CHAPTER 3
DATA SOURCE AND METHODOLOGY

The discussion of research methodology involves the scientific design of research process and covers the steps adopted by a researcher to pursue the objectives of the study. This chapter covers the data source and methodology used in examining the impact of exchange rate on the macro-economic indicators over the period of quarter 1 (q1) of the year 1991 to quarter 2 (q2) of the year 2009.

SECTION 3.1
DATA SOURCE:

Since the exchange rate is high frequency variable, it would be appropriate to undertake this study with high frequency data. In the present case, exchange rate is available with the monthly frequency, but other important variables such as GDP at factor cost, Foreign Direct Investment are available with quarterly frequency. It would have been better to use monthly time series, but due to above constraints quarterly time series have been used. The period of study chosen was 1991-2008 but the analysis is done from the first quarter of the year 1991 to second quarter of the year 2009.

The variables taken in the study are exchange rate against US dollar, the openness of the country which is the summation of exports as a percent of GDP and imports as a percent of GDP, GDP at factor cost with constant prices, international financial integration for which FDI as a percent of GDP has taken as the proxy variable, inflation rate for which WPI (all commodities) has been put for analysis. Data on India for these variables is available through the Reserve Bank of India website on Database on the Indian economy. The other data sources are Centre for Monitoring of Indian Economy (CMIE), Asian Development Bank (ADB), Internatio Monetary Fund (IMF) and United Nations Conference on Trade and Development (UNCTAD).
SECTION 3.2

METHODOLOGY

The collected data has been arranged in the form of tables so that meaningful inferences could be drawn. The aforementioned macroeconomic indicators can be measured by:

a. **Country Size**
   
   It is measured by GDP at factor cost

b. **Its Openness**
   
   It is measured by ratio of \([\text{imports} + \text{exports}]\) to GDP.

c. **The degree of international financial integration**
   
   It can be measured by the number of methods:
   
   1. Ratio of Gross Capital Flows to GDP.
   2. FDI to GDP Ratio (%).
   3. Private Capital Flows to GNP Ratio (%)

   In the present study this variable has been measured with FDI to GDP Ratio (%).

d. **Inflation Rate**

   For measuring inflation rate present study has taken Wholesale Price Index (WPI). Monthly data is available on these variables. We have converted it into quarterly form. We have processed the data by taking the same base for every data entry.

   The study adopts GARCH model for the long-run analysis and Granger-causality test in a multivariate VAR framework to examine the causal links between exchange rate and selected macro-economic indicators at aggregate level over the period 1991 q1 to 2009 q2. The estimation procedure involves three steps. The first step is to test for stationarity of the series with the help of unit root tests. In the presence of non-stationary variables, there is possibility of
spurious regression. The second step is to test for cointegration if the variables are non-stationary in their levels and stationary in first difference. Once the cointegration has been established amongst the variables, the third step is to formulate Error-Correction Model (ECM) to examine the causal relationship between exchange rate and macro-economic indicators. If the variables are not cointegrated of the same wavelength we had to apply ARDL procedure. If the problem still exists we had to switch over to GARCH model. For analyzing the results we have used statistical packages i.e. Eviews 5 and Microfit. For analyzing the results for ARDL model we have used Microfit and for analyzing unit root, GARCH and Granger Causality test we have used Eviews 5. All these testing procedures are outlined in the following sections:

3.2.1 Vector Auto-Regression (VAR) Framework

It is quite common in economics to have models where some variables are not only explanatory variables for a given dependent variable, but they are also explained by the variables that they are used to determine the dependent variable. In those cases, we have models of simultaneous equations, in which it is necessary to clearly identify which are the endogenous and which are the exogenous variables. The decision regarding such a differentiation among variables was strongly criticized by Sims (1980). According to him, if there is simultaneity among a number of variables, then all these variables should be treated in same way. In other words, there should be no distinction between endogenous and exogenous variables. Therefore, once this distinction is abandoned, all variables are treated as endogenous. This means that in its general reduced form each equation has the same set of regressors which leads to the development of VAR models. When we are not confident that a variable is really exogenous, we have to treat each variable symmetrically. For example, the time series $Y_t$ that is affected by current and past values of $X_t$ and, simultaneously, the time series $X_t$ to be a series that is affected by current and past values of the $Y_t$ series. In this case we will have the simple bivariate model given by:
Data Source and Methodology

\[ y_t = \beta_{10} - \beta_{12}x_t + \gamma_{11}y_{t-1} + \gamma_{12}x_{t-1} + u_{yt} \]  
(3.1)

\[ x_t = \beta_{20} - \beta_{21}y_t + \gamma_{21}y_{t-1} + \gamma_{22}x_{t-1} + u_{xt} \]  
(3.2)

Where we assume that both \( Y_t \) and \( X_t \) are stationary and \( u_{yt} \) and \( u_{xt} \) are uncorrelated white noise error terms. Equations (3.1) and (3.2) constitute a first order VAR model; because the longest lag length is unity. These equations are not reduced form equations since \( Y_t \) has a contemporaneous impact on \( X_t \) (given by \( -\beta_{21} \)), and \( X_t \) has a contemporaneous impact on \( Y_t \) (given by \( -\beta_{12} \)). Rewriting the system with the use of matrix algebra, we get:

\[
\begin{bmatrix}
1 & \beta_{12} \\
\beta_{21} & 1
\end{bmatrix}
\begin{bmatrix}
y_t \\
x_t
\end{bmatrix}
= \begin{bmatrix}
\beta_{10} \\
\beta_{20}
\end{bmatrix}
+ \begin{bmatrix}
\gamma_{11} & \gamma_{12} \\
\gamma_{21} & \gamma_{22}
\end{bmatrix}
\begin{bmatrix}
y_{t-1} \\
x_{t-1}
\end{bmatrix}
+ \begin{bmatrix}
u_{yt} \\
u_{xt}
\end{bmatrix}
\]  
(3.3)

Or

\[ Bz_t = \Gamma_0 + \Gamma_1z_{t-1} + u_t \]  
(3.4)

Where

\[ B = \begin{bmatrix}
1 & \beta_{12} \\
\beta_{21} & 1
\end{bmatrix}, \quad z_t = \begin{bmatrix} y_t \\ x_t \end{bmatrix}, \quad \Gamma_0 = \begin{bmatrix} \beta_{10} \\ \beta_{20} \end{bmatrix}, \quad \Gamma_1 = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}, \quad \text{and} \]

\[ u_t = \begin{bmatrix} u_{yt} \\ u_{xt} \end{bmatrix}. \]

Multiplying both sides by \( B^{-1} \), we obtain:

\[ z_t = A_0 + A_1z_{t-1} + e_t \]  
(3.5)

Where \( A_0 = B^{-1}\Gamma_0, \quad A_1 = B^{-1}\Gamma_1 \) and \( e_t = B^{-1}u_t \).

For purpose of notational simplification we can denote as \( a_{i0} \) the \( i^{th} \) element of the vector \( A_0 \), \( a_{ij} \) the element in row \( i \) and the column \( j \) of the matrix \( A_i \) and \( e_{it} \) as the \( i^{th} \) element of the vector \( e_t \). Using this we can rewrite the VAR model as:

\[ y_t = a_{10} + a_{11}y_{t-1} + a_{12}x_{t-1} + e_{1t} \]  
(3.6)

\[ x_t = a_{20} + a_{21}y_{t-1} + a_{22}x_{t-1} + e_{2t} \]  
(3.7)

To distinguish between the original VAR model and the system we have just obtained, we call the first a structural or primitive VAR system and the second
a VAR in standard (or reduced) form. It is important to note that the new error terms $e_{1t}$ and $e_{2t}$, are composites of the two shocks $u_{1t}$ and $u_{2t}$. Since $e_t = B^{-1}u_t$, we can obtain $e_{1t}$ and $e_{2t}$ as:

$$e_{1t} = (u_{1t} + \beta_{12}u_{2t})/(1 - \beta_{12}\beta_{21})$$

$$e_{2t} = (u_{2t} + \beta_{21}u_{1t})/(1 - \beta_{12}\beta_{21})$$

(3.8) (3.9)

Since $u_{1t}$ and $u_{2t}$ are white noise processes, it follows that both $e_{1t}$ and $e_{2t}$ are white noise processes as well.

**Testing for causality**

One of the good features of VAR models is that they allow us to test for the direction of causality. Causality in econometrics is somewhat different to the concept in everyday use; it refers more to the ability of one variable to predict (and therefore cause) the other. Suppose two variables, say $Y_t$ and $X_t$, affect each other with distributed lags. The relationship between those variables can be captured by a VAR model. In this case it is possible to have that a) $Y_t$ causes $X_t$, b) $X_t$ causes $Y_t$, c) there is bi-directional feedback (causality among the variables), and finally d) the two variables are independent. The problem is to find an appropriate procedure that allows us to test and statistically detect the cause and effect relationship among the variables.

1) The Granger causality test

Granger (1969) developed a relatively simple test that defined causality as follows: a variable $Y_t$ is said to Granger-cause $X_t$, if $X_t$ can be predicted with greater accuracy by using past values of the $Y_t$ variable rather than not using such past values, all other terms remaining unchanged. The Granger causality test for the case of two explanatory variables $Y_t$ and $X_t$, involves as a first step the estimation of the following VAR model:

$$y_t = a_t + \sum_{i=1}^{n} \beta_{yi}y_{t-i} + \sum_{j=1}^{m} \gamma_{ji}y_{t-j} + e_{1t}$$

(3.10)
Data Source and Methodology

\[ x_t = \alpha_2 + \sum_{i=1}^{n} \theta_i x_{t-i} + \sum_{j=1}^{m} \delta_j y_{t-j} + e_{2t} \]  

(3.11)

Where it is assumed that both \( e_{2t} \) and \( e_{it} \) are uncorrelated white-noise error terms. In this model we can have the following different cases:

**Case 1:** the lagged x terms in (3.10) may be statistically different from zero as a group, and the lagged y terms in (3.11) not statistically different from zero. In this case we have that \( x_t \) causes \( y_t \).

**Case 2:** the lagged y terms in (3.11) may be statistically different from zero, and the lagged x terms in (3.10) is not statistically different from zero. In this case we have that \( y_t \) causes \( x_t \).

**Case 3:** both sets of x and y terms are statistically different from zero in (3.10) and (3.11), so that we have bi-directional causality.

**Case 4:** both sets of x and y terms are not statistically different from zero in (3.10) and (3.11), so that \( x_t \) is independent of \( y_t \).

The Granger causality test, then, involves the following procedure. First, estimate the VAR model given by equations (3.10) and (3.11). Then check the significance of the coefficients and apply variable deletion tests first in the lagged x terms for equation (3.10), and then in the lagged y terms in (3.11). According to the results of the variable deletion tests we may conclude about the direction of causality based upon the four cases mentioned above.

II) The Sims causality test

Sims (1980) proposed an alternative test for causality making use of the fact that in any general notion of causality it is not possible for the future to cause the present. Therefore, when we want to check whether a variable \( y_t \) causes \( x_t \), Sims suggests estimating the following VAR model:

\[ y_t = \alpha_1 + \sum_{i=1}^{n} \beta_i x_{t-i} + \sum_{j=1}^{m} \gamma_j y_{t-j} + \sum_{\rho=1}^{k} \xi_{\rho} x_{t+\rho} + e_{1t} \]  

(3.12)
The new approach here is that apart from lagged values of x and y, there are also leading values of x included in the first equation (and similarly leading values of y in the second equation). Examining only the first equation, \( y_t \) causes \( x_t \), then we will expect that there is some relationship between \( y \) and the leading values of \( x \). Therefore, instead of testing for the lagged values of \( x_t \), we test for \( \sum_{\rho=1}^{k} \xi_{\rho} = 0 \). Note that if we reject the restriction then the causality runs from \( y_t \) to \( x_t \), and not vice versa, since the future cannot cause the present. To carry out the test we simply estimate a model with no leading terms (which is restricted version) and then the model as appears in (3.12) (which is unrestricted model), and then obtain the F statistic as in the Granger test. It is unclear which version of the two tests is preferable, and most researchers use both. The Sims test, however, using more regressors (due to the inclusion of the leading terms), leads to a bigger loss of degrees of freedom.

The testing procedure for the identification of causal directions becomes, however, more complex when, as is common in macroeconomic time series, the variables have unit-roots. In such a case- after testing for the existence of cointegration, it is useful to reparametrize the model in the equivalent ECM form (see Hendry et al., 1984; Johansen, 1988) as follows:

\[
\Delta y_t = \alpha_0 + \alpha_1 \sum_{\ell} \Delta x_{t-\ell} + \alpha_2 \sum_{k} \Delta z_{t-\ell} + \alpha_v y_{t-1} + u_t
\]

Where \( v_{t-1} = y_{t-1} - \alpha_1 x_{t-1} - \alpha_2 z_{t-1} \) is the residual of the cointegration equation.

The null hypothesis, now, that \( x \) does not Granger-cause \( y \), given \( z \), is \( H_0(\alpha_1 = \alpha_2 = 0) \). This means that there are two sources of causation for \( y \), either through the lagged terms or through the lagged cointegrating vector. This latter source of causation is not detected by a standard causality test. The null hypothesis can be rejected if either one or more of these sources affect \( y \) (i.e. the parameters are different from zero). The hypothesis is again tested using a
Data Source and Methodology

standard F test. Following Granger and Lin (1995), the conventional Granger causality test is not valid, because two integrated series cannot cause each other in the long run unless they are cointegrated. We therefore test for causality among the variables that are found to be cointegrated, using the VECM representations for the cointegrated variables. Causality in the long run exists only when the coefficient of the cointegrating vector is statistically significant and different from zero.

3.2.2 Iterating Towards Time-Series Stationarity

Stationarity is the first fundamental statistical property tested for in time series analysis, because most statistical models require that the underlying generating processes be stationary. A time series is covariance stationary when it has the following three characteristics:

1) Exhibits mean reversion in that it fluctuates around a constant long-run mean.
2) Has a finite variance that is time-invariant. And
3) Has a theoretical correlogram that diminishes as the lag length increases.

In its simplest terms a time series \( y_t \) is said to be stationary if:

a) \( E(y_t) = \) constant for all \( t \);
b) \( Var(y_t) = \) constant for all \( t \); and
c) \( Cov(y_t, y_{t+k}) = \) constant for all \( t \) and all \( k \neq 0 \), or if its mean, its variance and its covariances remain constant over time.

Stationarity is important because if the series is non-stationary then all the typical results of the classical regression analysis are not valid. Regression with non-stationary series may have no meaning and are therefore called spurious. In stationary time series, shocks will be temporary and over time their effects will be eliminated as the series revert to their long-run mean values. On the other hand, non-stationary time series will necessarily contain permanent components. Therefore, the mean and the variance of a non-stationary time
series will depend on time. Regression with non-stationary series may have no meaning and are therefore called spurious.

In general, a stationary series will follow a theoretical correlogram that will die out quickly as the lag lag-length increases, while the theoretical correlogram of a non-stationary time series will not die out (diminish or tend to zero) for increasing lag length. However this method of identifying non-stationary series is bound to be imprecise because a near unit-root process will have the same shape of autocorrelation function (ACF) with that of a real unit-root process. Thus, what might appear for one person may appear as a stationary process for another. So the formal tests for identifying non-stationarity (or, differently stated, the presence of unit roots) are needed. The next section explains what unit root is and then present a formal tests for the existence of unit roots and discusses the problems regarding the existence of unit roots in the regression models.

Consider the auto regressive model of order one:

\[
y_t = \phi y_{t-1} + u_t
\]

(3.14)

Where \( u_t \) is a white noise process and the stationarity condition is \( |\phi| < 1 \).

In general we have three possible cases:

Case 1: if \( |\phi| < 1 \), the series is stationary.

Case 2: if \( |\phi| > 1 \), in this case the series explodes.

Case 3: if \( |\phi| = 1 \), in this case the series contains a unit root and is non-stationary.

**Testing for the order of integration**

A test for the order of integration is a test for the number of unit roots, and it follows the steps described below:

**Step 1:** Test \( y_t \) to see if it is stationary. If yes then \( y_t \sim I(0) \); if no then \( y_t \sim I(n); n>0 \).

**Step 2:** Take first differences of \( y_t \) as \( \Delta y_t = y_t - y_{t-1} \), and test \( \Delta y_t \) to see if it is stationary. If yes then \( y_t \sim I(1) \); if no then \( y_t \sim I(n); n>0 \).
**Data Source and Methodology**

**Step 3:** Take second differences of $y_t$ as $\Delta^2 y_t = \Delta y_t - \Delta y_{t-1}$, and test $\Delta^2 y_t$ to see if it is stationary. If yes then $y_t \sim I(2)$; if no then $\Delta^2 y_t = \Delta y_t - \Delta y_{t-1}; n>0$ etc… till we find that it is stationary and then we stop. So, for example if $\Delta^3 y_t \sim I(0)$, then $\Delta^2 y_t \sim I(1)$ and $\Delta y_t \sim I(2)$, and finally $y_t \sim I(3)$; which means that $y_t$ needs to be differenced three times in order to become stationary.

**a) The Simple Dickey-Fuller test for unit roots**

Dickey and Fuller (1979, 1981) devised a procedure to formally test for non-stationarity. The key insight of their test is that testing for non-stationarity is equivalent to testing for the existence of a unit root. Thus the obvious test is the following which is based on the simple AR(1) model of the form:

$$y_t = \phi y_{t-1} + u_t \quad \text{(3.15)}$$

What we need to examine here is whether $\phi$ is equal to 1 (unity and hence ‘unit root’). Obviously, the null hypothesis is $H_0: \phi = 1$, and the alternative hypothesis is $H_1: \phi < 1$. A different or more convenient version of the test can be obtained by subtracting $y_{t-1}$ from both sides of above AR (1) model.

$$y_t - y_{t-1} = \phi y_{t-1} - y_{t-1} + u_t \quad \text{(3.16)}$$

$$\Delta y_{t-1} = (\phi - 1) y_{t-1} + u_t \quad \text{(3.17)}$$

$$\Delta y_{t-1} = \gamma y_{t-1} + u_t \quad \text{(3.18)}$$

Where of course $\gamma = (\phi - 1)$. Then, now the null hypothesis is $H_0: \gamma = 0$ and the alternative hypothesis is $H_1: \gamma < 0$, where if $\gamma = 0$ then $y_t$ follows a pure random walk model.

Dickey and Fuller (1979) also proposed two alternative regression equations that can be used for testing for the presence of a unit root. The first contains a constant in the random-walk process as in the following equation:

$$\Delta y_{t-1} = \alpha_0 + \gamma y_{t-1} + u_t \quad \text{(3.19)}$$
This is an extremely important case, because such processes exhibit a definite trend in the series when $y=0$, which is often the case for macroeconomic variables.

The second case is to also allow, a non-stochastic time trend in the model, so as to have:

$$\Delta y_{t-1} = \alpha_0 + \alpha_2 t + \gamma y_{t-1} + \epsilon_t$$  \hspace{1cm} (3.20)

The Dickey Fuller test for stationarity is then simply the normal ‘t’ test on the coefficient of the lagged dependent variable $y_{t-1}$ from one of the three models. This test does not however have a conventional ‘t’ distribution and so we must use special critical values which were originally calculated by Dickey and Fuller. MacKinnon (1991) tabulated appropriate critical values for each of the three above models and these are presented in Table 3.1

<table>
<thead>
<tr>
<th>Table 3.1 Critical values for the Dickey Fuller test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td>$\Delta y_{t-1} = \gamma y_{t-1} + \epsilon_t$</td>
</tr>
<tr>
<td>$\Delta y_{t-1} = \alpha_0 + \gamma y_{t-1} + \epsilon_t$</td>
</tr>
<tr>
<td>$\Delta y_{t-1} = \alpha_0 + \alpha_2 t + \gamma y_{t-1} + \epsilon_t$</td>
</tr>
<tr>
<td>Standard critical values</td>
</tr>
</tbody>
</table>


In all cases the test concerns whether $\gamma = 0$. The DF-test statistic is the t-statistic for the lagged dependent variable. If DF statistical value is smaller in absolute terms than the critical value then we reject the null hypothesis of a unit root and conclude that $y_t$ is a stationary process.

**b) The augmented Dickey-Fuller (ADF) test for unit roots**

As the error term is unlikely to be white noise, Dickey and Fuller extended their test procedure suggesting an augmented version of the test which includes extra lagged terms of the dependent variable in order to eliminate
Data Source and Methodology

autocorrelation. The lag length on these extra terms is either determined by the Akaike Information Criterion (AIC) or Schwartz Bayesian Criterion (SBC), or more usefully by the lag length necessary to whiten the residuals (i.e. after each case we check whether the residuals of the ADF regression are autocorrelated or not through LM tests and not the DW test).

The three possible forms of the ADF test are given by the following equations:

\[ \Delta y_t = \gamma_0 y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + u_t \]  
(3.21)

\[ \Delta y_t = \alpha_0 + \gamma_0 y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + u_t \]  
(3.22)

\[ \Delta y_t = \alpha_0 + \gamma_0 y_{t-1} + \alpha_d t + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + u_t \]  
(3.23)

The difference between the three regressions again concerns the presence of the deterministic elements \( \alpha_0 \) and \( \alpha_d t \). The critical values for the ADF tests are the same as those given in Table 3.1 for the DF test.

Unless the econometrician knows the actual data generating process, there is a question concerning whether it is most appropriate to estimate (3.21), (3.22) or (3.23). Dolado, Jenkinson and Sosvilla-Rivero (1990) suggest a procedure which starts from the estimation of the most general model given by (3.23) and then answering a set of questions regarding the appropriateness of each model and moving to the next model. This procedure is illustrated in Figure 3.1. It needs to be stressed here that, although useful, this procedure is not designed to be applied in a mechanical fashion. Plotting the data and observing the graph is sometimes very useful because it can clearly indicate the presence or not of deterministic regressor. However, this procedure is the most sensible way to test for unit roots when the form of the data-generating process is unknown.

c) The Phillips-Perron test

The distribution theory supporting the Dickey-Fuller tests is based on the assumption that the error terms are statistically independent and have a
constant variance. So, when using the ADF methodology we have to make sure that the error terms are uncorrelated and that they really have a constant variance. Phillips and Perron (1988) developed a generalization of the ADF test procedure that allows for fairly mild assumptions concerning the distribution of errors. The test regression for the Phillips-Perron (PP) test is the AR(1) process:

\[ \Delta y_{t-1} = \alpha_0 + \gamma y_{t-1} + \epsilon_t \]  

\[ (3.24) \]

\[ \Delta = \frac{\Delta Y}{\Delta t} \]

\[ \Delta Y = \alpha + \gamma Y + \epsilon_t + \sum \beta \Delta Y + \epsilon_t \]

Data Source and Methodology

3.2.3 Procedure for testing for unit root tests

While the ADF test corrects for higher order serial correlation by adding lagged differenced terms on the right hand side, the PP test makes a correction to the ‘t’ statistic of the coefficient $\gamma$ from the AR(1) regression to account for the serial correlation in $e_t$. So, the PP statistics are just modifications of the ADF ‘t’ statistics that take into account the less restrictive nature of the error process. The expressions are extremely complex to derive and are beyond the scope of this study. However, since many statistical packages (Eviews, EasyReg and Microfit) have routines available to calculate these statistics, it is good for the researcher to test the order of integration of a series performing the PP test as well. The asymptotic distribution of the PP ‘t’ statistic is the same as the ADF ‘t’ statistic and therefore the McKinnon (1991) critical values are still applicable. As with the ADF test, the PP test can be performed with the inclusion of a constant, a constant and linear trend, or neither in the test regression.

3.2.4 Cointegration and the Error Correction Mechanism (ECM)

One way of resolving the problem of non-stationarity in time series data is to difference the series successively until stationarity is achieved and then use stationarity series for regression analysis. However, this solution is not ideal. There are two main problems using first differences.

1) If the model is correctly specified as a relationship between variables (say Y and X) and we difference both variables then implicitly we are also differencing the error process in the regression. This would then produce a non-invertible moving average error process and would present serious estimation problems.

2) The second problem is that if we difference the variables the model can no longer give a unique long run solution.

The desire to have models which combine both short-run and long-run properties, and which at the same time maintain stationarity in all of the variables, has led to a reconsideration of the problem of regression using
variables that are measured in their levels. As we know that when we have non-stationary variables in a regression model then we may get results that are spurious. So if we have \( Y_t \) and \( X_t \) that are both I(1), then if we regress:

\[
Y_t = \beta_1 + \beta_2 X_t + u_t
\]  
(3.25)

We will not generally get satisfactory estimates of \( \beta_1 \) and \( \beta_2 \). One way of resolving this is to difference the data in order to ensure stationary of our variables. Therefore, after that we will have that \( \Delta Y_t \sim I(0) \) and, \( \Delta X_t \sim I(0) \) and the regression model will be:

\[
\Delta Y_t = a_1 + a_2 \Delta X_t + \Delta u_t
\]  
(3.26)

In this case the regression model may give us correct estimates of the \( a_1 \) and \( a_2 \) parameters and the spurious equation problem has been resolved. However, what we have from equation (3.26) is only the short-run relationship between the two variables. Remember that in the long run:

\[
Y_t = \beta_1 + \beta_2 X_t
\]  
(3.27)

So \( \Delta Y_t \) is bound to give us no information about the long run behavior of our model. Knowing that economists are mainly interested in long run relationship this constitutes a big problem, and in order to resolve this, the concept of cointegration and ECM are very useful. As \( Y_t \) and \( X_t \) are both are now integrated of order one i.e. \( Y_t \sim I(1) \) and \( X_t \sim I(1) \). In the special case that there is a linear combination of \( Y_t \) and \( X_t \), that is I(0), then \( Y_t \) and \( X_t \) are cointegrated. Thus, if this is the case the regression of equation (4.25) is no longer spurious, and it also provides us with the linear combination:

\[
\hat{u}_t = Y_t - \hat{\beta}_1 - \hat{\beta}_2 X_t
\]  
(3.28)

That connects \( Y_t \) and \( X_t \) in the long run.

**The Error-Correction Mechanism (ECM)**

If, \( Y_t \) and \( X_t \) are cointegrated, then by definition \( \hat{u}_t \sim I(0) \). Thus, we can express the relationship between \( Y_t \) and \( X_t \) with an ECM specification as:
\[ \Delta Y_t = a_0 + b_1 \Delta X_t - \Pi \hat{u}_{t-1} + Y_t \quad (3.29) \]

This will now have the advantage of including both long run and short run information. In this method, \( b_1 \) is the impact multiplier (the short run effect) that measures the immediate impact that a change in \( X_t \) will have on a change in \( Y_t \). On the other hand \( \Pi \) is the feedback effect, or the adjustment effect, and shows how much of the disequilibrium is being corrected, i.e. the extent to which any disequilibrium in the previous period effects any adjustment in \( Y_t \).

Of course \( \hat{u}_{t-1} = Y_{t-1} - \hat{\beta}_1 - \hat{\beta}_2 X_{t-1} \), and therefore from this equation we also have \( \beta_2 \) being the long run response (it is estimated by equation 3.26).

Equation (3.29) now emphasizes the basic approach of cointegration and error correction models. The spurious regression problem arises because we are using non stationary data but in equation (3.29) everything is stationary, the change in \( X \) and \( Y \) is stationary because they are assumed to be I(1) variables and the residual from the levels regression (3.28) is also stationary by the assumption of cointegration. So equation (3.29) fully conforms to our set of assumptions about the classical linear regression model and OLS should perform well.

3.2.5 Testing for Cointegration

There are two approaches to test the cointegration in a given equation:

1) Cointegration in single equations: the Engle-Granger approach

Granger (1981) introduced a remarkable link between non-stationary processes and the concept of long-run equilibrium; this link is the concept of cointegration defined above. Engle and Granger (1987) further formalized this concept by introducing a very simple test for the existence of cointegrating (i.e. long-run equilibrium) relationships.

In order to understand this approach (which is often called EG approach) consider the following two series \( X_t \) and \( Y_t \), and the following cases:
a) If \( Y_t \sim I(0) \) and \( X_t \sim I(1) \) then every linear combination of those two series

\[ \theta_1 Y_t + \theta_2 X_t \quad (3.30) \]

will result in a series that will always be \( I(1) \) or non-stationary. This will happen because the behavior of the non-stationary series will dominate the behavior of the \( I(0) \) one.

b) If we have both series are \( I(1) \) or non-stationary, then in general any linear combination the two series, say

\[ \theta_1 Y_t + \theta_2 X_t \quad (3.31) \]

will also be \( I(1) \). However, although this is the more likely case, there are exceptions to this rule, and we might find in the rare cases that there is a unique combination of the series as in (3.31) above that is \( I(0) \). If this is the case then we say that two series are cointegrated of order \((1, 1)\).

Now the problem is how can we estimate the parameters of the long run equilibrium relationship and make sure whether or not we have cointegration. Engle and Granger proposed a straightforward method which involves four steps:

**Step 1:** Test the variables for their order of integration

By definition, cointegration necessitates that the variables be integrated of the same order. The Dickey-Fuller and the ADF tests can be applied in order to infer the number of unit roots (if any) in each of the variables. We can differentiate three cases which will either lead us to the next step or will suggest stopping:

a) If both variables are stationary \((I(0))\), it is not necessary to proceed since standard time series methods apply to stationary variables.

b) If the variables are integrated of different order, it is possible to conclude that they are not cointegrated.

c) If both variables are integrated of the same order then we proceed with step two.
**Data Source and Methodology**

**Step 2:** Estimate the long run relationship

If the results of step 1 indicate that both variables are integrated of the same order, the next step is to estimate the long run equilibrium relationship of the form:

\[ Y_t = \beta_1 + \beta_2 X_t + e_t \]  

(3.32)

And obtain the residuals of this equation.

If there is no cointegration, the results obtained will be spurious. However, if the variables are cointegrated, then OLS regression yields ‘super consistent’ estimators for the cointegrating parameters \( \hat{\beta}_2 \).

**Step 3:** Check for (cointegration) the order of integration of the residuals

In order to determine if the variables are actually cointegrated, denote the estimated residual sequence from this equation by \( \hat{e}_t \). Thus, \( \hat{e}_t \) is the series of the estimated residuals of the long run relationship and if these deviations from long run equilibrium are found to be stationary, then \( X_t \) and \( Y_t \) are cointegrated.

In fact we perform a DF test on the residual series to determine their order of integration. The form of the DF test is the following:

\[ \Delta \hat{e}_t = a_t \hat{e}_{t-1} + \sum_{i=1}^{\delta} \Delta \hat{e}_{t-i-1} + \nu_t \]  

(3.33)

Note that because \( \hat{e}_t \) is a residual we do not include a constant nor a time trend.

The critical values differ from the standard ADF values, being more negative (typically around -3.5). Critical values are provided in Table 3.2. Obviously, if we find that \( \hat{e}_t \not\equiv I(0) \) then we can reject the null that the variables \( X_t \) and \( Y_t \) are not cointegrated. We repeat the same process if we have a single equation with more than just one explanatory variable.

<table>
<thead>
<tr>
<th>Description</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>No lags</td>
<td>-4.07</td>
<td>-3.37</td>
<td>-3.3</td>
</tr>
<tr>
<td>Lags</td>
<td>-3.73</td>
<td>-3.17</td>
<td>-2.91</td>
</tr>
</tbody>
</table>

**Note:** There are two sets of critical values; the first is for no lagged dependent variable terms in the augmentation terms (i.e. for simple DF test), and the second is for including lagged dependent variables (i.e. for the ADF test).
Step 4: Estimate the error correction model

If the variables are cointegrated, the residuals from the equilibrium regression can be used to estimate the error-correction model and to analyze the long run and short run effects of the variables as well as to see the adjustment coefficient, which is the coefficient of the lagged residual terms of the long run relationship identified in step 2. At the end we always have to check for the adequacy of the model by performing diagnostic tests.

2) Cointegration in multiple equations and the Johansen approach

If we have more than two variables in the model, then there is the possibility of having more than one cointegrating vector. By this we mean that the variables in the model might form several equilibrium relationships governing the joint evolution of all the variables. In general for n number of variables we can have only up to n-1 cointegrating vectors. Therefore, when n=2 which is the simplest case, we can understand that if cointegration exists then the cointegrating vector is unique.

Having n=2 and assuming that only one cointegrating relationship exists, where there are actually more than one, is a very serious problem that cannot be resolved by the EG single equation approach. Therefore, an alternative to the EG approach is needed and this is the Johansen approach for multiple equations.

In order to present this approach, it is useful to extend the single equation error correction model to a multivariate one, let’s assume that we have three variables, $Y_t$, $X_t$ and $W_t$ which can all be endogenous, i.e. we have that

\[ Z_t = A_1 Z_{t-1} + A_2 Z_{t-2} + \ldots + A_k Z_{t-k} + u_t \]  

(3.34)

It can be reformulated in a vector error correction model (VECM) as follows:

\[ \Delta Z_t = \Gamma_1 \Delta Z_{t-1} + \Gamma_2 \Delta Z_{t-2} + \ldots + \Gamma_{k-1} \Delta Z_{t-k+1} + \Pi Z_{t-1} + u_t \]  

(3.35)

Where $\Gamma_i = (I - A_1 - A_2 - \ldots - A_k)$ and $\Pi = -(I - A_1 - A_2 - \ldots - A_k)$.

\( i = 1, 2, \ldots, k-1 \)
**Data Source and Methodology**

Here we need to carefully examine the $\Pi$ matrix of order $3 \times 3$. The $\Pi$ matrix contains information regarding the long run relationships. We can decompose $\Pi = \alpha \beta'$ where $\alpha$ will include the speed of adjustment to equilibrium coefficients while $\beta'$ will be the long run matrix of coefficients. Therefore the $\beta'Z_{t-1}$ term is equivalent to the error correction term $(Y_{t-1} - \beta_0 - \beta_1X_{t-1})$ in the single equation case, except that now $\beta'Z_{t-1}$ contains up to $(n-1)$ vectors in a multivariate framework.

The steps of the Johansen approach in practice:

**Step 1:** Testing the order of integration of the variables

As with EG approach, the first step in the Johansen approach is to test for the order of integration of the variables under examination.

**Step 2:** Setting the appropriate lag length of the model

The issue of finding appropriate lag length is very important because we want to have Gaussian error terms (i.e. standard normal error terms that do not suffer from non-normality, autocorrelation, heteroskedasticity etc.). Setting the value of the lag length is affected by the omission of variables that might affect only the short run behavior of the model. This is due to the fact that omitted variables instantly become part of the error term. Therefore, very careful inspection of the data and the functional relationship is necessary before proceeding with estimation in order to decide whether to include additional variables.

The most common procedure in choosing the optimal lag length is to estimate a VAR model including all our variables in levels. This VAR model should be estimated for a large number of lags, then reducing down by re-estimating the model for one lag less until we reach zero lags (i.e. we estimate the model for 12 lags, then 11, then 10 and so on until we reach 0 lags). If there are no exogenous variables in the VAR, the lag starts at 1; otherwise the lag starts at 0. In each of these models we inspect the values of the AIC and the SBC criteria, as well as the diagnostic concerning autocorrelation,
Data Source and Methodology

heteroskedasticity, possible ARCH effects and normality of the residuals. In model that minimizes AIC and SBC is selected as the one with the optimal lag length. If there arise a conflict between the minimum values of AIC and SBC then one should prefer the minimum of SBC to select optimal lag length. This model should also pass all the diagnostic checks.

Step 3: Choosing the appropriate model regarding the deterministic components in the multivariate system

Another important aspect in the formulation of the dynamic model is whether an intercept and/or a trend should enter in either the short run or the long run model, or both models. The general case of the VECM including all the various options that can possibly happen is given by the following equation:

\[ \Delta Z_t = \Gamma_1 \Delta Z_{t-1} + \ldots + \Gamma_{k-1} \Delta Z_{t-k-1} + \alpha \begin{pmatrix} \beta \\ \mu_1 \\ \delta_1 \end{pmatrix} (Z_{t-1} | \theta \gamma) + \mu_2 + \delta_2 t + \epsilon_t, (3.36) \]

And for this equation we can see the possible cases. We can have a constant (with coefficient \( \mu_1 \)) and/or a trend (with coefficient \( \delta_1 \)) in the long run model (the cointegrating equation (CE)), and a constant (with coefficient \( \mu_2 \)) and/or a trend (with coefficient \( \delta_2 \)) in the short run model (the VAR model). In general five distinct models can be considered. Although the first and the fifth are not that realistic, we present all of them for reasons of complementarity.

Model 1: No intercept or trend in CE or VAR (\( \delta_1 = \delta_2 = \mu_1 = \mu_2 = 0 \)). In this case there are no deterministic components in the data or in the cointegrating relations. However, this is quite unlikely to occur in practice, especially as the intercept is generally needed in order to account for adjustments in the units of measurements of the variables in \( (Z_{t-1} | \theta \gamma) \).

Model 2: Intercept (no trend) in CE, no intercept or trend in VAR (\( k \times k \)). This is the case where there are no linear trends in the data, and therefore the first
Data Source and Methodology

Differenced series have a zero mean. In this case the intercept is restricted to the long run model (i.e. the cointegrating equation) to account for the unit of measurement of the variables in $\left( Z_{t-1} \right)$. 

**Model 3:** Intercept in CE and VAR, no trends in CE and VAR ($\delta_1 = \delta_2 = 0$). In this case there are no linear trends in the levels of the data, but we allow both specifications to drift around an intercept. In this case it is assumed that the intercept in the CE is cancelled out by the intercept in the VAR, leaving just one intercept in the short run model.

**Model 4:** Intercept in CE and VAR, linear trend in CE, no trend in VAR ($\delta_1 = 0$). In this model we include a trend in the CE as a trend stationary variable in order to take into account exogenous growth (i.e. technical progress). We also allow for intercepts in both specifications while there is no trend in the short run relationship.

**Model 5:** Intercept and quadratic trend in the CE intercept and linear trend in VAR. This model allows for linear trends in the short run model and thus quadratic trends in the CE. Thus, in this final model everything is unrestricted. However, this model is very difficult to interpret from an economics point of view, especially since the variables are entered as logs, because a model like this would imply an implausible ever-increasing or ever-decreasing rate of change.

So the problem is which of the five different models is appropriate in testing for cointegration. As we said before that the first and the fifth are not that realistic therefore the problem reduces to a choice of one of the three remaining models. Johansen (1992) suggests that we need to test the joint hypothesis of both the rank order and the deterministic components, applying the so called Pantula principle. The Pantula principle involves the estimation of all three models and the presentation of the results from the most restrictive hypothesis (i.e. $r=$number of cointegrating relations=$0$ and model 1) through...
the least restrictive hypothesis (i.e., \( r = \text{number of variables entering the VAR} = n-1 \) and model 4). The model selection procedure then comprise moving from the most restrictive model, at each stage comparing the trace test statistic to its critical value, stopping only when we conclude for the first time that the null hypothesis of no cointegration is rejected.

**Step 4: Determining the rank of \( \Pi \) or the number of cointegrating vectors**

According to Johansen and Juselius (1990), there are two methods for determining the number of cointegrating relations, and both involve estimation of the matrix \( \Pi \). This is a \( k \times k \) matrix with rank \( r \). The procedures are based on propositions about eigen values.

a) One method tests the null hypothesis, that \( \text{rank}(\Pi) = r \) against the hypothesis that the rank is \( r+1 \). So, the null in this case is that there is cointegrating vectors and that we have up to \( r \) cointegrating relationships, with the alternative suggesting that there is \( (r+1) \) vectors. The test characteristics are based on the characteristics roots (also called eigen values) obtained from the estimation procedure. The test consists of ordering the largest eigen values in descending order and considering whether they are significantly different from zero. To understand the test procedure, suppose we obtained \( n \) characteristic roots denoted by \( \hat{\lambda}_1 > \hat{\lambda}_2 > \hat{\lambda}_3 > ... > \hat{\lambda}_n \). If the variables under examination are no cointegrated, the rank of \( \Pi \) is zero and all the characteristic roots will equal to zero. Therefore \( 1 - \hat{\lambda}_1 \) will be equal to 1 and since \( \ln(1) = 0 \), each one of the expressions will be equal to zero for no cointegration. On the other hand, if the rank of \( \Pi \) is equal to 1, then \( 0 < \hat{\lambda}_1 < 1 \) so that the first expression \( 1 - \hat{\lambda}_1 \) <0, while all the rest will be equal to zero. To test how many of the numbers of the characteristic roots are significantly different from zero this test uses the following statistic:

\[
\lambda_{\text{max}}(r, r+1) = -T \ln(1 - \hat{\lambda}_{r+1})
\]  

(3.37)
Data Source and Methodology

As we said before, the test statistic is based on the maximum eigen value and because of that is called the maximum eigen value statistic (denoted by $\lambda_{\text{max}}$).

b) The second method is based on a likelihood ratio test about the trace of the matrix (and because of that it is called the trace statistic). The trace statistic considers whether the trace is increased by adding more eigen values beyond the rth eigen value. The null hypothesis in this case is that the number of cointegrating vectors is less than or equal to r. From the previous analysis it should be clear that when all $\hat{\lambda}_i = 0$, then the trace statistic is equal to zero as well. On the other hand, the closer the characteristic roots are to unity the more negative is the $\ln(1 - \hat{\lambda}_r)$ term and, therefore, the larger the trace statistic. This statistic is calculated by:

$$\hat{\lambda}_{\text{trace}}(r) = -T \sum_{i=r+1}^{n} \ln(1 - \hat{\lambda}_i)$$

The usual procedure is to work downwards and stop at the value of r which is associated with a test statistic that exceeds the displayed critical value. Critical values for both statistics are provided by Johansen and Juselius (1990).

Note that the trace statistic and the maximum eigenvalue statistic may yield conflicting results. For such cases, it is recommended that one should examine the cointegrating vector and base their choice on the interpretability of the cointegrating relations (Johansen and Juselius, 1990).

Step 5: Testing for weak exogeneity

After determining the number of cointegrating vectors we need to proceed with tests of weak exogeneity. Remember that the $\Pi$ matrix contains information about the long run relationships, and that $\Pi = \alpha \beta'$, where $\alpha$ represents the speed of adjustment coefficients and $\beta$ is the matrix of long run coefficients. From this it should be clear that when there are $r \leq n - 1$ cointegrating vectors in $\beta$, then this automatically means that at least $(n-r)$ columns of $\alpha$ are equal.
to zero. Thus, the typical problem faced, of determining how many \( r \leq n - 1 \) cointegration vectors exist in \( \beta \), amounts to equivalently testing which columns of \( \alpha \) are zero.

**Step 6:** Testing for linear restrictions in the cointegrating vectors

An important feature of the Johansen approach is that it allows us to obtain estimates of the coefficients of the matrices \( \alpha \) and \( \beta \), and then test for possible linear restrictions regarding those matrices. Especially for matrix \( \beta \), the matrix that contains the long run parameters, this is very important because it allows us to test specific hypothesis regarding various theoretical predictions from an economic theory point of view.

**SECTION 3.3**

**ARDL Approach:**

Economic literature has abundant techniques to investigate relationships among non-stationary macroeconomic variables and prominent among them are univariate cointegration technique (Engle-Granger (1987)), multivariate cointegration technique (Johansen (1988); Johansen & Juselius (1990)); and Johansen’s (1995)) and newly developed Autoregressive Distributed Lag (ARDL) model (Pesaran and Shin, 1995, 1998; Pesaran et al., 1996; Pesaran et al. 2001). The studies indicate that the ARDL approach to cointegration is preferable to other conventional approaches such as Engle-Granger (1987), Johansen (1988) etc. because of its applicability irrespective of whether the underlying regressors are purely I (0), purely I (1) or mutually cointegrated.

Another reason for preferring ARDL approach over other approaches is that it is more robust and performs better for small sample sizes.

Though ARDL is the better model for testing for small sizes, yet it is not the best as it is not free from the problem of heteroscedasticity and serial correlation. For correcting the problem of autoregression and heteroscedasticity simultaneously we have used GARCH model for finding long-run impact.

The specifications of the ARDL model is given below:
**The GARCH (1,1) Model:**

Generalised Autoregressive Conditional Heteroskedasticity (GARCH) models are specifically designed to model and forecast conditional variances. The GARCH is the extension of ARDL model for capturing volatility in the error term. The variance of the dependent variable is modeled as a function of past values of the dependent variable and independent or exogenous variables.

GARCH (1,1) model specifications:

\[
y_i = \sum_{k=1}^{p} \phi_k X_{it} + \varepsilon_t
\]

\[
\sigma^2_t = \omega + \alpha \varepsilon^2_{t-1} + \beta \sigma^2_{t-1}
\]

in which the mean equation given in equation (3.39) is written as a function of exogenous variables with an error term. Since is the one-period ahead forecast variance based on past information, it is called conditional variance. The conditional variance equation is specified in equation (3.40) is a function of three terms:

1. The constant term \(\omega\)
2. Volatility from the previous period, measured as the lag of the squared residual from the mean equation \(\varepsilon^2_{t-1}\) (the ARCH term).
3. Last period’s forecast variance \(\sigma^2_{t-1}\) (the GARCH term).

The (1,1) in GARCH (1,1) refers to the presence of a first-order autoregressive GARCH term (the first term in parentheses) and a first-order moving average ARCH term (the second term in parentheses). An ordinary ARCH model is a special case of a GARCH specification in which there are no lagged forecast variances in the conditional variance equation — *i.e.*, a GARCH (0, 1).
This specification is often interpreted in a financial context, where an agent or trader predicts this period’s variance by forming a weighted average of a long term average (the constant), the forecasted variance from last period (the GARCH term), and information about volatility observed in the previous period (the ARCH term).

There are two equivalent representations of the variance equation that may aid in interpreting the model:

- If we recursively substitute for the lagged variance on the right-hand side of equation (3.40) we can express the conditional variance as a weighted average of all of the lagged squared residuals:

\[ \sigma_t^2 = \frac{\omega}{(1-\beta)} + \alpha \sum_{j=1}^{\infty} \beta^{j+1} \varepsilon_{t-j}^2 \]  

(3.41)

We see that the GARCH (1,1) variance specification is analogous to the sample variance, but that it down-weights more distant lagged squared errors.

- The error in the squared returns is given by \( v_t = \varepsilon_t^2 - \sigma_t^2 \). Substituting for the variances in the variance equation and rearranging terms we can write our model in terms of the errors:

\[ \varepsilon_t^2 = \omega + (\alpha + \beta)\varepsilon_{t-1}^2 + v_t - \beta v_{t-1} \]  

(3.42)

Thus, the squared errors follow a heteroskedastic ARMA(1,1) process. The autoregressive root which governs the persistence of volatility shocks is the sum of \( \alpha \) plus \( \beta \) in many applied settings, this root is very close to unity so that shocks die out rather slowly.

SECTION 3.5
HYPOTHESES:

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Null Hypothesis</th>
<th>Alternative Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>No impact of exchange rate on economic growth.</td>
<td>Significant impact of exchange rate on economic growth.</td>
</tr>
<tr>
<td>II</td>
<td>No impact of exchange rate on openness.</td>
<td>Significant impact of exchange rate on openness.</td>
</tr>
<tr>
<td>III</td>
<td>No impact of exchange rate on FDI.</td>
<td>Significant impact of exchange rate on FDI.</td>
</tr>
<tr>
<td>IV</td>
<td>No impact of exchange rate on inflation rate</td>
<td>Significant impact of exchange rate on inflation rate.</td>
</tr>
</tbody>
</table>

Source: Framed Elaborations.
Data Source and Methodology

CONCLUSION:

In a nutshell, we can say that the simultaneity among the variables leads them to be treated in the same way as said by Sims (1980). In other words, there should not be any distinction between endogenous and exogenous variables. The features of VAR model have also been discussed i.e. test of Causality. There are two tests of causality namely Granger Causality Test and Sims Causality Test. To start with the VAR model, we have to see whether variables are stationary or not. If they are stationary then only we can apply model. If the variables are stationary of same level of integration the Engle Granger Procedure is used, otherwise other methods are used. All these procedures are discussed in detail.

All of the above mentioned time series analyses would be applied to simulate the impact of exchange rate on the selected macroeconomic indicators. The thorough analysis of these relationships has been examined in chapter 5 and 6.