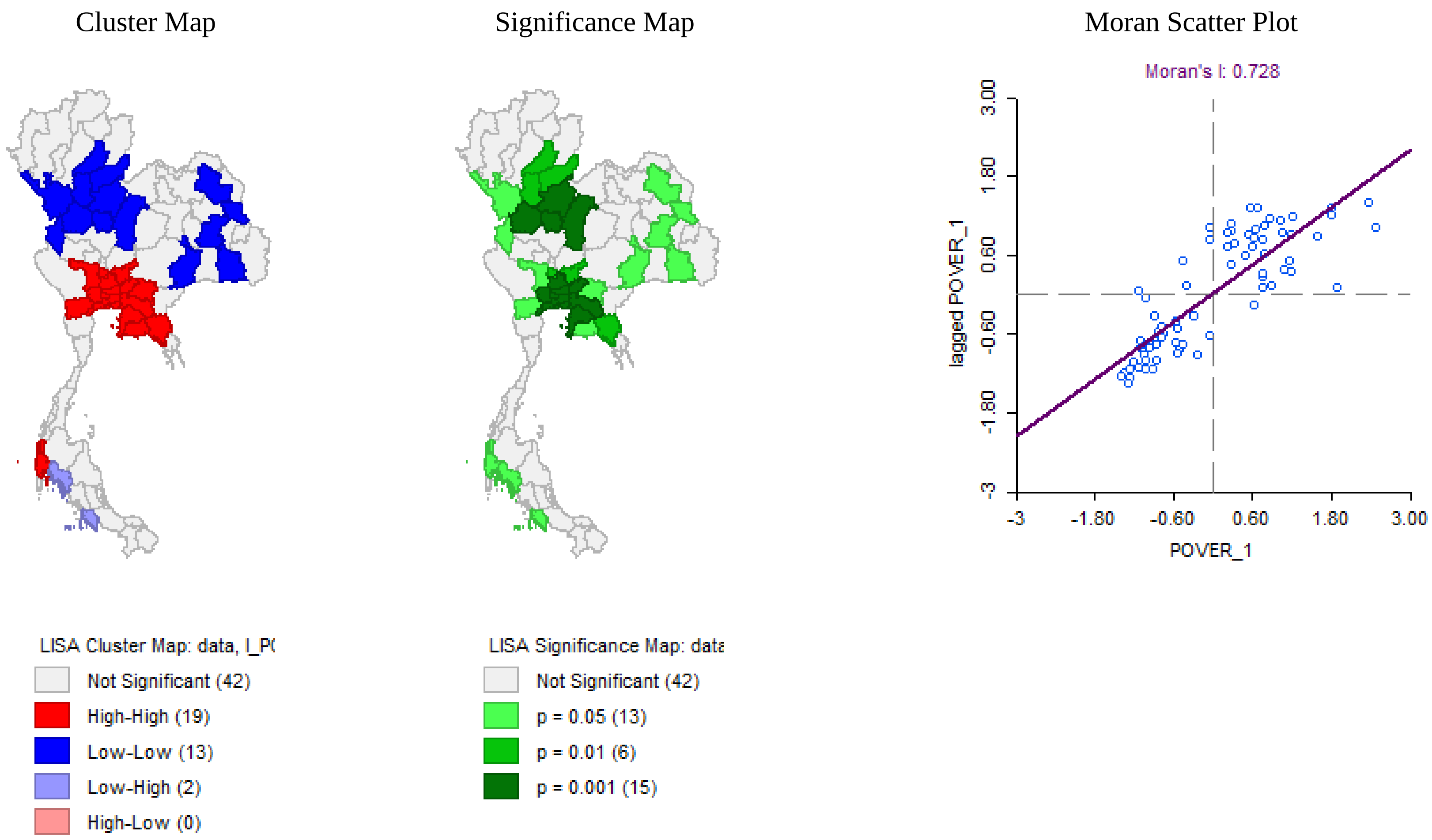


Moran Scatter Plot

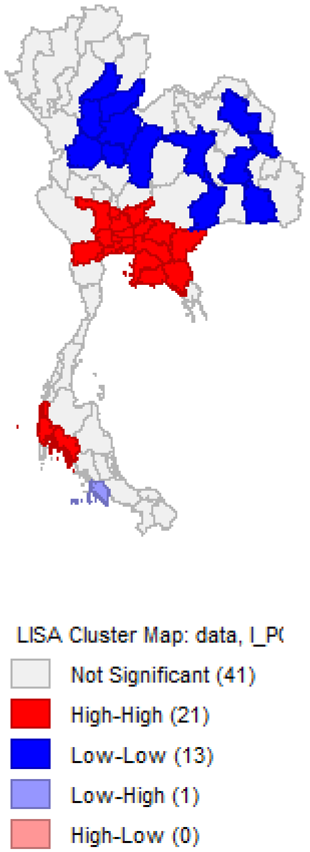


Source: Author's computation

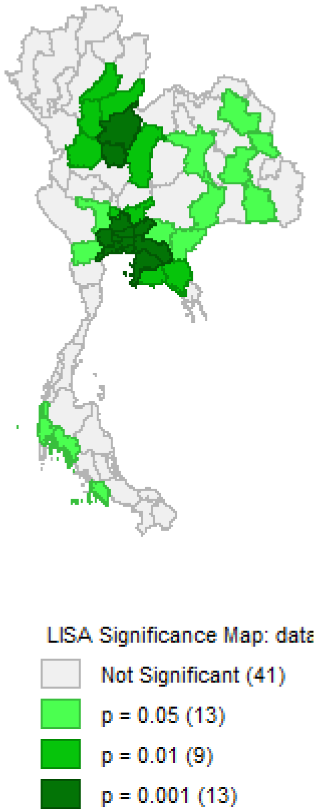
Figure 8
Univariate analysis of Poverty line 2016 and 2017



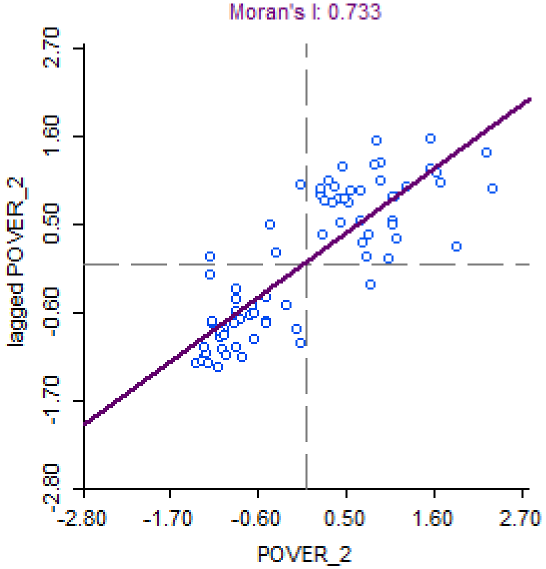
Cluster Map



Significance Map



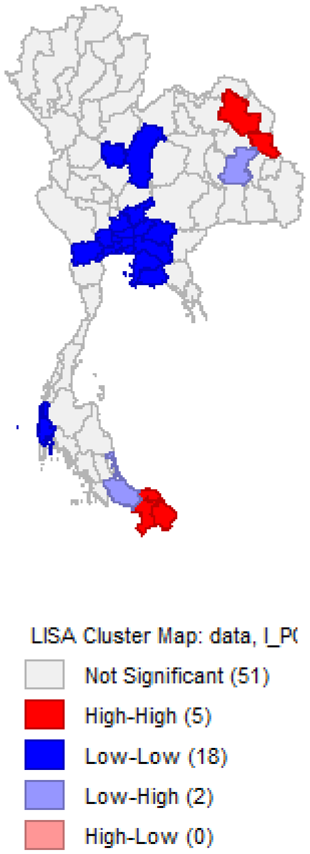
Moran Scatter Plot



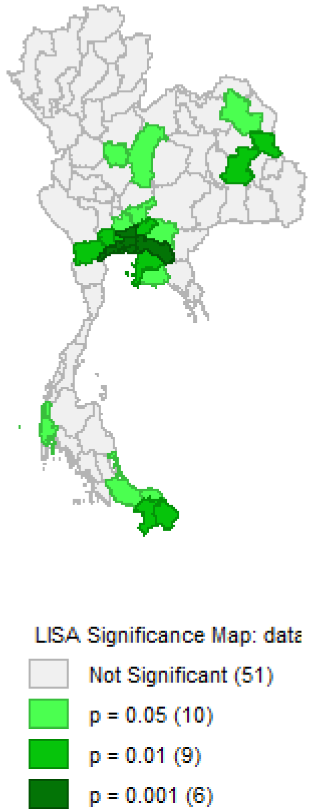
Source: Author's computation

Figure 9
Univariate analysis of Proportion of the poor as measured by expenditure 2016 and 2017

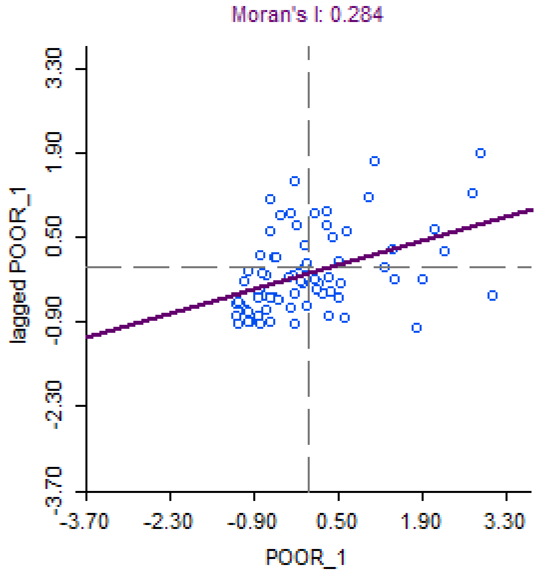
Cluster Map



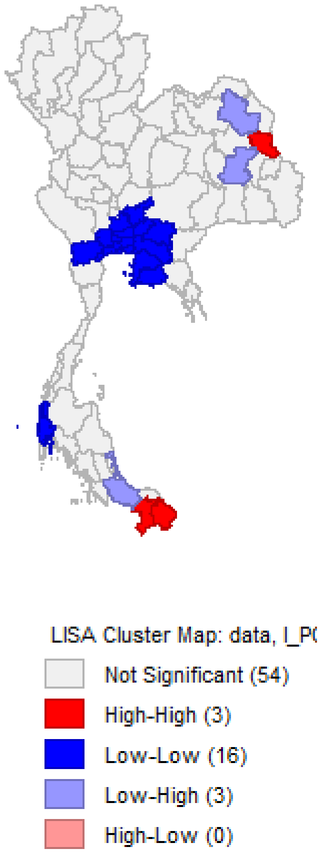
Significance Map



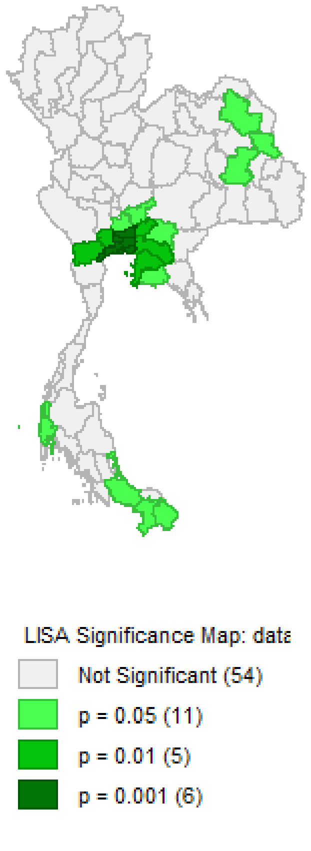
Moran Scatter Plot



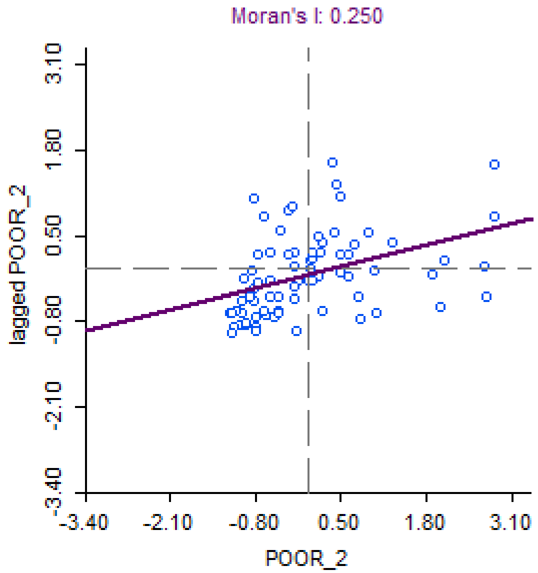
Cluster Map



Significance Map



Moran Scatter Plot



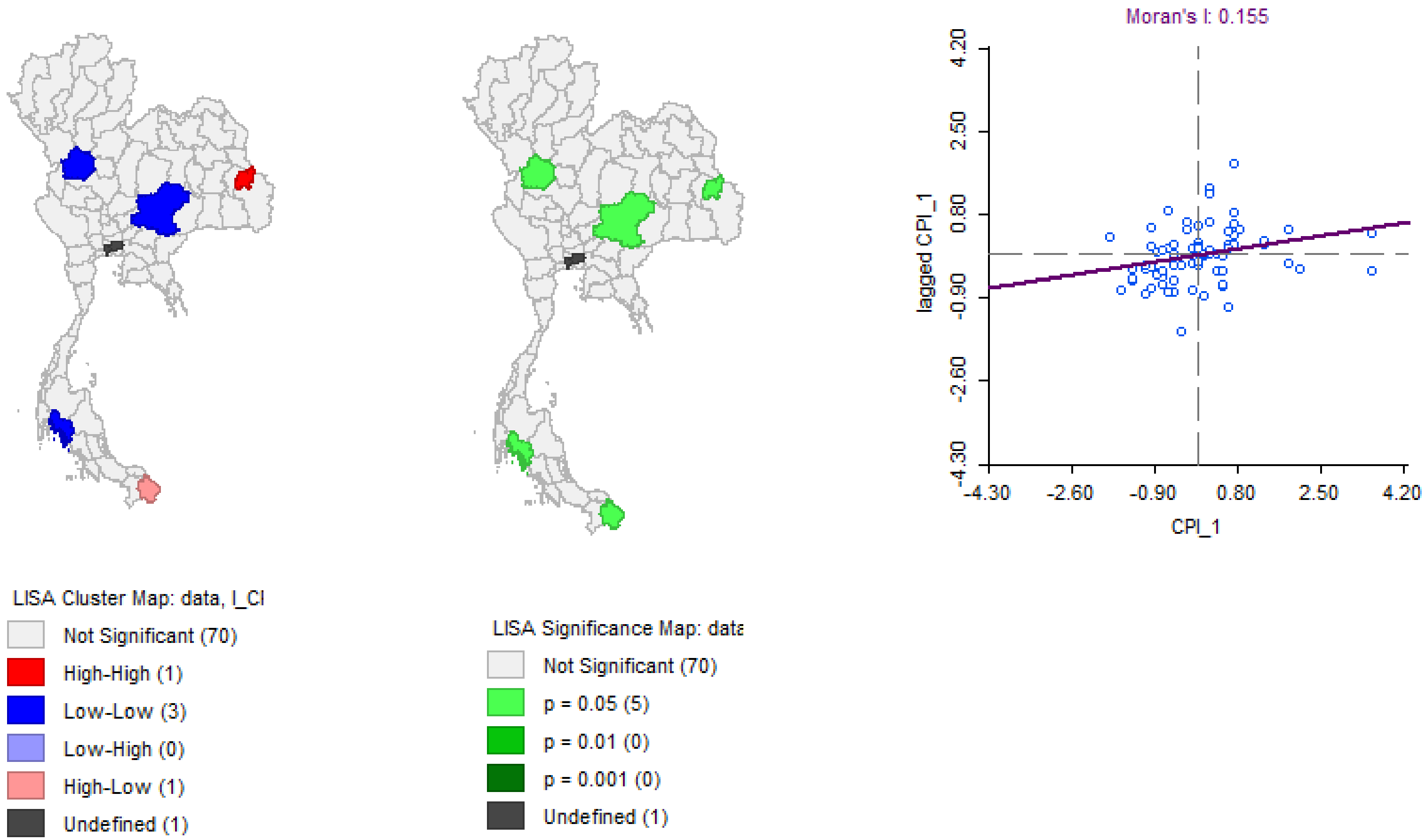
Source: Author's computation

Figure 10
Univariate analysis of Consumer price index 2016 and 2017

Cluster Map

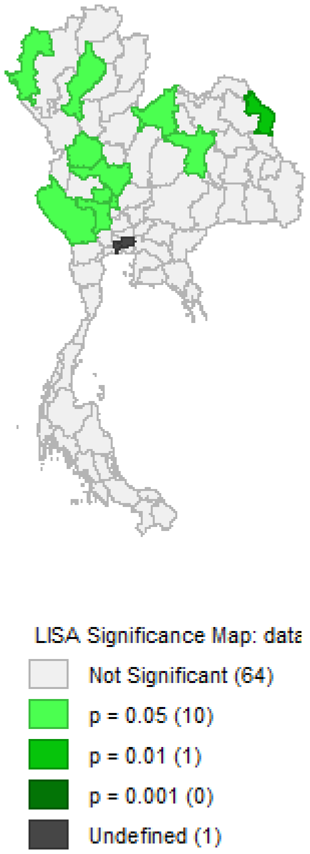
Significance Map

Moran Scatter Plot

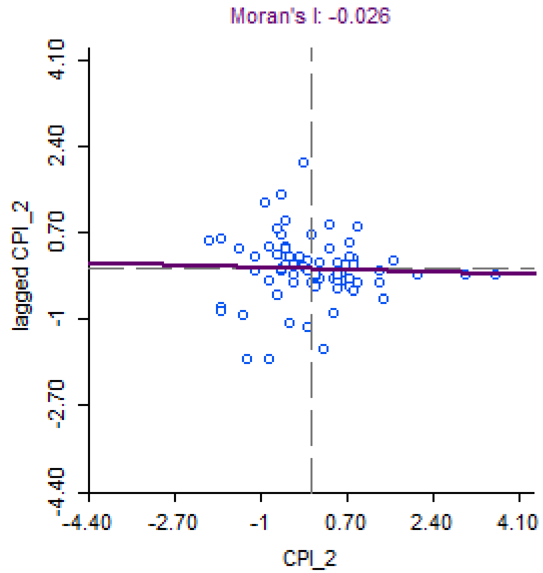


Cluster Map

Significance Map



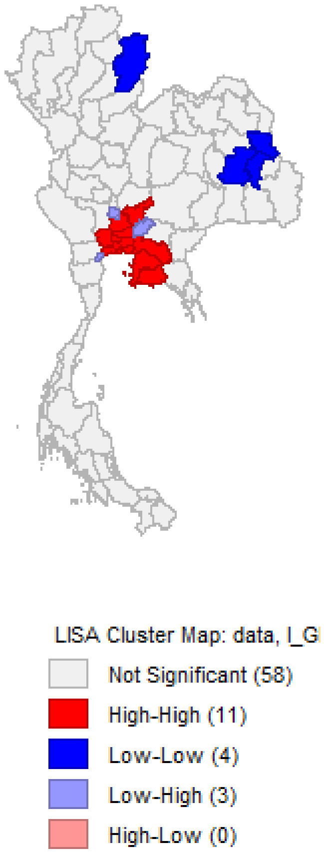
Moran Scatter Plot



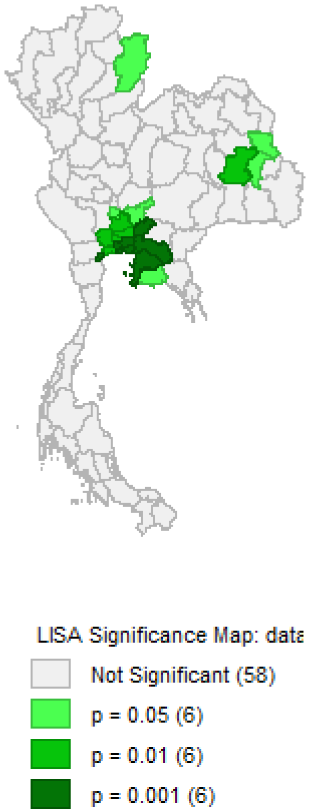
Source: Author’s computation

Figure 11
Univariate analysis of Gross provincial product 2016 and 2017

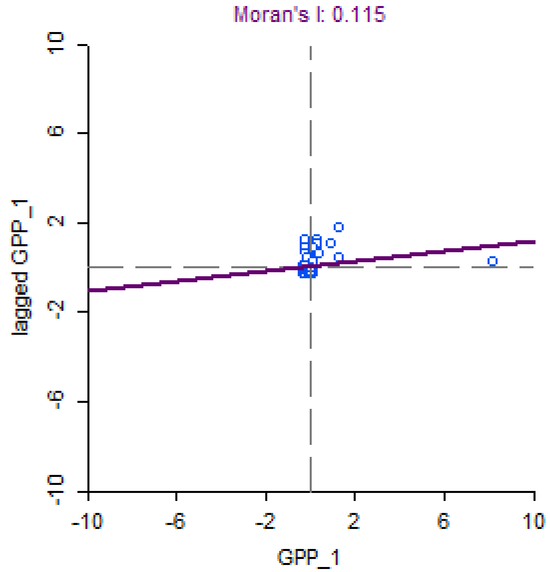
Cluster Map



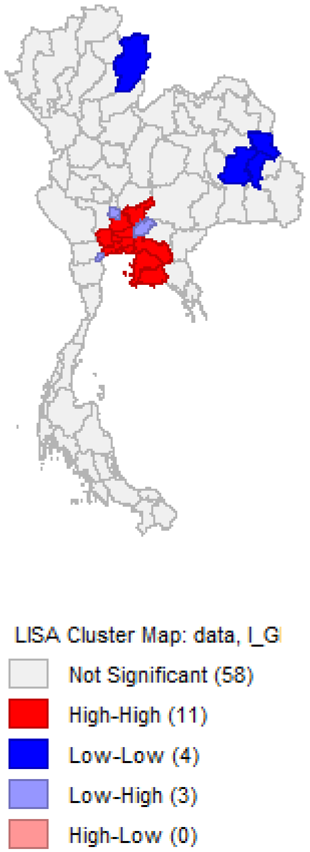
Significance Map



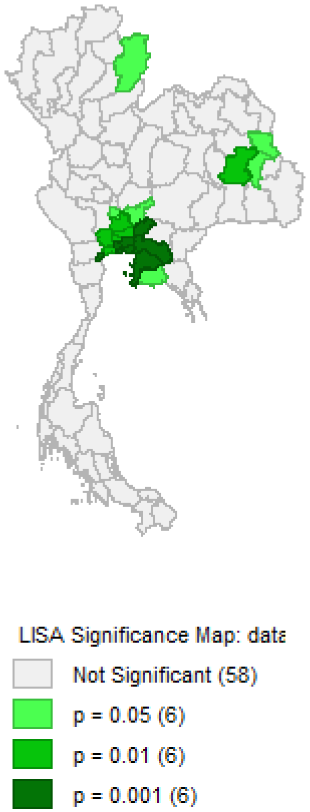
Moran Scatter Plot



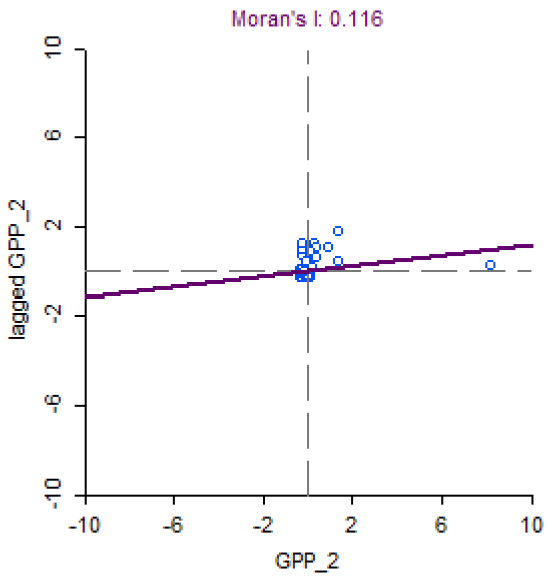
Cluster Map



Significance Map

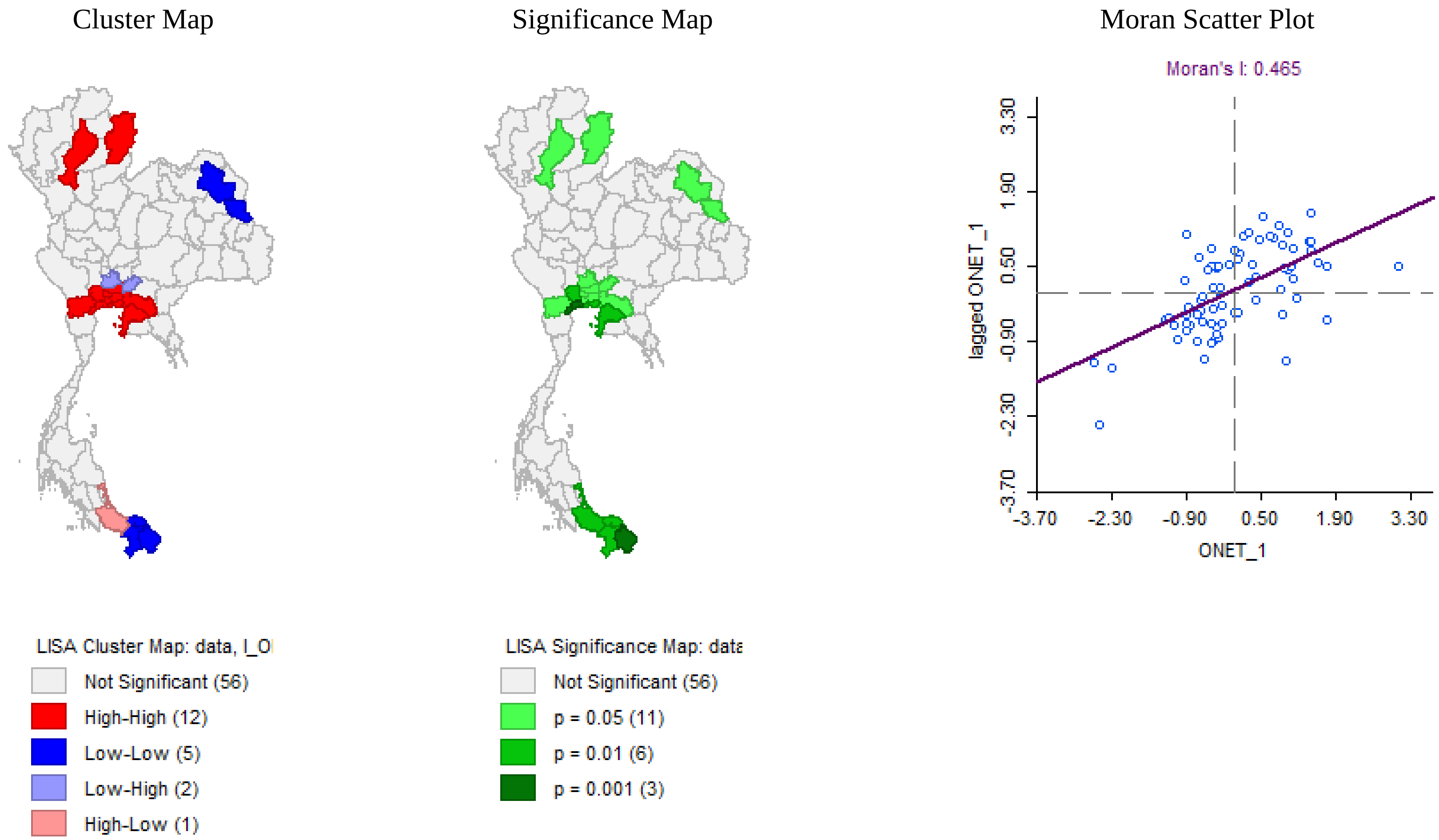


Moran Scatter Plot

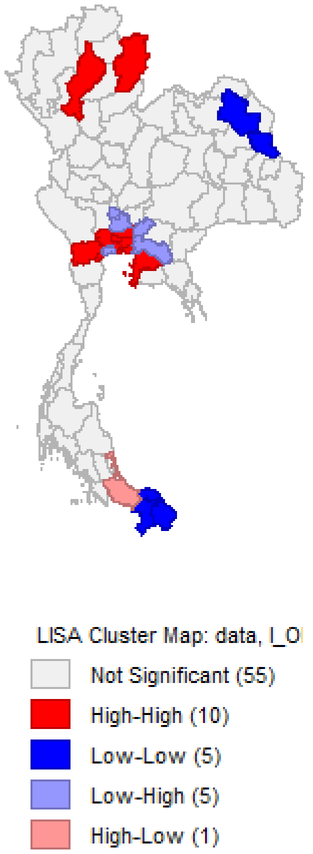


Source: Author’s computation

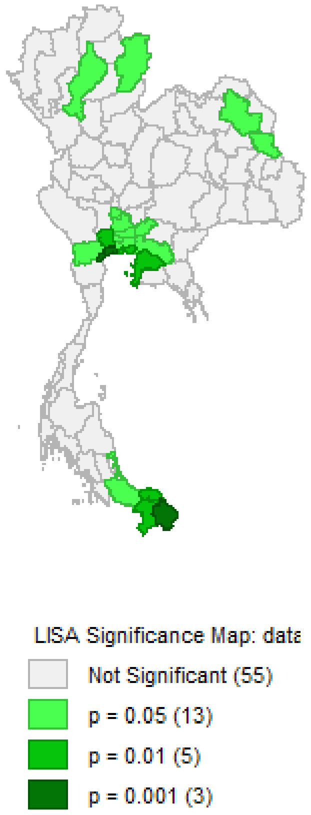
Figure 12
Univariate analysis of O-NET 2016 and 2017



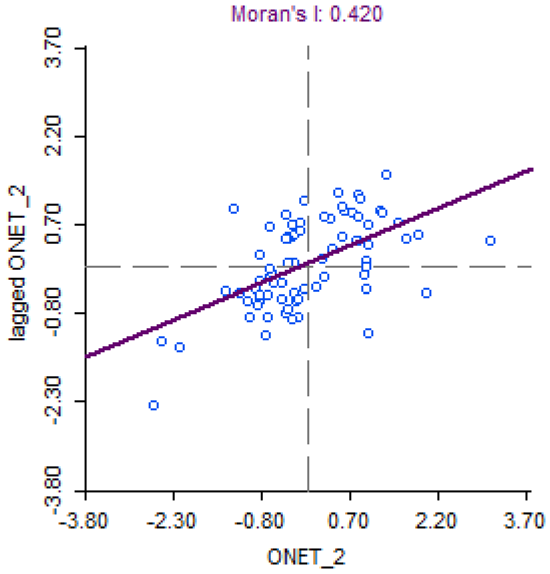
Cluster Map



Significance Map

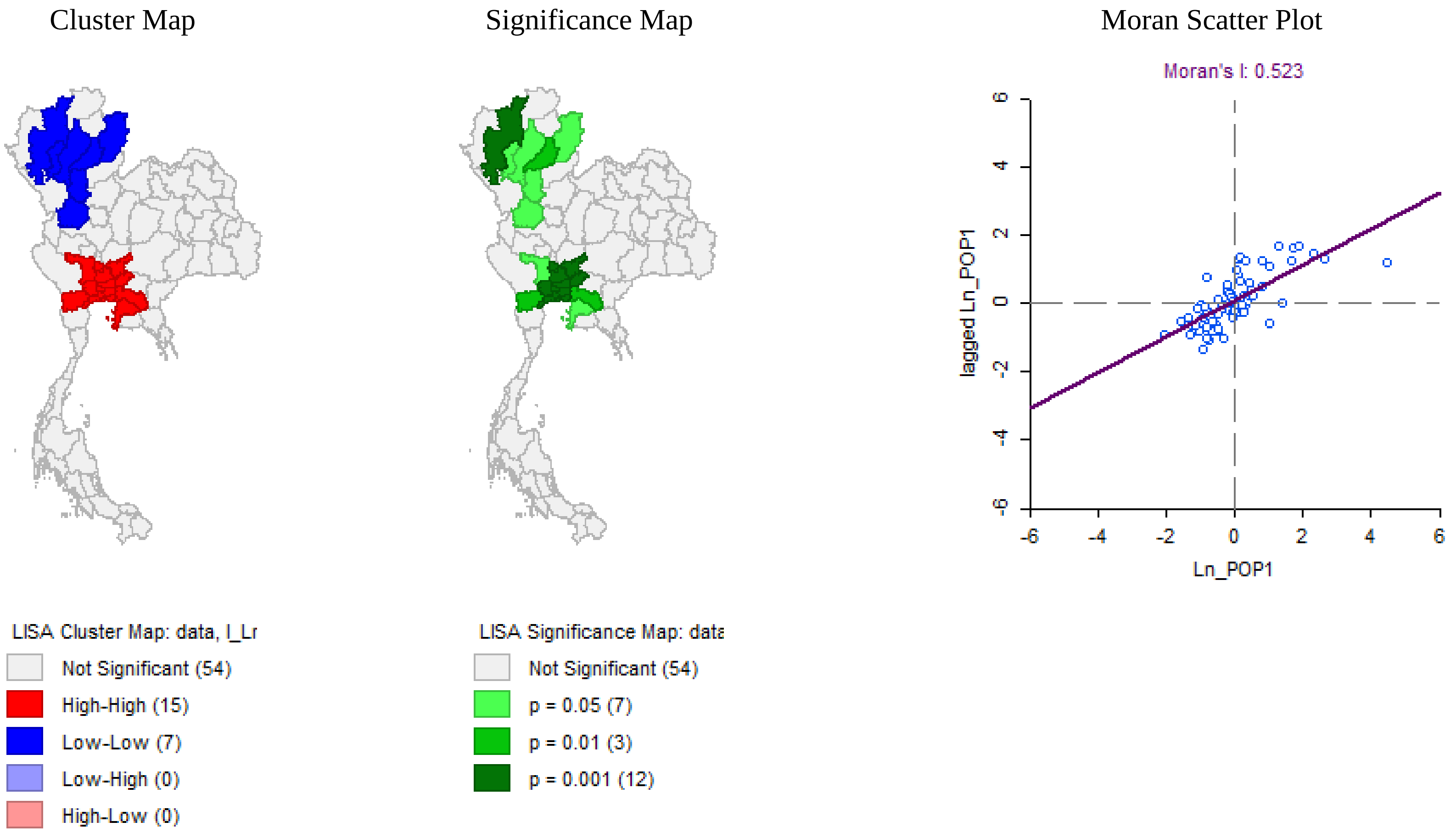


Moran Scatter Plot

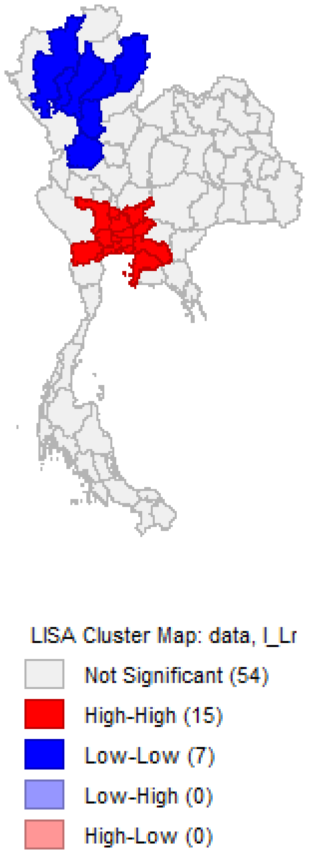


Source: Author's computation

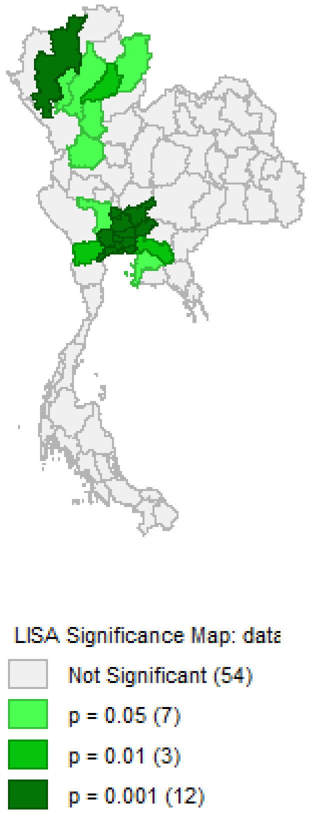
Figure 13
Univariate analysis of Population density 2016 and 2017



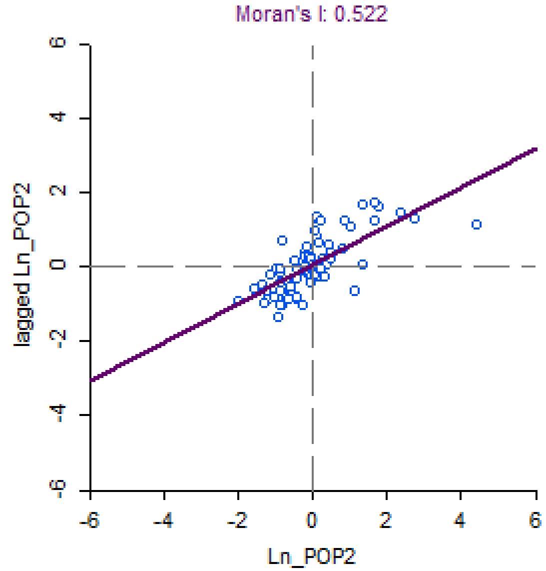
Cluster Map



Significance Map

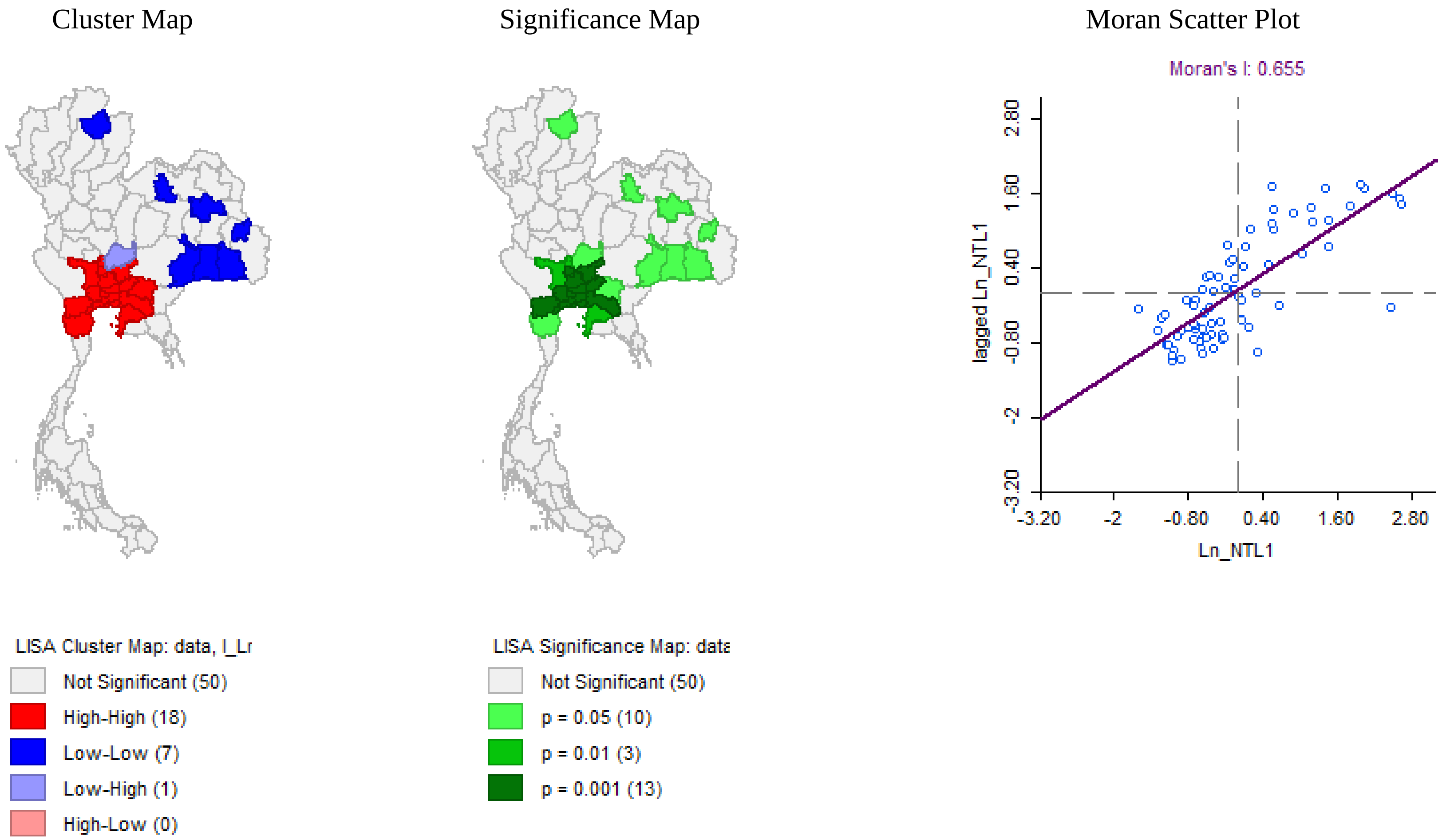


Moran Scatter Plot

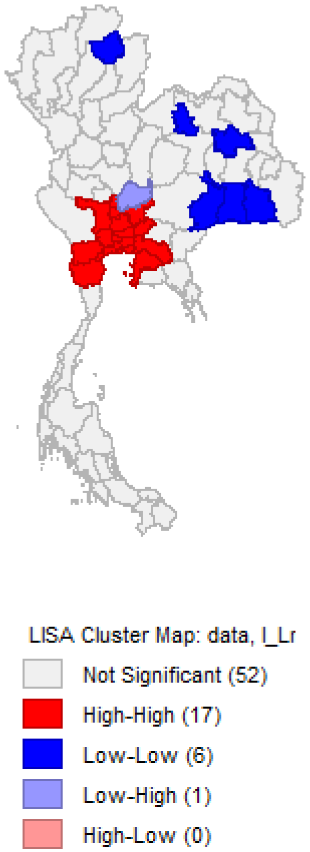


Source: Author's computation

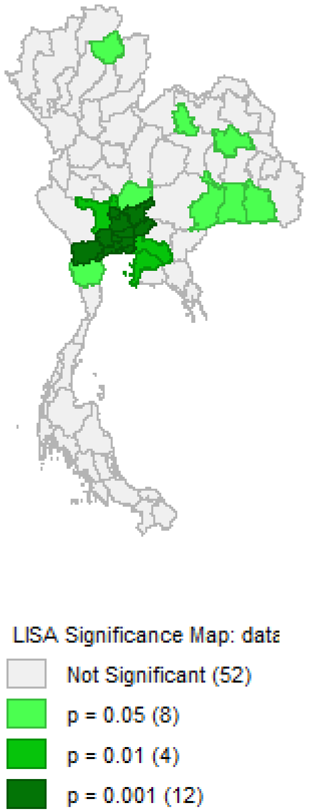
Figure 14
Univariate analysis of Nighttime light density 2016 and 2017



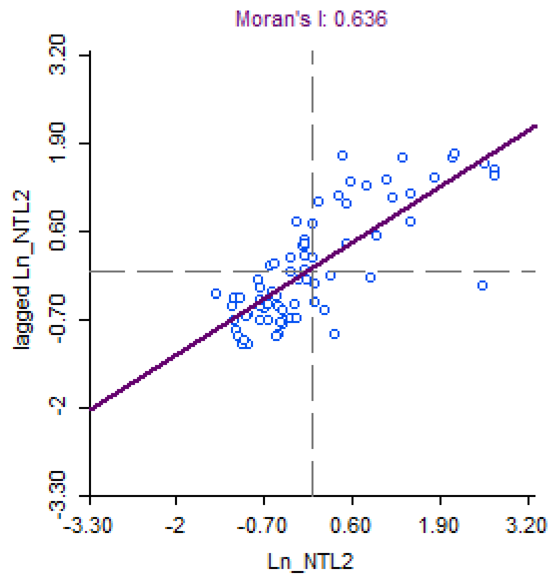
Cluster Map



Significance Map

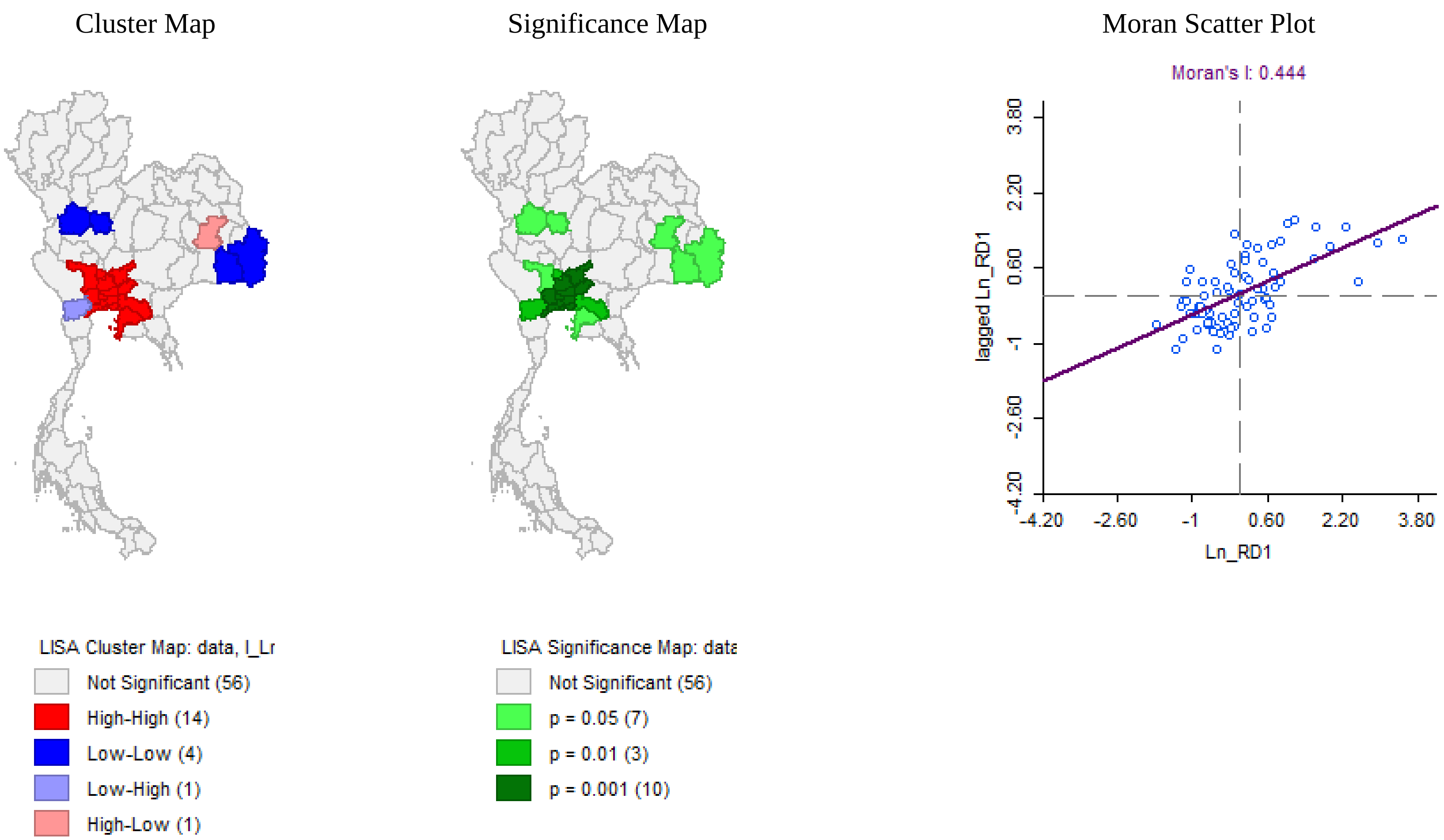


Moran Scatter Plot

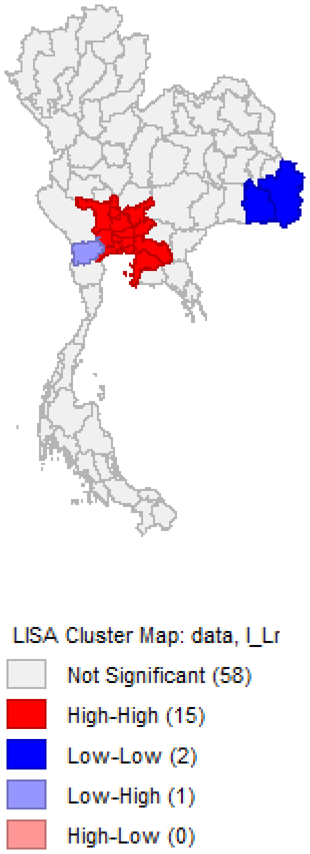


Source: Author's computation

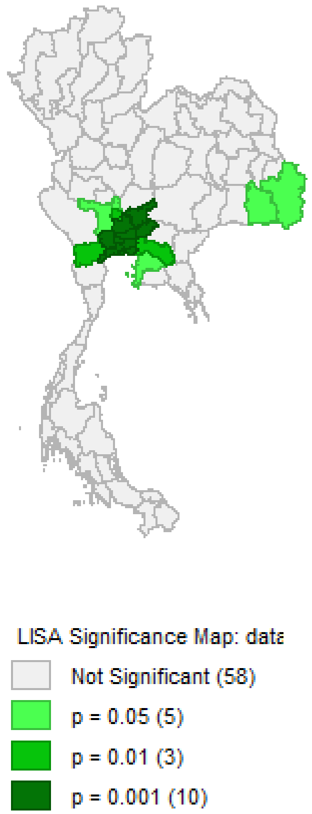
Figure 15
Univariate analysis of Road density 2016 and 2017



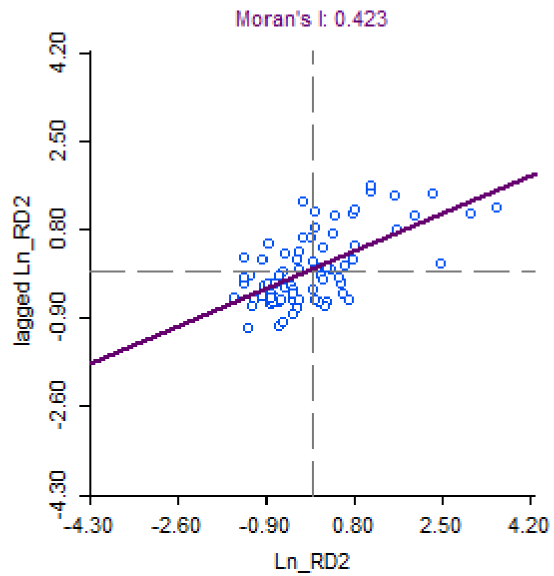
Cluster Map



Significance Map



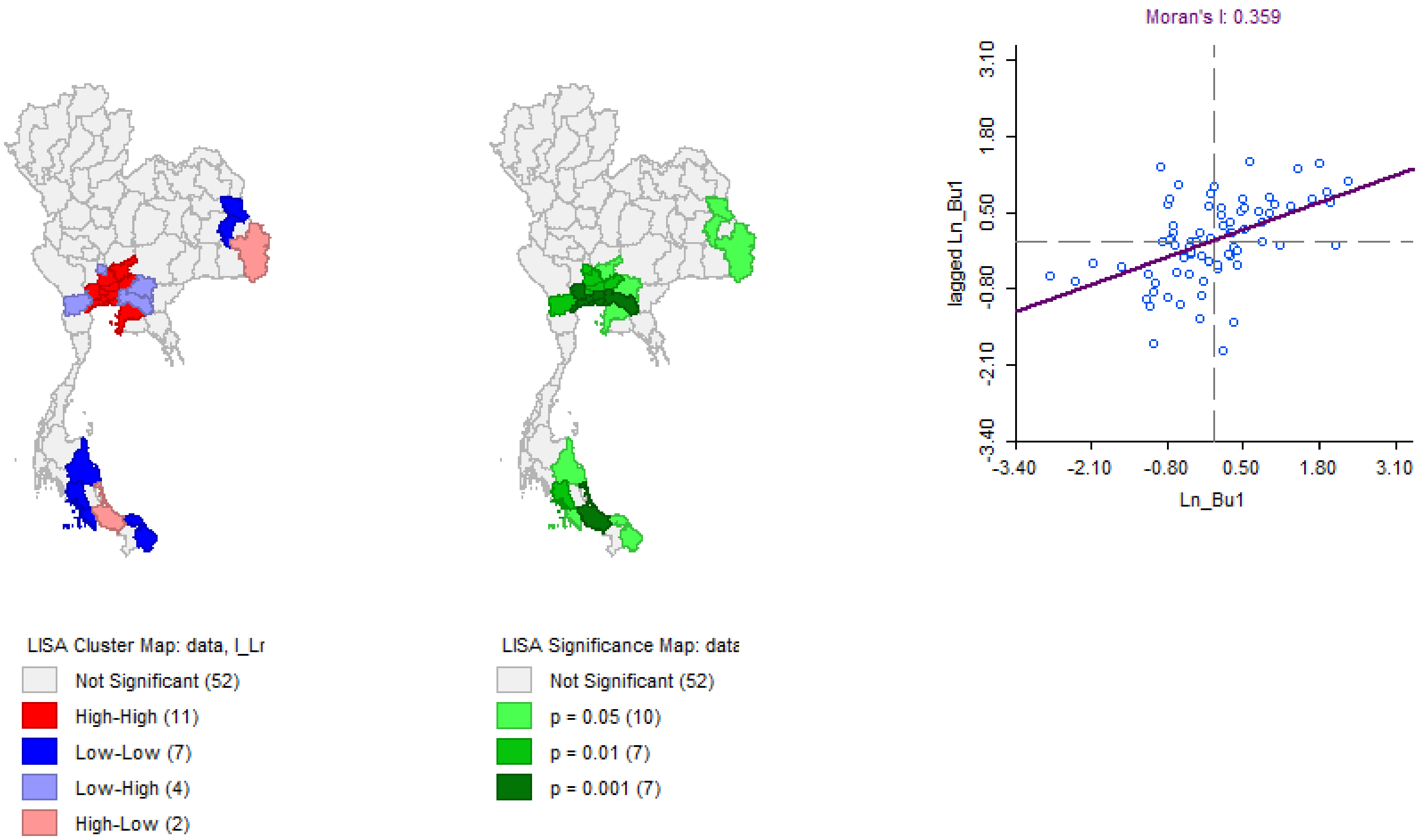
Moran Scatter Plot



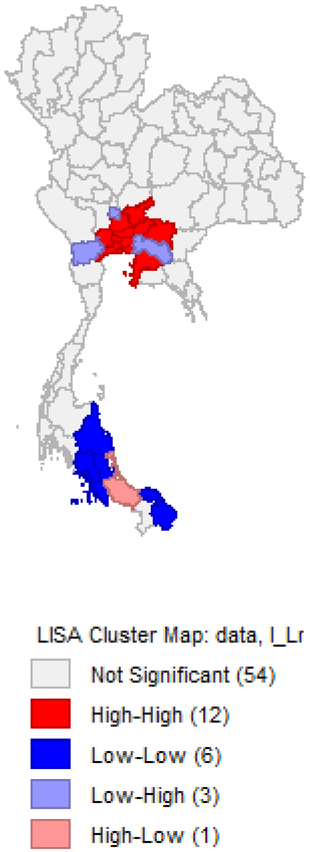
Source: Author's computation

Figure 16
Univariate analysis of Building density 2016 and 2017

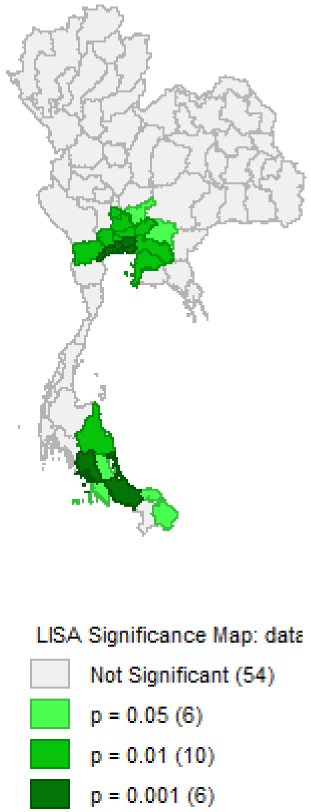
Cluster Map Significance Map Moran Scatter Plot



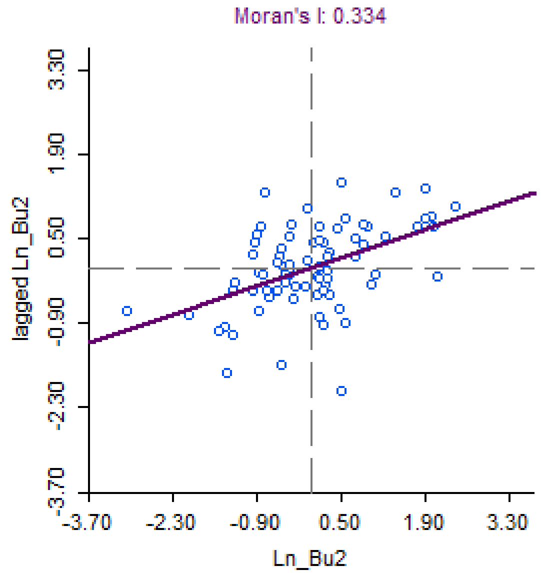
Cluster Map



Significance Map



Moran Scatter Plot



Source: Author’s computation

For NTL density, tt can be seen that there are some provinces in the northeast that the significance level is missing. And the density of the nighttime light is still concentrated in the central region of Thailand. Showing that the economic activity is still concentrated only the central region.

For Road density**,** the results are quite similar to the nighttime light which Road density is concentrated in only the central region, but it is better because the Kamphangphet and Phichit provinces which are not in the LOW-LOW group have shown better communication.

For average wage, Southern Thailand is relatively low-wage, as the World Bank said that the southern region has relatively low agricultural productivity. This point is an important part of thinking about policies that will help increase wages in the South.

For GPP, Provinces with a high GPP are still central to Thailand. Can still see the inequality and inequality of income distribution.

For the proportion of poverty, in the central region is quite low. Compared to other regions of Thailand. The provinces that have a clear proportion of the poor are 3 border provinces such as Pattani, Yala and Narathiwat.

For O-NET score rating has been damaged a lot because concentrated in the central region of the country comes from urban-rural student achievement differentials. It has been argued that such differentials exist because the school in the urban area enjoy more endowments than their rural counterparts; their students, therefore, are able to enjoy more benefits from these endowments.

5.3 Spatial Regression Model

We would like to see the factors that affect the enrollment rate and literacy rate measured through O-NET. The results show as follows.

5.3.1 Enrollment equation

For Ordinary Least Squared as follow :

$$\begin{aligned}\ln Enroll_i &= \beta_0 + \beta_1 \ln WAGE_i + \beta_2 \ln POVERTY_i + \beta_3 \,\mathrm{lnPOOR}_i + \beta_4 \,\mathrm{lnCPI}_i + \beta_5 \,\mathrm{lnGPP}_i \\ &+ \beta_6 ONET_i + \beta_7 \,\mathrm{lnPOP}_i + \beta_8 \,\mathrm{lnNTL}_i + \beta_9 \,\mathrm{lnROAD}_i + \beta_{10} \,\mathrm{lnBuilding}_i + u_i\end{aligned}$$

For Spatial Lagged Model (SLM) as follow :

$$\begin{aligned}\ln Enroll_i &= \rho W \ln Enroll_j + \beta_0 + \beta_1 \ln WAGE_i + \beta_2 \ln POVERTY_i + \beta_3 \,\mathrm{lnPOOR}_i + \beta_4 \,\mathrm{lnCPI}_i + \beta_5 \,\mathrm{lnGPP}_i \\ &+ \beta_6 ONET_i + \beta_7 \,\mathrm{lnPOP}_i + \beta_8 \,\mathrm{lnNTL}_i + \beta_9 \,\mathrm{lnROAD}_i + \beta_{10} \,\mathrm{lnBuilding}_i + u_i\end{aligned}$$

And For Spatial Error Model (SEM) as follow :

$$\begin{aligned}\ln Enroll_i &= \beta_0 + \beta_1 \ln WAGE_i + \beta_2 \ln POVERTY_i + \beta_3 \,\mathrm{lnPOOR}_i + \beta_4 \,\mathrm{lnCPI}_i + \beta_5 \,\mathrm{lnGPP}_i \\ &+ \beta_6 ONET_i + \beta_7 \,\mathrm{lnPOP}_i + \beta_8 \,\mathrm{lnNTL}_i + \beta_9 \,\mathrm{lnROAD}_i + \beta_{10} \,\mathrm{lnBuilding}_i + u_i \quad ; u_i = \lambda W u_j + \varepsilon_i\end{aligned}$$

Table 7
Estimation results of OLS, SLM and SEM

Variable	2016			2017		
	OLS	SLM	SEM	OLS	SLM	SEM
Ln_WAGE	-0.1934	-0.20878*	-0.20217*	-0.19319*	-0.22356**	-0.21447**
s.e.	0.124897	0.11401	0.110946	0.107684	0.096793	0.089742
Ln_POVER	-0.36291	-0.38746	-0.43931*	-0.27733	-0.29243	-0.33101*
s.e.	0.305265	0.278334	0.264149	0.241778	0.216843	0.192804
Ln_POOR	-0.01921	-0.01625	-0.01568	-0.01537	-0.00944	-0.00493
s.e.	0.017107	0.015617	0.015533	0.017524	0.015723	0.014619
Ln_CPI	-0.094	0.026195	0.183733	-0.30742	-0.44694	-0.62906
s.e.	1.76001	1.60401	1.59063	1.15982	1.03941	1.01348
Ln_GPP	0.002927	0.004942	0.009928	0.001064	0.002075	0.006842
s.e.	0.024015	0.021889	0.022172	0.021979	0.019697	0.018476
ONET	0.045477**	0.046507***	0.041313**	0.049416***	0.051993***	0.044337***
s.e.	0.019074	0.017539	0.016379	0.016613	0.015047	0.013274
Ln_POP	-0.05577	-0.06422*	-0.06316**	-0.02805	-0.03598	-0.0372
s.e.	0.036487	0.033727	0.030768	0.034079	0.030735	0.026117
Ln_NTL	0.013927	0.018414	0.020092	0.003282	0.008393	0.014208
s.e.	0.020961	0.01919	0.01856	0.020853	0.018718	0.017033
Ln_RD	0.066195	0.067669*	0.06864*	0.066087	0.072223*	0.071286*
s.e.	0.042313	0.038606	0.037276	0.045838	0.04114	0.038008
Ln_Bu	-0.02216*	-0.02211*	-0.02346**	-0.0234*	-0.02336*	-0.02453**
s.e.	0.012659	0.011537	0.011428	0.013674	0.012252	0.011853
CONSTANT	8.92719	9.39584	8.3024	9.02956	11.0393**	11.1409**
s.e.	8.63287	7.89882	7.74184	5.79925	5.2599	4.91294
W		-0.1589			-0.23237*	
s.e.		0.138696			0.136432	
LAMBDA			-0.17142			-0.31045**
s.e.			0.144779			0.140949
R-squared	0.180797	0.202641	0.201179	0.206149	0.25144	0.269627

*, **, and *** are significant at the levels of 0.1, 0.05, and 0.01 respectively.

Source: Author’s computation

Table 8
Diagnostic tests for spatial dependence

Test	2016			2017		
	MI/DF	VALUE	PROB	MI/DF	VALUE	PROB
Moran's I (error)	-0.0969	-0.4533	0.65035	-0.1692	-1.2362	0.21637
Lagrange Multiplier (lag)	1	1.6061	0.20504	1	3.3177	0.06854
Robust LM (lag)	1	0.9431	0.33149	1	0.1691	0.68091
Lagrange Multiplier (error)	1	1.0736	0.30013	1	3.1648	0.07524
Robust LM (error)	1	0.4106	0.52166	1	0.0162	0.89864
Lagrange Multiplier (SARMA)	2	2.0167	0.36482	2	3.3339	0.18882

Source: Author’s computation

Because Moran's I statistics are not statistically significant. Therefore only using OLS can explain the relationship of variables.

Table 9
Estimation results of OLS

Variable	2016	2017
Ln_WAGE	-0.1934	-0.19319*
s.e.	0.124897	0.107684
Ln_POVER	-0.36291	-0.27733
s.e.	0.305265	0.241778
Ln_POOR	-0.01921	-0.01537
s.e.	0.017107	0.017524
Ln_CPI	-0.094	-0.30742
s.e.	1.76001	1.15982
Ln_GPP	0.002927	0.001064
s.e.	0.024015	0.021979
ONET	0.045477**	0.049416***
s.e.	0.019074	0.016613
Ln_POP	-0.05577	-0.02805
s.e.	0.036487	0.034079
Ln_NTL	0.013927	0.003282
s.e.	0.020961	0.020853
Ln_RD	0.066195	0.066087
s.e.	0.042313	0.045838
Ln_Bu	-0.02216*	-0.0234*
s.e.	0.012659	0.013674
CONSTANT	8.92719	9.02956
s.e.	8.63287	5.79925
R-squared	0.180797	0.206149

*, **, and *** are significant at the levels of 0.1, 0.05, and 0.01 respectively.

Source: Author’s computation

5.3.2 O-NET equation

For Ordinary Least Squared as follow :

$$ONET_i = \beta_0 + \beta_1 \ln WAGE_i + \beta_2 \ln POVERTY_i + \beta_3 \,\text{lnPOOR}_i + \beta_4 \,\text{lnCPI}_i + \beta_5 \,\text{lnGPP}_i + \beta_6 \,\text{lnPOP}_i + \beta_7 \,\text{lnNTL}_i + \beta_8 \,\text{lnROAD}_i + \beta_9 \,\text{lnBuilding}_i + \beta_{10} \ln Enroll_i + u_i$$

For Spatial Lagged Model (SLM) as follow :

$$ONET_i = (\rho W) ONET_j + \beta_0 + \beta_1 \ln WAGE_i + \beta_2 \ln POVERTY_i + \beta_3 \,\text{lnPOOR}_i + \beta_4 \,\text{lnCPI}_i + \beta_5 \,\text{lnGPP}_i + \beta_6 \,\text{lnPOP}_i + \beta_7 \,\text{lnNTL}_i + \beta_8 \,\text{lnROAD}_i + \beta_9 \,\text{lnBuilding}_i + \beta_{10} \ln Enroll_i + u_i$$

And For Spatial Error Model (SEM) as follow :

$$ONET_i = \beta_0 + \beta_1 \ln WAGE_i + \beta_2 \ln POVERTY_i + \beta_3 \ln \mathrm{POOR}_i + \beta_4 \ln \mathrm{CPI}_i + \beta_5 \ln \mathrm{GPP}_i$$
$$+ \beta_6 \ln \mathrm{POP}_i + \beta_7 \ln \mathrm{NTL}_i + \beta_8 \ln \mathrm{ROAD}_i + \beta_9 \ln \mathrm{Building}_i + \beta_{10} \ln Enroll_i + u_i \quad ; u_i = \lambda W u_j + \varepsilon_i$$

Table 10
Estimation results of OLS, SLM and SEM

Variable	2016			2017		
	OLS	SLM	SEM	OLS	SLM	SEM
Ln_WAGE	1.04384	0.807746	1.08704*	1.00739	1.03606*	1.31323**
s.e.	0.788259	0.647457	0.613456	0.773662	0.624393	0.607714
Ln_POVER	4.04492**	3.4603**	4.9849***	2.94118*	2.6334*	3.39673**
s.e.	1.87119	1.54676	1.68014	1.69477	1.37587	1.62473
Ln_POOR	-0.0556	-0.08508	-0.05781	-0.03171	-0.07189	-0.04321
s.e.	0.108273	0.088501	0.078293	0.125138	0.10126	0.104643
Ln_CPI	11.7251	9.18046	5.32258	-6.61992	-3.88912	-1.77641
s.e.	10.9565	8.93615	8.16365	8.19866	6.61501	6.10781
Ln_GPP	-0.00264	0.035346	0.12334	-0.00877	0.020214	0.128828
s.e.	0.150848	0.123038	0.101073	0.15608	0.12599	0.120763
Ln_POP	-0.33114	-0.06677	-0.04401	-0.3263	-0.06535	-0.08789
s.e.	0.229603	0.192153	0.229408	0.239806	0.198533	0.245198
Ln_NTL	0.076508	-0.041	0.027789	0.110879	-0.03334	0.079224
s.e.	0.131754	0.110703	0.106419	0.147455	0.123507	0.126969
Ln_RD	-0.27037	-0.18893	-0.23815	-0.16902	-0.07995	-0.07957
s.e.	0.268669	0.220767	0.209662	0.330156	0.267112	0.259809
Ln_Bu	0.206301***	0.149095**	0.124104**	0.203832**	0.135861**	0.080394
s.e.	0.077194	0.063879	0.056399	0.095958	0.077931	0.072315
Ln_ENROLL	1.79386**	2.02294***	1.46268***	2.4921***	2.62373***	1.82051***
s.e.	0.75237	0.615807	0.500002	0.837799	0.678206	0.559586
CONSTANT	-97.183*	-83.1536*	-78.1741*	-6.89798	-21.6369	-37.4574
s.e.	53.3039	43.5279	40.2995	41.9591	33.8492	33.9144
W		0.402688***			0.426247***	
s.e.		0.095259			0.095422	
LAMBDA			0.625462***			0.610382***
s.e.			0.087159			0.089097
R-squared	0.528673	0.632617	0.697505	0.473449	0.597491	0.648658

*, **, and *** are significant at the levels of 0.1, 0.05, and 0.01 respectively.

Source: Author’s computation

Table 11
Diagnostic tests for spatial dependence

Test	2016			2017		
	MI/DF	VALUE	PROB	MI/DF	VALUE	PROB
Moran's I (error)	0.3675	4.8233	0	0.3291	4.3175	0.00002
Lagrange Multiplier (lag)	1	14.3744	0.00015	1	14.3337	0.00015
Robust LM (lag)	1	1.065	0.30208	1	2.6755	0.1019
Lagrange Multiplier (error)	1	15.4458	0.00008	1	11.9703	0.00054

Test	2016			2017		
	MI/DF	**VALUE**	**PROB**	**MI/DF**	**VALUE**	**PROB**
Robust LM (error)	1	2.1364	0.14384	1	0.3122	0.57636
Lagrange Multiplier (SARMA)	2	16.5108	0.00026	2	14.6458	0.00066

Source: Author’s computation

The test finds that Moran's I values were statistically significant. But when testing the LM test and Robust LM test, it cannot be concluded that between the Spatial Lagged Model (SLM) or the Spatial Error Model (SEM) the model is more suitable.

Table 12
Estimation results of SLM and SEM

Variable	2016		2017	
	SLM	**SEM**	**SLM**	**SEM**
Ln_WAGE	0.807746	1.08704*	1.03606*	1.31323**
s.e.	0.647457	0.613456	0.624393	0.607714
Ln_POVER	3.4603**	4.9849***	2.6334*	3.39673**
s.e.	1.54676	1.68014	1.37587	1.62473
Ln_POOR	-0.08508	-0.05781	-0.07189	-0.04321
s.e.	0.088501	0.078293	0.10126	0.104643
Ln_CPI	9.18046	5.32258	-3.88912	-1.77641
s.e.	8.93615	8.16365	6.61501	6.10781
Ln_GPP	0.035346	0.12334	0.020214	0.128828
s.e.	0.123038	0.101073	0.12599	0.120763
Ln_POP	-0.06677	-0.04401	-0.06535	-0.08789
s.e.	0.192153	0.229408	0.198533	0.245198
Ln_NTL	-0.041	0.027789	-0.03334	0.079224
s.e.	0.110703	0.106419	0.123507	0.126969
Ln_RD	-0.18893	-0.23815	-0.07995	-0.07957
s.e.	0.220767	0.209662	0.267112	0.259809
Ln_Bu	0.149095**	0.124104**	0.135861**	0.080394
s.e.	0.063879	0.056399	0.077931	0.072315
Ln_ENROLL	2.02294***	1.46268***	2.62373***	1.82051***
s.e.	0.615807	0.500002	0.678206	0.559586
CONSTANT	-83.1536*	-78.1741*	-21.6369	-37.4574
s.e.	43.5279	40.2995	33.8492	33.9144
W	0.402688***		0.426247***	
s.e.	0.095259		0.095422	
LAMBDA		0.625462***		0.610382***
s.e.		0.087159		0.089097
R-squared	0.632617	0.697505	0.597491	0.648658

*, **, and *** are significant at the levels of 0.1, 0.05, and 0.01 respectively.

Source: Author’s computation

6. Conclusion

O-NET scores are spatial inequality as can be seen from the LISA, the inequality of the school in the urban area enjoy more endowments than their rural counterparts; their students, therefore, are able to enjoy more benefits from these endowments. And the factors that affect the score increase are average wage rate, the density of the building and enrollment rate. We will see the knowledge spillover of the o-net scores, From the SLM model, it can be seen that the spatial autocorrelation coefficient is positive, which means that if any province has an O-NET score, it will also result in higher scores in the surrounding provinces. Besides, infrastructure It is important to get a high o-net score, as seen from the building density is statistically significant.

For enrollment, it is probably not related to spatial involvement because student enrollment always considers quality schools in order to get a good education. Naturally moved themself to study in different areas. The word quality may be seen from the said school, has a high o-net score or many students have graduated.

Poverty line affect O-NET scores. The direct relationship is because the poverty line can measure the lowest consumption expenditures that consumers must have in order to stay in this province. Which the said expenses include education expenses show that the more the line of higher poverty may have high educational expenses causing to affect the score as well.. Furthermore, indirect results is provinces with high poverty lines exhibits that there is a lot of money that can be consumed and having a good social welfare also results in good socio-economic status of the society and affects the score.

7. Reference

Ahmed, S. (2011). Does economic geography matter for Pakistan? A spatial exploratory analysis of income and education inequalities. *The Pakistan Development Review*, 929-952.

Brasington, D., Flores-Lagunes, A., & Guci, L. (2016). A spatial model of school district open enrollment choice. *Regional Science and Urban Economics*, *56*, 1-18.

De la Fuente, H., Rojas, C., Salado, M., Carrasco, J., & Neutens, T. (2013). Socio-spatial inequality in education facilities in the Concepción metropolitan area (Chile). *Current Urban Studies*, *1*(4), 117-129.

Gordon, I., & Monastiriotis, V. (2007). Education, location, education: a spatial analysis of English secondary school public examination results. *Urban studies*, *44*(7), 1203-1228.

Karahasan, B. C., & Uyar, E. (2009). Spatial distribution of education and regional inequalities in Turkey.

Lounkaew, K. (2013). Explaining urban–rural differences in educational achievement in Thailand: Evidence from PISA literacy data. *Economics of Education Review*, *37*, 213-225.

Senadza, B. (2012). Education inequality in Ghana: gender and spatial dimensions. *Journal of Economic Studies*.

Zhang, X., & Kanbur, R. (2009). Spatial inequality in education and health care in China. In *Regional Inequality in China* (pp. 92-110). Routledge.