CHAPTER III
RESEARCH METHODOLOGY

A. Subject and Object Research

The covered research area is the entire Yogyakarta province which consists of Bantul, Gunung kidul, Kulon progo, Sleman, Bantul and Yogyakarta. An object in this study is using the dependent variable profitability (ROA) whereas the independent variables used are CAR, LDR, BOPO and NPL.

B. Type and Source of Data

The type of data that is needed in this study includes secondary data. Secondary data is data obtained indirectly from the source, such as quotes from books, literature, scientific literature, journals and so on that have peculiar relevance to the theme research. Secondary data in this study is also obtained from a variety of required documents, sourced from Bank Indonesia and supported by the journal, or the previous studies and the relevant literature. This research is conducted in Yogyakarta which covers six districts items, namely from Bantul, Gunung kidul, Kulon progo, Sleman, Bantul, and Yogyakarta. This secondary data is in the form of time series and cross section.

The study period covers the period from 2012 to 2015. The data needed in the research are:

1. Profitability (ROA) Rural Bank in entire Yogyakarta province includes Bantul, Gunung kidul, Kulon progo, Sleman, and Yogyakarta city 2012-2015.
2. CAR of Rural Bank in entire Yogyakarta province includes Bantul, Gunung kidul, Kulon progo, Sleman, and Yogyakarta city 2012-2015
3. LDR of Rural Bank in entire Yogyakarta province includes Bantul, Gunung kidul, kulon progo, Sleman, and Yogyakarta city 2012-2015
4. BOPO of Rural Bank in entire Yogyakarta province includes Bantul, Gunung kidul, Kulon progo, Sleman, and Yogyakarta city 2012-2015.
5. NPL of Rural Bank in entire Yogyakarta province includes Bantul, Gunung kidul, KulonProgo, Sleman, and Yogyakarta city 2012-2015.

B. Population and Sample

Population is the subject of research. If someone wants to examine all elements within the study area, the research is the study population. The research subject covering all contained in the population (Arikunto, 2002).

The population of this research is the BPR in 5 regions in the province of Yogyakarta (Bantul, Gunung kidul, Kulon progo, Sleman and Yogyakarta), from the population, the researchers took data for the research sample that is from 2012 to 2015.

C. Data Collection Techniques

Data collection in this study is intended to obtain materials that are relevant and accurate. The data used is secondary data by using the method of data collection study in the documentation that has been published by the

agency who published the data. The data is derived from Bank of Indonesia, as well as other literature sources associated with this research. The collected data are Development Profitability (ROA), CAR, LDR, BOPO and NPL in five districts in Yogyakarta province consisting of Bantul, Gunung kidul, Kulon progo, Sleman, and Yogyakarta.

D. Definition Variable Operational Research

This study uses two variables: the dependent variable and independent variables. The dependent variable in this study is Rural Bank of profitability (ROA), whereas the independent variables are the CAR, LDR, BOPO and NPL. Here is an operational definition of each variable.

Table 3.1.
Variables Operational Definition

Variable	Concept of Variables	Formula
Dependent :		
Profitabilitas (ROA)	ROA is the ratio that measures the ability of the company's assets in the company's profit from operations, if ROA increased, indicating that the performance of a company will get better.	$\frac{\text{Debit}}{\text{Total Assets}}100\%$
Independent :		
CAR *(Capital Adequacy Ratio)*	CAR is the capital adequacy ratio by demonstrating the ability of banks while maintaining sufficient capital and the ability of bank management to identify, measure, monitor and control risks may arise due to the influence of the performance of a bank at the time of generating a profit and keep the amount of capital owned banking company ,	$CAR\frac{\text{Modal}}{\text{ATMR}}100\%$
BOPO *(BiayaOperasionalTerhadapPendapatanOperasional)*	ROA ratio is used to measure the efficiency and ability of banks to carry out operations..	$BOPO\frac{\text{Operational Cost}}{\text{Operational Income}}100\%$

NPL *(Non Performing Loan)*	NPL according to the credits that are experiencing congestion in callable that occur because of intentional or external factors.	$NPL \frac{Non\ performing\ loans}{\text{Total Credit}} 100\%$

E. Analysis Methode

1. Panel Data Regression Model

To achieve the objectives of research and testing hypothesis, then in this research used multiple regression analysis to determine whether there is a significant effect of the independent variable to dependent variable or not. Data used in this research is panel data then, it used multiple linear regression panel data models which formulated as follows:

Y= α + bXit + b2X2it + b3X3it + b4X4it + e

Explanation:

Y= Dependent variable (ROA)

α = Constanta

X1 = CAR

X2 = LDR

X2 =BOPO

X3 = NPL

$b_{(1\ldots3)}$ = Regression coefficient on each independent variable

e = Error term

t = Time

i = Company

2. Metode Estimasi Model Regresi Panel

In the method of estimation models for the data panel can be done through two approaches that are fixed effect and random effect.

a. Common Effect Model

Panel data model approach is the simplest because it combines data time series and cross section. In this model neglected dimension of time as well as individuals, so it is assumed that the behavior of the data the same company in different periods. This method can use the approach Ordinary Least Square (OLS) or a least squares technique to estimate the panel data model.

Panel data model approach is the simplest because it combines data time series and cross section. In this model neglected dimension of time as well as individuals, so it is assumed that the behavior of the data the same company in different periods. This method can use the approach Ordinary Least Square (OLS) or a least squares technique to estimate the panel data model.

b. Fixed Effect

This model assumes the differences among individual can be accommodate from the difference intercept. To estimate panel data fixed effect model is using dummy variable to capture the intercept differences between companies, intercept differences

may occur due to differences in work culture, managerial, and intensive. However, the slop is same among the companies. This model is also called as *Least Squares Dummy Variable* technique (LSDV).

c. Random Effect

This model will estimate panel data where the disturbance variable may be related among the time and among the individual. In random effect model the intercept differences accommodate by error terms of each companies. The advantage using random effect model is to eliminate heteroscedasticity. This model is also called as *Error Component Model* (ECM) or Generalized Least Square technique (GLS).

3. The Selection of Estimation Method

In this research, data panel is the model to test whether independent variables influence the dependent variable. There are two approaches to estimate the regression model namely Fixed Effect Model and Random Effect Model. To select the best model, it is needed to do the specification test namely Chow Test and Hausman Test.

a. Uji Chow

Chow test is testing to determine the appropriate model among Fixed Effect or Random Effect that used to estimate Panel data. This testing is done with the hypothesis as follows:

Ho : Model PLS (Restricted)

H1 : Model Fixed Effect (Unrestricted)

From rejected to H0 is used F-statistic like this formula

Where:

RSS = Restricted Residual Sum Square (Sum Square Residual is obtained from panel data estimation with pooled least square method / common intercept).

N = Number of cross section data

T = number of time series data

K = The number of explanatory variables

This test follows the distribution of the F statistic is FN-1, NT-NK if the value of F-test or Chow Statistic (F statistic) test results greater than F-table, it is enough to reject the null hypothesis that the model to be used is the model fixed effect.

b. Uji Hausment

Hausman test aimed to compare the fixed effect method with random method effect. The result of this test is to determine which method should be selected. The hypotheses that used in this test are:

H_0: Random effects method

H_1: Fixed effects method

If *p*-value > 0.05, then H_0 accepted and H_1 rejected

If *p*-value < 0.05, then H_0 rejected and H_1 accepted

c. Uji Lagrange Multiplier

To find out if the model is better than the Random Effect Effect Common method (OLS) test was used Lagrange Multiplier (LM). LM test is based on the distribution of Chi-Squares with degrees of freedom (df) equal to the number of independent variables. Nulnya hypothesis is that the right model for panel data regression is Common Effect, and the alternative hypothesis is the right model for panel data regression is Random Effect. If the value of LM count is greater than the critical value, the Chi-Squares null hypothesis is rejected, which means the right model for panel data regression model is Random Effect. And conversely, if the value of LM count is smaller than the critical value, the Chi-Squares null hypothesis is accepted which means the right model for panel data regression model is Common Effect.

4. Mechanical Estimating Model

In economic research, a researcher often face constraints data. If the estimated regression with time series data, the observations are not sufficient. If the regression is estimated with data from a cross-sectoral too little to produce estimates that efficient. One solution to produce estimates that efficiently is by using panel data regression model. Data panel (pooling data) is a model that combines cross-sectoral observation and time series data. The goal is that the number of observations increases. If the observation increases, will reduce collinearity between explanatory variables and then it will improve the efficiency of the econometric estimation (Insukindro, 1999).

To test the effect estimation CAR, LDR, BOPO to ROA Yogyakarta area used tool panel data regression model. There are two approaches used in mengalisis panel data. Approach Fixed Effects and the Random Effect. Before the estimation model with an appropriate model, first tested the specification whether the Fixed Effects and the Random Effect or both give the same result.

Featuring a translation for GLS method (Generated Least Square) selected in this study because of the value that is owned by GLS than OLS in estimating the regression parameters. Gujarati (2003) mentions that the OLS common method assumes that the variable variance is heterogeneous, in fact variations on pooling the data tend to be heterogeneous. GLS method already take into account the heterogeneity contained in the independent variables explicitly that the method is able to produce an estimator which meet the criteria BLUE (Best Linear Unbiased Estimator).

Apart from the method of GLS (Generated Least Square) selected in this study because of the value that is owned by GLS than GLS in estimating the regression parameters. Gujarati (2003) mentions that the GLS common method assumes that the variable variance is heterogeneous, in fact variations on pooling the data tend to be heterogeneous. GLS method already take into account the heterogeneity contained in the independent variables explicitly that the method is able to produce an estimator which meet the criteria BLUE (Best Linear

Unbiased Estimator).GLS (Generated Least Square) selected in this study because of the value that is owned by GLS than GLS in estimating the regression parameters. GLS method already take into account the heterogeneity contained in the independent variables explicitly that the method is able to produce an estimator which meet the criteria BLUE (Best Linear Unbiased Estimator).

From some of the variables used in this study, it can be made a model research which is then transformed into the equation,

namely: $Yit = \beta0 + \beta1X1it + \beta2X2it + \beta3X3it + \beta4X3it\ \varepsilon$

Information :

Yit = ROA

β0 = Konstanta

β123 = Koefisien variabel 1,2,3,4

X1 = Car

X2 = Ldr

X3 = Bopo

X4 = Npl

i = Regency / City

t = Period Waktu ke-t

ε = Error Ter

5. Quality Test Data

a. Multicollinearity Test

Multikolinearitas is a state where the independent variables in a multiple regression model found a correlation (correlation) between one another. Multicolinearity test aims to test whether the regression found the correlation. In the event of multikolinearitas, the regression coefficients of the independent variables will not be significant and have a high standard error. One way to detect their multikolinieritas namely:

R2 is quite high (0.7 to 0.1), but the t-test for each of its regression coefficient R2 signifikan.Tingginya not a sufficient condition (sufficent) but not a necessary condition (Necessary) for the multikolinearitas, because on the lower R2 <0.5 can also occur multikolineraritas.

- Regressing independent variable X with the independent variables to another, then in his R2 calculated by F test;
- If F* > F table means H0 is rejected, there multikolinearitas
- If F*< F table means H0 is accepted, there is no multicollinearity.

There are several ways to overcome multicollinearity in the model, one is to look at the correlation coefficient of the output of the computer. If there is a greater correlation coefficient is 0.9 then there multikolinearitas symptoms.To overcome the problem of multicollinearity, the independent variables were correlated with other independent variables should be removed. In terms of the

GLS method, this model has been anticipated from multikolienaritas.Uji Heteroskedastisitas

Heteroskidastity is detection to see whether a variable is not constant interruption or change. Heteroskedastisitas test aims to test whether the regression model occurred inequality residual variance from one observation to another observation. If the variance of the residuals of the observations to other observations remain, then called Homoskedastisitas and if the variance is not constant or variable is called Heteroskidastity.

A good regression model is that homoskedastisitas or did not happen heteroskedastisitas. To detect the problem of heteroscedasticity in the model, the authors use the park test is often used in reference. In the method, Park suggests a specific function between the variant form of the error and the independent variables were expressed as follows:

= ……. (1)

linear equations

Ln = α + β Ln Xi + vi ……..(2)

Because the error variance () was not observed, it is used as a replacement. So the equation becomes:

Ln = α + β Ln Xi + vi ……...(3)

If the parameter β coefficients of the regression equation is statistically significant, meaning there is a problem in the data

heteroscedasticity. Conversely, if β is not significant, then the assumption homokedastisitas the data can be received. (Park in Sumodiningrat, 2010).

- Significant correlations> 0.05 means free from heterokedastisitas.
- Significant correlations <0.05 means exposed heterokedastisitas

6. Test Statistic Regression Analysis

Significance test is a procedure used to examine errors or correctness of the results of the null hypothesis of the sample.

a. Test Coefficient of Determination (R-Square)

The coefficient of determination R2 essentially measures how far the ability of the model to explain variations in the independent variables. The coefficient of determination between 0 and 1 (0 <R2 <1), the value (R2) is small means that the ability of independent variables in explaining the variation of independent variables is very limited. Value close to 1 means that the independent variable provide almost all the information needed to predict the variations of the model dependent (Gujarati, 2003).The fundamental weakness of the use of the coefficient of determination is biased against the number of dependent variables (R2) definitely improved, no matter whether these variables significantly influence the dependent variable or not. Therefore, many researchers advocate for the use of adjusted R2 value

when evaluating the best regression model. Unlike the value of R2, adjusted R2 value can rise can be dropped if the independent variables are added in the model. This test is essentially the measure of how far the ability of the model to explain variations in the independent variables.

A model has goodness and weakness when applied to different problems. To measure the goodness of a model (goodnes of fit) used the coefficient of determination (R2). The coefficient of determination is a measure of the great contribution of the independent variable on the dependent variable, or in other words the variation coefficient of determination shows the decline in Y explained by the influence of linear X.

Determinant coefficient value between 0 and 1. The determinant coefficient value close to 0 (zero) means the ability of all the independent variables in explaining the dependent variable is very limited. Determinant coefficient value close to 1 (one) means the independent variables described nearly giving information to predict what the variation of the dependent variable.

b. The F-statistics

F-statistic test is done to see how the influence of the independent variables in whole or jointly on the dependent variable. To test the hypothesis is done as follows:

1) H0: b1: b2: b3 = 0, meaning that together there is no independent effect on the dependent variable.

2) H0: b1: b2: b3 ≠ 0, meaning that together there is the influence of the independent variable on the dependent variable.

This test was conducted to compare the value of the F-count the F-table. If the F-count is greater than F-table then H0 is rejected, which means that the independent variables jointly affect the dependent variable.

If the probability of independent variables> 0.05 then H0 accepted, meaning that the independent variables simultaneously (together) do not significantly affect the dependent variable.

If the probability of independent variables <0.05 then H0 is rejected, meaning that the independent variables simultaneously (together) influence on the dependent variable.

c. T-statistics (Partial Test)

T statistical test is basically to show how far the influence of the independent variables individually in explaining the variation of the dependent variable with the following hypotheses (Imam Ghozali in Bethany, 2014). This test can be done by comparing t arithmetic with t table.

As for the formula to get the t is as follows:

$$t = (bi - b) / SBI$$

Where:

bi = coefficients of the independent variables to-i

b = the value of the null hypothesis

Sbi = standard deviation of the independent variable to-i

At the 5% significance level with the testing criteria are used as follows:

1) If t <t table then H0 is accepted and H1 rejected which means that one of the independent variables (independent) does not affect the dependent variable (dependent) significantly.

2) If t> t table then H0 rejected and H1accepted, which means that one of the independent variables (independent) affects the dependent variable (dependent) significantly.