

CHAPTER III RESEARCH METHOD

3.1. Types and Model of Panel Data

The type of study conducted by researcher was quantitative research. This research used quantitative methods by generating numerical data or data that can be transformed into useable statistics. The data model used in this study were secondary data, as: the Gross Regional Domestic Product (PDRB) report of Indonesia based on the 2010 constant prices by province, government expenditure, Human Development Index (HDI), worker and investment. Those data were issued by the Central Statistics Agency (BPS) and the Directorate General of Financial Balance (DJPB).

Data required in this research were:

1. Indonesia's provincial Gross Regional Domestic Product (GRDP) based on constant Prices 2010 by Province, the year 2013-2018.
2. The size of government expenditure data in Indonesia by Province in 2013-2018.
3. Human Development Index (HDI) data in Indonesia by Province in 2013-2018.
4. The number of worker data in Indonesia by Province in 2013-2018.
5. The volume of investment covers domestic and foreign direct investment realization by Province in 2013-2018.

3.2. Data Sources and Definition

In collecting data this study uses a literature study by processing data and analyzing literature publications. In this case, the data is collected to obtain accurate and precise information. The data used were secondary data using the original data of Indonesia's Gross Regional Domestic Product (GRDP) in Indonesia by Provinces as well as other literature sources related to research. The tool used for statistical testing was Microsoft Excel 2013, while for Panel Data testing was EViews 8 by entering the data into Microsoft Excel 2013 software in the .xlsx format, then imported into EViews 8 software to be tested.

3.3. Research Variable

This research contained independent variable and dependent variable. The dependent variable in this research was Indonesia's Gross Regional Domestic Product (GRDP) by Provinces and the independent variables were Government expenditure, worker, Human Development Index (HDI), and investment, those could be defined as follows:

3.3.1 Dependent Variable

The dependent variable is a variable of magnitude, which influenced by other variables. This study used Indonesia's Gross Regional Domestic Product (GRDP) as a dependent variable (Y). GRDP is all goods and services as a result of economic activities operating in the domestic area (BPS, 2018). Based on these explanations, the researcher used GRDP as an indicator of economic growth, while the data used

are data according to constant prices because the effect of price changes or inflation has been eliminated so that it is more representative of the economy in real term.

3.3.2 Independent Variable

The independent variable is the variable that can affect another variable.

Independent variables used in this study are:

A. Government expenditure

Government expenditure is government routine expenditure every year in the context of organizing, implementing and maintaining government activities. According to Sukirno (2006) government expenditure is part of fiscal policy, namely a government action to regulate the economy by determining the amount of revenue and annual government expenditure, which is reflected in the National Budget (APBN) and the Provincial Budget (APBD) for the region or region. The purpose of this fiscal policy is to stabilize prices, output levels, and employment opportunities and stimulate or encourage economic growth. Hence, researcher used the summary of actual expenditures of Provincial government in period 2013-2018.

B. Human Development Index (HDI)

The Human Development Index (HDI) as an indicator of human capital is a comparative measurement of life expectancy, education, and living standards for all countries. HDI is used as an indicator to assess the quality aspects of development and classify whether a country is a developed country, a developing country, or an underdeveloped country and also to measure the effect of economic

policies on quality of life. (BPS, 2015). The range used is the number from 0 to 100, where the number 0 is the lowest number and 100 is the highest number.

In this study, researcher used the latest method of HDI because it uses indicators that are more precise and can distinguish well (discriminatory) by doing the following:

- Replacing literacy rates with an average length of schooling and long-term expectation rates, it is hoped that a more relevant picture can be obtained in education and the changes that occur.
- Replacing GDP with GNP because it is more representative of people's income in an area.

C. Worker

In-Law No. 14 of 1969 concerning "Basic Labor Provisions" states that workers are any person who is able to do work both inside and outside the employment relationship to produce goods or services to meet the needs of the community and receive similar wages or rewards. In this case, the author used data on the number of workers per province in Indonesia each year from 2013-2018 for all types of workforce whether educated, trained, or uneducated and trained. In this research the authors use the worker who involved in economic activities as independent data because its contribution is more significant to gross regional domestic product (GRDP) of a region.

D. Investment

Investment activities have a very important role in economic growth that is as a driver of income in a region. The role of investment in the economy in addition

to spurring economic is to absorb labor and can be as an expansion of business opportunities. The author used the combination of both realization of DDI and FDI because it measures the investment as a whole. The middle exchange rate of the United States Dolar (USD) against the Indonesian Rupiah (IDR) issued by BPS is used in converting the USD value of FDI to IDR.

3.4. Analysis Technique

In order to be the result of the relationship between dependent and independent variables, the stages are as follows:

3.4.1 Panel Data Method

Panel data is the combination of from both time series and cross-section data. According to Gujarati (2003) there is some advantages in the use of panel data, namely; can consider heteroskedasticity by introducing specific variables, reduces inter-variable collinearity, of data panels also makes a greater degree of freedom (*df*), where the estimation results are better use cross-section analysis or time series and panel data integration makes the data more efficient, informative, less collinearity and minimize bias. The regression data panel has three approach of estimation models, namely the common effect model (CEM), the fixed effect model (FEM), and the random effect model (REM).

A) Common Effect model

In this approach, all cross-section units and time series are treated the same and then regressed using the ordinary least square method which will produce

equations with constant intercepts and coefficients of independent variables for each unit. The following is a regression model for this model:

$$Y_{it} = \beta_0 + \beta_1 X + \beta_{2it} X_{it} + e_{it} \dots \dots \dots (3.1)$$

This method is the simplest method but the results are not adequate because each observation is treated as a stand alone observation, so it is quite likely that the error term correlates with several independent variables in the model. Another obstacle that is owned by this model is an assumption that considers the same intercept and slope coefficients for each cross section and time series unit. Overcome this use Fixed Effect Model or can be said as Least Square Dummy Variable.

B) Fixed Effect Model

The Fixed Effect model approach assumes that the intercept of each individuals are different while the slope between individuals is fixed (the same). In other words, in the fixed effect model there is no difference in time variant but there are differences in intercepts between cross sections. The estimation model using the Fixed Effect Model as follows:

$$Y_{it} = \beta_0 + \beta_1 X_1 + \beta_{2it} X_{2it} + e_{it} \dots \dots \dots (3.2)$$

For estimating the fixed effect model (FEM) needed a dummy variable to satisfy Different intercepts - differences between individuals, intercept differences can occur because of differences in work culture, management and incentives. Hence, this estimation models often called technically Least Square Dummy Variable (LSDV).

C) **Random Effect Model**

The approach used in the Random Effect assumes that each company has different intercepts, which intercepts are random or stochastic variables. This model is very useful if the individuals (entities) taken as a sample are chosen randomly and are representative of the population. The following is a regression model for this model:

$$Y_{it} = \beta_0 + \beta_1 X_{1i} + \beta_{2it} X_{2it} + \beta_0 + \beta_1 X_{1it} + \beta_3 X_{1i} + \dots + e_{it} \dots\dots\dots (3.3)$$

According to Widarjono (2009) the random effect model used to overcome the weaknesses of the fixed effect model that it uses dummy variable. Panel data analysis method with a random effect model must be meet the requirements, namely the number of cross sections must be greater than the amount research variable. The advantage of using the random effect model is eliminate heterokesdasticity. This model is often called the Error model Component Model (ECM) or Generalized Least Square (GLS) technique.

3.4.2 Selection Panel Data Estimation

This study uses panel data regression in analyzing the influence of government expenditure, HDI, worker, and investment on Gross Regional Domestic Product (GRDP) in Indonesia by provinces from 2013-2018, where cross-section data are from the three estimation techniques, one of the most appropriate techniques will be chosen to estimate panel data regression. The selection is based on the following tests:

D) Chow Test

Chow test is used to test the best model in explaining data between the Common Effect Model (CEM) and Fixed Effect Model (FEM). In this test the hypothesis is as follows:

H₀: Common Effect Model (CEM) is better than the Fixed Effect Model (FEM).

H₁: Fixed Effect Model (FEM) is better than Common Effect Model (CEM).

If $F_{stat} > F_{table}$ then H₀ is rejected and it can be concluded that the best model is the Fixed Effect Model (FEM). Conversely, if $F_{stat} < F_{table}$ then H₀ is accepted and it can be concluded that the best model is the Common Effect Model (CEM).

E) Hausman Test

Hausman test was conducted to determine between Fixed Effect Model (FEM) and Random Effect Model (REM) as the appropriate model that should be used. The Hausman test is calculated using the equation as follows:

$$H = X_{(k)}^2 = (\beta_{re} - \beta_{fe}) (\Sigma_{re} - \Sigma_{fe})^{-1} (\beta_{re} - \beta_{fe}) \dots\dots\dots(3.4)$$

Where:

β_{re} : Random method estimator

β_{fe} : Fixed effect estimator method

Σ_{re} : Covariance coefficient matrix on the random effect method

Σ_{fe} : Covariance coefficient matrix on the fixed effect method

k : free degrees (number of parameters)

Hausman test statistic follows the Chi-Square statistic distribution with a degree of freedom is k , where k is the number of independent variable. Thus, we can see the result of Chi-square. The hypotheses proposed are the following:

H_0 : Random Effect Model (REM) is better than the Fixed Effect Model (FEM).

H_1 : Fixed Effect Model (FEM) is better than Random Effect Model (REM).

When chi-square table is greater than chi-square statistic means accept H_0 then Random Effect Model (REM) is better and reversely if the chi-square statistic is greater than the chi-square table, Fixed Effect Model (FEM) is better while rejecting H_0 .

3.4.3 Hypothesis Testing

A) Coefficient Determinants (R^2)

Coefficient determination (R^2) is an important measurement in the regression because it will determine that the regression model is good or not. The coefficient of determination (R^2) is used to measure how far the model's ability to explain the variation of the dependent variable (Gujarati, 2003). If it finds that R^2 is zero then the variation of the Y cannot be explained by X altogether. Otherwise, if R^2 is one then a variation of the Y can be explained by X altogether. For that reason, it can be concluded that the greater R^2 is the better regression model.

B) t-Test

The t-test is used to know the effect of the significance of independent variable individually over the dependent variable, T-test Hypothesis is:

H_0 = independent variable does not influence the dependent variable significantly

H_1 = independent variable influenced the dependent variable significantly.

When probability value is greater than alpha means reject H_0 and accept H_1 . Reversely, when the result of alpha greater than probability means accept H_0 , which indicates no significant influence of independent variable in dependent variable.

C) F-test

The F-test is used to explain the effect of independent variables on the dependent variable. F-test has hypothesis below:

H_0 = No independent variable influenced significantly the dependent variable.

H_1 = At least one independent variable influenced significantly the dependent variable.

If F-test is greater than F critical, H_0 is rejected. Rejected H_0 means that there is at least one independent variable that is influenced by the dependent variable. And conversely if the F critical is greater F-test then there is no independent variable (X) influenced significantly to the dependent variable (Y) and the study cannot continue further.

3.5. Model

The influence of independent variable toward the dependent variable systematically can be described in the following formula:

$$\log(Y_{it}) = \beta_0 + \beta_1 \log(X_1) + \beta_{2it} \log(X_{2it}) + \beta_{3it} \log(X_{3it}) + \beta_{4it} \log(X_{4it}) + e_{it} \dots \dots \dots (3.4)$$

Where:

Y: Gross Regional Domestic Product (GRDP)

X₁, X₂, X₃, and X₄: government expenditure (X₁), human development index (X₂), worker (X₃) and investment (X₄).

β_0 : Constanta

$\beta_1, \beta_2 \dots \beta_n$: The magnitude of the influence of the independent variable toward the dependent variable

i: 34 Provinces in Indonesia

t: Series 2013-2018

e_{it} : Error or residual term