CHAPTER III

RESEARCH METHODOLGY

A. Subject of Research

This study will analyze the relationship between the variable military costs with variables of economic growth, unemployment and investment in 5 countries in Southeast Asia. The time of the study includes the development of these variables for 10 years, from 2009 to 2018.

B. Type of Data

The analysis technique used in this study is qualitative analysis and quantitative analysis. Qualitative analysis was carried out using a variety of literature studies, books, and articles in accordance with this research topic which were used as a reference. Furthermore quantitative analysis uses the econometrics model to explain the relationship between variables. This study uses data from 5 ASEAN countries including Indonesia, Malaysia, Singapore, Thailand, the Philippines, the observation period chosen was 2009 to 2018 and this study uses panel data.

The data used in this study is panel data. Referring to Gujarati and porters (Gujarati And Porter, 2009), Baltagi explained several advantages in using panel data, namely:

1. Individual heterogeneity can be controlled by panel data
2. Data are more informative, more varied, degrees of freedom are more efficient and can avoid colinearity between variables due to the high

3. number of observations. Panel data can estimate each individual characteristic and time characteristic separately
4. Can identify and measure the effect that normally cannot be detected by cross section or time series data

This research will use one of the two panel data estimation model approaches, namely the Fixed Effect Model (FEM) approach and the Random Effect Model (REM) approach. Here are some criteria or considerations that can be used as a reference in choosing the use of the fixed effect model or random effect model according to (Gujarati and Porter, 2009):

1. Time series data is greater than the amount of cross section data. In this case, the fixed effect model is a better choice
2. The time series data is smaller than the amount of cross section data, the estimation generated by the fixed effect model and the random effect model can differ significantly. Individually or by cross section, the units in the study sample do not provide a random picture of a larger sample, so the fixed effect model method is most appropriate.
3. If the error component has a correlation with one or more independent variables, it will happen biased on the results of the REM method, while the results of FEM will be unbiased
4. If the amount of cross section data is greater than the amount of time series data, and the assumptions based on the random effect model, the random effect model estimator is more efficient than the fixed effect model estimator

C. Sampling Technique

The analytical method used in this research is Regression Data Panel Data processing techniques performed using the Eviews7 computer program Hypothesis testing is done, unit root test REM and FEM models that will be used free from violations of classical assumptions (normality, heteroscedasticity, multicollinearity and autocorrelation test) so that the test results can be interpreted appropriately.

D. Operational Definition of Variables

Operational definition is given to variable by giving meaning or specifying the activities needed to measure these variables. The operational definitions used in this study are:

1. Military Spending

   Military expenditure is one of the most important parts in a country where each country has an allocation of funds for military expenditure and The cost of defense for security is one of the important things that must be managed by the state. The data of military spending in this study was measured by the large number of Military Spending in 5 Asian country in World Bank in 2009-2018 which were reported in USD and annual data.

2. Economic growth

   Economic growth is a process of changing a country's economic conditions on an ongoing basis towards better conditions for a certain

period. Economic growth data is taken from the World Bank and Economic growth in this study is expressed in USD and uses annual data.

3. Unemployment

Unemployment is the number of workers in the economy who are actively looking for work but have not found one. The data of Unemployment is taken from the world bank and in this study is expressed in units of percent and uses annual data.

4. Investment

Investment is defined as expenditures or expenditures from capital investors or companies to buy capital goods and also equipment to improve capabilities producing goods and services available in the economy. The data of Investment is taken from the world bank and in this study is expressed in units of USD and uses annual data.

E. Data Analysis Instrument

In this study, the author used several statistical programs for processing secondary data that has been collected from several sources, such as Microsoft Excel 2010 programs and Eviews7. Microsoft Excel 2010 is used for processing data concerning making tables and analysis as well as log transformation. Meanwhile, regression processing uses the writer Eviews7 computer program.

F. Analysis and Hypothesis Test

The model used in this study is the smallest economic square model that refers to research conducted by (Korkmaz, 2015). Research examines the relationship between economic growth, unemployment and investment in military costs. This study uses panel data of 10 ASEAN member countries with a period of 2009-2018.

To examine the impact of defense costs on economic growth and unemployment, the authors adapted the literature. This research model becomes:

$$ME\iota\tau = \beta 0 + \beta GG\iota\tau + \beta 2UE\iota\tau + \beta 3INV\iota\tau + \mu\iota\tau$$

Where :
ME= Military Expenditure
β0 = Coefficient
GG= GDP Growth
UE= Unemployment
INV= Investment
μ= Error Term

There are two types of variables in this study, depending on the variable and the independent variable. Sugiono revealed that the research variable is an attribute or value of people, objects or activities that have certain variations determined by researchers to be studied and concluded (Sugiyono, 2009). While the definition of dependent and independent variables according to Sugiono is an independent variable (independent) is a variable that affects or is the cause of changes or the emergence of a dependent variable (dependent) while the dependent variable (dependent) is a variable that is

influenced or which is due, because of the independent variable (Sugiyono, 2009).

The dependent variable in this study is the military costs of 5 ASEAN countries including Indonesia, Malaysia, Thailand, Thailand, Thailand and the Philippines, data obtained from the World Bank. There are two independent variables that are used, namely the GG of GDP growth of 5 ASEAN countries, Unem of unemployment of 5 ASEAN countries, investment of 5 ASEAN countries investment. All data on independent variables were obtained from the World Bank from 5 countries in ASEAN. The variables that will be used in this study can be seen specifically in the following table.

Table 3.1

Operasionalisasi Variabel

Variabel	**Notasi**	**Definition**	**Source**
Dependent	ME	Indicate the percentage of the amount of military expenditure in 5 ASEAN countries in 2009-2018	World Bank
Independent	GG	Indicate the percentage of the amount of military expenditure in 5 ASEAN countries in 2009-2018	World Bank
Independent	UE	Shows the percentage of GDP Growth in 5 ASEAN countries in 2009-2018	World Bank
Independent	INV	Shows the percentage of Unemployment in 5 countries ASEAN 2008-2019	World Bank

G. Hypothesis Test

In this research using data panel data regression method, Panel data is a combination of time series and cross section data, according to Agus Widarjono (2009) the use of panel data in an observation has several advantages: First, panel data is a combination of two time series and cross data. section is able to provide more data so that it will result in greater degrees of freedom Second, combining information from time series data and cross sections can overcome problems that arise when there are omitted-variable problems.Hsiao (1986), notes that the use of panel data in economic research has several main advantages over cross section and time series data types. First, it can provide researchers with a large number of observations, increase degrees of freedom, data have large variability and reduce collinearity between explanatory variables, which can produce efficient econometric estimates. Second, panel data can provide more information that cannot be provided only by cross section or time series data. And Third, panel data can provide better resolution in dynamic change inference compared to cross section data.According to Wibisono (2005) advantages of panel data regression include: First. The data panel is able to calculate individual heterogeneity by allowing individual specific variables. Second. This ability to control heterogeneity further makes panel data usable for testing and building more complex behavioral models. Third, panel data is based on repeated cross-section observations (time series), so that the panel

data method is suitable for use as a study of dynamic adjustment. Fourth, the high number of observations has implications for data that are more informative, more varied, and collinierity (multico) between data decreases, and higher degrees of freedom (df) so that more efficient estimation results can be obtained. Fifth, panel data can be used to study complex behavioral models. And Sixth, panel data can be used to minimize the bias that might be caused by aggregation of individual data , Panel Regression Model from the title above is as follows:

$$Y = \alpha + b1X1it + b2X2it+b3x3 + e$$

Where :

Y = Military Expenditure
a = *Coefficient*
X1 = GDP Growth
X2 = Unemployment
X3 = Investment
e = Error Term
t = Time
i = Country

In the regression model estimation method using panel data can be done through three approaches, including:

1. Common Effect Model

   It is the simplest panel data model approach because it only combines time series data and cross sections. In this model, time and individual dimensions are not considered, so it is assumed that corporate data behavior is the same over various time periods. This method can use the Ordinary Least Square (OLS) approach or the least squares technique to estimate the panel data model.

The regression equation in the common effects model follows:

$$Yit = \alpha + Xit\beta + \varepsilon it$$

Where:

i = Indonesia,Malaysia,Singapore,Thailand and Philipphine

t = 2009,2010,2011,2012,2013,2014,2015,2016,2017,2018

where i shows the cross section (individual) and t indicates the time period. Assuming an error component in ordinary ordinary least square processing, a separate estimation process for each cross section unit can be carried out.

2. Fixed Effect Model

This model assumes that differences between individuals can be accommodated from their intercept differences. To estimate the Fixed Effects model panel data using variable dummy techniques to capture intercept differences between companies, intercept differences can occur due to differences in work culture, managerial, and incentives. However, the slopes are the same between companies. This estimation model is often called the Least Squares Dummy Variable (LSDV) technique (Basuki, 2014).

The formula for FEM it self was :

$$\begin{bmatrix} y^1 \\ y^1 \\ yn \end{bmatrix} = \begin{bmatrix} \alpha \\ \alpha \\ \alpha \end{bmatrix} + \begin{bmatrix} i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & i \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha n \end{bmatrix} + \begin{bmatrix} x_{11} & x_{21} & xp_1 \\ x_{12} & x_{22} & xp_2 \\ x_1 n & x_2 n & xpn \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta n \end{bmatrix} + \begin{bmatrix} \mathcal{E}_1 \\ \mathcal{E}_2 \\ \mathcal{E}_n \end{bmatrix}$$

This technique is called the Least Square Dummy Variable (LSDV). Besides being applied to the effects of each individual, this LSDV can also accommodate time effects that are systemic. This can be done by adding dummy time variables to the model.

3. Random Effect Model

This model will estimate panel data where interruption variables may be interconnected between time and between individuals. In the Random Effect model, intercept differences are accommodated by the error terms of each company. The advantage of using the Random Effect model is to eliminate heteroscedasticity. This model is also called the Error Component Model (ECM) or the Generalized Least Square (GLS) technique (Basuki, 2014).

Thus, the equation of the random effects model can be written as follows:

$$Yit = \alpha + Xit\beta + \varepsilon it$$

Where:

i = Indonesia,Malaysia,Singapore,Thailand and Philipphine

t = 2009,2010,2011,2012,2013,2014,2015,2016,2017,2018

Where:

wit = εit + u1; E (wit) = 0; E (wit2) = $a^2 + au^2$;

E (wit, wjt-1) = 0; i ‡ j; E (ui, εit) = 0;

E (εi, εis) = E (εit, εjt) = E (εit, εjs) = 0

Although the component of error wt is homoscedastic,, namely:

$$\text{Corr (wit, wi (t-1))} = au^2 / (a^2 + au^2)$$

Therefore, the OLS method cannot be used to obtain an efficient estimator for the random effects model. The right method for estimating random effects models is Generalized Least Squares (GLS) with a homokedastic assumption and no cross-sectional correlation.

To choose the most appropriate model to use in processing panel data, there are several tests that can be carried out, namely:

a. Chow Test (likelihood test)

Chow Test is a test to determine the best model between the Fixed Effect Model and the Common / Pool Effect Model. If the results state that they accept the null hypothesis, then the best model to use is the Common Effect Model. However, if the results state that they reject the null hypothesis then the best model used is the Fixed Effect Model, and the test will continue to the Hausman test (Basuki, 2014)

Chow test is a test to determine the Common Effect or Fixed Effect model that is most appropriate to be used in estimating panel data. The hypothesis in the chow test is:

H0: Common Effect Model or pooled OLS

H1: Fixed Effect Model

The basis for rejecting the above hypothesis is to compare the F-statistic calculation with the F-table. Comparison is used if the calculated F is greater (>) than the F table then H0 is rejected, which means the most appropriate model to use is

FixedEffectModel. Vice versa, if the calculated F is smaller (<) than the F table then H0 is accepted and the model used is the Common Effect Model (Widarjono, 2009).

b. Hausman Test

Hausman test is a test to determine the Fixed Effet or Random Effect model that is best used in estimating panel data. The hypothesis in the chow test is:

H0: Random Effect Model

H1: Fixed Effect Model

If the Hausman Test results state that it accepts the null hypothesis, the best model to use is the Random Effect model. However, if the results state that they reject the null hypothesis, the best model used is the Fixed Effect model.

H. Classic Assumptions Test

The classic assumption test is performed to test the feasibility of the regression model so that it can be used for estimation purposes and to reduce data bias. This test is also intended to ensure that in the regression model there is no multicollinearity and heterocedasticity and the data generated are normally distributed.

1. Multicollinearity Test

Multicollinearity (multicollinearity) is a linear relationship between independent variables in multiple regression. If the correlation coefficient is high enough, ie above 0.85, we can expect multicollinearity to occur in

the model. Conversely, if the correlation coefficient is less than 0.85, it is assumed that the model does not contain multicollinearity problems (Widarjono, 2010: 77).

2. Heteroscedasticity Test

This test aims to analyze whether the variance of the error is fixed / constant (homokeditastik) or changeable (heteroscedastic). Detection of heteroscedaticity can be done graphically by seeing whether there are non-random patterns of residual or quadratic residual plots of a dependent variable Y (with an estimated model). Formally, it can also be done by testing the hypothesis: (Rosadi, 2012: 53)

H0: Assumption of homokedasticity is fulfilled

H1: Assumption of homokedasticity is not fulfilled

To detect whether there is heteroscedasticity can use white test with the help of Eviews software. The white test uses squared residuals as the dependent variable, and the independent variable consists of the independent variables that already exist, plus the squares of the independent variables (Winarno, 2015: 15.17).

Heteroscedasticity test means that the variable variance of the disorder is not constant. Heteroscedasticity problems thus appear more often in cross section data than in time series data. One assumption of the OLS method is that the variance of the same interruption variable

(homoskedasticity). To detect whether heteroscedasticity occurs with the White test can be seen from the probability value. (Widarjono, 2010: 84).

In the heteroscedasticity test how to detect with the White test can be seen from the value of the coefficient of determination (R2) and the value of the calculated Chi square obtained from the information value on Obs * R-squared ie the number of observations multiplied by the coefficient of determination seen from the critical value of chi squared (X2) at α = 5% with df (independent variable) of the results seen in the chi squared distribution table. And can be seen also for ensure the absence of heteroscedasticity seen from the value of the chi squares probability must be greater than α = 5%, which means it is not significant and not affected by heteroscedasticity (Widarjono, 2009: 129).

I. Testing Statistics

Below this is a regression equation used in this research, namely:

1. Test T (Partial Significant Test)

   According to Widarjono (2010: 25), t test is used to prove whether the independent variables individually affect the dependent variable. There are two hypotheses proposed by each researcher, namely the null hypothesis (H0) and the alternative hypothesis (H1). To do this test, you can compare the calculated t value with its critical t.

   If rejecting H0 or accepting H1 means statistically the independent variable significantly influences the dependent variable and if accepting

H0 or rejecting Ha means statistically the independent variable does not significantly affect the dependent variable. The decision to reject H0 or accept Ha can also be explained through the probability distribution t. The t value is obtained from the critical t value from the distribution of t table with a certain α and degree of freedom (df). Degree of freedom (df) the number of observations is reduced by the number of estimated parameters (k) and in this case n-k, where n is the number of observations and k the number of estimated parameters (Widarjono: 2010).

According to Iqbal (2005), t test results can be seen from the probability value. If the value of prob. t arithmetic (shown in Prob.) is smaller than the error level (α) 0.05 (which has been determined) then it can be said that the independent variable has a significant effect on the dependent variable, whereas if the value of prob. t count is greater than level error of 0.05, it can be said that the independent variable does not significantly influence the dependent variable.

After knowing that the independent variable influences the dependent variable, the magnitude of the effect of each independent variable can be seen from the value of the regression coefficient (Widarjono, 2010: 29).

2. Test F (Simultaneous Significant Test)

The model reliability test or the model feasibility test or more popularly referred to as the F test (some also call it the simultaneous test

of the model) is the initial stage of identifying a regression model that is estimated to be feasible or not. Decent (reliable) here means that the estimated model is feasible to use to explain the effect of independent variables on the dependent variable. If the value of prob. F count is smaller than the error rate (α) 0.05 (which has been determined) then it can be said that the estimated regression model is feasible, whereas if the value of prob. F arithmetic greater than the error rate of 0.05, it can be said that the estimated regression model is not feasible (Iqbal, 2005).

The F test is used to evaluate the effect of all the independent variables on the dependent variable. F test results can be seen by using the probability value. The probability value (significance) is smaller than α = 5%, meaning that simultaneously the independent variable does not affect the dependent variable. To find the calculated F value and the critical F value from the distribution table F. The critical value is based on the magnitude of α and df where the magnitude is determined by the numenator (k-1) and df for the denominator (n-k). where n = number of observations and k = number of estimated parameters including constants (intercepts) (Widarjono, 2010: 24).

3. Determination coefficient (R^2)

According to Widarjono (2010: 19) this coefficient of determination measures the percentage of the total dependent variable (Y) which is explained by the independent variables in the regression line. The coefficient of determination is applied in this study because there are

more than two independent variables. The coefficient of determination is only between 0 and 1, if a result> 0.5 is obtained, the model used can be said to be convincing in estimating. If the number generated is large, the better the model used in describing the relationship between independent and dependent variable.