

A Panel Data Regression Analysis for Economic Growth Rate In Bengkulu Province

Filo Supianti^{1*}¹ BPS Kota Bengkulu, Bengkulu* Corresponding Author: filo@bps.go.id

Article Info**Article History:**

Received: 04 27 2023

Accepted: 05 26 2023

Available Online: 06 06 2023

Key Words:

Panel Data

Panel Data Regression

Economic Growth

Abstract

Panel data is a combination of time series data and cross section data. The analytical method used for panel data is panel data regression. One of the advantages of analysis using panel data regress One of the indicators to measure the development of the production of goods and services in an economic area in a given year against the value of the previous year which is calculated based on GDP/GRDP at constant prices is Economic Growth. The dependent variable in this study is the growth rate of GRDP. The independent variable in this study is IPM, TPAK, TPT. This study uses panel data regression analysis with the Common Effect Model (CEM), Fixed Effect Model (FEM) and Random Effect Model (REM). The data processing in this study uses the R Studio application.

1. INTRODUCTION

In the international economic sphere, countries are divided based on their economy, starting from developed, developing, underdeveloped or poor countries. One indicator of this division is economic growth. Economic growth is a process of changing economic conditions that occur in an area on an ongoing basis towards a condition that is considered better for a certain period of time. An economy can be said to have improved if the level of regional economic income that has been achieved is increasing from year to year [4]. By knowing the level of economic growth, the government of a region can make plans regarding state revenues and development planning in the region in the future. In addition, economic growth for business sector actors can be used as a basis for planning product development and its resources. In order to see the growth in the production of goods and services in an economic area within a certain time interval, it can be monitored through indicators of the rate of economic growth. A positive economic growth rate illustrates the state of the economy in a region that is experiencing an increase and vice versa when a negative economic growth rate indicates that the economic condition of that region is experiencing a decline.

The central government always tries to keep the rate of economic growth at a positive rate. Likewise with the regional government of Bengkulu Province. Various development programs are attempted to maintain a positive rate of economic growth. Based on BPS data in 2020 the rate of economic growth in Bengkulu Province has decreased by -0.02. This decline needs to be addressed so that it does not further decrease in the following year. Seeing this, an analysis is needed to determine the factors that influence the rate of economic growth. The analysis can be done with panel data regression. This method is used to analyze data involving time variables.

To find out the factors that influence the rate of economic growth in Bengkulu Province can be analyzed using panel data regression. In this paper, panel data regression analysis will be reviewed using three method approaches, namely the Common Effect Model (CEM), Fixed Effect Model (FEM), Random Effect Model (REM) methods. The processing of data in this study uses the R Studio application.

2. DATA AND METHOD

The data used in this study is the economic growth rate of 9 regencies and 1 city in Bengkulu Province from 2017 to 2020. The data used comes from the publication of BPS Bengkulu Province. The data is a combination of time

series and cross section data which is called panel data. This data will be modeled using panel data regression. The response variable used in this study is the rate of economic growth while the predictor variables used are the labor force participation rate (TPAK) as X_1 in percent, the open unemployment rate (TPT) as X_2 in percent, and the human development index (IPM) as X_3 in numeric digits.

Panel data regression is one of the developments of regression analysis methods [8]. Panel data regression is a regression technique that combines cross section data and time series data, so of course there will be more observations compared to cross section data or just time series data [3]. In general, the use of panel data can provide many advantages statistically and in economic theory. One of them is that panel data is able to explicitly account for individual heterogeneity by allowing individual-specific variables to be used in econometric equations [9]. The general model of panel data regression is expressed in the following equation:

$$Y_{it} = \alpha + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \dots + \beta_k X_{kit} + \varepsilon_{it} \tag{1}$$

with $i = 1, 2, \dots, n$ is the number of cross section data and $t = 1, 2, \dots, T$ is the amount of time series data, α is the intercept coefficient, $\beta_1, \beta_2, \dots, \beta_k$ is the slope coefficient with the predictor variable and ε_{it} is the regression error from the i -th data at the t -th time.

Before carrying out data analysis using the panel data regression model, a multicollinearity test was first performed. This test aims to determine whether in the regression model there is a perfect or near perfect linear relationship between some or all of the predictor variables. This test is carried out by calculating the VIF (Variance Inflation Factor) value for each variable.

After carrying out the multicollinearity test, modeling was then carried out with panel data regression. There are three methods that can be used to estimate the parameters of the panel data regression model, that is:

2.1 Common Effect Model (CEM)

The CEM method uses the same α for each individual and at all times, in other words, the behavior of the data for each individual is the same in various time periods. The regression equation using the CEM approach can be stated as follows:

$$Y_{it} = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k + \varepsilon_{it} \tag{2}$$

To estimate parameters in the CEM method, the Ordinary Least Square (OLS) method is used.

2.2 Fixed Effect Model (FEM)

The FEM method can be expressed by the equation:

$$Y_{it} = \alpha_1 + \sum_{k=2}^N \alpha_k D_{ki} + \beta X_{it} + \varepsilon_{it} \tag{3}$$

This method assumes that the intercept is different for each individual while the slope remains the same between individuals [1]. In distinguishing one individual from another, a dummy variable is used [6]. This model is often called the least square dummy variables (LSDV) model.

2.3 Random Effect Model (REM)

The REM method assumes that there are differences in intercept for each individual. So that there are two components of error, namely the overall model error and individual error. The overall model error is a combination of time series data and cross section, while individual errors are errors from each cross section data [5]. This method can be expressed by the equation:

$$Y_{it} = \alpha + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \dots + \beta_k X_{kit} + \varepsilon_{it} + \mu_i \tag{4}$$

where μ_i is the error component of the cross section data. To estimate the random effects model parameters can use Generalized Least Squares (GLS) [7].

Selection of the panel data regression model with the best approach is determined by using the Chow test, Hausman test, and Lagrange multiplier test.

1. Chow test

The Chow test is used to select the two models between the Common Effect Model and the Fixed Effect Model [2]. Chow test statistics are expressed by the equation [1]:

$$F_{hitung} = \frac{(SSE_1 - SSE_2)(nT - n - K)}{SSE_2(n - 1)} \tag{5}$$

SSE_1 is the sum square error of the common effect model, SSE_2 is the sum square error of the fixed effect model, n is the number of individuals, nT is the number of multiplications of the time series by cross section, and K represents the number of independent variables [1].

The hypothesis used is:

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_n = 0 \text{ (this indicates a suitable CEM model)}$$

$$H_1: \alpha_i \neq 0; i = 1, 2, \dots, n \text{ (this indicates a suitable FEM model)}$$

with the rejection criteria, reject H_0 if $F_{count} > F_{table}$

2. Hausman test

This test is a follow-up test that is used to find out which FEM model or REM model is the most appropriate after carrying out the chow test [10]. The test statistics are [1]:

$$W = \hat{q}' [var(\hat{q}')]^{-1} \hat{q}$$

$$W = (\hat{\beta}_{MET} - \hat{\beta}_{MEA})' [var(\hat{\beta}_{MET} - \hat{\beta}_{MEA})]^{-1} (\hat{\beta}_{MET} - \hat{\beta}_{MEA}) \tag{6}$$

where $\hat{\beta}_{MET}$ is the slope estimation vector of the fixed effect model and $\hat{\beta}_{MEA}$ is the slope estimation vector of the random effects model. The hypothesis used is:

$$H_0: \text{Appropriate REM model}$$

$$H_1: \text{Corresponding FEM model}$$

with the rejection criterion reject H_0 if $W > \chi^2_{(\alpha, K)}$

3. Lagrange Multiplier Test

When the Chow test and Hausman test cannot provide the most suitable model conclusion, the Lagrange Multiplier test is carried out with the test statistical equation [5]:

$$LM = \frac{nT}{2(T - 1)} \left[\frac{\sum_{i=1}^n (\sum_{t=1}^T e_{it})^2}{\sum_{i=1}^n \sum_{t=1}^T e_{it}^2} - 1 \right]^2 \tag{7}$$

Testing using the hypothesis:

$$H_0: \sigma_u^2 = 0 \text{ (this shows CEM is better than REM)}$$

$$H_1: \sigma_u^2 \neq 0; i = 1, 2, \dots, n \text{ (this shows REM is better than CEM)}$$

with the rejection criteria $LM > \text{chi-square table}$, this H_0 rejection indicates that the REM model is better than CEM.

3. RESULTS AND DISCUSSION

Bengkulu is a province on the island of Sumatra, which is directly adjacent to the provinces of West Sumatra, Jambi, South Sumatra and Lampung. Bengkulu Province has an area of 19,919 km² consisting of 9 regencies and 1 city. Bengkulu Province is directly adjacent to the Indian Ocean. Its strategic location because it borders the sea and has a national port makes the economy in Bengkulu Province continue to grow. This can be seen in the graphic image below:

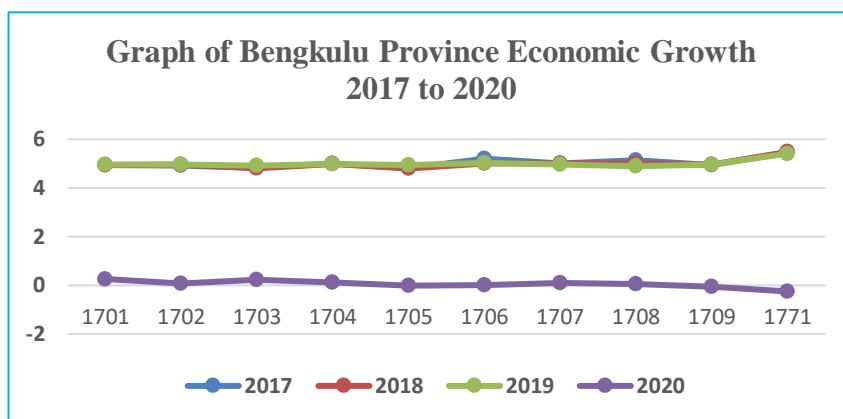


Figure 1. Graph of Bengkulu Province Economic Growth 2017 to 2020

The Covid-19 pandemic has caused the Bengkulu province's economic growth rate to decline in 2020. This is due to the impact on the community's economy and the mobility of the community in carrying out economic activities is hampered. To find out what factors have an influence and how much influence they have on the rate of economic growth, research can be done by modeling the data into the panel data regression model.

Before constructing the model, a multicollinearity test was first carried out. This is useful to see the relationship between predictor variables. Multicollinearity testing is done by calculating the VIF (Variance Inflation Factor) value. The results of data processing show the VIF values obtained for each variable presented in the table below;

Table 1. VIF value of each variable

Value	X_1	X_2	X_2
VIF	2,134464	1,198424	1,924412

Based on Table 1, the VIF value generated for each variable is less than 10. So it can be concluded that multicollinearity does not occur in the data and can be carried out to the next stage, namely the model formation stage.

The first step in building a model with panel data regression is to do a test to determine the appropriate approach method. The results of the R application can be seen in Table 1. The Chow test produces an H_1 acceptance so that the FEM model is appropriate. The Hausman test produces an acceptable H_1 so that the FEM model is appropriate. The Lagrange Multiplier test does not need to be continued because this test is carried out to determine the appropriate CEM or REM model while it has been found that the FEM model is suitable.

Table 2. Test Results Using the R Application

Test Statistics	$p - value$	Criterion
Chow Test	$2,123e^{-5} < \alpha = 5\%$	Accept H_1 : The corresponding FEM model
Hausman Test	$4,101e^{-13} < \alpha = 5\%$	Accept H_1 : The corresponding FEM model

From the results of the Chow and Hausman tests, the appropriate FEM model was selected. The model is

$$Y = -0,22994X_1 - 0,215619X_2 - 2,816363X_3$$

The resulting model shows that every 1 point increase in the labor force participation rate (X_1) will reduce the rate of economic growth by -0,22994. Every 1 increase in the open unemployment rate X_2 will reduce the rate of economic growth by -0,215619. For every 1 point increase in the human development index, the economic growth rate will decrease by -2,816363.

4. CONCLUSION

The results of data processing calculating the VIF value of the data are not greater than 10, this indicates that there is no multicollinearity in the data. Based on the statistical results of the Chow and Hausman tests that have been carried out, it is obtained that the FEM model is suitable for panel data regression in the case of the rate of economic growth in Bengkulu Province. The modeling results show that labor force participation (TPAK), open unemployment rate (TPT), and human development index (IPM) have an influence on the rate of economic growth in Bengkulu Province. However, to sharpen the analysis in this study, it is necessary to add several other variables such as inflation and other supporting variables.

REFERENCES

- [1] D. A. Nandita, L. B. Alamsyah, E. P. Jati, and E. Widodo, "Regresi Data Panel untuk Mengetahui Faktor-Faktor yang Mempengaruhi PDRB di Provinsi DIY Tahun 2011-2015," *Indonesian Journal of Applied Statistics*, vol. 2, no. 1, pp. 42-52, 2019.
- [2] D. Apriliawan, T. Tarno, and H. Yasin, "Pemodelan Laju Inflasi di Provinsi Jawa Tengah Menggunakan Regresi Data Panel," *Jurnal Gaussian*, vol. 2, no. 4, pp. 311 - 321, 2013.
- [3] D. N. Gujarati, *Dasar-Dasar Ekonometrika*. Jakarta: Erlangga, 2004.
- [4] I. Suswita, D. Damanik, and P. D. Panjaitan, "Pengaruh Infrastruktur terhadap Pertumbuhan Ekonomi di Kabupaten Simalungun," *Jurnal Ekuilnomi*, vol. 2, no. 1, pp. 1-11, 2020.
- [5] M. J. Hidayat, A. F. Hadi, and D. Anggraeni, "Analisis Regresi Data Panel Terhadap Indeks Pembangunan Manusia (Ipm) Jawa Timur Tahun 2006-2015," *Majalah Ilmiah Matematika dan Statistika*, vol. 18, no. 2, pp. 69-80, 2018.
- [6] M. Kuncoro, *Metode Riset Untuk Bisnis & Ekonomi*. Jakarta: PT Gelora Aksara Pratama, 2012.
- [7] M. Srihardianti, M. Mustafid, and A. Prahutama, "Metode Regresi Data Panel Untuk Peramalan Konsumsi Energi di Indonesia," *Jurnal Gaussian*, vol. 5, no. 3, pp. 475-485, 2016.
- [8] M. Y. Matdoan and M. W. Talakua, "Analisis Faktor-Faktor Yang Mempengaruhi Penyerapan Tenaga Kerja Pada Sektor Perhotelan di Provinsi Maluku dengan Menggunakan Metode Regresi Data Panel," *Jurnal Matematika Sains dan Teknologi*, vol. 22, no. 2, pp. 1-12, 2021.
- [9] Rahmadeni and N. Wulandari, "Analisis Faktor-Faktor yang Mempengaruhi Inflasi pada Kota Metropolitan di Indonesia dengan Menggunakan Analisis Data Panel," *Jurnal Sains Matematika dan Statistika*, vol. 3, no. 2, pp. 34-42, 2017.
- [10] Y. M. Sitorus and L. Yuliana, "Penerapan Regresi Data Panel Pada Analisis Pengaruh Infrastruktur Terhadap Produktivitas Ekonomi Provinsi-Provinsi Di Luar Pulau Jawa Tahun 2010-2014," *Media Statistika*, vol. 11, no. 1, pp. 1-15, 2018.