

Factors Affecting Economic Growth in West Sumatera Province Using Panel Data Regression Analysis

Sherly Helma Putri^{1,*} and Helma¹

¹ Departement of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Padang, Indonesia

*sherlyhelma07@gmail.com

Abstract. Economic development is considered successful if the economic growth rate of its people reaches a high level. Indonesia has positive economic growth, with growth rates above 5% each quarter. However, the high economic growth of Indonesia does not mean that all regions have the same growth rate. The increase in the number of goods and services received, or the added value of production factors is often called economic growth. Regional economic growth in West Sumatera Province tends to be negative. This study aims to obtain an overview of panel data regression models and factors that influence economic growth in West Sumatera Province from 2018 to 2022. The best regression model obtained is the fixed effect model (FEM), where at a significant level of 5%, the factors that have an influence and positive relationship on economic growth are the human development index and government spending.

Keywords: Economic Growth, Panel Data Regression, Fixed Effect Model (FEM).

1 Introduction

A country's economy is one of the indicators of realizing the welfare and prosperity of society. Therefore, its growth is very concerning. To measure the level of economic growth is to look at the country's productivity level each year, where this productivity measure uses Gross National Product (GNP) and Gross Domestic Product (GDP). Economic growth can be positive if economic activity has increased from the previous year, while economic growth is negative if economic activity has decreased from the previous year. Growth activities are the production of various types of goods and services increased by using enlarged production capacity to produce greater output (goods and services). Growth that includes various types of activities (sectoral) carried out in a region is called regional economic growth. To measure the region's economic growth level, its productivity economically uses the Gross Regional Domestic Product (GRDP) [1].

During 2022, Indonesia showed its success in dealing with various economic problems. This can be seen from the economic growth rate, which is above 5% in each quarter. Indonesia's economic growth remained high amidst slowing global economic growth. Various groups consider Indonesia to have successfully overcome the Covid-19 pandemic and handled the price spike of various commodities due to the Russia-Ukraine war. However, the high economic growth of Indonesia does not mean that all of its regions have the same growth rate.

In Indonesia in 2022, based on data from [2], the spatial structure of the economy is dominated by a group of provinces in Java Island, Kalimantan Island, and Sumatera Island. Sumatera Island is ranked third, with South Sumatera Province dominating the economic growth recorded above 5%. However, it differs from West Sumatera Province, the fourth lowest province experiencing economic growth on the island of Sumatera. Regional economic growth in West Sumatera Province has economic growth that tends to be negative. Economic growth in West Sumatera Province, based on data from [3] in 2018, was recorded at 5.14% and experienced a decline in 2019, reaching 5.01% until 2020, which experienced a significant decline recorded at -1.61%; the Covid-19 pandemic caused this.

Furthermore, in 2022, economic growth in this region has increased, reaching 4.36% or up 1.07% compared to 2021. However, this figure is still below the average national economic growth rate of 5.31%. This illustrates that the economic growth of West Sumatera Province is slow in Indonesia. Therefore, it is necessary

to observe the factors that influence economic growth in West Sumatera Province in 2018-2022 so that it experiences a significant increase.

As a result of these problems, systematic research is carried out in the form of statistical analysis using data from several objects and a period. The analysis is a panel data regression analysis [4], which is expected to provide an overview of the factors affecting West Sumatera Province's economic growth in 2018-2022. The factors used in this study are the human development index [5], percentage of poor people [6], labor force participation rate [7], literacy rate [8] and government spending [9]. Therefore, panel data regression analysis is used to analyze the factors affecting economic growth in West Sumatera Province in 2018-2022.

2 Theoretical Basic

2.1 Panel Data Regression Model

Panel data regression models are more complex and not as simple as ordinary regression models. The general model of panel data regression is expressed in the form of the following equation [10]:

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \beta_4 X_{4it} + \beta_5 X_{5it} + u_{it} \quad (1)$$

Where:

Y_{it}	: Dependent variable of the "i" th individual in period "t"
X_{it}	: Independent variable of the "i" th individual in period "t"
i	: Unit cross-section/individual (1,2,...,19)
t	: Unit time series/period (1,2,...,5)
β_0	: Intercept
β	: Slope/ Independent variable parameters
u_{it}	: Galat atau error of individual "i" of period "t"

Panel data regression consists of three parameter estimates used in determining the panel data regression model, which are as follows [10]:

Common Effect Model (CEM)

In CEM, it is assumed that the intercept and slope are constant. In other words, the data behavior of each individual is the same in various periods.

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \beta_4 X_{4it} + \beta_5 X_{5it} + u_{it} \quad (2)$$

Fixed Effect Model (FEM)

In FEM, it is assumed that the intercept shows differences between individuals but does not change over time, while the slope is fixed for each individual and various time periods.

$$Y_{it} = \beta_{0i} + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \beta_4 X_{4it} + \beta_5 X_{5it} + u_{it} \quad (3)$$

Random Effect Model (REM)

In REM it is assumed that the intercept shows differences between individuals but does not change over time, where the intercept is assumed to be a random variable.

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \beta_4 X_{4it} + \beta_5 X_{5it} + w_{it} \quad (4)$$

2.2 Selection of Panel Data Regression Estimation Model

There are three types of specific tests used to select a panel data regression model, which are as follows [11]:

Chow Test

It is necessary to determine the best panel data regression model between the model obtained based on the CEM approach and the model obtained using the FEM approach.

$H_0: \beta_i = 0$, Common Effect Model

$H_1: \beta_i \neq 0$, Fixed Effect Model

With test statistics:

$$F = \frac{\frac{SSE_1 - SSE_2}{N-1}}{\frac{SSE_2}{NT-N-K}} \quad (5)$$

Where:

- SSE_1 : Sum square error dari model common effect
- SSE_2 : Sum square error dari model fixed effect
- N : Number of cross section units/individual
- T : Number of time series units/period
- NT : Multiplication of time series by cross section
- K : Number of units of the independent variable
- α : Significance level

Lagrange Multiplier Test

It is necessary to know the best panel data regression model between the model obtained based on the CEM approach and the model obtained with the REM approach.

$H_0: \sigma_u^2 = 0$, Common Effect Model

$H_1: \sigma_u^2 \neq 0$, Random Effect Model

With test statistics:

$$LM = \frac{NT}{2(T-1)} \left[\frac{\sum_{i=1}^N [\sum_{t=1}^T e_{it}]^2}{\sum_{i=1}^N \sum_{t=1}^T e_{it}^2} - 1 \right] \quad (6)$$

Where:

- e_{it} : error of individual "i" of period "t"

Hausman Test

It is necessary to know the best panel data regression model between the model obtained based on the REM approach and the model obtained using the FEM approach.

$H_0: \text{correlation } X_{it}, u_{it} = 0$, Random Effect Model

$H_1: \text{correlation } X_{it}, u_{it} \neq 0$, Fixed Effect Model

With test statistics:

$$W = (\hat{\beta}_{MET} - \hat{\beta}_{MEA})' [var(\hat{\beta}_{MET} - \hat{\beta}_{MEA})]^{-1} (\hat{\beta}_{MET} - \hat{\beta}_{MEA}) \quad (7)$$

Where:

- $\hat{\beta}_{MET}$: Fixed effects model slope estimation vector
- $\hat{\beta}_{MEA}$: Random effect model slope estimation vector

2.3 Panel Data Regression Model Assumption Test

Normality Test

In the normality test based on [12] if the CEM model and FEM model are selected, it is not mandatory to use the normality test, otherwise if the REM model is selected, it is mandatory to use the normality test using the Jarque-Bera (JB) test [13].

$$JB = \frac{n}{6} \left[s^2 + \frac{(k-3)^2}{4} \right] \quad (8)$$

Where:

- n : Sample size
 s : Skewness
 k : Kurtosis

Multicollinearity Test

The multicollinearity test aims to test whether the regression model found a correlation between the independent variables. A good regression model does not correlate with the independent variables. If there is a reasonably high correlation between independent variables (generally above 0.90), then this is an indication of multicollinearity [14].

Heteroscedasticity Test

The heteroscedasticity test is one of the factors that support the coefficient and accuracy of the regression model, where there are residuals from each variable that are not the same. Therefore, it is essential to know whether or not heteroscedasticity exists. One method that can be used to detect heteroscedasticity is the Glejser Test Method [4].

H_0 : There is no heteroscedasticity problem in the regression model

H_1 : There is a heteroscedasticity problem in the regression model

2.4 Panel Data Regression Model Equation Check

Concurrent Test (F-test)

The F test is used to evaluate the effect of the independent variables on the dependent variable together. Therefore, the F test procedure can be stated as follows [15]:

$H_0 : \beta_1 = \beta_2 = \beta_3 = \dots = \beta_n = 0$

$H_1 : \text{there is } \beta_k \neq 0, k=1,2,\dots,n$

With test statistics:

$$F = \frac{\frac{R^2}{K-1}}{\frac{1-R^2}{N-K}} \quad (9)$$

Where:

- R^2 : Determination Coefficient

Partial Test (t-test)

The t test is used to partially evaluate the effect of the independent variable on the dependent variable. Therefore, the t-test procedure can be stated as follows [15]:

$H_0 : \beta_1 = 0$

$H_1 : \beta_1 < 0$

With test statistics:

$$t = \frac{\hat{\beta}_k}{se(\hat{\beta}_k)} \quad (10)$$

Where:

- $\hat{\beta}_k$: Regression coefficients
 $se(\hat{\beta}_k)$: Standard error of regression coefficients

Determination Coefficient

The coefficient of determination (Goodness of fit) is the goodness of fit of a regression line fitted to a data set to determine how well the sample regression line fits the data. Therefore, the coefficient of determination is an overview measure that says how well the sample regression line fits the data [10].

With test statistics:

$$R^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} \quad (11)$$

Where:

- \hat{Y}_i : Estimated value of Y
 \bar{Y} : The average value of Y
 Y_i : Actual value of Y

3 Method

This research uses applied research with data collection tools in the form of documents with secondary data types. The data source in this study is from the Central Bureau of Statistics in West Sumatera Province. Where GRDP data is used to describe the economic growth that occurred in West Sumatera Province for the period 2018 to 2022. The steps of data analysis techniques in this study are as follows:

1. Determining research data variables

In this study used variables, namely:

- a. GRDP at constant prices as the dependent variable.
- b. Human Development Index (X_1), Percentage of Poor Population (X_2), Labor Force Participation Rate (X_3), Literacy Rate (X_4), and Government Spending (X_5) as independent variables.

2. Determining the best panel data regression model

In this study, to determine the best panel data regression model that can show the factors that affect economic growth is as follows:

- a. Determine the model of the three estimation models between the common, fixed, and random effect models.
- b. Select the best model using the Chow test, Lagrange multiplier test, and Hausman test.
- c. Test the panel data regression model's assumptions with normality, multicollinearity, and heteroscedasticity tests.

3. Determining the factors affecting

To determine the factors that influence economic growth, namely:

- a. Testing the significance of panel data regression model parameters.
- b. Interpret the model and conclude the panel data analysis.

4 Result and Discussion

In this study, to analyze and understand the factors that influence economic growth in West Sumatera Province, data on Gross Regional Domestic Product (GRDP) at constant prices is used as the dependent variable, and Human Development Index, Percentage of Poor Population, Labor Force Participation Rate, Literacy Rate, and Government Spending as independent variables. Economic growth data and factors affecting economic growth in West Sumatera Province are 95 data from 19 regencies/cities in West Sumatera Province in 2018-2022. The characteristics of each variable are shown in Table 1. Based on what can be seen in Table 1, the economic growth of 19 regencies/cities in West Sumatera Province during the 2018-2022 time span was highest achieved by Padang City in 2022, reaching 47,185,098.81 million rupiah, and the lowest was achieved by Padang Panjang City in 2018 amounting to 2,444,773.8 million rupiah.

Table 1. Descriptive Statistics of Economic Growth and Factors that Affect It in West Sumatera Province in 2018-2022

Variable	N	Descriptive Statistics		
		Mean	Maximum	Minimum
Y	95	9,165,784.96	47,185,098.81	2,444,773.8
X ₁	95	72.5	83.29	60.28
X ₂	95	6.33	14.84	2.16
X ₃	95	69.41	83.1	61.98
X ₄	95	99.59	100	98.1
X ₅	95	1,007,155,202.93	2,351,499,268.94	494,650,066.51

4.1 Determining the Panel Data Regression Estimation Model

Determining the estimation of panel data regression models that describe the factors affecting economic growth using three estimation methods, as follows:

Chow Test

The results obtained based on data processing using Eviews 12 software are shown in Table 2.

Table 2. Chow test

Effects Test	Statistic	df	Prob.
Cross-section F	243.303362	(18.71)	0.0000
Cross-section Chi-square	393.117883	18	0.0000

Based on Table 2, it can be seen that the Cross-section F probability value of 0.000 is smaller than the significance level of 5% ($0.000 < 0.05$). So H_0 is rejected, which means that the most appropriate model used using the Chow test is the fixed effect model. Then, the next test is carried out.

Lagrange Multiplier Test

The results obtained based on data processing using Eviews 12 software are shown in Table 3.

Table 3. Lagrange Multiplier Test

	Test Hypothesis		
	Cross-section	Time	Both
Breusch Pagan	109.3094 (0.0000)	0.030998 (0.8602)	109.3404 (0.0000)

Based on Table 3, it can be seen that the Breusch Pagan probability value of 0.000 is smaller than the significance level of 5% ($0.000 < 0.05$). So that H_0 is rejected, which means that the most appropriate model used using the Lagrange multiplier test is the random effect model. So, to ensure that the more appropriate model between the model obtained with the random effect model approach and the fixed effect model approach is used, the Hausman test is used.

Hausman Test

The results obtained based on data processing using Eviews 12 software are shown in Table 4.

Table 4. Hausman Test

Test Summary	Chi-Sq. Statistic	Chi-Sq. df	Prob.
Cross-section random	87.511625	5	0.0000

Based on Table 4, it can be seen that the cross-section random probability value of 0.000 is smaller than the significance level of 5% ($0.000 < 0.05$). So that H_0 is rejected, which means that the most appropriate model used using the Hausman test is the fixed effect model. So, to ensure that the more appropriate model between

the model obtained with the random effect model approach and the fixed effect model approach is used, the Hausman test is used.

Thus the best model to describe the factors affecting economic growth that is appropriate for use in this study is the fixed effect model.

4.2 Panel Data Regression Model Assumption Test

After determining the best model, the best model obtained in this study is the fixed effect model. The next step is to test the assumptions of the panel data regression model. For testing the assumptions of the panel data regression model, because the best model chosen is the fixed effect model, it is unnecessary to use the normality test in this study. So, the panel data regression model assumption test used in this study is only a multicollinearity test and a heteroscedasticity test. The assumption tests are as follows:

Multicollinearity Test

The results obtained based on data processing using Eviews 12 software are shown in Table 5.

Table 5. Multicollinearity Test

Variable	X_1	X_2	X_3	X_4	X_5
X_1	1	-0.74	-0.62	0.45	0.05
X_2	-0.74	1	0.59	-0.53	0.10
X_3	-0.62	0.59	1	-0.27	-0.27
X_4	0.45	-0.53	-0.27	1	0.05
X_5	0.05	0.10	-0.27	0.05	1

Based on Table 5, it can be seen that the correlation coefficient between independent variables is smaller than 0.9. This indicates that the resulting regression model is free from multicollinearity.

Heteroscedasticity Test

The results obtained based on data processing using Eviews 12 software are shown in Table 6.

Table 6. Heteroscedasticity Test

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	10186315	12820878	0.794510	0.4295
X_1	-7872.324	66965.37	-0.117558	0.9067
X_2	-3562.994	106545.4	-0.033441	0.9734
X_3	-1194.564	14107.16	-0.084678	0.9328
X_4	-92036.61	120932.3	-0.761059	0.4491
X_5	-0.000146	0.000386	-0.377789	0.7067

Based on Table 6, it can be seen that the probability value of each independent variable is greater than the significance level of 5%. So that H_0 is accepted, showing that the resulting regression model is free from heteroscedasticity problems.

4.3 Panel Data Regression Model Equation Check

Then, the next step is to check the panel data regression model equation by knowing the significance of the panel data regression model parameters. The parameter significance test is as follows:

Concurrent Test (F-test)

The results obtained based on data processing using Eviews 12 software are shown in Table 7.

Table 7. *F* test

Variable	F-Statistic	Probabilitas F	Description
X ₁			
X ₂			
X ₃	1,676.876	0.0000	Significant
X ₄			
X ₅			

Based on Table 7, it can be seen that the value of *F* is 1,676.876 greater than F_{table} is 2.317 and the *F* probability value of 0.000 is smaller than the significance level of 5%. So H_0 is rejected, which means that the independent variables simultaneously have a significant influence on the economic growth variable in West Sumatera Province.

Partial Test (t-test)

The results obtained based on data processing using Eviews 12 software are shown in Table 8.

Table 8. *t* test

Variable	Coefficient	t-Statistic	Prob	Description
X ₁	545589.6	4.608309	0.0000	Significant
X ₂	33670	0.178745	0.8586	Insignificant
X ₃	-7858.5	-0.315084	0.7536	Insignificant
X ₄	-191005.6	-0.893366	0.3747	Insignificant
X ₅	0.001445	2.115207	0.0379	Significant

Based on Table 8, it can be seen that the human development index variable (X₁) and the government spending variable (X₅) are variables that have a significant influence partially or individually on economic growth in West Sumatera Province.

Determination Coefficient

The results obtained based on data processing using Eviews 12 software are shown in Table 9.

Table 9. Test Coefficient of Determination

Variable	R-square
X ₁	
X ₂	
X ₃	0.998162
X ₄	
X ₅	

Based on Table 9, the results of the calculation on the fixed effect model show that the effect of the independent variables on economic growth in West Sumatera Province is 0.998162. This means that the independent variables in this study can explain 99.8162% of economic growth, with the remaining 0.1838% of other variables not included in the variables studied.

Thus, this study obtained a panel data regression model with the best model, namely the fixed effect model (FEM). With the panel data regression equation model is as follows:

$$\begin{aligned} \hat{Y}_{it} = & -54,397.51 + 989,017.7D_{1i} + 2,357,720D_{2i} + 304,145.7D_{3i} + \\ & 503,718.2D_{4i} + 4,057,760D_{5i} + 5,015,301D_{6i} + 3,356,187D_{7i} + \\ & 404,111.5D_{8i} - 2,973,171D_{9i} - 876,025.5D_{10i} + 4,309,334D_{11i} + \\ & 28,026,408D_{12i} - 8,811,567D_{13i} - 5,707,275D_{14i} - 8,877,220D_{15i} - \\ & 6,946,187D_{16i} - 7,700,080D_{17i} - 7,377,780D_{18i} + 545,589.6X_1 + \end{aligned}$$

$$0.001445X_5$$

In the fixed effect equation model above, it can be seen that the factors that affect economic growth in West Sumatera Province from 2018 to 2022 are the human development index (X_1) and government spending (X_5).

5 Conclusion

Based on the description of the results and discussion in this study, the following conclusions can be obtained:

- a. The regression model used in the panel data regression analysis of this study is the fixed effect model (FEM) which describes the factors that influence economic growth in West Sumatera Province in 2018-2022. With the equation model formed, namely:

$$\begin{aligned} \hat{Y}_{it} = & -54,397.51+989,017.7D_{1i}+2,357,720D_{2i}+304,145.7D_{3i}+ \\ & 503,718.2D_{4i}+4,057,760D_{5i}+5,015,301D_{6i}+3,356,187D_{7i}+ \\ & 404,111.5D_{8i} - 2,973,171D_{9i} - 876,025.5D_{10i}+4,309,334D_{11i}+ \\ & 28,026,408D_{12i} - 8,811,567D_{13i} - 5,707,275D_{14i} - 8,877,220D_{15i} - \\ & 6,946,187D_{16i} - 7,700,080D_{17i} - 7,377,780D_{18i}+545,589.6X_{1i}+ \\ & 0.001445X_5 \end{aligned}$$

- b. Factors that have a significant influence and have a positive relationship to economic growth in West Sumatera Province in 2018-2022 are the human development index (X_1) and government spending (X_5).

References

- [1] A. Rahardjo, "Pertumbuhan Ekonomi Wilayah," in *Pengelolaan Pendapatan dan Anggaran Daerah*. Yogyakarta: Graha Ilmu, 2014.
- [2] Badan Pusat Statistik, *Pertumbuhan Ekonomi*, 2021. Indonesia: Badan Pusat Statistik. Accessed: Sept. 15, 2023. [Online]. Available: <https://www.bps.go.id/id>
- [3] Badan Pusat Statistik Provinsi Sumatera Barat. *Provinsi Sumatera Barat Dalam Angka 2022*, 2022. Sumatera Barat: : Badan Pusat Statistik.
- [4] A. Widarjono, "Model Regresi Data Panel," in *Ekonometrika Teori dan Aplikasi*, 2nd ed. Yogyakarta: Ekonisia FE UII, 2007.
- [5] M. Hasan, and M. Azis, *Pembangunan Ekonomi & Pemberdayaan Masyarakat*, 2nd ed. CV. Nur Lina. 2018.
- [6] G. Ranis, *Human Development and Economic Growth*. Center Discussion Paper No. 887. Amerika Serikat: Yale University. 2004.
- [7] V. Octaviany, "Pengaruh Kualitas Produk Pariwisata Terhadap Keputusan Berkunjung Di Bale Seni Barli-Kota Baru Parahyangan," In *Tourism Scientific Journal*, 2016, Vol. 1 No. 2, pp. 184-195. doi: 10.32659/tsj.v1i2.11.
- [8] L. V. Yuhendri, "Pengaruh Kebutuhan Akan Prestasi, Lokus Kendali, dan Efikasi Diri Terhadap Minat Berwirausaha Fakultas Ekonomi Universitas Negeri Padang," in *Jurnal Penelitian Pendidikan*, 2014, Vol. 14 No. 1. doi: 10.17509/jpp.v14i1.3212.
- [9] J. Sodik, "Pengeluaran Pemerintah dan Pertumbuhan Ekonomi Regional: Studi kasus data panel di Indonesia," *Jurnal Ekonomi Pembangunan*, 2007, Vol. 12 No 1. doi: 10.20885/ejem.v12i1.516.
- [10] D. N. Gujarati and D. C. Porter, "Panel Data Regression Models," in *Basic Econometrics*, Fifth Edition. McGraw-Hill, New York, 2009.
- [11] B. H. Baltagi, *Econometric Analysis of Panel Data*, Third Edition. USA: John Wiley & Sons, 2005.
- [12] M. Kuncoro, *Metode Riset untuk Bisnis dan Ekonomi*, Fourth Edition. Jakarta: Erlangga, 2013.

- [13] W. W. Winarno, *Analisis Ekonometrika dan Statistika dengan EViews*, 2nd ed. Yogyakarta: UPP STIM YKPN 1, 2006.
- [14] I. Ghozali, *Aplikasi Analisis Multivariate dengan Program IBM SPSS19*, Fifth Edition. Semarang: Program Doktor Ilmu Ekonomi Universitas Diponegoro. CV. NUR LINA, 2011.
- [15] D. C. Montgomery, E. A. Peck, and G. G. Vining. (2012). "Hypothesis Testing on The Slope and Intercept," in *Introduction to Linear Regression Analysis*. Fifth Edition. US: John Wiley & Sons.