CHAPTER III

RESEARCH METHODOLOGY

3.1 Scope of research

This research is analyzing the Demographic Factors to Foreign Direct Investment: Case Study in five developing ASEAN countries specifically Indonesia, Malaysia, Philippines, Thailand and Vietnam. The processed data used in this research come from World Bank Data. Demographic factors used in this research is Total population, Life expectancy, and Population Density, meanwhile for Economic factors is growth rate of GDP.

3.2 Data

The type of data used in this study is secondary data. This study covers five countries in ASEAN from 2008-2015. Secondary data in this research is retrieved from World Bank.

3.3 Method Analysis

This research uses descriptive and quantitative analysis. Descriptive analysis is used for finding the result of previous studies related to the variables used in this research. The reason behind this analysis is to support the result of research by referencing studies beforehand, creating a compatibility proof for research result. Meanwhile, Quantitative analysis use econometric model for determine the results of the study mathematically.

3.3.1 Econometric Model

This research use panel regression model with data from five countries. Therefore, the econometric models used can be described as follows:

- $ln\_fdi = \quad \beta_0 \quad + \beta_1 \ln\_Totpop_{it} + \beta_2 life\_ex_{it} +$

$$+ \beta_4 Density\_pop_{it} + \beta_4 gdpgrowth_{it} + \varepsilon_{it} \qquad (1)$$

Explanation :

- *ln_fdi*: foreign direct investment inflow for each country from 2008-2015 in US Dollars in natural log form.

- *gdpgrowth*: Annual percentage growth rate of GDP for each country from 2008-2015

- *Ln_Totpop*: Total Population for each country from 2008-2015 in natural log form.

- Life_ex: Life expectancy at birth for each country from 2008-2015

- Density_pop: Population density (people per sq. km of land area) for each country from 2008-2015

3.4. Operational Variables

Table 3.1 Operational Variables

Variable	Notes	Definition	Data source
Foreign Direct Investment	Ln_fdi	Foreign direct investment inflow for each country from 2008-2015 in US Dollars in natural log form.	World Bank
Annual percentage growth rate of GDP	Gdpgrowth	Annual percentage growth rate of GDP from 2008-2015	World Bank
Population Density	Density_pop	Population Density for each country from 2008-2015	World Bank
Life expectancy	Life_ex	life expectancy at birth for each country from 2008-2015	World Bank
Total Population	Ln_Totpop	Total Population for each country from 2008-2015 in natural log form	World Bank

3.5 Statistical test

3.5.1 Hausman Test

In Hausman test, if random effect is rejected it means fixed effect is more suitable model for this research. The purpose of this test is to determine which model is the best for panel data regression (Jefrey M Wooldridge, 2010). In this research, Hausman test is applied by STATA. Hypothesis for the Hausman test is:

- H0 : Random Effect Model

- H1 : Fixed Effect Model

  If P-Value smaller than α value, H0 is rejected. In the other hand, H0 cannot be rejected if P-value bigger than α value. α value can be varied between 1%, 5%, and 10%.

3.5.2 Coefficient Determination

Coefficient determination or R^2 test is a measuring instrument to find out how much contribution of independent variables toward dependent variables. Coefficient determination (R^2) value is reflect how big a variation from y (Dependent variable) can be explained from x (Independent variable). If the coefficient determination value is 0, it concludes that variance of y cannot be explained by x at all. Meanwhile if the value of $R^2 = 1$, the variance of y is explained fully by x. the value of R^2 have value between 0 and 1.

3.5.3 Classical assumption

Classical assumption is a procedure for determine autocorrelation, heteroskedasticity, and multicollieanirity problem. If the research model is using generalized least square (GLS), all classical assumption is no need to be done. Generalized least square is usually found in random effect (Gujarati, 2004). In conclusion, classical assumption isn't needed in random effect. If the regress equation is fixed or common effect, classical assumption is should be done.

3.5.4 Heteroskedasticity

Heteroskedasticity appear when a residual value of model did not have any constant variance. This problem usually occurred in cross section data and hard to be avoided at panel data (Gujarati, 2004). Heteroskedasticity job is to makes estimator not have any minimum and efficient variance, but still does not eliminate the irregularity and consistency of estimators. If the research model use fixed effect, Wald test must be concluded. Meanwhile for random effect, no heteroskedasticity test need to be examined because of GLS estimation. Hypothesis for heteroskedasticity test is shown below:

Heteroskedasticity test hypothesis

- Ho: No heteroskedasticity problem
- H1:   there is a heteroskedasticity problem

Criteria:

- Ho is rejected when (Prob>chi) value is less than alpha, so there is a heteroscedasticity problem in the regression model.

- Ho is not rejected when (Prob>chi) value is bigger than alpha, so there is not any heteroscedasticity problem in the regression model.

3.5.5 Autocorrelation

Autocorrelation happens because error for each individual tend to affect other same individual in the next period. Autocorrelation is a problem that concur because residuals that are not free between one observations to another observation (Kuncoro, 2011), This phenomenon usually happens at time series data. The hypothesis for autocorrelation test problem is shown below:

Autocorrelation test hypothesis

- Ho: no autocorrelation problem
- H1:   there is an autocorrelation problem

Criteria:

- Ho is rejected when (Prob>chi) values is less than alpha, so there is an autocorrelation problem in the regression model.

- Ho is not rejected when (Prob>chi) values is bigger than alpha, so there is an autocorrelation problem in the regression model.

3.5.6 Multicollinearity

Multicollinearity test shows if there is any relationship between each variables that explained in the regression model. If the independent variables correlates with each other, then the variable that correlate strongly with other variables in the model, will have weak predictive strength and unstable. How to detect multicollieanirity can be seen from the following cases:

1. When a model have high R^2 value, but the t-statistic parameter coefficient significance is low.

2. The coefficient variable is not fit with hypothesis, when the variable should have a positive impact but shown as negative value.

3. If a model have correlation value bigger than 0.8 (Rule of Thumb) it means that variable have Multicollinearity problem (Gujarati, 2004).

3.6   Research object

3.6.1   Foreign Direct Investment (FDI)

Foreign direct investment in this context is foreign direct investment net inflows in current U.S. Dollars. This FDI data is retrieved from World Bank data from 2008-2015 in U.S Dollars currency unit ($).

3.6.2   Annual percentage growth rate of GDP

Growth rate GDP data are taken from five ASEAN countries in 2008-2015. This data is retrieved from World Bank data.

3.6.3   Population Density

Population density is midyear population divided by land area in square kilometers. Population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship--except for refugees not permanently settled in the country of asylum, who are generally considered part of the population of their country of origin. Land area is a country's total area, excluding area under inland water bodies, national claims to continental shelf, and exclusive economic zones. In most cases the definition of inland water bodies includes major rivers and lakes. Population density expressed as people per sq. km of land area. Population Density data is taken from 5 ASEAN countries in 2008-2015. This data is retrieved from World Bank data.

3.6.4   Life Expectancy

Life expectancy is the average number of life years carried out by babies at the birth. Life expectancy can reflect the quality of health of one country. Life expectancy data is retrieved from World Bank data from 2008-2015 in 5 ASEAN Countries. This data estimate unit in Percentage (%)

3.6.5   Total Population

Total Population, expressed as a unit. Population data is retrieved from World Bank data with a span of time 2008 to 2015 in 5 ASEAN Countries