CHAPTER 4

DATA AND METHODOLOGY

4.1 The Data

The sample period selected for this study is from the first quarter of 1978 to the fourth quarter of 2002. This time period makes for an interesting study sample as the Federal government experienced budget deficits in 68% of the 100 observations. Besides that, interest rates movements seem to be quite volatile. It is highest between 1998 and 1999 and it hit a low as recently as 2001.

Data in this study is mainly collected and taken from the various publications of Bank Negara Malaysia (BNM). They include the Monthly and Quarterly Statistical Bulletins from 1978 right up to 2002. In addition to that, all the figures are quantified in terms of RM million except for the interest rates and the prices.

Besides that, all the data, except for the interest rates are divided by prices in 1987 so that the variables are in real terms. This is to enhance accuracy and comparability in this study. As there is no continuity in the price series in terms of a single base year, it is necessary to make certain calculations in order to obtain the price series based on prices in 1987.

For example, the available data on price level of the first quarter of 1980 is based on prices in 1967. This is denoted by $P^{1967}_{Q1/1980}$. In order to obtain the price level of the first quarter of 1980 based on prices in 1987, the following is calculated;

$$P^{1967}_{Q1/1980} \times \frac{P^{1987}_{1987}}{P^{1967}_{1987}}$$

whereby $P^{1987}_{1987} = 100$ and P^{1967}_{1987} is the price level of 1987 based on prices in 1967.

4.2 The Model and the Variables

The model in this study is based on the many previous researches on the relationship between budget deficits and interest rates. In order to effectively test this relationship, other variables that directly or indirectly influence the dynamics of deficits and interest rates are included in the equation. They are selected based on the previous studies reviewed, *a priori* information and macroeconomic theory. Therefore, a system comprising of the main components, interest rates and budget deficits; and other variables needs to be determined.

4.2.1 Interest Rates and Budget Deficits Definitions

As mentioned in the earlier section, the definition of the interest rates and budget definitions used is important. This particular study will examine the effects of budget deficits on average 3-month and 1-year Treasury bill nominal discount rates. 3-month rates are specified as $INTR3MON_t$ in the regression equation while 1-year rates are $INTR1YR_t$.

The budget deficits definition in this study is represented by a proxy ($RDEFP_t$);

$$RDEFP_t = \frac{\text{Real seasonally adjusted government expenditure}}{\text{Real seasonally adjusted government revenue}}$$

This measure is taken in order to reduce the possibility of heteroscedasticity. A value exceeding 1 (> 1) indicates deficits for a particular quarter. The higher the value, the larger is the deficit. A value of less than 1 (< 1) indicates a surplus and the lower the value, the larger is the surplus for the particular quarter. This definition of deficit is not cyclically-adjusted and contains both the structural and cyclical components.

Interest rates of two different maturities are used here for comparative reasons. 3-month Treasury bill rates represent short-term rates while 1-year rates represent medium-term rates. Quarterly figures are normal averages of monthly figures published in BNM's periodical publications.

4.2.2 Specification of the Other Variables

i) Inflation Rate ($INFL_t$)

The inflation rate, as the percentage rise in the price level, is an important factor that influences budget deficits and interest rates. According to theory, a high inflation rate will induce a high nominal interest rate and vice versa. Besides that, budget deficits are also believed to cause inflation, whether deficits are money-financed or to a certain extent, debt-financed.

The inflation rate is calculated by;

$$\frac{P_t - P_{t-1}}{P_{t-1}} \times 100$$

whereby P_t is the price level at period t. The price level is measured by the Consumer Price Index (CPI). Quarterly figures are normal averages of monthly figures available in BNM publications.

ii) Real M1 Growth ($RM1GROW_t$)

The monetary aggregate used in this study is M1. It is defined as the sum of currency in circulation and demand deposits. According to BNM, 'currency in circulation refers to the notes and coins issued by BNM less the amount held by commercial banks and Islamic banks while demand deposits refer to the current accounts of the non-bank private sector placed with the commercial banks and Islamic banks' (BNM Monthly Statistical Bulletin). Quarterly figures are normal averages of the monthly figures, not quarter-end figures.

These values are then used to compute the rate of growth of M1, using the formula below;

$$\frac{RM1_t - RM1_{t-1}}{RM1_{t-1}} \times 100$$

with $RM1_t$ being seasonally adjusted. Based on theory, higher money growth leads to lower nominal interest rates in the short run. However, it also leads to higher nominal interest rates in the medium run (Blanchard, 2003). Therefore, it is reasonable to include this variable into the study.

iii) Real Gross National Product

Another essential variable that has always been considered in previous investigations of the relationship between budget deficits and interest rates is the measure of aggregate output of a country. The measure chosen for this study is the Gross National Product (GNP) as it takes into account the total output value owned by Malaysians.

As quarterly GNP figures are not available, the real GNP (RGNP) data for a particular quarter, q in year t is obtained by the calculation below;

$$RGNP_t^q = \frac{IPP_t^q}{\sum_{q=1}^{4} IPP_t^q} \times RGNP_t$$

where IPP_t^q is the Index of Industrial Production data for year t and quarter q = 1, 2, 3, 4.

iv) Real Government Purchases

Government spending is divided into two categories, government expenditure on goods and services ($RGEXP_t$) which is permanent and government transfers ($RGTRAN_t$), which is transitory. According to Seater (1993), government purchases needs to be decomposed into permanent and transitory components as to avoid methodological problems because 'both the deficit and debt variables may have elements of simultaneity bias in them' (Seater, 1993: 183).

The government transfers series in this study is defined as the sum of transfer payments, debt service charges and pensions and gratuities. Although transfer payments in Malaysia contribute very little to the overall government spending compared to other countries like the US, it is still being considered in this study so as not to incur biased results.

v) Proxy for Economic Openness ($RECON_t$)

As Malaysia is a small open economy, the level of economic openness has to be taken into account. Although it is suggested that the current account balance is the best proxy to represent economic openness, quarterly figures are not available. Therefore, an alternative representation (as suggested by Dr Tan) which considers the imports and imports of the country is used instead;

$$\frac{\text{Real Imports} + \text{Real Exports}}{\text{Real GNP}} \times \frac{1}{2}$$

The three series; Real Imports, Real Imports and Real GNP are deseasonalized before this calculation is made.

Logs are taken for real government expenditure on goods and services ($LNRGEXP_t$), real government transfer ($LNRGTRANS_t$) and real GNP ($LNRGNP_t$) because of the large figures for each of the series. On top of that, log transformation of the series reduces the large variation and therefore, enhances stability in the data.

4.2.3 Seasonal Adjustment

Time series, especially monthly and quarterly data, often contain seasonal patterns. Figures for a particular quarter may be especially high and low for another quarter of the year. According to Greene (2000), observations of a particular quarter may be more closely related to the same quarter in previous years than to the direct previous quarter. Therefore, adjustments have to be made to the data in order to remove these seasonal components and reduce the probability of autocorrelation. Seasonal adjustment is a procedure of eliminating these seasonal effects.

All the data in this study are seasonally adjusted. The method used to 'deseasonalize' the series in this research is presented as an option in the EViews 3.1 software and is known as the differencing from moving average (additive) method.

At first, the centered moving average is calculated by applying the following formula;

$$x_t = \frac{0.5y_{t+2} + y_{t+1} + y_t + y_{t-1} + 0.5y_{t-2}}{4}$$

whereby y_t is the series at period t for quarterly data. Then, the difference between y_t and x_t is calculated.

In order to obtain the seasonal indices, the average of all the values for each quarter is computed. These four final estimates (for each quarter), denoted by $\bar{S}_q$, where q = 1, 2, 3, 4; are normalized so that they all add up to zero. This is done by first obtaining the adjustment factor, given by $f = \sum \bar{S}_q / 4$. The adjusted seasonal indices are therefore given by $\bar{S}_q - f$.

Subsequently, to obtain the seasonally adjusted series, the adjusted seasonal indices are subtracted from y_t according to its respective quarter. If y_t is of the first quarter, then $\bar{S}_1 - f$ is subtracted from it and so on. In this paper, this procedure is done by using EViews 3.1 on all the variables.

4.3 Methodology

4.3.1 The Unit Root Test

Prior to more elaborate examinations in this research paper, the knowledge of stationarity of the data is very important. In this context, a time series data is stationary if its mean and variance do not vary systematically over time (Gujarati, 2003). Non-stationary data need to be identified and will be treated to an alternative method of investigation as they cannot be regressed by using the usual Ordinary Least Squares (OLS) method.

The method used in this study to determine the stationarity of the data is through detecting the presence of unit roots for each of the variables. As a prelude to the actual method used, it is appropriate to review the brief origin of the method.

Considering the first-order autoregressive model below;

$$y_t = \rho y_{t-1} + u_t \qquad -1 \leq \rho \leq 1 \qquad (1)$$

whereby y_t is a random variable and u_t is the error term. A unit root problem exists if $\rho = 1$. This indicates non-stationarity in the data.

In order to test for unit roots, equation (1) is manipulated into equation (2) below;

$$y_t - y_{t-1} = \rho y_{t-1} - y_{t-1} + u_t$$

$$\Delta y_t = \delta y_{t-1} + u_t \qquad (2)$$

with $\delta = \rho - 1$. As in equation (1), y_t will be non-stationary if $\rho = 1$. This implies that for y_t to be stationary, the condition of $\delta < 0$ must be satisfied. Equation (2) is then estimated by using OLS and can then be tested with the following hypotheses;

$$H_0 : \delta = 0$$

$$H_0 : \delta < 0$$

In this case, an alternative test statistic, the τ statistic, is used instead of the usual t test because of the changing variances in the process. The τ statistic is actually an analogy of the 't statistic' regression that had been first introduced by Dickey and Fuller. The τ test statistic is formulated as below;

$$\tau = \hat{\delta} / s.e.(\hat{\delta})$$

The entire procedure is known as the Dickey-Fuller test.

If H_0 is rejected, y_t does not contain a unit root and is stationary. It is also classified as integrated of order zero, and is denoted as $I(0)$. If H_0 is not rejected, y_t contains at least a unit root and the process is repeated for d times until H_0 is rejected.

In that case, y_t is said to be integrated of order *d*, or simply $I(d)$. This essentially implies that y_t needs to be differenced *d* times to become stationary.

For the Dickey-Fuller test, it has to be assumed that the error term, u_t is uncorrelated. Otherwise, a different test has to be used. This alternative procedure, which is the one is used in this study, is known as the Augmented Dickey-Fuller (ADF) test.

The ADF test is marginally different in the sense that lagged values of Δy_t are added to the original equation. Equation (3) is the modified equation to Equation (2).

$$\Delta y_t = \delta y_{t-1} + \alpha_i \sum_{i=1}^{m} \Delta y_{t-1} + u_t \tag{3}$$

According to Gujarati again, lagged terms are added to ensure that the error term is not serially correlated. The lag length, *m* of the equation is determined by choosing the value that minimizes the Akaike Information Criterion (AIC). The AIC value is given as

$$AIC = n \sum \hat{u}_t^2 + 2k \tag{4}$$

where $\hat{u}$ represents the residuals and *k* represents the number of parameters in the system, including the intercept.

Based on the description above, this procedure is used on a process which only has an intercept and without a deterministic time trend,

$$\Delta y_t = \mu + \delta y_{t-1} + \alpha_i \sum_{i=1}^{m} \Delta y_{t-1} + u_t \tag{5}$$

The ADF test also uses the same test statistic, which is the τ statistic, to test whether $\delta = 0$, as in the Dickey-Fuller test.

4.3.2 Cointegration Analysis

If all the tested variables are found to be non-stationary (at least one unit root exists) and in addition to that are integrated to the same order, cointegration tests can be undertaken. Generally, the cointegration technique is an approach to test for long-run or equilibrium relationships among the variables. These non-stationary variables are said to be cointegrated if there exists a long-run relationship between them.

According to the definition presented by Engle and Granger (1987), the components of the vector $y_t = [y_{1t}, y_{2t}, \ldots, y_{mt}]'$ are cointegrated of order *d*, *b* if

i) all components of y_t are $I(d)$

ii) there is a vector β so that $\beta' y_t \sim I(d-b), b > 0$

with β known as the co-integrating vector.

This step is essential to avoid regressing non-stationary variables by using the usual OLS method, which will incidentally lead to misleading and biased results, also known as 'spurious regressions'.

There are many cointegration tests but the one chosen for this study is the Johansen methodology as it is based on the Vector Autoregressive Regression (VAR) approach and it utilizes the likelihood-ratio statistics.

Let $y_t = [y_{1t}, y_{2t}, \ldots, y_{mt}]'$ represent *m* number of variables that will be included for this research. If the all the variables are $I(1)$, the following equation is formulated;

$$\Delta \mathbf{y}_t = \boldsymbol{\mu}_m + \boldsymbol{\pi} \mathbf{y}_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-1} + \boldsymbol{e}_t \quad (6)$$

whereby

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{bmatrix}, \quad \boldsymbol{\pi} = \boldsymbol{\alpha\beta'} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} [\beta_1 \quad \beta_2 \quad \cdots \quad \beta_m],$$

$$\Gamma_i = \begin{bmatrix} \alpha_{11,j} & \alpha_{12,j} & \cdots & \alpha_{1m,j} \\ \alpha_{21,j} & \alpha_{22,j} & \cdots & \alpha_{2m,j} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1,j} & \alpha_{m2,j} & \cdots & \alpha_{mm,j} \end{bmatrix} \quad \text{and } \boldsymbol{e}_t = [e_{1t} \quad e_{2t} \quad \cdots \quad e_{mt}].$$

The number of cointegrating relations, r, is given by the rank of $\boldsymbol{\pi}$. Equation (6) is known as the Vector Error Correction (VEC) model.

In order to determine the number of cointegrating relations, the system of variables is tested with the null hypothesis of there are r cointegrating relations against the alternative of there are more than r cointegrating relations. The test begins for $r = 0$. The test statistic used is the likelihood-ratio trace test statistic given as in Equation (7);

$$Q_r = -n \sum_{j=r+1}^{m} \log(1 - \hat{\lambda}_j) \quad (7)$$

The critical values are based on the one developed by Osterwald-Lenum. The hypotheses used are

H_0 : there are r cointegrating relations

H_1 : there are more than r cointegrating relations

If H_0 is rejected at $r = 0$, there is at least one cointegrating relation in the system and the procedure is repeated on H_0 with $r = 1$ and so forth until H_0 is not rejected at a certain r. This would indicate the number of cointegrating relations in the system. If H_0 is not rejected at $r = 0$, then it is concluded that there are no cointegrating relations in the system.

In this study, the Johansen cointegration analysis is applied to two systems. The test is performed on the first system which consists of eight variables ($INTR1YR_t$, $RDEFP_t$, $LNRGNP_t$, $LNRGEXP_t$, $LNRGTRAN_t$, $INFL_t$, $RECON_t$ and $RM1GROW_t$) to determine the number of cointegrating relations for the relationship between budget deficits and medium-term (1-year) interest rates. This set of variables is named as Group 1. The test on the second system of variables ($INTR3MON_t$, $RDEFP_t$, $LNRGNP_t$, $LNRGEXP_t$, $LNRGTRAN_t$, $INFL_t$, $RECON_t$ and $RM1GROW_t$) is to determine the number of cointegrating relations for the relationship between budget deficits and short-term (3-month) interest rates. This set of variables is named as Group 2.

4.3.3 The Error Correction Model Equations

If a system is cointegrated, the error correction (EC) model will exist. The EC model consists of lagged EC terms nested in the vector autoregressive (VAR) model. It represents the speed of adjustment and it measures the time taken by the stationary dependent variables to return to the long-run equilibrium. Equation (8) is the general EC model with all the variables integrated of order 1, $I(1)$.

$$\begin{bmatrix}\Delta y_{1,t}\\ \Delta y_{2,t}\\ \vdots \\ \Delta y_{m,t}\end{bmatrix} = \begin{bmatrix}\mu_1\\ \mu_2\\ \vdots\\ \mu_n\end{bmatrix} + \begin{bmatrix}\sum_{j=1}^{p}\alpha_{11,j}\Delta y_{1,t-j} & \sum_{j=1}^{p}\alpha_{12,j}\Delta y_{2,t-j} & \cdots & \sum_{j=1}^{p}\alpha_{1m,j}\Delta y_{m,t-j}\\ \sum_{j=1}^{p}\alpha_{21,j}\Delta y_{1,t-j} & \sum_{j=1}^{p}\alpha_{22,j}\Delta y_{2,t-j} & \cdots & \sum_{j=1}^{p}\alpha_{2m,j}\Delta y_{m,t-j}\\ \vdots & \vdots & \ddots & \vdots\\ \sum_{j=1}^{p}\alpha_{m1,j}\Delta y_{1,t-j} & \sum_{j=1}^{p}\alpha_{m2,j}\Delta y_{2,t-j} & \cdots & \sum_{j=1}^{p}\alpha_{m.n,j}\Delta y_{m,t-j}\end{bmatrix} + \begin{bmatrix}\lambda_{11}\\ \lambda_{21}\\ \vdots\\ \lambda_{m1}\end{bmatrix} EC_{1,t-1} +$$

$$\begin{bmatrix}\lambda_{12}\\ \lambda_{22}\\ \vdots\\ \lambda_{m2}\end{bmatrix} EC_{2,t-1} + \cdots + \begin{bmatrix}\lambda_{1r}\\ \lambda_{2r}\\ \vdots\\ \lambda_{mr}\end{bmatrix} EC_{r,t-1} + \begin{bmatrix}e_1\\ e_2\\ \vdots\\ e_m\end{bmatrix}$$

(8)

μ represents the intercepts while α and λ are coefficients. $EC_{n,t-1}$ represents the error correction term. The number of error correction terms in a system depends on the number of cointegrating relations, *r*.

For this study, the analysis is focused on the effects of budget deficits on the two definitions of interest rates and vice versa. The following are the four EC model equations that will be investigated. The first model, Equation (9) is used to test for the effects budget deficits on 1-year Treasury bill rates.

$$\Delta^c INTR1YR_t = \alpha_1 + \sum_{i=1}^{p}\alpha_{11,i}\Delta^c INTR1YR_{t-i} + \sum_{i=1}^{p}\alpha_{12,i}\Delta^c RDEFP_{t-i} + \sum_{i=1}^{p}\alpha_{13,i}\Delta^c LNRGNP_{t-i} +$$
$$\sum_{i=1}^{p}\alpha_{14,i}\Delta^c INFL_{t-i} + \sum_{i=1}^{p}\alpha_{15,i}\Delta^c RM1GROW_{t-i} + \sum_{i=1}^{p}\alpha_{16,i}\Delta^c RGEXP_{t-i} + \sum_{i=1}^{p}\alpha_{17,i}\Delta^c RGTRAN_{t-i} +$$
$$\sum_{i=1}^{p}\alpha_{18,i}\Delta^c RECON_{t-i} + \lambda_{11}EC_{11,t-1} + \lambda_{12}EC_{12,t-1} + \ldots + \lambda_{1r}EC_{1r,t-1} + \varepsilon_{1t}$$

(9)

Equation (10) on the following page is the model used to test for the effects of 1-year Treasury bill rates on budget deficits.

$$\Delta^c RDEFP_t = \alpha_2 + \sum_{i=1}^{p} \alpha_{21,i} \Delta^c INTR1YR_{t-i} + \sum_{i=1}^{p} \alpha_{22,i} \Delta^c RDEFP_{t-i} + \sum_{i=1}^{p} \alpha_{23,i} \Delta^c LNRGNP_{t-i} + \sum_{i=1}^{p} \alpha_{24,i} \Delta^c INFL_{t-i} + \sum_{i=1}^{p} \alpha_{25,i} \Delta^c RM1GROW_{t-i} + \sum_{i=1}^{p} \alpha_{26,i} \Delta^c RGEXP_{t-i} + \sum_{i=1}^{p} \alpha_{27,i} \Delta^c RGTRAN_{t-i} + \sum_{i=1}^{p} \alpha_{28,i} \Delta^c RECON_{t-i} + \lambda_{21} EC_{11,t-1} + \lambda_{22} EC_{12,t-1} + \ldots + \lambda_{2r} EC_{1r,t-1} + \varepsilon_{2t} \quad (10)$$

Note that the error correction terms are the same for models in Equation (9) and Equation (10).

The equation shown below is used to test for the effects of budget deficits on 3-month Treasury bill rates.

$$\Delta^c INTR3MON_t = \alpha_3 + \sum_{i=1}^{p} \alpha_{31,i} \Delta^c INTR3MON_{t-i} + \sum_{i=1}^{p} \alpha_{32,i} \Delta^c RDEFP_{t-i} + \sum_{i=1}^{p} \alpha_{33,i} \Delta^c LNRGNP_{t-i} + \sum_{i=1}^{p} \alpha_{34,i} \Delta^c INFL_{t-i} + \sum_{i=1}^{p} \alpha_{35,i} \Delta^c RM1GROW_{t-i} + \sum_{i=1}^{p} \alpha_{36,i} \Delta^c RGEXP_{t-i} + \sum_{i=1}^{p} \alpha_{37,i} \Delta^c RGTRAN_{t-i} + \sum_{i=1}^{p} \alpha_{38,i} \Delta^c RECON_{t-i} + \lambda_{31} EC_{21,t-1} + \lambda_{32} EC_{22,t-1} + \cdots \lambda_{3r} EC_{2r,t-1} + \varepsilon_{3t} \quad (11)$$

Equation (12) below is the model specification to test for the effects of 3-month Treasury bill rates on budget deficits.

$$\Delta^c RDEFP_t = \alpha_4 + \sum_{i=1}^{p} \alpha_{41,i} \Delta^c INTR3MON_{t-i} + \sum_{i=1}^{p} \alpha_{42,i} \Delta^c RDEFP_{t-i} + \sum_{i=1}^{p} \alpha_{43,i} \Delta^c LNRGNP_{t-i} + \sum_{i=1}^{p} \alpha_{44,i} \Delta^c INFL_{t-i} + \sum_{i=1}^{p} \alpha_{45,i} \Delta^c RM1GROW_{t-i} + \sum_{i=1}^{p} \alpha_{46,i} \Delta^c RGEXP_{t-i} + \sum_{i=1}^{p} \alpha_{47,i} \Delta^c RGTRAN_{t-i} + \sum_{i=1}^{p} \alpha_{48,i} \Delta^c RECON_{t-i} + \lambda_{41} EC_{21,t-1} + \lambda_{42} EC_{22,t-1} + \cdots + \lambda_{4r} EC_{2r,t-1} + \varepsilon_{4t} \quad (12)$$

The error correction terms are the same for models in Equations (11) and (12). For all four models, Δ^c is the difference operator of order c and ε is the disturbance term.

4.3.4 Granger Causality

After determining the number of cointegrating relations for this system and constructing the error correction models, the next step is to look into the effects of budget deficits on interest rates and vice versa. While some of the studies that include the cointegration analysis investigated this relationship directly from the equation (correlation-based analysis), this paper emulates other papers that study this relationship with a causality-based analysis. This is because theoretically, interest rates could also possibly have an effect on budget deficits besides the conventional view that budget deficits raise interest rates.

Granger's theory of causality is defined as 'y_t causes x_t if x_t is better predicted using all available information than if the information apart from y_t had been used' (Zellner, 1979: 32). If both y_t and x_t causes each other, this relationship is known as 'feedback'. According to Zellner again, there are two important points that Granger had brought up in his theory. Firstly, his definition of causality is only relevant to stochastic variables and secondly, he assumed that the future cannot cause the past. However, this concept is not unanimously agreed. In an empirical testing point of view, the concept proposed by Granger, is used to investigate whether lagged values of an exogenous variable have any explanatory effect in a regression that includes other variables and their lagged values.

In order to explain the procedure, Equation (8) is used as reference. Suppose if a Granger Causality analysis is undertaken to investigate the relationship between Δy_2 and Δy_1. The variable Δy_2 Granger causes Δy_1 if $\alpha_{12,j} \neq 0$ for any *j* or $\lambda_{1r} \neq 0$ for any *r*. Similarly, Δy_1 will Granger cause Δy_2 if $\alpha_{21,j} \neq 0$ for any *j* or $\lambda_{2r} \neq 0$ for any *r*.

To test if Δy_2 Granger causes Δy_1, two hypotheses are built as below;

H_0: $\alpha_{12,1} = \alpha_{12,2} = \ldots = \alpha_{12,p} = \lambda_{11} = \lambda_{12} = \ldots = \lambda_{1r} = 0$

H_1: at least one of the restrictions in H_0 is not true

In other words, it implies that the null hypothesis states that Δy_2 does not Granger-cause Δy_1 and the alternative is set as Δy_2 Granger-causes Δy_1.

Then the F-statistic with the formula below,

$$F = \frac{(RSS_R - RSS_U)/p}{RSS_U/(n - mp - 1)}$$

whereby RSS_R is the residual sum of squares obtained from the restricted model, RSS_U is the residual sum of squares for the unrestricted (original) model, *p* is the number of lags, *n* is the number of observations and *m* is the number of variables; is compared with the critical value. The null hypothesis is rejected if $F > F_{\alpha;p,n-2m-1}$ at significance level $\alpha = 5\%$. The entire process is then repeated to test if Δy_1 Granger-causes Δy_2.

As mentioned in the literature review, there are four possible outcomes to this investigation of Δy_1 and Δy_2. The four are summarized below;

i) Δy_2 Granger-causes Δy_1 but not the other way round (unidirectional causality from Δy_2 to Δy_1).

ii) Δy_1 Granger-causes Δy_2 but not the other way round (unidirectional causality from Δy_1 to Δy_2).

iii) Δy_1 and Δy_2 Granger-cause each other (bidirectional causality between Δy_1 and Δy_2)

iv) Δy_1 and Δy_2 do not Granger-cause each other (no causality between Δy_1 and Δy_2)

The whole Granger Causality procedure is run by applying the Wald Test that is available in EViews 3.1. It is a test that allows for the computation of a regression with the coefficient restrictions specified by the null hypothesis and compare it with the unrestricted regression. Besides that, applying the Wald's Test to run this procedure is quite simple as the restrictions are easily assigned.

4.3.5 Diagnostic Tests

Diagnostic tests are also conducted on the system estimations to ensure that the testing procedure is valid. The first test is the Jarque-Bera (JB) Normality Test. This test is used on large-sample test to determine whether the residuals are normally distributed. The test statistic used is as the following;

$$JB = n\left[\frac{S^2}{6} + \frac{(K-3)^2}{24}\right]$$

whereby n is the sample size, S is the skewness coefficient and K is the kurtosis coefficient and it follows the chi-squared distribution.

The null hypothesis of the disturbances being normally distributed is not rejected if the computed p value is high, which in this study refers to a value exceeding 0.05. Otherwise, the null hypothesis is rejected.

Besides that, the histogram of all the EC models are also studied along with the kurtosis and skewness statistics. A kurtosis statistic value of about 3 would indicate that the data is normally distributed. A negative skewness statistic value would indicate a longer left tail in the histogram while a positive value would indicate a longer right tail.

The second test used to validate the findings in this study is by applying the cumulative sum (CUSUM) of squares test. It is based on the squares of the recursive residuals, w_r^2 in the model and uses the test statistic of

$$S_t = \frac{\sum_{r=k+1}^{t} w_r^2}{\sum_{r=k+1}^{n} w_r^2}, \qquad t = k+1, k+2, \ldots, n$$

whereby n is the sample size and k is the number of coefficients in the model. The mean value line is given by

$$E[S_t] = (t-k)/(n-k) \tag{13}$$

This test requires analysis of a plot of the CUSUM of w_r^2 with confidence bounds at the significance level of 5%. If the cumulative sum moves outside the bounds defined by two critical lines, there is a possibility of parameter instability in the model investigated.