CHAPTER III
RESEARCH METHODOLOGY

Research methodology represents the guidance for a researcher in conducting a research (Moh. Nasir: 1998). Sequences being performed within this research cover type and sources of data, operational data and data analysis methods.

3.1. Type and Sources of Data

This research makes use of secondary data that is collected from literatures, governmental and international agency as well as from a private organization. The data are taken from newspapers, websites, International Financial Statistics, Central Bureau Statistics of Indonesia (CBSI), Indonesia's Directorate General of Estates Crop (DGEC), Indonesian Rubber Trade Association GAPKINDO, and other reports and publication related to this research. Meanwhile, the research uses time series data pertain to Indonesia for the period 1980 to 2003.

3.2. Operational Data

To avoid from misinterpretation against both independent and dependent variables, it is necessary to make them obvious. These annual data pertain to Indonesia for the period 1980 to 2003. The variables are defined as follows:

A. C, consumption for natural rubber (NR), the dependent variable (in 000 metric tonne), is obtained from total production in a year minus export.

B. **PNR**, price of NR (in Rp/Kg), the independent variable, is initially weighted (and/or divided) by consumer price index. (1995=100).

C. **GDP**, total real GDP (in billion USD), the independent variable, is initially weighted (and/or divided) by the average of ASEAN[4] (without Indonesia) total real GDP. (1995=100).

D. **PSR**, world price of synthetic rubber (SR) at London (in USD/tonne), the independent variable, is obtained by firstly transforming Pound-sterling to USD and then divided by UK's CPI to get the real price. (1995=100).

E. **D**, dummy variable, the independent variable, "1" for NR consumption greater than (or close to) a hundred and "0" otherwise ("1" if $C_t > 100$ and "0" if not).

3.3.   Data Analysis Methods

Indonesia has long been experienced in the "rubber world" since the colonial system government up to now. However, this research uses annually time series data for the year between 1980 and 2003 as an object of observation.

Both descriptive and quantitative analysis will be applied to give details about the effects raised by explanatory variables, the perennial question whether NR could be "traded on" in our own domestic market, and other findings obtained from the analysis.

---

[4] ASEAN includes Brunei Darussalam, Cambodia, Laos, Malaysia, Myanmar, Philippines, Singapore, and Thailand.

This section will provide steps in conducting the analysis of demand model of economy as a guide to obtain the empirical findings of the study. The steps are as follows:

3.3.1. Externalities Identification

Because the data of dependent variable are so fluctuate, there may be externalities affecting the data observed. It is useful to know how significant the influence is. Such qualitative factors are not observable and it is necessary to create a new variable called dummy variable to quantify the factors. The different possible outcomes of the dummy variable are then coded by numerical values of 1 and 0.

This study enables the dummy variable to receive value 1 if data value of independent variable is greater than its average value and 0 if data value of independent variable is lower that the average value.

3.3.2. Choosing the Proper Model

Considering the data is time series, naturally, it faces the problem of stationary. It is necessary to know the degree of the stationary of each variable— stationary, weakly nonstationary and/or strongly nonstationary. The detection is as follows:

A. The equation resulted in OLS multiple regression model, particularly R^2 and DW-stat, indicates whether the data is spurious or not. Data is spurious if $R^2 > DW$-stat and not spurious if $R^2 < DW$-stat. Hence, it has not to apply the error correction model (ECM) if the data is not spurious.

B. Then, the detection of stationary is necessary to know whether this present regression models are stationary or not by committing the application of *unit root tests*. Regression of a nonstationary time series on another nonstationary time series may produce a spurious regression. Unit root tests determine in what level the stationary of the factors are, whether it is in original level, 1st or 2nd level. It thus indicates the degree of stationary, weakly nonstationary and/or strongly nonstationary. The summary is as follows:

1. OLS multiple regression model is a proper model if all factors, the independent variable, are stationary in original level of unit root tests.

2. Partial Adjustment Model (PAM) is a proper model to recognize the weak nonstationary.

3. Error Correction Model (ECM) is a proper model to recognize the strongly nonstationary.

3.3.3. Choosing Between Linear and Log-Linear Regression Model

Choosing between a linear regression model (the regressand is a linear function of the regressors) and a log-linear regression model (the log of the regressand is a function of the logs of the regressors) is essential in empirical analysis. In this study, *MWD test*, a test proposed by MacKinnon, White, and Davidson, will be used to choose between two models.

3.3.4. Statistical Hypotheses Testing

The problem of statistical hypothesis testing may be stated simply as: *Is a given observation or finding compatible with some stated hypothesis (H_0)?* The word "compatible," as used here, means "sufficiently" close to the hypothesized value so that we do not reject the stated hypothesis (H_0).

Test of significance approach, one-tail and two-tail tests, is a procedure by which samples results are used to verify the truth or falsity of a null hypothesis. The objective of this test is to know the relationship between independent and dependent variable individually.

3.3.5. F-Statistic Testing

F-test is a test of overall significance of the observed or estimated regression line, that is, whether *Y* as dependent variable is linearly related to independent variable X_1 X_2 and X_3. In other words, that is whether all independent variable has a joint impact on dependent variable.

3.3.6. Testing Violation of Classical Assumption

This section discuss about the testing of whether or not the estimated equation contains *multicollinearity, autocorrelation, heteroscedasticity* and *specification errors.* The appropriate tools will be conducted to detect them.

Multicollinearity arises when the explanatory variables in an equation are highly or perfectly correlated. Although this only violates a model assumption in an extreme case of perfect multicollinearity, the existence of even moderate multicollinearity can cause problems with a sample regression function.

Autocorrelation is most likely to occur in time-series data. When the data are temporally ordered, the error from one time can affect the error in other time periods. For example, unexpected surge in consumer confidence can cause a consumer durable goods equation to underestimate durables consumption for two or more periods.

One of the important assumptions of the classical linear regression model is that the variance of each disturbance term u_i is some constant number equal to σ^2. This is the assumption of homoscedasticity. However, there is *Heteroscedasticity*, symbolically $E(u_i^2) = \sigma_i^2$, notice the subscript of σ_i^2, which indicates that the conditional variances of u_i are no longer constant.

The *specification error* detection is whether the model commits (1) omission of important variable(s), (2) inclusion of superfluous variable(s), (3) adoption of the wrong function form, (4) incorrect specification of the error term ut, and (5) errors of measurement in the regressand and regressors.

The appropriate tools will be used to detect these violations.

3.3.7. Causality in Economics

Regression analysis that deals with the dependence of one variable on other variables does necessarily imply causation. In other words, the existence of a relationship between variables does not prove causality or the direction of influence. However, in regression involving time series data, the situation may be somewhat different. To test such phenomenon, this study will present the Granger Causality Test.