ISSN 2229-5518

Modeling of Gross Regional Domestic Product in Central Java, Indonesia

Lydia Mirna Wening Handayani, Anik Djuraidah, Utami Dyah Syafitri

Abstract—Gross Regional Domestic Product (GRDP) at constant prices is the main indicator to measure the region's economic growth. Java Island still contributes the largest of Indonesia's Gross Domestic Product (GDP). Central Java is one of the three largest provinces in Java Island, but the GRDP average of Central Java is still far behind the other two others. Therefore, it is necessary to examine the factors affecting GRDP in Central Java. The addition of time elements in this data as consequence of the GRDP were measured every year. The analysis used was panel data regression analysis. This aims of this paper was to investigate the factors which affected GRDP by panel data regression analysis in Central Java. The observed data were recorded from 2011-2015. The results showed that the factors which were affecting the GRDP were original local goverment revenue, regional minimum wage, and human development index.

Index Terms— Panel data, gross regional domestic product, economic growth, Central Java.

1 Introduction

ECONOMIC growth is an indicator of the development success. Based on data from Statistics of Indonesia (BPS), Indonesia's economic growth in 2011 was 6.5%. The number continues to decrease until 2014 to 4.88%. The decline in economic growth reflects that Indonesia's economic is slowing down.

An indicator to measure a country's economic growth was Gross Domestic Product (GDP), meanwhile the economic growth of a region was measured by Gross Regional Domestic Product (GRDP). According to Statistics of Indonesia (BPS), GRDP is the gross added value of all goods and services created or produced in an area obtained from various economic activities within a certain period. GRDP is the primary indicator to measure the economic growth of a region. The region could be a province or a district. There are two types of GRDP, namely GRDP at current prices and GRDP at constant prices. There are three approaches used in GRDP calculation i.e. production approach, income approach, and expenditure approach. In this paper used GRDP at constant prices.

Based on data from BPS, Java was the maiin contribution of Indonesia's Gross Domestic Product (GDP). The average GRDP in Java in the period 2011-2015 contributed 57.73% to the GDP of Indonesia. Central Java is one of the third largest provinces in the Java island, however the average GDP of Central Java contributeds 8.93% for Indonesia's GDP. On other side, the average GRDP of West Java and East Java contributed 13.33% and 14.61%, respectively. The contribution of Central Java was the lowest among West Java and East Java Province. A study needed to evaluate factors which contributed to GRDP in Central Java. The result was expected to be an information to local government to make poicy so the GRDP is increasing.

- *Student of masters degree program in applied statistics in Bogor Agriculture University, Indonesia,. E-mail:eena.the@gmail.com*
- *Main Supervisor, Lecturer of Departement of Statistics, Bogor Agriculture University, Bogor, Indonesia. E-mail:anikdjuraidah@gmail.com*
- *Co-supervisor, Lecturer of Departement of Statistics, Bogor Agriculture University, Bogor, Indonesia. Emai:utamids@gmail.com*

The data used in this paper were GRDP all districts/cities in Central Java for periods 2011 until 2015. As the observations were over time, the analysis used was panel data regression analysis. Panel data is a combination of cross-sectional and time series data. According to Hsiao [4], advantages of panel data are first, more accurate model parameter estimates of the model, and, second, more allows researchers to analyze issues that can not be solved with cross-sectional or time series data only.

2 Literature Reviews

Panel Data Regression

Panel data are a combination of cross-sectional data and time series data. A time series data is a series of data in which the observations are taken sequencing with equally space points in time. Cross-sectional data are a type of data in which the observations are consisted of several or many objects in a period of time. According to Gujarati [5], panel data are the cross-sectional data which observed several times. In general, according to Baltagi [3], the panel data regression model is expressed as follows:

$$y_{ij} = \alpha + \sum_{k=1}^{p} \beta_k x_{kij} + \mu_i + \varepsilon_{ij} \tag{1}$$

where $i = 1, 2, ..., n$; $j = 1, 2, ..., t$; $k = 1, 2, ..., p$; with i as cross sectional dimension, j index shows the time dimension, and k indicates the number of independent variable used. y_{ij} is the *i-th* response on location and *j-th* period and x_{kij} is the *k-th* independent variable for the *i-th* location and *j-th* period, β_k is the coefficient of *k-th* independent variable, α is a constant, μ_i is the effect of unobservable individual-spesific effect and ε_{ij} is *i-th* residual of location and *j-th* period.

There are three forms of data panel models, namely:

a) Pooled Model
   This model does not concern on the effect of location and time. The resulting panel data regression will be applied to each location. This model is the simplest model.

b) Fixed Effect Model

In this model, the intercept of the model can be distinguished between locations because each location is considered to have its characteristics. The assumptions of the model are:

(i) μ_i is assumed to be fixed so that it can be expected

(ii) ε_{ij} has normal distribution, with mean 0 and variance σ_{μ}^2.

(iii) $\mathbf{X}_{ij}$ is mutually independent with ε_{ij} for each *i* and *j*.

One of several ways to calculate parameter estimation on the fixed effect model is within transformation method. This method is done by subtracting the equation of the fixed effect model with the average equation. The equation model for the average is as follows:

$$\bar{y}_i = \alpha + \sum_{k=1}^{p} \beta_k \bar{x}_{ki} + \mu_i + \bar{\varepsilon}_i \tag{3}$$

The equation of the fixed effect model with the fixed effect estimator is obtained by subtracting equation (1) by equation (3) as follows:

$$y_{ij} - \bar{y}_i = \sum_{k=1}^{p} \beta_k \left(x_{kij} - \bar{x}_{ki} \right) + \left(\varepsilon_{ij} - \bar{\varepsilon}_i \right) \tag{4}$$

So the general equations of the fixed effect model with the fixed effect estimator are as follows:

$$\tilde{y}_{ij} = \sum_{k=1}^{p} \beta_k \tilde{x}_{kij} + \tilde{\varepsilon}_{ij} \tag{5}$$

c) *Random Effect Model*

The object used in this model is usually a randomly selected object from a large population. The assumptions that must be met in this model are:

(i) μ_i has normal distribution with mean 0 and variance σ_{μ}^2

(ii) ε_{ij} has normal distribution with mean 0 and variance σ_{ε}^2.

(iii) $\mathbf{X}_{ij}$ mutually independent with μ_i and ε_{ij} for each *i* and *j*.

3 METHODS

3.1 Data

The data used were secondary data from Statistics of Central Java Province. The data consisted of all districts/cities in Central Java in period 2011-2015. The respon variable was the GRDP of each districs/cities in million Rupiah meanwhile the explanatory variables can be seen in Table 1.

Table 1 The explanatory variables used in the study

Variables/Data	Units
1. Amount of Labor (X_1)	People
2. Regional Income (X_2)	Million Rupiah
3. Regional Minimum Wage (X_3)	Rupiah
4. Human Development Index (X_4)	Percent

3.2 Method of Analysis

The steps of the analysis data of this research are:

a. Conducting data exploration
b. Conducting multicollinearity test
c. Regressing the panel data:
   1) Conducting Lagrange Multiplier Test
      (i) Testing of time or cross-sectional effects
      The hypothesis is:
      $H_0 : \sigma_{\mu}^2 = \sigma_{\lambda}^2 = 0$ (no time or location effects)
      H_1 : at least one of the value is nonzero (there is a time or cross-sectional effect)
      The test statistic used is as follows:

$$\mathrm{LM} = \frac{nt}{2(n-1)(t-1)} \left[(n-1)\left[1 - \frac{\tilde{u}'(I_n \otimes J_t)\tilde{u}}{\tilde{u}'\tilde{u}} \right]^2 + (r-1)\left[1 - \frac{\tilde{u}'(J_n \otimes I_r)\tilde{u}}{\tilde{u}'\tilde{u}} \right]^2 \right] \sim \chi^2_{(2)} \tag{6}$$

      (ii) Testing of cross-sectional effects
      The hypothesis is:
      $H_0 : \sigma_{\mu}^2 = 0$ (Location has no effect)
      $H_1 : \sigma_{\mu}^2 \neq 0$ (Location has an effect)

      The test statistic used is as follows:

$$\mathrm{LM}_1 = \frac{nr}{2(t-1)} \left[1 - \frac{\tilde{u}'(I_n \otimes J_t)\tilde{u}}{\tilde{u}'\tilde{u}} \right]^2 \sim \chi^2_{(1)} \tag{7}$$

      (iii) Testing of time effect
      The hypothesis is:
      $H_0 : \sigma_{\lambda}^2 = 0$ (Time has no effect)
      $H_1 : \sigma_{\lambda}^2 \neq 0$ (Time has an effect)

      The test statistic used is as follows:

$$\mathrm{LM}_2 = \frac{nt}{2(n-1)} \left[1 - \frac{\tilde{u}'(J_n \otimes I_t)\tilde{u}}{\tilde{u}'\tilde{u}} \right]^2 \sim \chi^2_{(1)} \tag{8}$$

   2) Conducting Chow test to test whether the model is pooled model or fixed effects model
   The Chow test is used to test the significance between the pooled model and the fixed effects model. The hypothesis of the Chow test is as follows Baltagi [3]:
   $H_0 : \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_{n-1} = 0$ (pooled model)
   H_1 : at least one of the μ_1 is zero (fixed effect model)

   The test statistic used is as follows:

$$F_{hit} = \frac{(RRSS - URSS)/(n-1)}{URSS/(nt-n-k)} \tag{9}$$

   with RRSS as the restricted residual sum of squares estimation of pooled models. URSS is an unrestricted residual sums of squares predicted by a fixed effects model. Chow test statistic follows F distribution with the degree of freedom (n-1, nt-n-k).

   3) Conducting Hausman test to test whether the model

is random or fixed effects model

The Hausman test is used to test the significance between the random effects model and the fixed effects model. The hypothesis used is as follows Baltagi [3]:

$H_0 : E(\mu_i, \varepsilon_{ij} \mid X_{ij}) = 0$ (random effects model)

$H_1 : E(\mu_i, \varepsilon_{ij} \mid X_{ij}) \neq 0$ (fixed effects model)

Test statistic used is as follows:

$$\chi^2_{hit} = (\mathbf{b} - \boldsymbol{\beta})' Var(\mathbf{b} - \boldsymbol{\beta})^{-1} (\mathbf{b} - \boldsymbol{\beta}) \quad (10)$$

with:

β : vector coefficient of variable of fixed model

b : vector coefficient of variable of random effects model

The test statistic value folows Chi Square distribution with the degree of freedom *k*, with *k* indicating the number of explanatory variables.

4) Conducting test of normality of residual. If residual does not follow the normal distribution, the response variable needs to transform. The Box-Cox method was used to choose the appropriate transformation of the data.
5) Doing interpretation of the model

4 Result and Discussion

4.1 Data Description

Central Java Province consists of 35 districts/cities, namely 29 districts and 6 cities. GRDP of Central Java Province in 2015 amounted to 805.84 trillion rupiahs. Semarang city had 109.14 trillion rupiahs of GRDP, the highest value in Central Java. Gross Regional Domestic Product (GRDP) of Cilacap Regency and Kudus Regency became the second and third highest, i.e., 88.78 trillion rupiahs and 65.18 trillion rupiahs. The lowest GRDP in 2015 is Magelang City, which amounted to 5.24 trillion rupiahs.

Semarang City had the highest value of the amount of labor (X_1) in 2015 and Magelang City was the lowest. The highest value of the variable of original local government revenue (X_2) was Semarang City and the lowest was Banjarnegara District. The highest value of the variable of Regional Minimum Wage (X_3) was also owned by Semarang City while the lowest was Banyumas District. The highest value of HDI (X_4) was obtained by Salatiga City and the lowest was Brebes District.

4.2 Multicolinearity Test

The multicollinearity test was performed to determine the correlation among the explanatory variables. This test was performed by calculating the VIF (Variance Inflation Factor) value. Table 2 shows the VIF values of each explanatory variable. VIF value > 5 indicates that there is multicollinearity in the explanatory variables. Based on Table 2, the VIF values of the four explanatory variables are less than 5, so it can be concluded that there were no multicollinearity in all explanatory variables used in this study.

Table 2 VIF value of explanatory variables

Variable	VIF
X_1	2.223
X_2	3.110
X_3	1.720
X_4	2.160

4.3 Determination of Panel Model

Lagrange Multiplier test was conducted to test the effect of time, cross-sectional or both. The result of Lagrange Multiplier test can be seen in Table 3. Based on Table 3, it concluded that only the cross-sectional effect had a significant effect on GDP data of districts/cities in Central Java in 2011-2015, while the effect of time was not significant.

Table 3 Lagrange Multiplier, Chow, and Hausman test results

Test	Effect	Chi-square	F-count	Df	p-value
Lagrange Multiplier	Time and cross-sectional	309.06	-	2	0.000
	Time	0.1599	-	1	0.689
	Cross-sectional	308.9	-	1	0.000
Chow	Cross Section Fixed	-	600.43	(34,136)	0.000
Hausman	Cross Section Random	86.73	-	4	0.000

The Chow test was performed to choose between pooled models or fixed effect models. In Table 3 the F value was 600.43 with the p-value of 0.000, it was less than α = 5%. It concluded that model in this research was a fixed effects model. Next, the Hausman test was performed to determine the choice between a random model and a fixed effect model. The value of Chi-Square of the Hausman test was 86.73 with a p-value of 0.000, it less than α = 5%. It concluded that the final model used in this study was a fixed effects model.

4.4 The estimation of fixed effects panel model with cross-sectional effects

The parameter estimation of the fixed effects model is shown in Table 4. The R-square value of the model was 77.81% and R-square adjusted value of the model was 60.47%. This means that the explanatory variables used in the model were able to explain the variance of GRDP in Central Java by 77.81%, while the remainder was explained by other factors outside the model.

Tabel 4 shows that only X_2 (the variable of original local government revenue) was Significant to GRDP at α = 5%. Other explanatory variables were not significant at α = 5%. However, X_4 (HDI) was significant at α = 10%. Hence, two explanatory variables were significant at α = 10%, namely X_2 (the variable of original local government revenue) and X_4 (HDI). The coefficient of X_4 (HDI) was negative. It is supposed to be positive.

Table 4 The estimation of fixed effects model with cross-sectional effect

Variable	Coefficient	Standard Error	p-value	
X_1	6.147	4.778	0.200	
X_2	25.52	2.531	0.000	*
X_3	1.877	1.588	0.239	
X_4	-412900	232530	0.078	**

*) significant at α = 5%, **) significant at α = 10%

The validation of the model was done by doing the normality test of residual. The normality test was done by Kolmogorov-Smirnov (KS) test. The value of statistic KS was 0.112 with the p-value was 0.000. As the p-value was lower than α = 5%, the null hypothesis was rejected. It means that the data did not follow the normal distribution. It might be a cause why the coefficient of X_4 (HDI) was negatif.

To overcome the problem, the transformation of the response variable was conducted by Box-Cox transformation. The optimum lambda of Box-Cox transformation was -0.23 (Figure 1). As the resulting lambda value was close to 0, so the selected transformation was a natural logarithm function (ln).

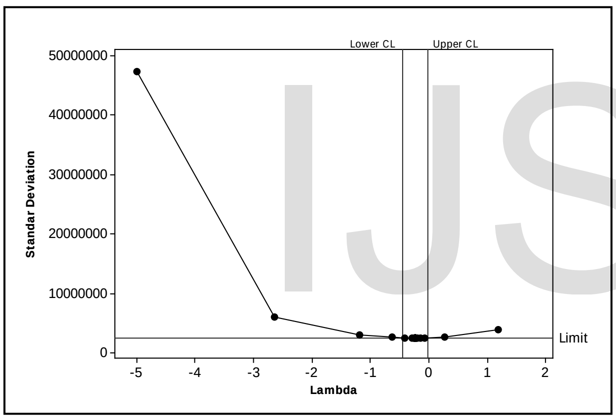

Figure 1 Graphic of Box-Cox Transformation

The estimation of parameter model of the transformation GRDP are shown in Table 5. The table shows that there were three explanatory variables that significantly affect at α = 5%, i.e. X_2, X_3 and X_4. One additional explanatory variable was significant compared to the original model (see Table 4). Furthermore, X_1 effected ln (GRDP) at α = 10%. The good news is all the coefficient of the model had a positive coefficient value. This means that all explanatory variables had a positive relationship to ln (GRDP). The more amount of labor, regional income, Regional Minimum Wage and the higher the HDI led to the growth of ln (GRDP). Hence, the transformation also overcame the magnnitude of coefficient problem.

The value of R-square was higher than the R-square of the original model. It was 95.09% and R^2 adjusted of 73.90%. This means that the explanatory variables used in the model were able to explain the variability of GRDP by 95.09%, while the remainder was explained by other factors outside the model. The Kolmogorov-Smirnov test confirmed that residual of the model followed the normal distribution (KS statistic = 0.060 and p-value = 0.123).

Table 5 Model estimation with transformation

Variable	Coefficient	Standard Error	p-value	
X_1	1.598×10^{-7}	8.179×10^{-8}	0.053	*
X_2	1.256×10^{-7}	4.333×10^{-8}	0.004	**
X_3	1.222×10^{-7}	2.719×10^{-8}	0.000	**
X_4	4.580×10^{-2}	3.981×10^{-3}	0.000	**

*) significant at α = 5%, **) significant at α = 10%

5 Conclusion

In conclusion, three of four explanatory variables had significant effect variables had significant effect to GRDP. Those were original local government revenue, regional minimum wage, and human development index. The model was also satisfied because the R-square achieve 95.09%.

References

[1] Badan Pusat Statistik, *Produk Domestik Regional Bruto Kabupaten/Kota di Indonesia 2011-2015*. Jakarta, Indonesia: BPS, 2016.

[2] Badan Pusat Statistik, "Konsep PDRB," available at https://www.bps.go.id/Subjek/view/id/52#subjekViewTab2, March. 2017.

[3] B.H. Baltagi, *Econometrics Analysis of Panel Data, Third Edition.* Chicester, England: John Wiley & Sons, 2005.

[4] C. Hsiao, *Analysis of Panel Data, Second Edition*. California, California: Cambridge University Press, 2014.

[5] D.N. Gujarati, *Applied Econometrics, Fourth Edition.* Singapore, Singapore: Mc. Graw-Hill International Editions, 2004.

[6] N. Hikmah, "Pemodelan Tingkat Produk Domestik Regional Bruto Kabupaten/Kota Jawa Barat dengan Spasial Data Panel,"undergraduate thesis, Dept. of Statistics, Bogor Agricultural Univ., Bogor, Indonesia, 2012.

[7] S. Sukirno, *Pengantar Teori Makroekonomi, Edisi Kedua*. Jakarta, Indonesia: PT RajaGrafindo Persada, 2002.

[8] Y. Sumardi, "Bayesian Model Averaging (BMA) dalam Pemodelan PDRB ADHK di Indonesia Tahun 2013," thesis, Padjadjaran Univ., Bandung, Indonesia, 2016.

[9] W.B. Suryono, "Analisis Pengaruh Pendapatan Asli Daerah, Tingkat Investasi dan Tenaga Kerja Terhadap PDRB Jawa Tengah," undergraduate thesis, Dept. of Economics, Diponegoro Univ., Semarang, Indonesia.,2010.