Working Paper
in Economics and
Development Studies

Department of Economics
Padjadjaran University

No. 201001

The Spatial Dimension of Human Development Index in Indonesia

Rullan Rinaldi
Eva Nurwita

Department of Economics,
Padjadjaran University

June, 2010

Center for Economics and Development Studies,
Department of Economics, Padjadjaran University
Jalan Cimandiri no. 6, Bandung, Indonesia.
Phone/Fax: +62-22-4204510
http://www.lp3e-unpad.org

The Spatial Dimension of Human Development Index in Indonesia

Rullan Rinaldi and Eva Nurwita

*Center for Economics Development Studies, Padjadjaran University**

Abstract

As the new paradigm of economic development pioneered by UNDP and Mahbub Ul-Haq undertaken, development processes no longer viewed as mono-dimensional process of economic growth indicated by GDP growth solely. Human Development Index on the other side offer an indicator that takes into account other aspecta as proxies of life quality such as life expectancy and literacy rate wrapped as a composite index. Several previous researches has try to explain the determinant of HDI, but as HDI was start to calculated at sub national level, the complexity of the task to explain the determinants was escalating due the fact that sub national data has geographical information attached in it.

This paper tries to explain the spatial pattern on HDI achievement at sub national level in Indonesia, and estimate the determinants of HDI using spatial econometrics method. The use of the tools based on the necessity to put into account spatial dependence as special form of cross-sectional serial correlation, which is a common situation in observations that has geographical information.

Keywords : Human Development Index, Spatial Econometrics,
Sub National Data

JEL Classification : O15, R58, R11

Acknowledgement

This is unrevised version of paper prepared for SRG-10

*Center for Economic Development Studies (CEDS), Faculty of Economics, Padjadjaran University. Jalan Cimandiri No.6, Bandung, 40115. Telp: +62224204510. Contact email: (rullan.rinaldi@fe.unpad.ac.id);(eva.nurwita@gmail.com)

1 Introduction

As a brainchild of Mahbub Ul Haq (1990), Human Development Index brings new paradigm on economic development measurement. Economic development on this paradigm was not solely indicated by income per capita as its measurement, but also takes into account life quality of the society as a whole.

During times, as a composite index, HDI was not a non-revisable concept, it open itself for criticism and revision. Compared with earlier version from UNDP (1990), contemporary version of HDI underwent several changes, especially on stipulating threshold from income level with a purpose to create diminishing return effect from enhancing income towards HDI (Lüchter and Menkhoff, 1996).

HDI as a simple measure yet multidimensional, has been criticized whether on its eligibility as statistical measurement (Srinivasan,1994; Fukuda,2003), or of its simplistic proxies of life quality, where (Sagar and Najam, 1997) came with the idea to include the environment aspect and environment depreciation as component of HDI. Critics also came from Noorbakhsh (1998) who proposed the possibility of a new reformulation that count the utility of life standard using Atkinson Formulation.

Despite the controversy about eligibility of HDI measurement, several empirical research has underwent previously to observe the relationship between HDI and economic growth using cross country data (Ranis et al., 2000).

In further development, HDI estimation at sub national level (Ivanov, 2005), has reveal interesting facts. Sub national estimation brings empirical facts that deviation has occurred in HDI achievement within sub national level from national number. For Indonesian case, figure 1 represent the differences of HDI achievement at district (*kabupaten/kota*) level, where darker polygon indicate higher HDI achievement compared with the brighter one.

From Figure 1 below, we found out that there is disparity in HDI achievement between prominently between western part of Indonesia (Kawasan Barat Indonesia/KBI) and eastern part of Indonesia (Kawasan Timur Indonesia/KTI), where western part tends to have higher achievement than eastern does. Spatial pattern of HDI achievement in this study recognized as an outcome from the disparity in sub national development process.

This research tries to explain the spatial aspect of HDI achievement using spatial univariate analysis and also its determinants using spatial econometrics method. Spatial analysis methods became necessity in this case, due to the fact that cross sectional data where the observation unit has a geographical information attached into it due to the existence of spatial autocorrelation. Analogously to serial correlation in time series data, spatial autocorrelation as a special form of cross sectional serial correlation.

Further more, Kwan et al. (2003) emphasize on the importance of regional scale and spatial effect in special case named spatial dependence, as a factor that has to be observe when analyzing regional level.

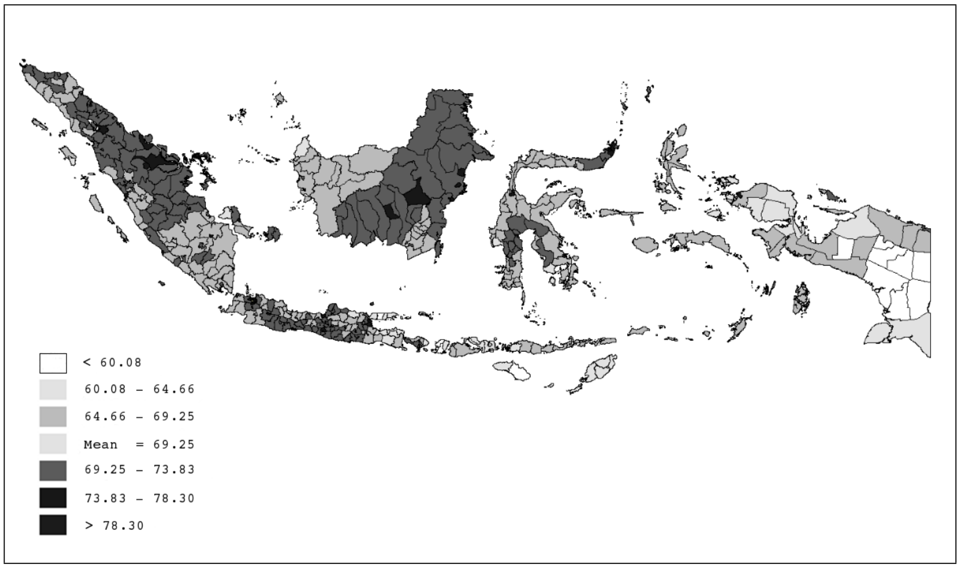

Figure 1: HDI Achievement at District/City Level in Indonesia for 2006

2 Research Questions

- Spatial Autocorrelation Identification
  In regression analysis, exclusion of spatial aspect when the data has geographical information attached into it could lead to bias and inefficient estimation result due to omission of spatial autocorrelation from the model (LeSage, 1999).
  In this paper we are testing the existence of spatial autocorrelation Indonesia HDI achievement at district level.

- Geographical Factors Effect
  Without disregarding the possibility of redudancy problem with the inclusion of geographical factor in spatial regression, we tries to estimate the relationship between geographical variables such as Islands, Western Indonesia, and Urban Agglomeration (further description of the variables can be found in Section 4 and HDI achievement. The fact that the variables are qualitative, the redudancy of geographical aspect and distance (represented by spatial weight matrix) could be avoided.
  In this paper we are interested in testing the effect of governance fators toward HDI, we expect that Islands negatively affect HDI achievement, because if a district situated in Islands it will be more remote that district not situated in Island. Western part of Indonesia expected to be perform better in HDI achievement compared to others, and districts in Urban Agglomeration expected to have highet HDI achievement due to its extensive infrastructure and regional integration.

- Governance Factors Effect
  The inclusion of governance factor in HDI determinants, it to examine the relationship

between governance and HDI achievement. Specifically the test was conduct on span of control with the expectation that shorter span of control can promote higher HDI achievement due to ease of public services and facility access by peoples in each districts. The role of political party also being questioned here, we want to test whether highest share of majority party affect HDI or not. Policy aspect was represented by spatial plan, with the expectation that districts having spatial plan will better maintain its regional development and could benefit them higher HDI achievement.

- Endowment Factor Effect
  Endowment factor was exogenous variable, the nature was given to specific districts. We include this factor with the expectation that districts blessed with mining and forestry sector will have the given benefit for their economy, enables them to better promote HDI achievement.

- Input Factor Effect
  Inputs factor are represented by poverty relative to province and education enrollment rate. We expect that these variables have direct impact to HDI achievement, poverty expected affect HDI negatively because due to negative relationship between poverty and life quality (health and education). While education enrollment expected to affect HDI positively due to positif relationship between education and specific component of HDI (literacy rate).

3 Policy Relevance

The outcome from this paper, could become relevance input in policy formulation regarding HDI achievement in national and sub national level, to the fact that HDI in national level are aggregation of HDI achievement in districts level. Spatial autocorrelation identification, could shares the benefit in identifying spillover effect of HDI achievement throughout district, univariate analysis (i.e. LISA) could specifically informed us more localized information about the cluster pattern of "*hot-spot*"-s of HDI.

Spatial regression could inform us which determinants affect and how they affect HDI in district level. More specifically, the Geographically Weighted Regression enable us to observe spatial heterogeneity of each determinant in affecting HDI achievement in each district, lead us to have specific information which determinants affects HDI the most and the least throughout districts, enables policymaker to have a more specific HDI improvement agenda.

4 Data

We collected secondary data at district level, mostly from Indonesian Statistical Bureau for 2006, the observation consist of 440 districts. We classified the independent variables into 6 categories of HDI determinant; Geographic Factor, Governance Factor, Scale Effect, Infrastructure, Endowment Factor and Input Factor.

Geographical Factor consists of 3 qualitative variables; Islands, Western Indonesia, and Urban Agglomeration. Islands was use to represent districts which one third of its polygon shape are consist of separate polygon, in broader sense, consist of islands (for example are districts in Lesser Sunda Islands). Western Indonesia variable in this study presage districts that included into western Indonesia area, this variable used as proximity to the priory hypothesis that western Indonesia has higher economic development than other region. Formally, the value of the variable is 1 for districts regarded as part of KBI (Kawasan Barat Indonesia) and 0 otherwise. The definition of KBI refer to Indonesian Outline of State Policy (GBHN) year 1993, that islands included into KBI are Java, Sumatera, Borneo and Bali Islands.

Urban Agglomeration can be define as groups of municipalities in certain region that shares neighboring location and economic growth similarities mostly driven by innovation activity and industrial production (Martin and Ottaviano : 1996). As determinants of HDI achievement, in this paper our model take into account urban agglomeration which refer to Indonesian General Construction Department definition, Governor Decree, RTRWN, RTRWP and Government Regulation No.26 in 2008 where the agglomerate cities/districts classified (List of District can be seen on appendix). The value of the variable is "1" if regions classified as part of agglomeration and "0" if otherwise.

Spatial Plan is a qualitative dummy variable; formally having value "1" for district that has legalized RTRWS and 0 otherwise. We expect that districts with legalized spatial plan (RTRW-Rancangan Tata Ruang dan Wilayah) could better manage their regional development goals, that could trickle down to the HDI achievement.

In altering HDI, local government has the responsibility to bring public service regarding human development aspects as close as possible to the people. As proxy of governance factor, we use span of control as a latent variable derived from factor analysis of district size, topographical classification (elevation/height above sea level), number of village, and number of sub district (that can be seen in Table 1). The variable itself intent to foresee how close is the district government to its people, we perceived that sub district and village level governance was the frontier of public service, district size and topographical classification has its part to control remoteness in each districts. In order to control scale effect, we use population relative to province as dependent variable in the model.

As variable representing Infrastructure factor, we choose Sea Port and Trans Highway as regressor of HDI achievement, both of them are qualitative variables. Formally, seaport variable's value equal "1" if the districts has a seaport and "0" if otherwise.

Mining and Forestry Share employed as representation of initial endowment of each

Table 1: Factor Analysis of Span of Control

	Eigenvalue	Variable	Loading Factor	Coefficient
Factor1	1.38285	Number of Sub District	0.815	0.45394
Factor2	0.09791	Number of Village	0.8188	0.46601
Factor3	-0.08696	District Size	0.1333	0.02739
Factor4	-0.18384	Topography (height above sea level)	0.1747	0.03879

districts. On input factors, we have 3 variable; Primary School Enrollment Rate, Secondary School enrollment rate and Poverty Rate (Head Count Index). These three variables use to asses quality of human capital itself, the assumption of three variables above are the higher the enrollment rate could lead the higher the HDI achievement of each district, conversely the higher the poverty rate could lead to lower HDI achievement. As it takes time for inputs to affect HDI, for school enrollment rate, we take the value from 2004 enrollment rate (t-2) instead of 2006. Comprehensive information, desctiptive statistics, and sources of data are provided in the appendix.

5 Spatial Analysis and Regression

5.1 Spatial Weight Matrix

To test the existence of spatial dependence/spatial autocorrelation [1], we have to built a weight matrix representing spatial location of the observations. Following Tobbler's First Law of Geography (1970), we employ spatial weighting matrix using row standardized continues distance decay function, whereas the distance measurements were taken from the district's centroid points. The spatial weight that we use ($\boldsymbol{W}$) is based on kilometer-converted Euclidean Distance (d_{ij}) between districts (i and j) on the sphere :

$$d_{ij} = \arccos[(sin\varphi_i sin\varphi_j) + (cos\varphi_i cos\varphi_j cos|\delta\gamma|)] \tag{1}$$

Where i and j are the centroid's latitude of district i and j, respectively $|\delta\gamma|$ denotes the absolute value of the difference in longitude and latitude between i and j. The distance obtain from Equation 2, than substituted into a distance-decay function :

$$w_{ij} = (d_i j)^{-1} \tag{2}$$

The spatial weights matrix $\boldsymbol{W}$ above is a $\boldsymbol{N}$ by $\boldsymbol{N}$ non-negative matrix, which expresses for each region (row) those regions (columns) that belong to its neighborhood set as nonzero elements. By convention, the diagonal elements of weight matrix are set to zero, since no

[1]The term "Spatial Dependence" and "Spatial Autocorrelation" can be used interchangeably, in the rest of the paper we will use the term "Spatial Autocorrelation"

regions can be viewed as its own neighbor. For the ease of interpretation, it is common practice to normalize $\boldsymbol{W}$ such that the elements of each row sum to one. Since W is nonnegative, this ensures that all weights can be interpreted as an averaging of neighboring values.

As prelimenary evidence of the nexus between HDI achievement and geographical location, Figure 2(a) displays every districts HDI achievement plotted against distance to the districts with highest HDI achievement, and to districts with lowest HDI achievement (Figure 2(b)), the figure shows negative relationship between distance to district with lowest HDI (Teluk Womdana) and HDI achievement of each districts, and shows positive relationship between distance of each districts to the district with highest HDI achievement(South Jakarta).

As a formal test for spatial autocorrelation, we calculate Moran's I and Geary's c statistics of HDI ini districts level using spatial weight matrix derived from distance decay function in Equation 2. The null hypothesis of the test for spatial autocorrelation is that there is no spatial autocorrelation, we compare both Moran's I and Geary's c statistics with their expected value $E(I)$ and $E(c)$. If Moran's I are greater than its expected value and Geary's c statistics are smaller than their expected value, than we can conclude that positive spatial autocorrelation existed in HDI achievement at district level in Indonesia, and vice vers. The statistical inference is computed on the basis if z-statistics. We can see the result of the test in Tabel 2 below :

Table 2: Moran's I and Geary's c for Human Development Index

Moran's I	E(I)	SD(I)	z-stat	
0.162	-0.002	0.007	24.685	***

Geary's c	E(c)	SD(c)	z-stat	
0.778	1.000	0.015	-14.719	***

* 1-tail test

As the outcome we have found evidence of the existence of positive spatial autocorrelation in HDI achievement at district level in Indonesia. In other words, HDI does spill across districts border, and districts with high HDI achievement will have their neighbors also shares the simmilarities.

5.2 LISA : Local Indicator of Spatial Association

The idea of Local Indicator of Spatial Association (LISA) is to identify every single observations coefficient of local indicator of spatial autocorrelation. Local Indicator of Spatial Association (LISA) is statistics that satisfies the following two requirements: (1) The LISA for each observation gives an indication of the extent of significant spatial clustering of similar values around that observation. (2) The Sum of LISAs for all observations is proportional to a global indicator of spatial autocorrelation (Anselin, 1995). Local spatial cluster sometimes referred to as a hot-spot, may be identified as those locations or sets of contiguous locations

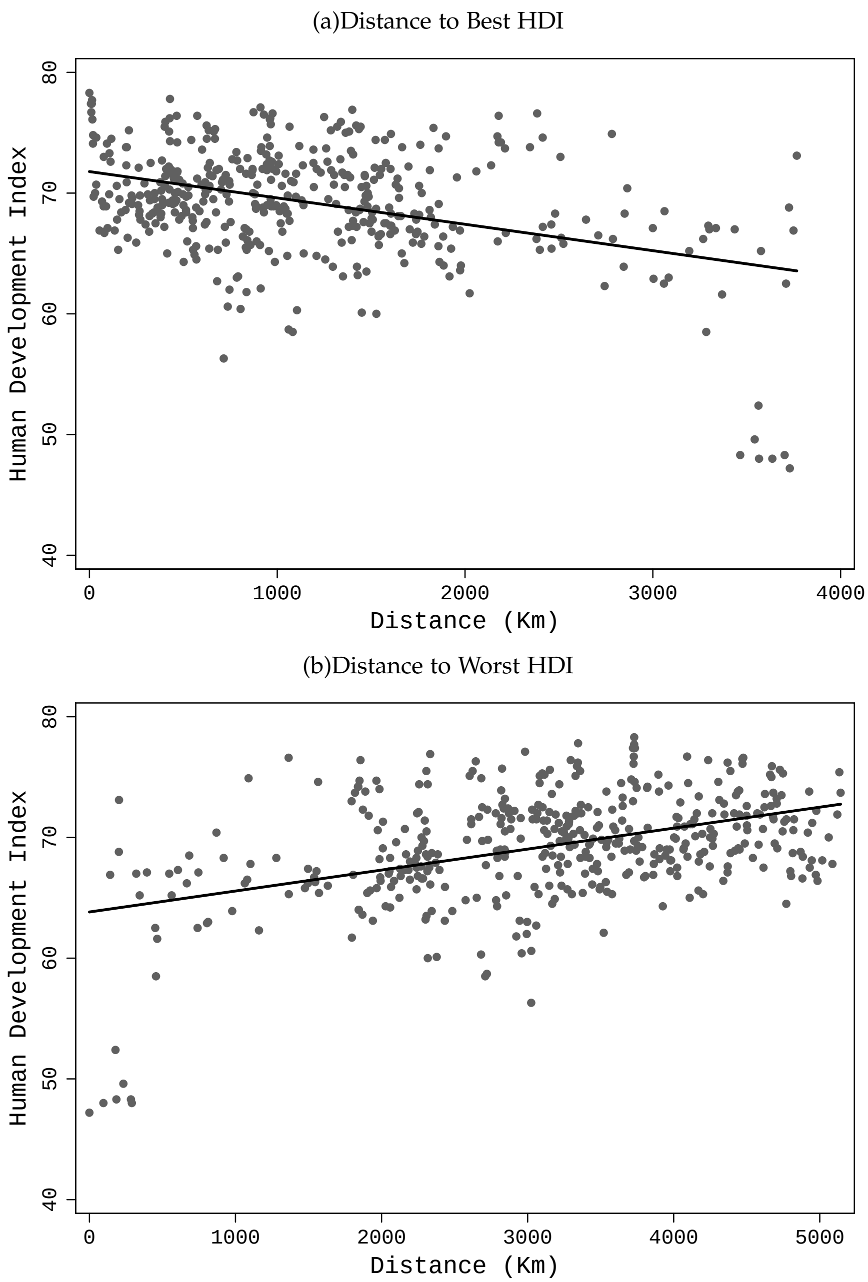

Figure 2: Distance to Best(Worst) HDI

for which the LISA is significant. Similar to the rationale behind the significance test for G_i and G_j statistics of Getis and Ord (1992), the general LISA can be used as the basis for a test on the null hypothesis of no local spatial autocorrelation.

The basic methodology of LISA calculation, use local Moran's I decomposes global spatial pattern and indicates to which extent the geographic region is surrounded by similar or dissimilar value, forming the geographical pattern. This implies that some structure is present in the data, which can be considered as additional information. At the second step we calculate the Moran in order to get more precise result, because LISA cluster maps provide information on statistically significant cluster and outliers compared to global mean, but they may very well show spatial dependence when compare to a local mean (Anselin, 1995).

Local Moran's I is can be calculated using the equation below :

$$I_i = \frac{Z_i \sum W_{ij} Z_i}{\sum Z_i^2 / n} = \frac{(x_i - \mu_x) \sum_i W_{ij}(x_i - \mu_x)}{\sum i (x_i - \mu_x)^2 / n)} \tag{3}$$

Where $(x_i - \mu_x)$ is deviation of region i value from the mean, $\sum_i W_{ij}$ is deviation from neighboring area j values from the mean and $\sum(x_i - \mu_x)^2/n$ is average area squared deviations from the mean. The outcome of LISA calculation allow us to determine whether the formation of certain variable in a certain location has High-High, High-Low, Low-High or Low-Low relationship with the formation of same variable in the surround location, clustered in four quadrants named quadrant 1 for high-high group of region, region that has a highest coefficient of LISA also surrounded by other highest region (Positive among positive). Quadrant 2 for High-Low group of regions, region that has high coefficient of LISA but surrounded by low coefficient of LISA (Positive among negative). Quadrant 3 for Low-Low group area of region that has low coefficient of LISA surrounded by low coefficient of LISA also (negative among negative) and last for quadrant 4 that indicates Low-High area of region group that has low coefficient of LISA but surrounded by high coefficient of LISA (negative among positive).

In summarize for this HDI case, Local indicator Spatial Autocorrelation use to identify regions specific spatial autocorrelation coefficient (whether it is in a negative or positive form). From an overall applicability perspective, LISAs detect significant spatial clustering around individual location and pinpoint region that contribute most to an overall pattern of spatial dependence of HDI's achievement in Indonesia. Particular attention will be given to the positive (negative) spatial autocorrelation, significant LISAs located near other positive (negative) significant LISAs (Hot Spot) (Khomiakova, 2008).

This subsection study first focus on a comparison of the identification of local spatial clusters provided by Moran Scatter Plot classification. Using the same row standardized matrix as for the global measure, the result for indicators of LISA reported on the appendix D show districts list based on the cluster of the HDI's achievement, resulted from local Moran's I identification and Scatter Plot in Figure 3.

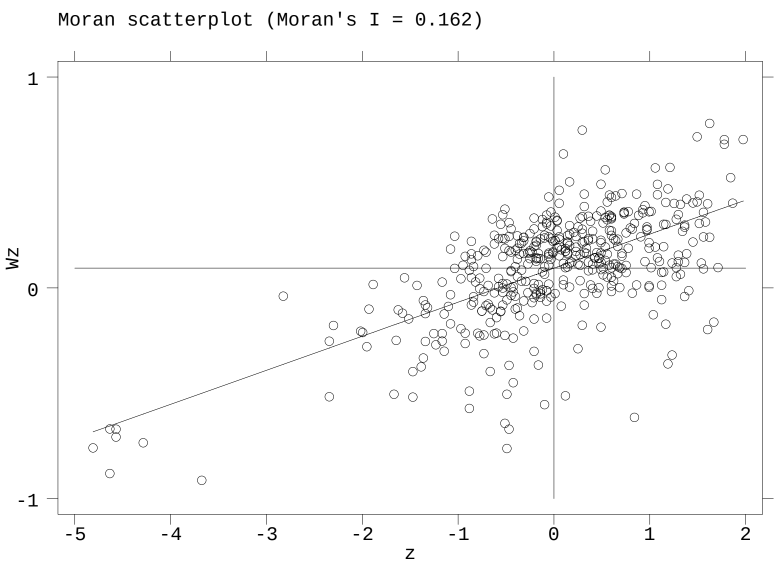

Figure 3: LISA Scatter Plot

Cluster maps allows us to detect the possible geographical patterns, Moran *I* and Geary's *c* test for the presence of "global" spatial autocorrelation, but do not individuate local cluster. Hence, to identify the local clusters we use the local indicators of spatial association (Cracolici and Uberti, 2009). From 440 observations there is 100 regions recognized belong to quadrant 1 on Moran's *I* scatter plot, in the map, the polygons indicated by red area, also known as hot spots. Most of northern Sumatera, West Java, Central Java, Banten, Central Kalimantan, East Kalimantan and North Sulawesi detected as hot spots. Pointed on west Java Case, where Jakarta and regions surround also colored by red declare that the result of LISA mapping also support the fact that Jakarta as capital city of Indonesia known as region with highest HDI achievement.

In comparison for quadrant 1, we recognize quadrant 3 as Low-Low (negative spatial autocorrelation amongst negatives) group of regions. At least there are 65 regions belongs here, we can identified the groups by polygons indicated with blue area in the Moran's *I* cluster map. The groups lies on East Java and lesser sunda islands (West Nusa Tenggara, East Nusa Tenggara), West Kalimantan, South Sulawesi, Southeast Sulawesi, Maluku and Papua (list of districts are available in appendix D). From the map, the dark blue area cover most of eastern Indonesia, apart from other variables explanation about this circumstances it is also the fact that most of eastern Indonesia Region has the smallest number of HDI Achievement. All of the illustration above, clearly clarify that the farther a districts from the node of the highest HDI achiever, the lower they are. For quadrant 2 and quadrant 4, most of local indicator coefficients are not statistically significant (see Signifance Map in Figure 5), thats why there are only few polygons indicated belongs to quadrant 2 and quadrant 4.

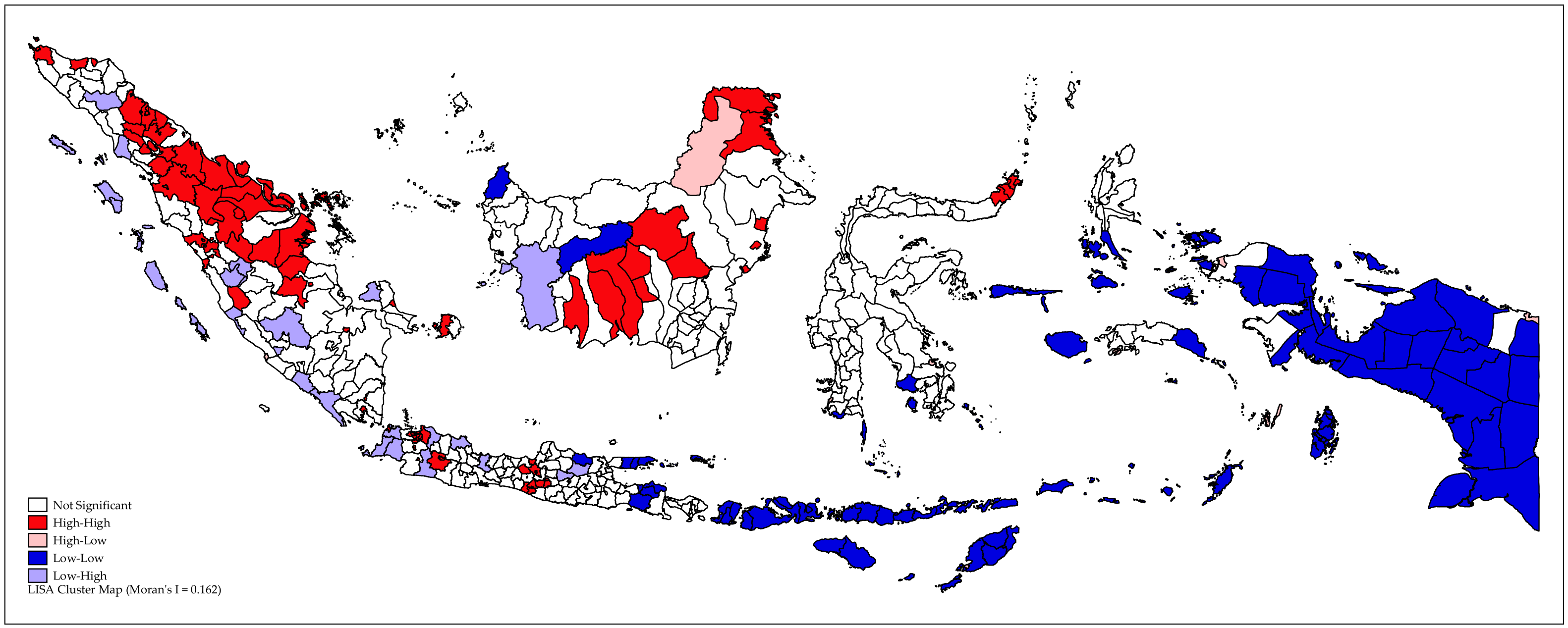

Figure 4: LISA Cluster Map

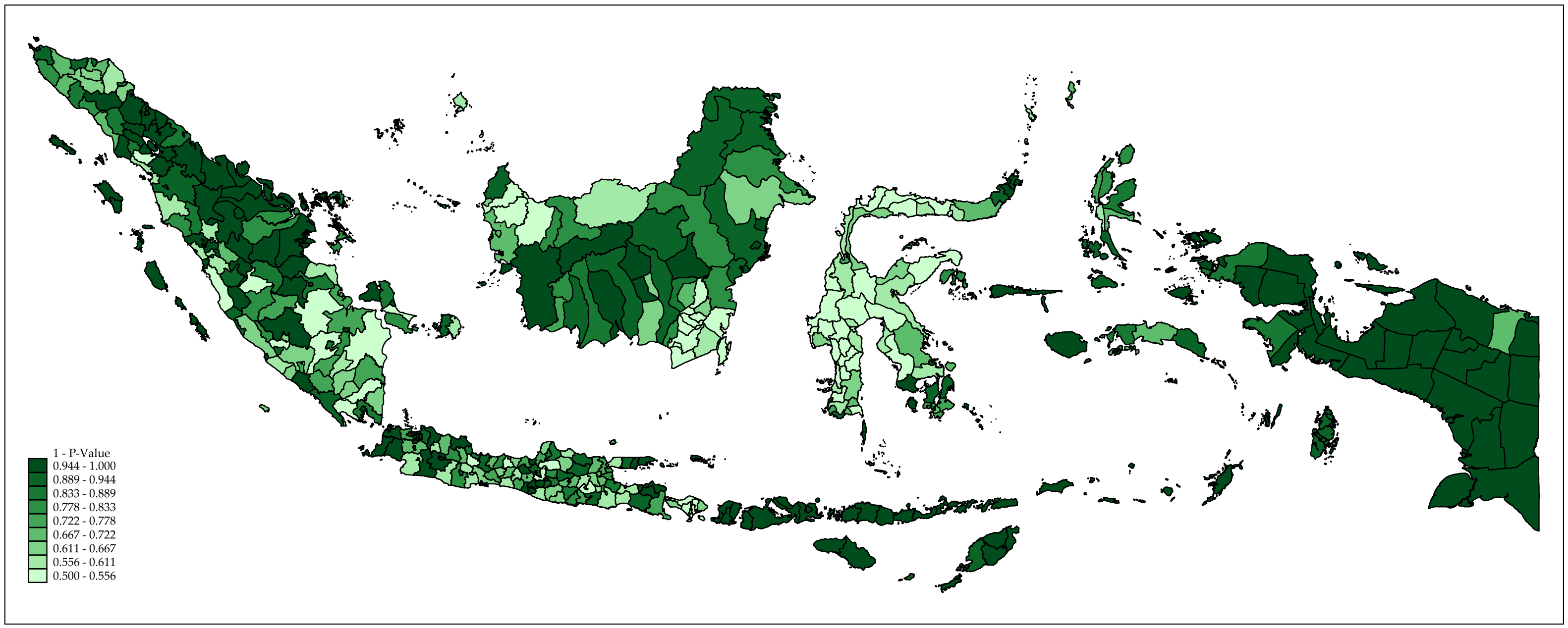

Figure 5: LISA Significance Plot

5.3 Spatial Regression

There is two kinds of spatial regression model that usually use on spatial analysis. The first is Spatial Auto Regressive (SAR) and the second is Spatial Error Model (SEM) respectively :

5.3.1 Spatial Auto Regressive Model

This model focuses on spillover effect existence. This model tries to estimate whether "x" variable in a region affect or affected by "x" variable in the neighbor region. The model noted as :

$$y = \rho W_y + X_{ij} + \epsilon_{ij} \tag{4}$$

Where ρ is a spatial autoregressive coefficient, and ϵ is a vector of error terms and the other notation are the independent and dependent variables. Unlike what holds for time series counterpart of this model, the spatial lag term W_y is correlated with the disturbances, even when the later are i.i.d. This can be seen on a reduce form :

$$y = (I - \rho W_y)^{-1} X\beta + ((I - \rho W)^{-1}\epsilon \tag{5}$$

In which inverse can be expanded into an infinite series including both the explanatory variables and the error terms at all location (the spatial multiplier). Consequently, the spatial lag term must be treated as an endogenous variable and proper estimation methods must account for this endogeneity (OLS will be biased and inconsistent due to the simultaneity bias) (Anselin, 1980).

5.3.2 Spatial Error Model

Spatial Error model estimates correction model with large error. A spatial model error is a special case of regression with non-spherical error term, in which the off-diagonal elements of the covariance matrix express the structure of spatial dependence. Consequently, OLS remain unbiased, but it is no longer efficient and the classical estimator for standard errors will be biased. In the form of spatial Durbin or spatial common factor model (Anselin, 1980). The Spatial Error Model is

$$y = X\beta + \epsilon \;\; and \; \epsilon = \lambda W_\epsilon + u \tag{6}$$

Since $\epsilon = (I - \lambda W)^{-1} u$ and thus $y = X\beta + (1 - \lambda W)^{-1}$ is equivalent to

$$y = \lambda W_y + X\beta - \lambda W X\beta + \epsilon \tag{7}$$

Where y is dependent variables, λ is Spatial Autocorrelation Parameter W is spatial Weight Matrix X_{ij} is Explanatory Variable and ϵ and u is Error Term i.i.d.

Which is spatial lag model with an additional set of spatially lagged exogenous variables (WX) and set of "k" nonlinear (common factor) constrains on the coefficient (the product of the spatial autoregressive coefficient β should equal the negative of the coefficient of WX. The similarity between the error model and the spatial lag model will complicate specification testing in practice, since tests designed for a spatial lag alternative will also have power against a spatial error alternative will also have power against a spatial error alternative, and vice versa.

5.4 Estimation Result from Spatial Regression

For specification identification based on robust Lagrange multiplier test (LM) test for each SEM (spatial error model) and SAR (spatial autoregressive model or spatial lag model), according to the procedure suggested by Florax and Nijkamp (2003). In our estimation result (as can be seen in Table 3), the robust LM statistic for spatial lagged dependent variable(LM_{ρ}^{r}) was equal to 33.316 and significant at 99% level of significance. The number was higher than those of SEM (LM_{λ}^{r}), which is 16.153, although it is also significant at 99% level of significance. This leads us to the conclusion that spatially lagged dependent variable model more preferable than spatial error model.

From the regression, several variable could be noticed as significant determinants of HDI such Islands (as qualitative variable), urban agglomeration, span of control, poverty, primary education enrollment rate, marine transportation infrastructure (sea port), spatial plan and natural resources endowment.

As SAR gave efficient estimation result compared to OLS in HDI cases (due to the problem of spatial autocorrelation), the figure in SAR estimation result better depict the determination of HDI in Indonesia. Variables that are significant in OLS estimation came out as significant determinant in SAR models, such as marine transportation infrastructure (harbor). Meanwhile, several other turns out to be insignificant in SAR estimation albeit they are significant in the linear model, such as western part of Indonesia (qualitative variable), majority party's share in the parliament, and population relative to province.

Table 3: OLS Regresion and Spatial Regression (Lag and Error) of HDI Determination

Determinants	OLS			Spatial Lag			Spatial Error		
	Coef.	Robust s.e.		Coef.	s.e.		Coef.	s.e.	
Geographic Factor :									
Island	-1.401	0.525	***	-1.302	0.435	***	-0.991	0.485	**
Western Indonesia	1.757	0.448	***	0.359	0.347		1.801	0.567	***
Urban Agglomeration	1.137	0.360	***	0.835	0.390	**	1.245	0.465	***
Governance Factor :									
Span of Control	-1.249	0.287	***	-1.137	0.181	***	-1.029	0.192	***
MajorityŠs Party Share	-2.310	1.856		-1.983	1.573		-2.306	1.731	
Spatial Plan	0.812	0.345	**	0.710	0.357	**	0.843	0.386	**
Scale Effect :									
Population Relative to Province	5.457	2.696	**	2.589	2.408		3.092	2.654	
Infrastructure :									
Sea Port	0.576	0.406		0.983	0.349	***	0.835	0.361	**
Trans Highway	0.454	0.367		0.168	0.329		0.016	0.353	
Endowment Factor :									
Mining and Forestry Share	2.367	1.021	**	1.990	0.833	**	2.074	0.873	**
Input Factor :									
Poverty	-2.100	0.514	***	-2.233	0.455	***	-2.270	0.469	***
Primary School Enrollment Rate (t-2)	0.185	0.045	***	0.169	0.024	***	0.168	0.025	***
Secondary School Enrollment Rate (t-2)	0.073	0.014	***	0.061	0.011	***	0.070	0.013	***
Constant	48.500	4.154	***	-16.046	2.774	***	46.458	7.192	***
ρ				0.975	0.025	***			
λ							0.973	0.026	***
R^2	0.491								
(Pseudo) R^2				0.601			0.486		
F-Statistic	26.91		***						
Log Likelihood				-1098.91			-1108.10		
Spatial Lag :									
LM_ρ				192.978		***			
LM_ρ^r				33.316		***			
Spatial Error :									
LM_λ							175.814		***
LM_λ^r							16.153		***

***, **, and *: Significant at 1%, 5%, and 10%

6 Geographically Weighted Regression

6.1 The Basic Methodology

Models estimated before whether the OLS or Spatial Regression, has been assumed that each coefficient of HDI determinants are the same for all observations. Geographically Weighted Regression in the other hand, enable us to examine wheter each coefficient of HDI determinants ate vary over space.

For this purpose, we apply Geographically Weighted Regression (GWR) techniques of Fotheringham et al. (2002) that allows for variability in the parameters. Consider a global regression model written as :

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \epsilon_i \tag{8}$$

GWR extend the global regression estimated by traditional Ordinary Least Square by allowing local rather than global parameters to be estimated, so that the model is rewritten as :

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_{ik} + \epsilon_i \tag{9}$$

where (u_i,v_i) denotes the coordinates of the ith point in space and $\beta_k(u_i, v_i)$ is a realization of the continous function $\beta_k(u)$ at point i. That is, we allow there to be continous surface of parameter values, and measurements of thus surface are taken at certain point to denote the spatial variability of the surface. Equation 9 is a special case of equation 8 in which the parameters are assumed to be spatially invariant. Thus the geographically weighted regression equation in 9 recognises that spatial variations in relationships might exist and provides a way in which they can be measured.

The inclusion of index i implies that equation 9 is not a single equation, but a set of n equations where the dimensions of β are $n \times \boldsymbol{J}$ localized regression estimates, where each observations is given a certain weight such that neighboring districts have more influence on the parameters than those located farther away (Seldadyo, 2007).

6.2 Testing Individual Parameter Stationarity

An earlier perhaps more pragmatic, approach to inference about GWR models was outlined in Fotheringham et al. (1996), that allows us testing the stationarity of individual parameter based on measuring their variability over space when estimated using GWR. The method is carried out as follows : a GWR estimate of the coefficient of interest in takes at each of th n data points and the variance (or standard deviation) of these estimates is computed. If the variance for parameter k is termed V_k then

$$V_k = \frac{1}{n} \sum_{i=1}^{n} \left(\hat{\beta_{ik}} - \frac{1}{n} \sum_{i=1}^{n} \hat{\beta_{ik}} \right) \tag{10}$$

Of course, even if the parameter of interest did not vary geographically, one would expect to see some variation in the estimated local values of the parameter. The question here is whether the observed variation is sufficient to reject the hypothesis that the parameter is globally fixed. To do this, consider the null distribution of the variance under this hypothesis. If there is no spatial pattern in the parameter, the any permutation of the regression variables against their locations is equally likely and on this basis we can model the null distribution of the variance.

6.3 GWR Estimation Result

Since the result of GWR provide us variation of variable coefficient for each observation, this study will do further analysis about suitability among GWR estimation and priori information. Analysis of Geographical Weighted Regression (GWR) result will helped by map appearance of each variable coefficient spread in Indonesia. There are 13 maps will represent 13 independent variables in the model (can be seen on appendix E) ; Island, Urban Agglomeration, Western Part of Indonesia, Span of Control, Majority Party Share, Population, Poverty, Elementary School Enrollment Rate, Junior High School Enrollment Rate, Marine Infrastructure, Trans Highway Infrastructure, Spatial Plan, Share of Mining and Oil in district's GRDP.

From GWR output and its stationarity test in Table 4, several variable are recognized to be involved in non-stationarity process over space, such as in district situated in islands and western part of Indonesia, span of control, scale effect of popoulation relative to province, transportation infrastrucure (both marine and land transportation infrastructure), and secondary school enrollment rate.

Geographical Factor Variables such as Island, Western Indonesia and Urban Agglomeration Variable show that eastern Indonesia tends to have lower coefficient than western Indonesia area. Similar result also showed by Governance Factor Variable, where span control show eastern the district the less the coefficient they have, meanwhile, GWR result on spatial plan indicate that the role of spatial plan are more crucial in the east due to the fact that they have higher coefficient of spatial plan compare to other area, so is the majority party's share in local parliament role. The same indication also showed by Infrastructure variable, the role of infrastructure more needed in eastern Indonesia than in western or central area. Endowment factor coefficient tends to higher in central Indonesia than western and eastern area, even western area does not less than eastern. Poverty role negatively affect HDI achievement, since poverty use poverty relative to province number the higher the poverty rate relative to province's poverty rate, the lowest the HDI will achieve. Precisely, that result also show in GWR coefficients, western Indonesia incline to have higher coefficient than eastern area. Another variable in input factor, school enrolment rate disposed differences between elementary and junior high school, junior high school enrollment rate play more important character in HDI achievement, the map of junior high school coefficient show that higher coefficient are in western and central Indonesia.

Table 4: Significance Test for GWR Output

Independent Variables	$\bar{\beta}_j$	β_j^{min}	β_j^{max}	Non-Stationarity Test		% Outliers
				$\sigma^2_{\beta_j}$	Sign.	
Significance Test For Non-Stationarity						
Geographic Factor :						
Island	-1.401	-4.778	5.313	2.082	***	0.23
Western Indonesia	1.757	-0.497	8.371	1.443	***	6.36
Urban Agglomeration	1.137	-1.894	4.078	0.752		2.05
Governance Factor :						
Span of Control	-1.249	-3.511	0.033	0.726	**	7.50
MajorityŠs Party Share	-2.310	-19.403	3.315	2.926		4.09
Spatial Plan	0.812	0.306	4.711	0.670		5.68
Scale Effect :						
Population Relative to Province	5.457	-8.236	42.517	8.027	**	5.92
Infrastructure :						
Sea Port	0.576	-0.079	4.534	0.847	*	6.82
Trans Highway	0.454	-1.578	7.953	1.753	***	7.05
Endowment Factor :						
Mining and Forestry Share	2.367	-4.902	4.197	1.763		1.14
Input Factor :						
Poverty	-2.100	-8.474	-1.055	0.853		2.95
Primary School Enrollment Rate (*t*-2)	0.185	0.012	0.193	0.052		10.91
Secondary School Enrollment Rate (*t*-2)	0.073	-0.153	0.114	0.037	**	1.59
Constant	48.500	44.112	66.796	4.775		0.00
Significance Test For Bandwidth						
Bandwidth	7.557				***	

Outlier are defined as those outside a range of $\bar{\beta}_j + 1.5\sigma^2_{\beta_j}$

7 Conclusion

In this paper we have found the existence of spatial autocorrelation in HDI achievement in sub national districts level in Indonesia. Due to the fact, we have undertaken exploration of HDI's spatial aspect using univariate analysis, spatial regression, and its spatial heterogeneity.

Taking into account the spatially lagged dependent variable, statistically speaking, there is no significance difference in HDI achievement in western part compared to eastern part of Indonesia in spatial lag model, despite it is significant in others including GWR. Although the result was shocking, univariate analysis could explain the phenomenon by the clustering pattern of districts with high HDI achievement out of western area due to the high initial endowment (especially in Eastern and Central Kalimantan and most Northern part of Sulawesi) which affecting HDI achievement with significance estimation result. Despite the area mentioned before, the clustering pattern of eastern part of Indonesia was in the low-low quadrant. In addition, member districts of agglomeration area, tends to have higher HDI achievement.

In governance factors, less heterogen (more homogeny) political climate does not lead to higher HDI achievement. Inspite of it, government role in shortening the span of control and administering spatial plan could lead to better HDI achievement. Input factors, such as poverty and primary education enrollment in the other hand plays as significance role in altering HDI achievement.

GWR output shows that transportation infrastructure plays important role in HDI achievement especially in the eastern part of Indonesia. We also might want to emphasize even more on marine infrastructure in the area, due to the significance of the estimation result of sea ports compared to trans highway in the spatial regression model. The result was also confirmed by the negatively significance estimation coefficient of districts located in islands in spatial lag model.

References

Anselin, L. (1995). Local Indicator of Spatial Association (LISA). *Geographical Analysis*, 27:93–115.

Cracolici, M. F. and Uberti, T. E. (2009). Geographical Distribution of Crime in Italian Province : A Spatial Econometrics Analysis. *Jahrbuch für Regionawissenchaft*, 29:1–28.

Florax, R. J. and Nijkamp, P. (2003). Misspecification in Linear Spatial Regression Models. *Tinbergen Institute Discussions Papers*, (2003-081/3).

Fotheringham, A. S., Brunsdon, C., and Charlton, M. (1996). Geographically Weighted Regression : A Method of Exploring Spatial Non-Sationarity. *Geographical Analysis*, 28:281–298.

Fotheringham, A. S., Brunsdon, C., and Charlton, M. (2002). *Geographically Weighted Regression : The Analysis of Spatially Varying Relationship*. West Sussex : John Wilye & Sons Ltd.

Fukuda, S. P. (2003). Rescuing Human Development Concept from The HDI : Reflection on The New Agenda. *Readings in Human Development, United nations Development Programme (UNDP) Europe and CIS Regional Centre*.

Getis, A. and Ord, K. (1992). The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis*, 24:189–206.

Ivanov, A. (2005). How to Make Local Development Work. United Nations Development Programmer (UNDP) RBEX Bratislava Regional Centre Working Paper.

Khomiakova, T. (2008). Spatial Analysis of Regional Divergence in India : Income and Economic Structure Perspective. *The International Journal of Economic Policy Studies*, 3(7):2008.

Kwan, M. P., Janelle, D. G., and Goodchild, M. F. (2003). Accesibility in Space and Time : A Theme in Spatially Integrated Social Science. *Journal of Geographical System*, 5.

Lüchter, G. and Menkhoff, L. (1996). Human Development as Statistical Artifact. *World Development*, 24(8):1385–1392.

LeSage, J. P. (1999). The Theory and Practice of Spatial Econometrics. *Department of Economics, University of Toledo*.

Noorbakhsh, F. (1998). A Modified Human Development Index. *World Development*, 26(3):517–528.

Ranis, G., Frances, S., and Ramirez, A. (2000). Economics Growth and Human Development. *World Development*, 28(2):197–219.

Sagar, A. D. and Najam, A. (1997). The Human Development Index : A Critical Review. *Ecological Economics*, 25:249–264.

Seldadyo, H. (2007). *Geography and Governance: Does Space Matter?* PhD thesis, Chapter in Disserations, University of Groningen.

Srinivasan, T. N. (1994). Human Development : A New Paradigm or Reinvention of The Wheel? *The American Economic Review, Papers and Proceedings of The Hundred and Sixth Annual Meeting of The American Economic Association*, 84:238–243.

Tobbler, W. R. (1970). A Computer Movie Simulating Urban Growth in The Detroit Region. *Economics Geography (Supplement : Proceedings, International Geographical Union, Commision on Quantitative Methods)*, 46(2):234–240.

UNDP (1990). Human Development Report. *New York : Oxford University Press*.

A Data Sources and Definition

Table 5: Data Definition and Sources

Determinants	Description	Sources
Geographic Factor :		
Island	1 if Districts is island region and 0 otherwise	*.shp map of Indonesia obtain from KPU CD-ROM*
Western Indonesia	1 if Districts constitute KBI and 0 otherwise	*.shp map of Indonesia obtain from KPU CD-ROM
Urban Agglomeration	1 if District constitute agglomeration and 0 otherwise	Indonesia General Contruction Department
Governance Factor :		
Span of Control	Factor Analysis of Size, Topography, and number of sub district and villages Population and local Government Unit	BPS**
Majority's Party Share	Share from majority party in DPRD as a proxy ofdistricts political complexity.	KPU CD-ROM
Spatial Plan	1 if districts have legalized Spatial Plan and 0 otherwise	http://www.kimpraswil.go.id
Scale Effect :		
Population Relative to Province	Population	BPS
Infrastructure :		
Sea Port	1 if Districts have port and 0 if not	http://www.dephub.go.id/hubla/, http://www.pelni.co.id, http://www.inaport1.co.id, http://www.inaport2.co.id, http://www.pp3.co.id, http://pelabuhan4.co.id, http://www.pertamina.com
Trans Highway	1 if Districts situated at trans highway and 0 otherwise	http://www.dephub.go.id
Endowment Factor :		
Mining and Forestry Share	Share from mining sector, oil and forestry to Domestic Product as a proxy of initial endowment	BPS
Input Factor :		
Poverty Relative to Province	Head Count Index	BPS
Primary School Enrollment Rate	Primary School Enrollment Rate 2004	BPS
Secondary School Enrollment Rate	Secondary School Enrollment Rate 2004	BPS

* KPU stands for Komisi Pemilihan Umum (The Commision of General Election *Indonesia General Election 2004 Result*, Interactive CD with maps 2006

** Badan Pusat Statistik (Central Bureau of Statistics

B Descriptive Statistics

Table 6: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Human Development Index	440	69.248	4.5853	47.20	78.30
Islands	440	0.184	0.3880	0.00	1.00
Urban Agglomeration	440	0.182	0.3861	0.00	1.00
Western Indonesia	440	0.620	0.4858	0.00	1.00
Span of Control	440	0.000	0.8729	-1.20	4.51
Majority's Party Share	440	0.299	0.0960	0.10	0.65
Population Relative to Province	440	0.075	0.0646	0.00	0.48
Poverty Relative to Province	440	1.000	0.3666	0.17	2.45
Primary School Enrollment Rate (t-2)	440	93.057	6.5450	23.15	99.98
Secondary School Enrollment Rate (t-2)	440	56.091	14.5765	21.57	99.47
Sea Port	440	0.280	0.4493	0.00	1.00
Trans Highway	440	0.343	0.4753	0.00	1.00
Spatial Plan	440	0.239	0.4267	0.00	1.00
Mining and Forestry Share	440	0.101	0.1789	0.00	0.97

C Urban Agglomeration List

Table 7: Urban Agglomeration Definition

Urban Agglomeration	Member District	Legal Basis
Mebidang	Medan	Government Regulation No.26, 2008, Appendices II
	Binjai	*Lampiran II Peraturan Pemerintah Republik*
	Deli Serdang	*Indonesia Nomor : 26 Tahun 2008*
Jabodetabekjur	Jakarta	Presidential Decree No. 54, 2008 regarding Spatial Plan
	Depok	In Region Jakarta, Bogor, Depok, Tangerang, Puncak, and
	Bogor	Cianjur (Jabodetabekpunjur)
	Tangerang	*(Perpres) No. 54 Tahun 2008 tentang Tata Ruang Kawasan*
	Bekasi	*Jakarta, Bogor, Depok, Tangerang, Bekasi, Puncak, dan*
	Cianjur	*Cianjur (Jabodetabekpunjur).*
	Bogor (City)	
	Tangerang (City)	
	Bekasi (City)	
Bandung Raya	Bandung (City)	West Java Spatial Plan, Local Regulations No.47, 1997
	Cimahi (City)	*RTRWP Jawa Barat, PP No.47 Tahun 1997 Mengenai RTRWN*
	Bandung	
	Sumedang	
	Subang	
Semarang Metropolitan	Semarang (City)	Semarang Spatial Plan (Central Java Development Planning Agency
	Demak	Cooperating with Diponegoro University)
	Kendal	
	Semarang	
Mamminasata	Sungguminasa	Governor of South Sulawesi Decree, 2003
	Maros	*SK Gubernur Propinsi Sulawesi Selatan Tahun 2003*
	Makassar	
	Takalar	
Gerbang Kertasusila	Gresik	East Java Spatial Plan
	Bangkalan	*RTRWP Jawa Timur*
	Mojokerto	
	Surabaya	
	Sidoarjo	
	Lamongan	
Sarbagita	Denpasar	Government Regulation No.26, 2008, Appendices II
	Bangli	*Lampiran II Peraturan Pemerintah Republik*
	Gianyar	*Indonesia Nomor : 26 Tahun 2008*
	Tabanan	

Sources : http://www.pu.go.id, Government Regulation No.26, 2008, Appendices II

D Districts List Based on LISA Scatterplot Quadrant

Table 8: LISA Quadrant 1

District	Province	District	Province
Aceh Besar	Nanggroe Aceh Darussalam	Kota Pangkal Pinang	Bangka Belitung
Bireuen	Nanggroe Aceh Darussalam	Karimun	Riau Islands
Banda Aceh (city)	Nanggroe Aceh Darussalam	Batam (city)	Riau Islands
Sabang (city)	Nanggroe Aceh Darussalam	Tanjung Pinang (city)	Riau Islands
Lhoksumawe (city)	Nanggroe Aceh Darussalam	Jakarta Selatan (city)	Jakarta
Dairi	North Sumatra	Jakarta Pusat (city)	Jakarta
Karo	North Sumatra	Jakarta Timur (city)	Jakarta
Tanjung Balai (city)	North Sumatra	Jakarta Barat (city)	Jakarta
Pematang Siantar (city)	North Sumatra	Jakarta Utara (city)	Jakarta
Tebing Tinggi (city)	North Sumatra	Bandung	West Java
Medan (city)	North Sumatra	Bekasi	West Java
Binjai (city)	North Sumatra	Bogor (city)	West Java
Simalungun	North Sumatra	Sukabumi (city)	West Java
Tapanuli Selatan	North Sumatra	Bandung (city)	West Java
Labuhan Batu	North Sumatra	Bekasi (city)	West Java
Langkat	North Sumatra	Depok (city)	West Java
Sibolga (city)	North Sumatra	Cimahi (city)	West Java
Padang Sidempuan (city)	North Sumatra	Klaten	Central Java
Deli Serdang	North Sumatra	Sukoharjo	Central Java
Tapanuli Utara	North Sumatra	Semarang	Central Java
Toba Samosir	North Sumatra	Temanggung	Central Java
Samosir	North Sumatra	Surakarta (city)	Central Java
Serdang Bedagai	North Sumatra	Salatiga (city)	Central Java
Tanah Datar	West Sumatra	Semarang (city)	Central Java
Agam	West Sumatra	Kulon Progo	Yogyakarta
Padang (city)	West Sumatra	Bantul	Yogyakarta
Solok (city)	West Sumatra	Sleman	Yogyakarta
Sawah Lunto (city)	West Sumatra	Yogyakarta (city)	Yogyakarta
Padang Panjang (city)	West Sumatra	Blitar (city)	East Java
Bukittinggi (city)	West Sumatra	Tangerang (city)	Banten
Payakumbuh (city)	West Sumatra	Cilegon (city)	Banten
Pariaman (city)	West Sumatra	Kotawaringin Barat	Central Kalimantan
Kuantan Singingi	Riau	Kotawaringin Timur	Central Kalimantan
Indragiri Hulu	Riau	Barito Utara	Central Kalimantan
Pekan Baru (city)	Riau	Katingan	Central Kalimantan
Kampar	Riau	Gunung Mas	Central Kalimantan
Dumai (city)	Riau	Murung Raya	Central Kalimantan
Rokan Hulu	Riau	Palangka Raya (city)	Central Kalimantan
Bengkalis	Riau	Bulungan	East Kalimantan
Indragiri Hilir	Riau	Nunukan	East Kalimantan
Siak	Riau	Balikpapan (city)	East Kalimantan
Rokan Hilir	Riau	Samarinda (city)	East Kalimantan
Kerinci	Jambi	Tarakan (city)	East Kalimantan
Batang Hari	Jambi	Bontang (city)	East Kalimantan
Tanjung Jabung Barat	Jambi	Minahasa Selatan	North Sulawesi
Jambi (city)	Jambi	Manado (city)	North Sulawesi
Palembang (city)	South Sumatra	Bitung (city)	North Sulawesi
Bandar Lampung (city)	Lampung	Tomohon (city)	North Sulawesi
Metro (city)	Lampung	Minahasa	North Sulawesi
Belitung	Bangka Belitung	Minahasa Utara	North Sulawesi

Table 9: LISA Quadrant 2

District	Province
Bengkulu (city)	Bengkulu
Kupang (city)	East Nusa Tenggara
Malinau	East Kalimantan
Makassar (city)	South Sulawesi
Kendari (city)	Southeast Sulawesi
Kota Ambon	Maluku
Maluku Tenggara	Maluku
Sorong (city)	West Irian Jaya
Jayapura (city)	Papua

Table 10: LISA Quadrant 3

District	Province	District	Province
Bondowoso	East Java	Jeneponto	South Sulawesi
Tuban	East Java	Wakatobi	Southeast Sulawesi
Pamekasan	East Java	Bombana	Southeast Sulawesi
Jember	East Java	Maluku Tenggara Barat	Maluku
Sampang	East Java	Buru	Maluku
Situbondo	East Java	Kepulauan Aru	Maluku
Sumenep	East Java	Seram Bag. Timur	Maluku
Lombok Barat	West Nusa Tenggara	Halmahera Selatan	North Maluku
Lombok Tengah	West Nusa Tenggara	Kepulauan Sula	North Maluku
Lombok Timur	West Nusa Tenggara	Manokwari	West Irian Jaya
Dompu	West Nusa Tenggara	Kaimana	West Irian Jaya
Bima	West Nusa Tenggara	Sorong Selatan	West Irian Jaya
Bima (city)	West Nusa Tenggara	Raja Ampat	West Irian Jaya
Sumbawa Barat	West Nusa Tenggara	Teluk Bintuni	West Irian Jaya
Sumbawa	West Nusa Tenggara	Teluk Wondama	West Irian Jaya
Belu	East Nusa Tenggara	Merauke	Papua
Lembata	East Nusa Tenggara	Jayawijaya	Papua
Sumba Barat	East Nusa Tenggara	Nabire	Papua
Timor Tengah Selatan	East Nusa Tenggara	Yapen Waropen	Papua
Timor Tengah Utara	East Nusa Tenggara	Paniai	Papua
Alor	East Nusa Tenggara	Puncak Jaya	Papua
Ende	East Nusa Tenggara	Mimika	Papua
Flores Timur	East Nusa Tenggara	Boven Digoel	Papua
Kupang	East Nusa Tenggara	Mappi	Papua
Manggarai	East Nusa Tenggara	Asmat	Papua
Manggarai Barat	East Nusa Tenggara	Yahukimo	Papua
Ngada	East Nusa Tenggara	Pegunungan Bintang	Papua
Rote Ndao	East Nusa Tenggara	Tolikara	Papua
Sikka	East Nusa Tenggara	Sarmi	Papua
Sumba Timur	East Nusa Tenggara	Keerom	Papua
Sambas	West Kalimantan	Waropen	Papua
Melawi	West Kalimantan	Supiori	Papua
Selayar	South Sulawesi		

Table 11: LISA Quadrant 4

District	Province
Simeulue	Nanggroe Aceh Darussalam
Aceh Singkil	Nanggroe Aceh Darussalam
Gayo Lues	Nanggroe Aceh Darussalam
Nias	North Sumatra
Nias Selatan	North Sumatra
Kepulauan Mentawai	West Sumatra
Solok Selatan	West Sumatra
Dharmas Raya	West Sumatra
Musi Rawas	South Sumatra
Muko-Muko	Bengkulu
Kaur	Bengkulu
Kepahiang	Bengkulu
Lampung Barat	Lampung
Bangka Barat	Bangka Belitung
Cianjur	West Java
Indramayu	West Java
Karawang	West Java
Brebes	Central Java
Ngawi	East Java
Bojonegoro	East Java
Lebak	Banten
Pandeglang	Banten
Serang	Banten
Ketapang	West Kalimantan

E Geographically Weighted Regression Output

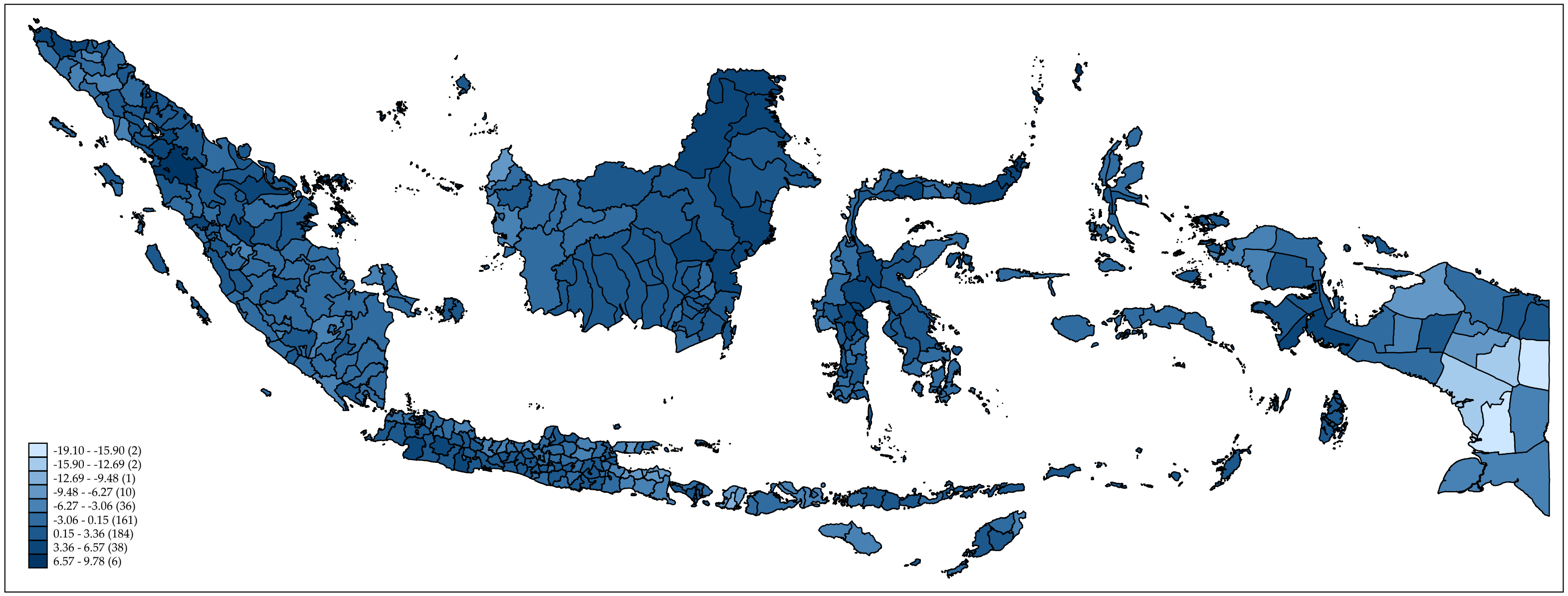

Figure 6: Residuals from Global Model

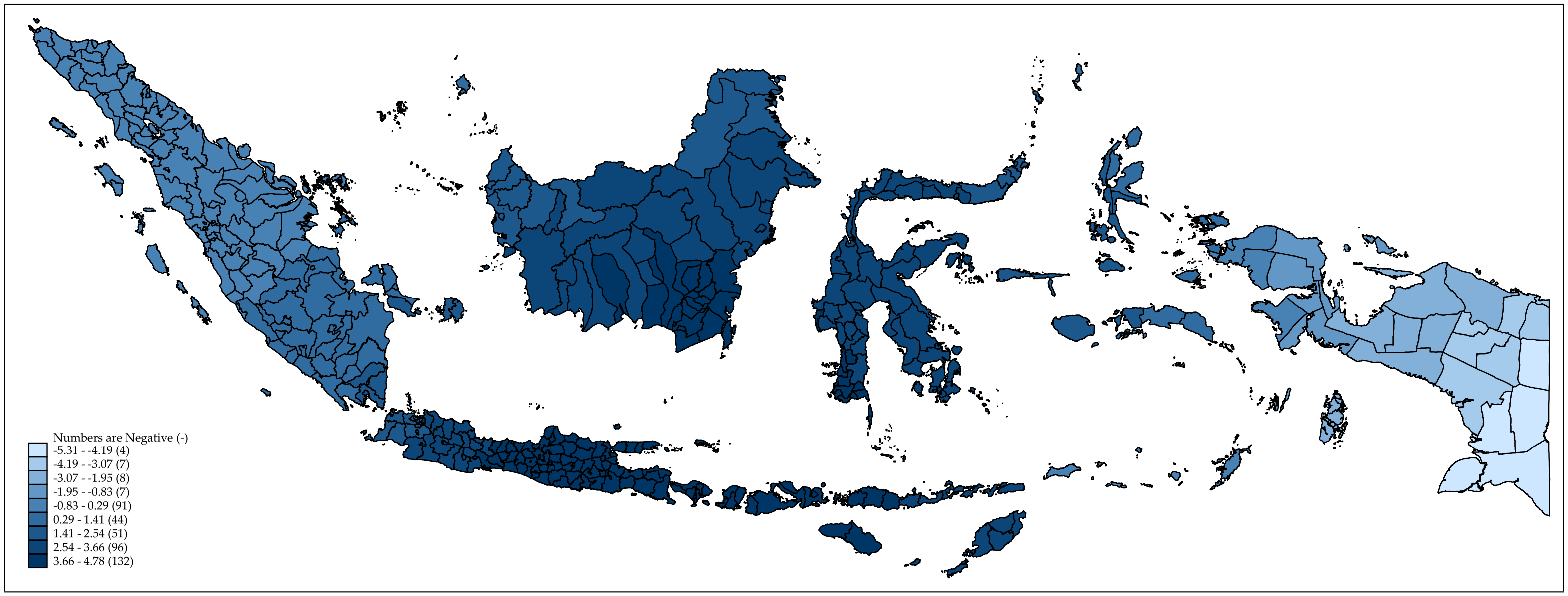

Figure 7: Coefficient of Islands

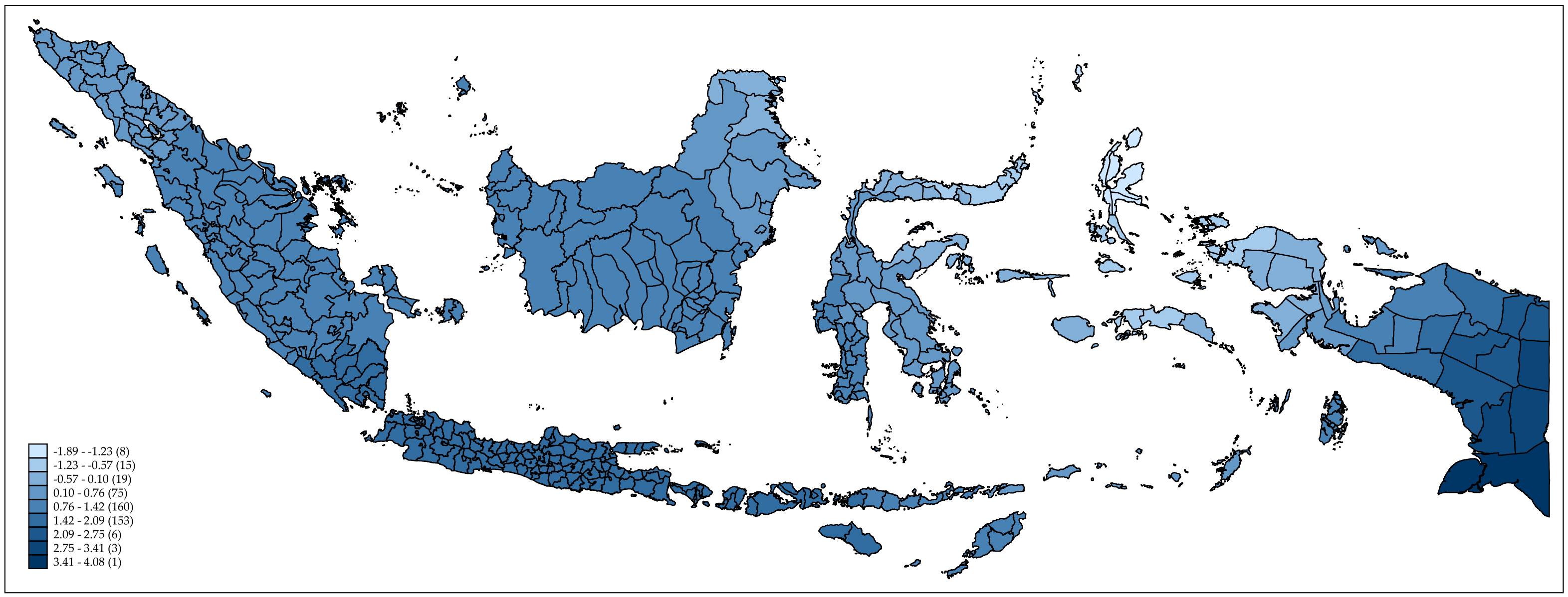

Figure 8: Coefficient of Urban Agglomeration

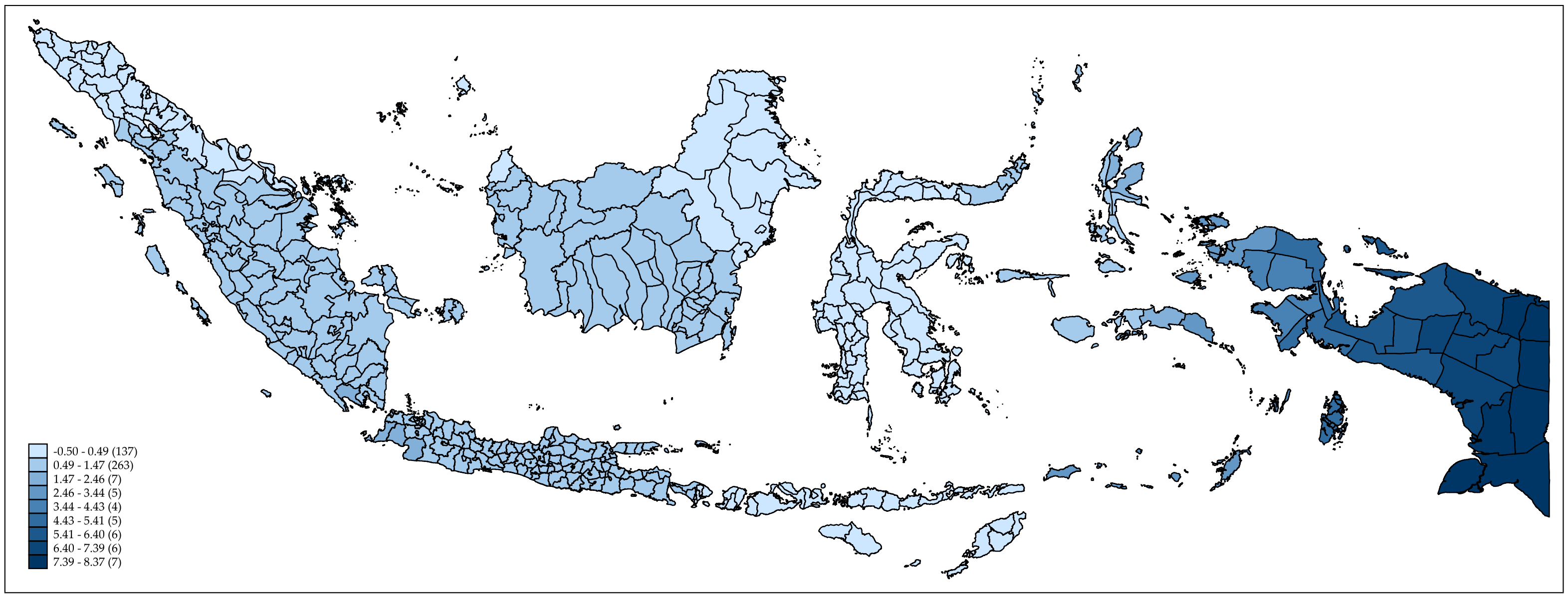

Figure 9: Coefficient of "Western Part of Indonesia"

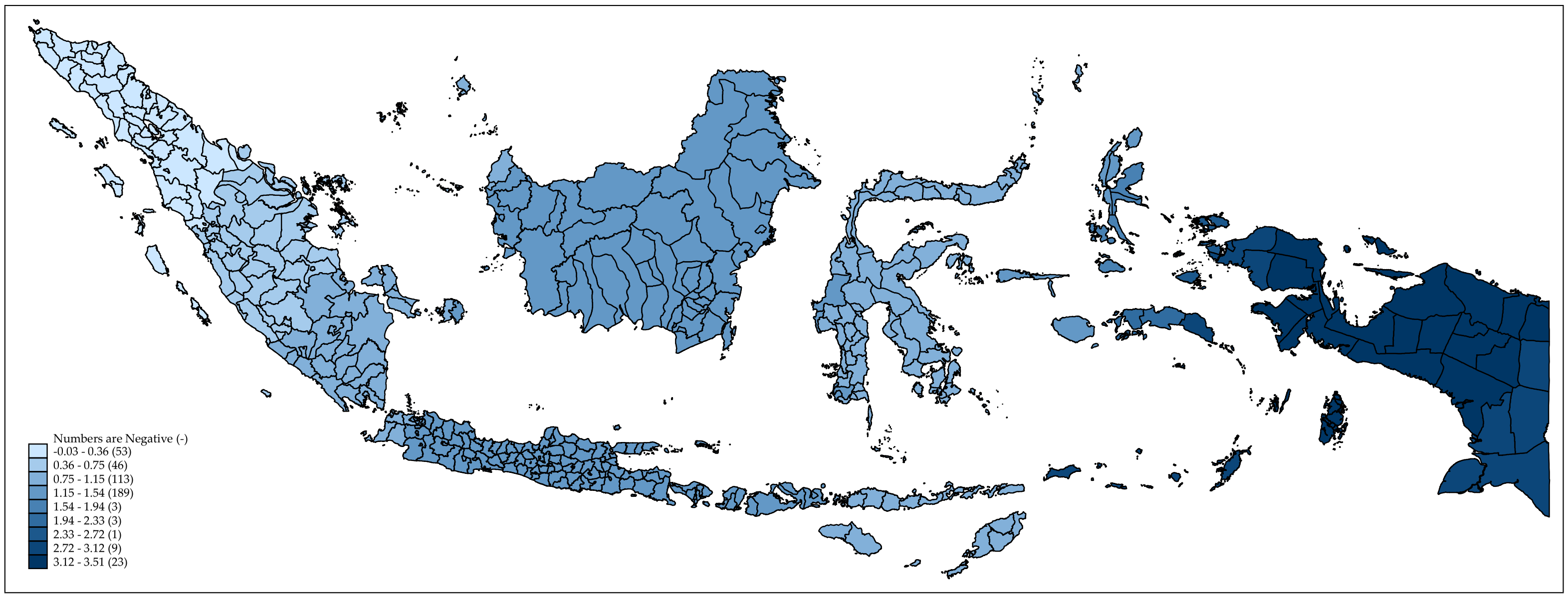

Figure 10: Coefficient of Span of Control

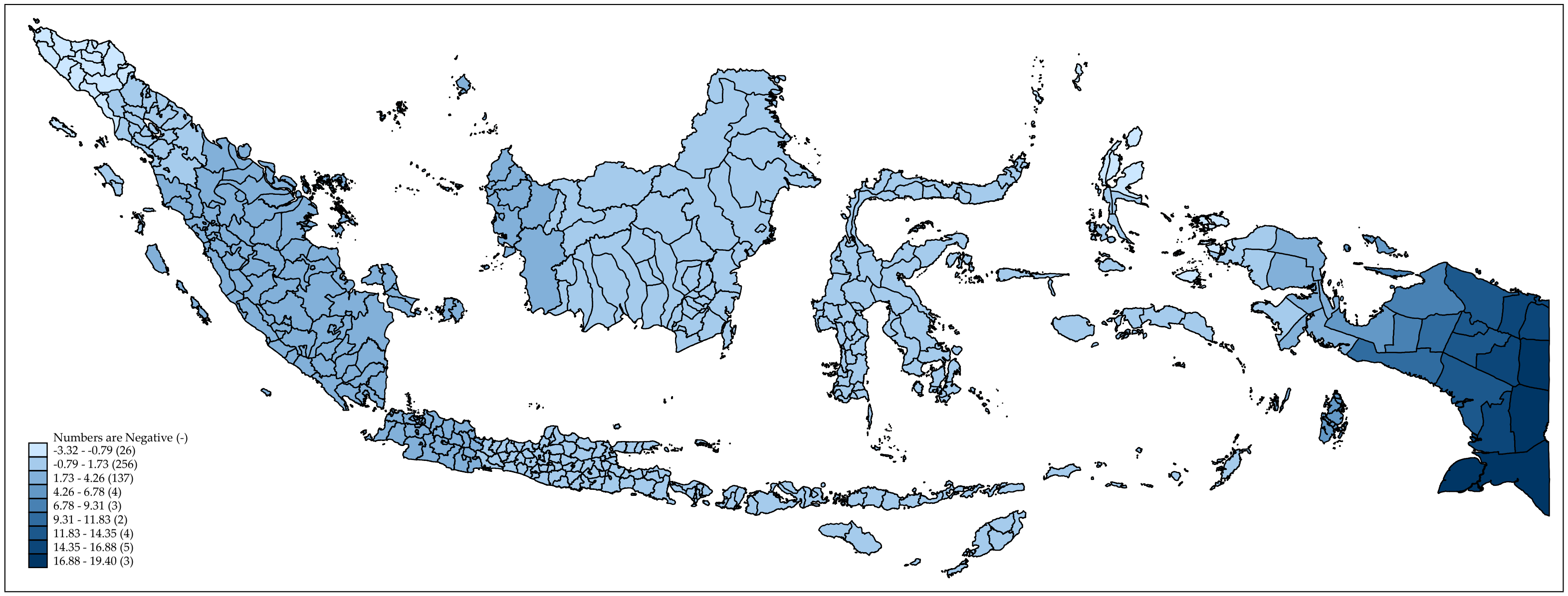

Figure 11: Coefficient of Majority Party's Share in Local Parliament

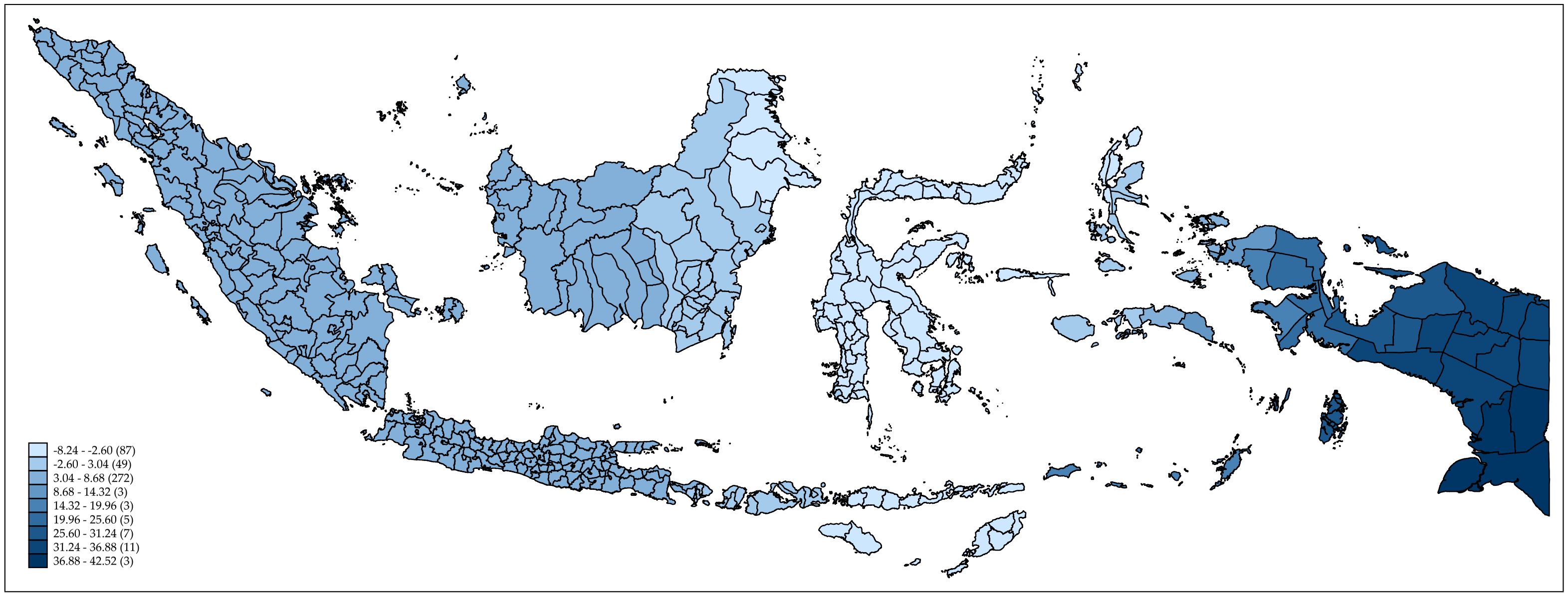

Figure 12: Coefficient of Population Relative to Province

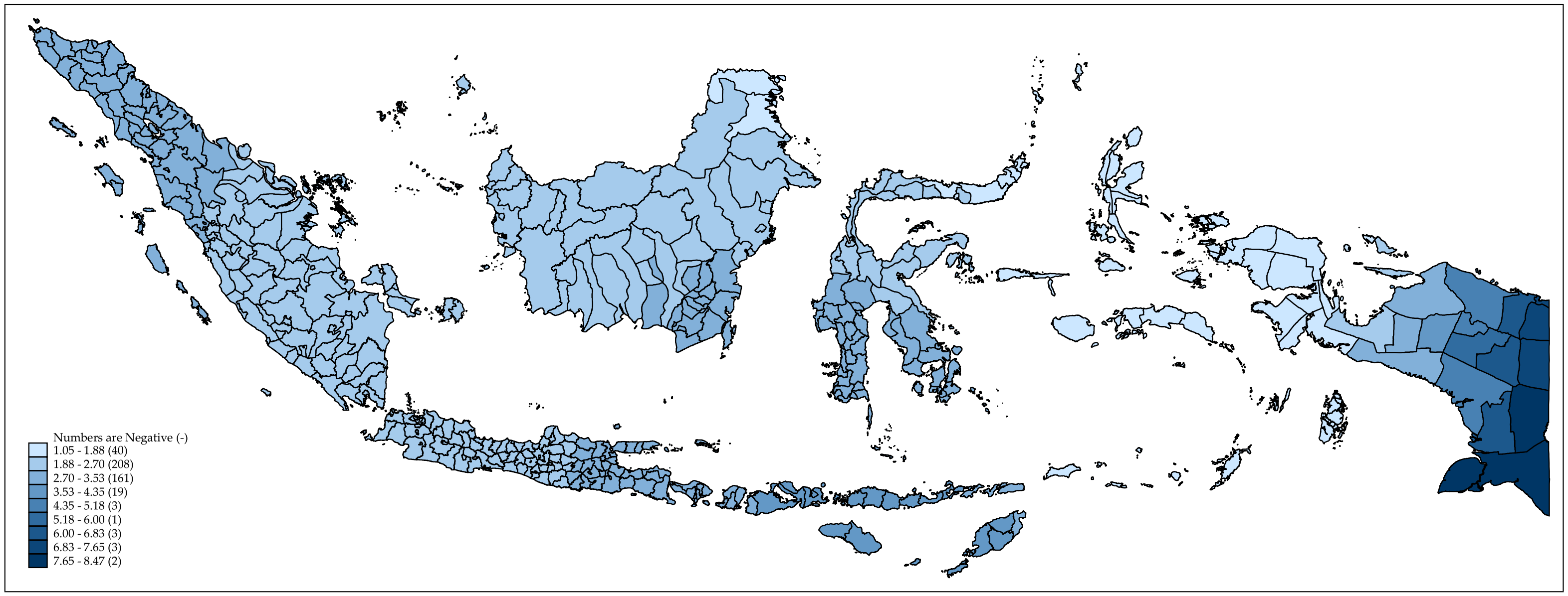

Figure 13: Coefficient of Poverty Relative to Province

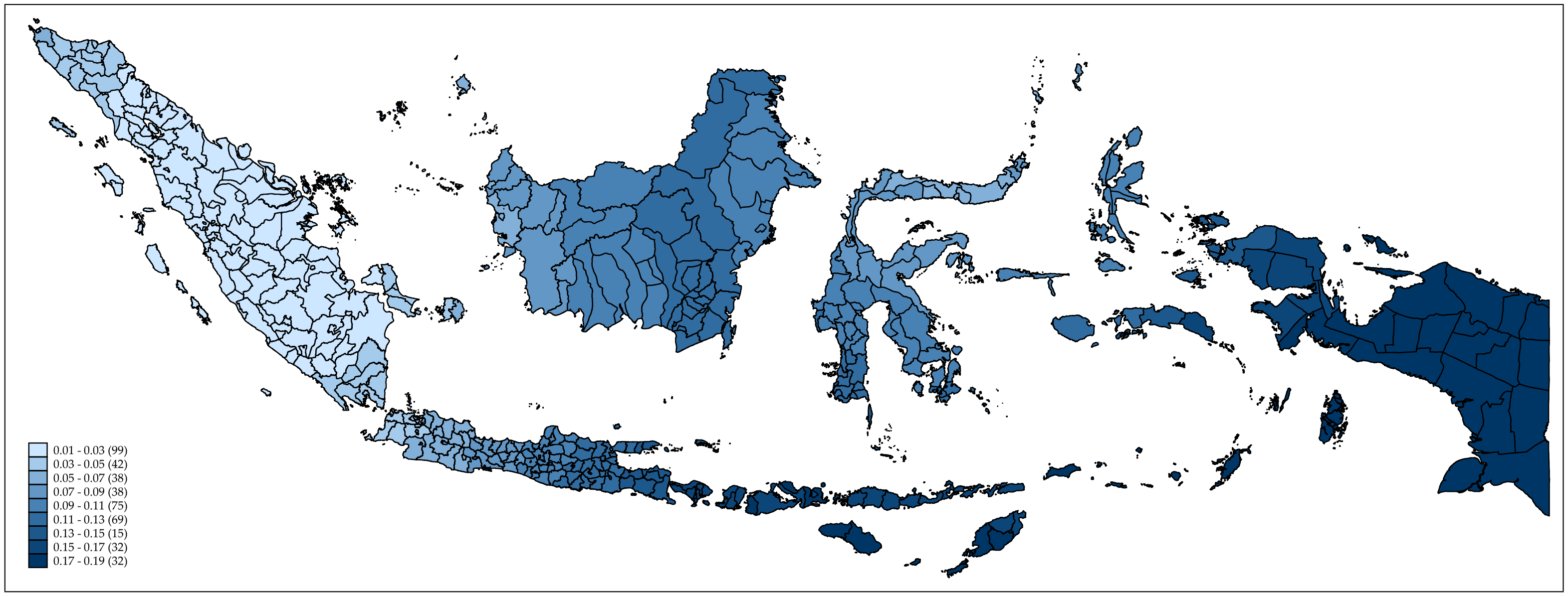

Figure 14: Coefficient of Elementary School Enrollment Rate in 2004 (*t*-2)

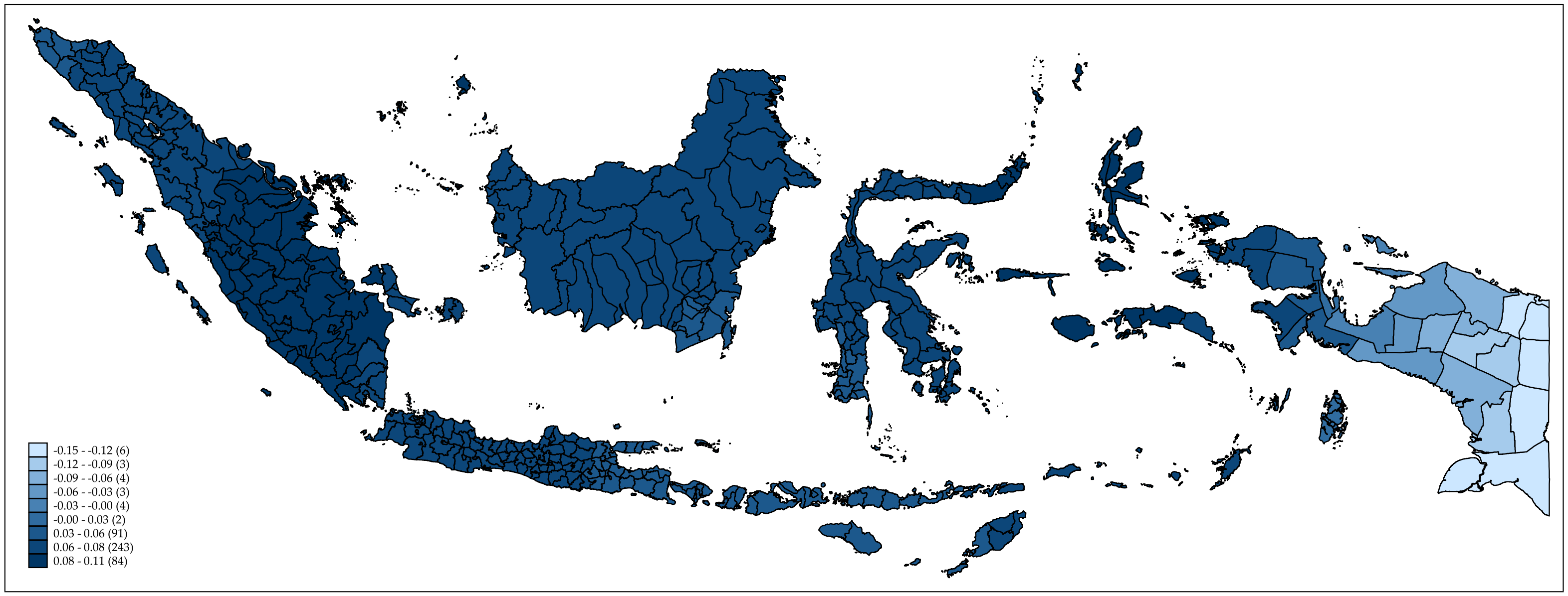

Figure 15: Coefficient of Junior Highschool Enrollment Rate in 2004 (*t*-2)

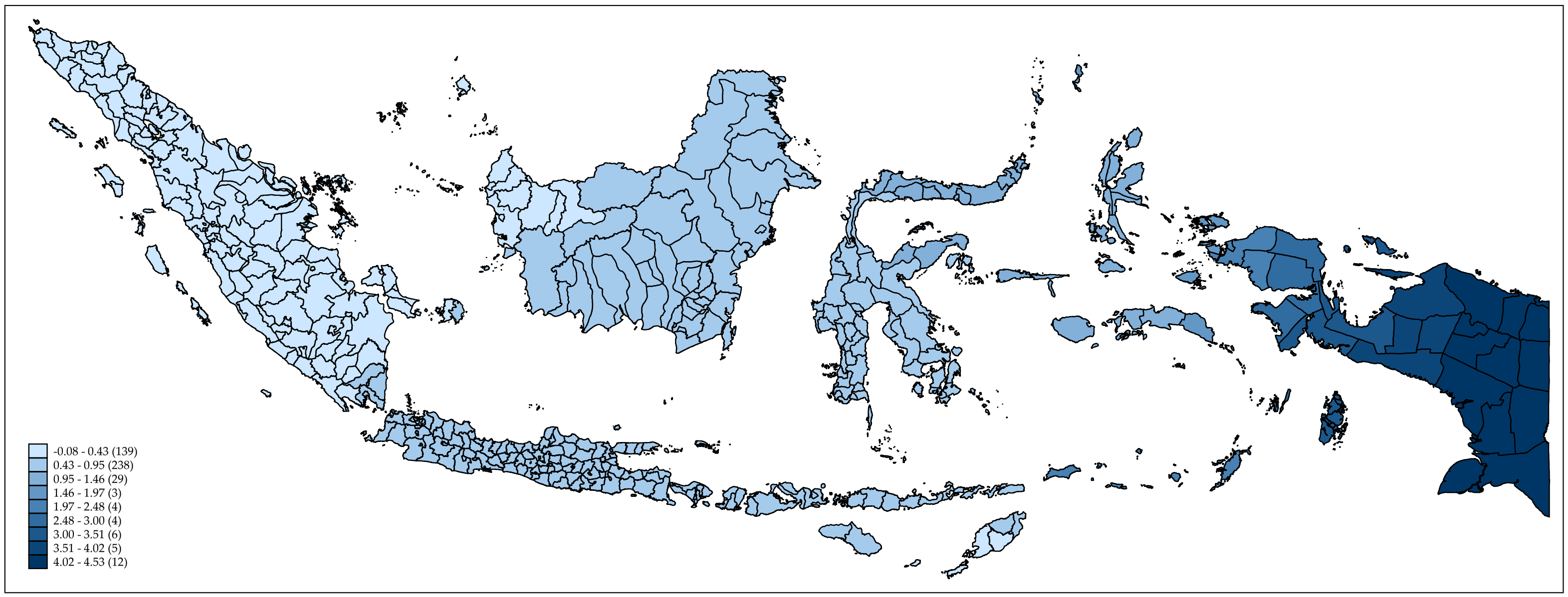

Figure 16: Coefficient of Sea Ports

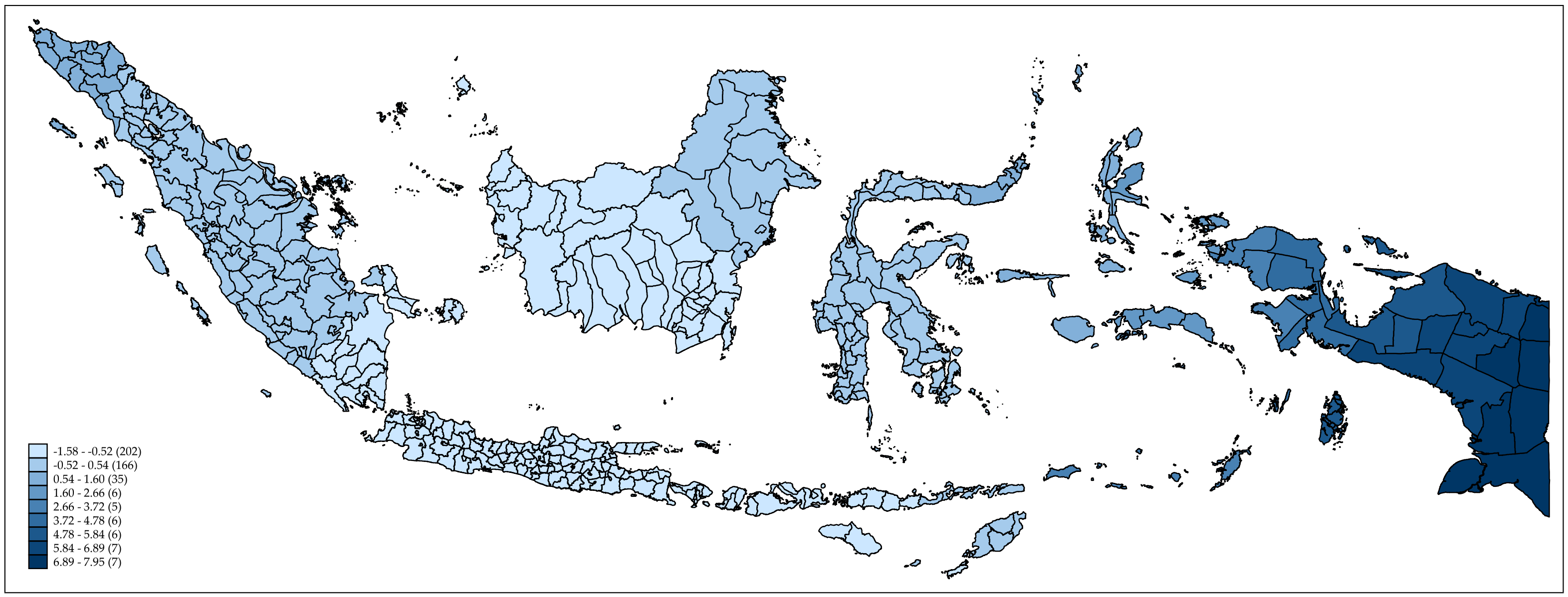

Figure 17: Coefficient of Trans Highway

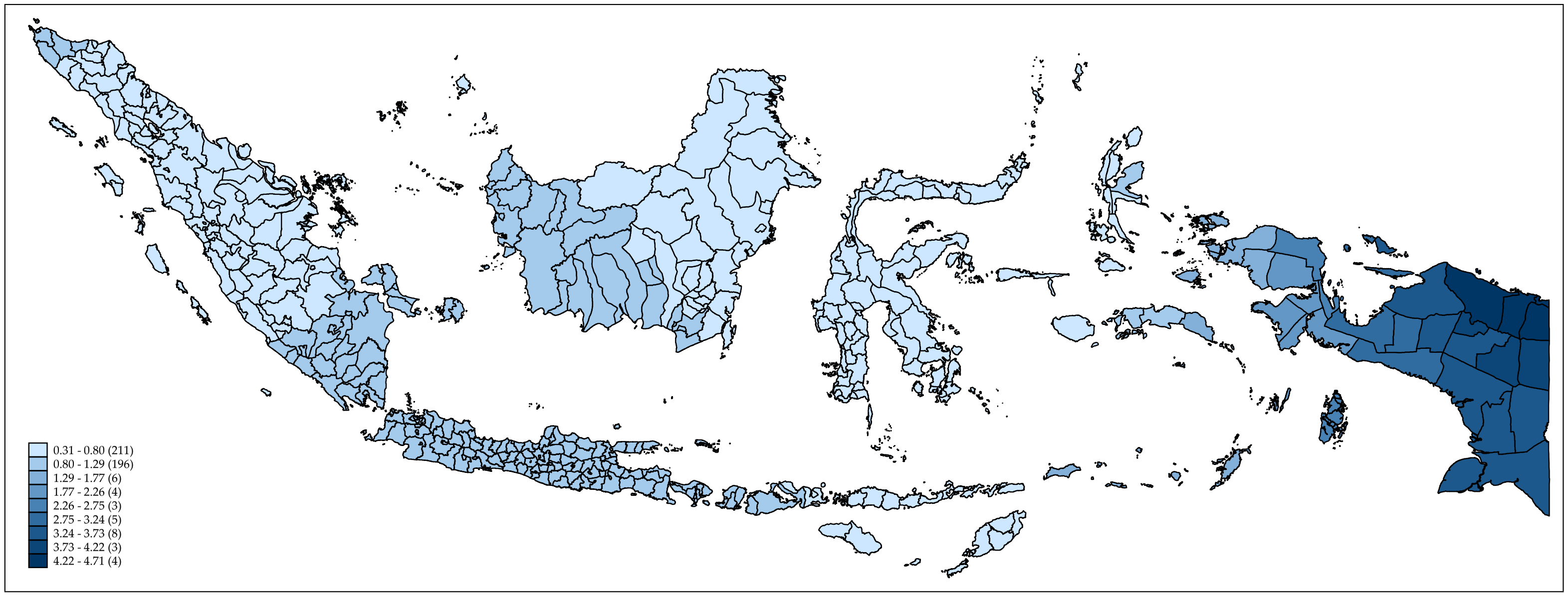

Figure 18: Coefficient of Spatial Plan

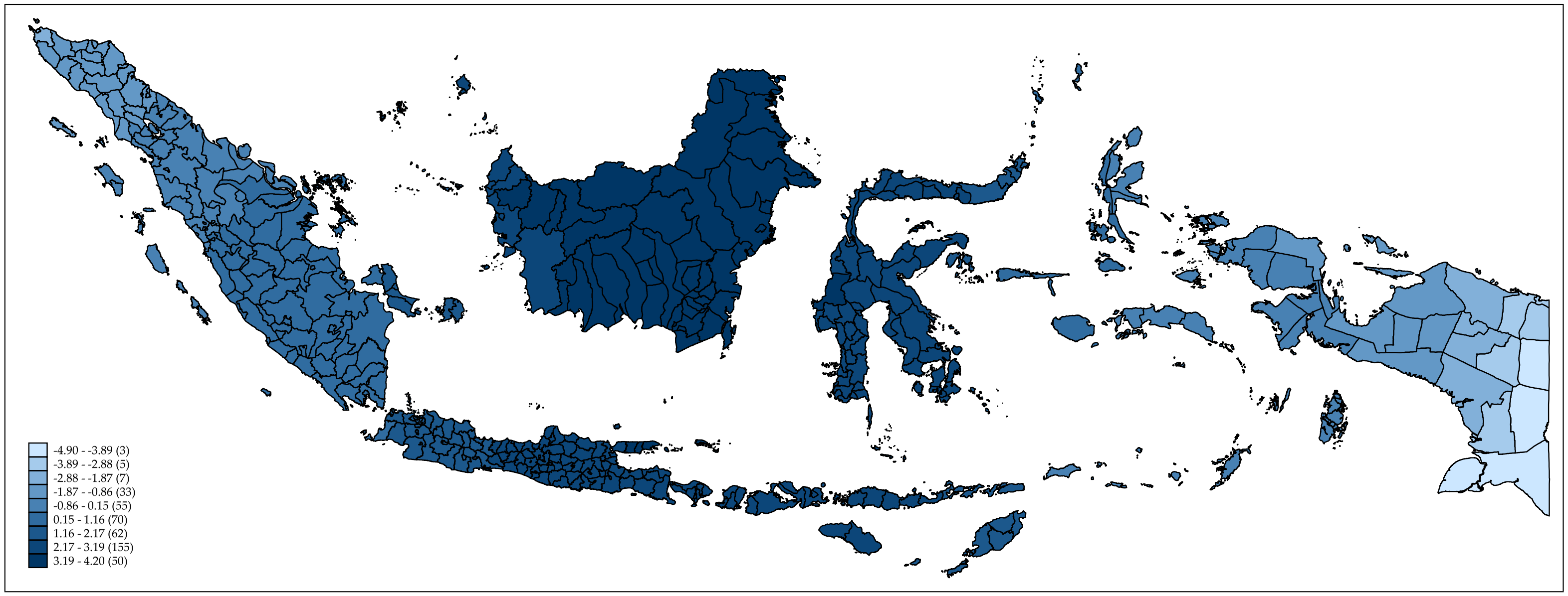

Figure 19: Coefficient of Share of Mining and Oil in District's GDRP

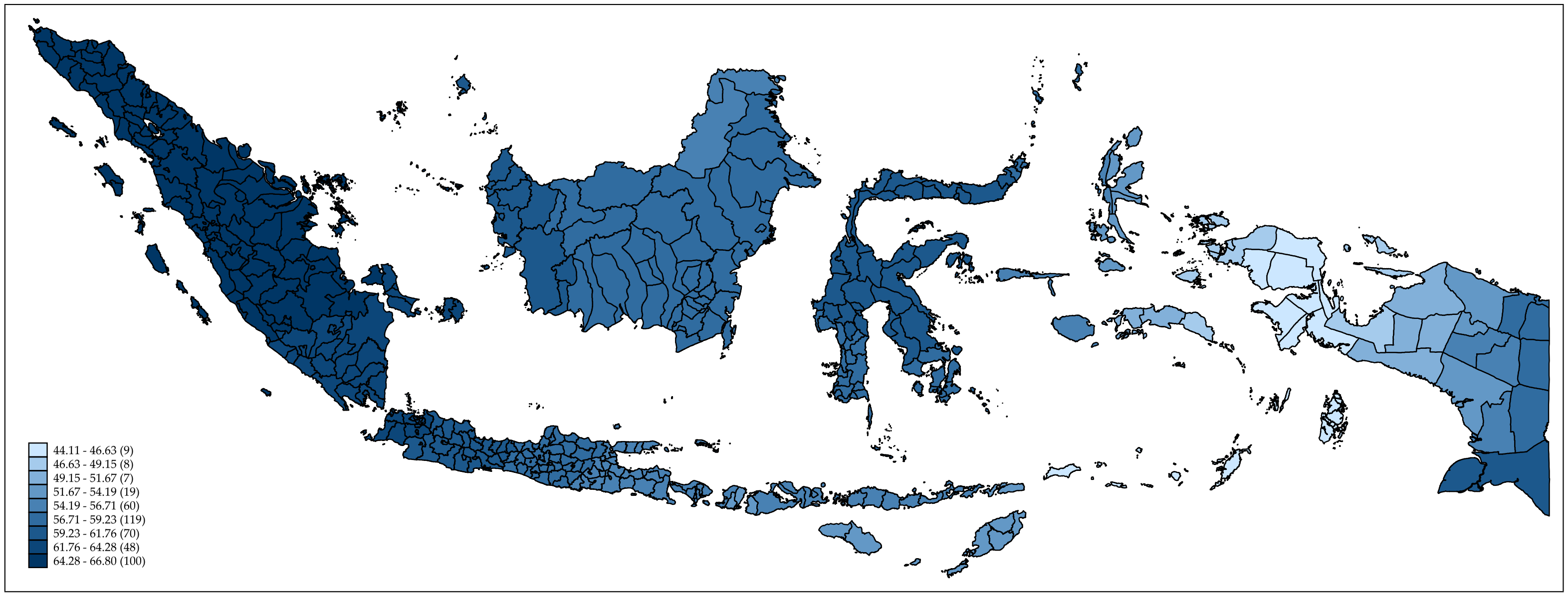

Figure 20: Coefficient of Constant