5.1 Introduction

This chapter focuses on the methodology adopted to reach to a conclusion or finding. The process involves defining data, selecting appropriate econometric technique and the necessary assumptions or properties of the technique to be adopted. As the study is a causal one, that is, the aim is to find a causal relationship between Foreign Direct Investment Inflows and components of India’s Balance of Payments, due consideration has been given to elucidate the techniques employed. In this chapter, all necessary information is explained that will be employed on the data in the next chapter. Wherever data transformation has been done the method and the manner of doing so has also been explained and argued upon in this chapter.

5.2 Sample Period, Data Description and Data Sources

5.2.1 Sample Period

The study is based on sample period from 1970-71 to 2014-15. Though the researcher was interested in a longer period, but due to the unavailability of data, particularly related to Foreign Direct Investment Inflows the study has to be restricted to the following period. Further, from this sample period, periods of short run would also be identified for an analysis in the Short Run period.

5.2.2 Data Description

Data for three variables are used in the study i.e. FDI Inflows (FDII), Current Account Balance (CAB) and Capital Account Balance (KAB). All the three variables are macroeconomic in nature. The data for FDII was available on annual basis while the data for CAB and KAB was available on fiscal year basis. The fiscal year in India is from April 1 of one year to 31st March of next year. Thus, there was no such problem in the KAB and CAB data but together the three variables were not in synchronisation. To avoid this problem and have data of all three variables in terms of fiscal year, the data for FDII has been transformed from annual to fiscal year of India. The practice has been adopted by IMF and World Bank. Two types of statistical practices are employed by IMF. In the first one, the date measuring FDII is the key to deciding the annual FDII. If the timing of FDII is after June 30, it is to be considered in the next annual year but if it is before June 30, it is to be counted for the current
year FDI. This practice does not concern with the data needed for the study. The second technique is relevant to transform the default data to get data for this study.

The default yearly data is divided into two groups, one of 3 months duration and the other of 9 months duration. For this the annual data is divided by 12 and multiplied by the respective months. Though, there will not arise any problem of seasonality still it is assumed that the Inflow of the FDI throughout the year was even. This transformation is done with the help of EViews9 command method. The equation used for transforming the data is as follows:

\[ fdii_t = \frac{fdii_{t+1}}{12} \times 3 + \frac{fdii_{t-1}}{12} \times 9 \ldots (5.1) \]

Where:

- \( t \) = period, e.g. 1991-1992
- \( t+1 \) = latter year, e.g. 1992 for the above period
- \( t-1 \) = previous year, e.g. 1991 for the above period

In order to make it clear a table of data transformation for 3 years with actual data is presented in Table 5.1.

**Table 5.1**

*FDII Data Transformation to Fiscal Year*

<table>
<thead>
<tr>
<th>Year</th>
<th>FDII (US$ millions)</th>
<th>3 months data</th>
<th>9 months data</th>
<th>Fiscal Year</th>
<th>FDII (US$ millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>75</td>
<td>18.75</td>
<td>56.25</td>
<td>Nil</td>
<td>Nil</td>
</tr>
<tr>
<td>1992</td>
<td>252</td>
<td>63</td>
<td>189</td>
<td>1991-92</td>
<td>156.25+63 = 119.25</td>
</tr>
<tr>
<td>1993</td>
<td>532</td>
<td>133</td>
<td>399</td>
<td>1992-93</td>
<td>189+133= 322</td>
</tr>
</tbody>
</table>

Source: Prepared by the research scholar

Table 5.1 shows how the fiscal year data is calculated with the help of the command in the EViews through the equation 3.1. The data for CAB is both positive and negative; while positive indicates Current Account Surplus, negative denotes Current Account Deficit. There are observations for CAB which are positive and therefore it was not possible to use the terminology of Current Account Deficit. The data for KAB is positive throughout and there is no transformation required in both CAB and
KAB. The data for all the variables is in US$ millions. Appendix I (Data Set 4) presents the matrix of the data for total of 45 fiscal years.

### 5.2.3 Data Sources

Secondary data has been collected from three sources i.e. UNCTAD Statistics Database, Department of Industrial Policy and Promotion (DIPP), Government of India and Reserve Bank of India (RBI) Database. The data for FDII has been collected from UNCTAD and DIPP. FDII data from 1970 to 2001 has been collected from UNCTAD and from fiscal year 2000-01 to 2014-15 from DIPP. The balance of payments statistics is published by RBI and thus the data for CAB and KAB is collected from RBI and are represented in fiscal year.

### 5.3 Research Hypotheses

In this section all the hypothesis used in the study are formulated along with the argument for their formulation. For an ease of understanding, this section is further divided into FDII and CAB and FDII and KAB.

#### 5.3.1 FDII and CAB

As it has been discussed that the first objective is to measure the impact of FDI Inflows on Current Account Balance of India’s BOP. The theoretical relation between FDI Inflows and CAB is well known. It states that the Deficit in CAB is financed through the Capital Account and FDII is one of the component of Capital Account. The studies are divided on the issue of causality from unidirectional causality to bidirectional causality and even of no causality as such. The variances in the results may be due to differences in sample time period and even in the previous studies Short Run period was not identified with the help of any of the structural break unit root tests. Thus, the study divided the analysis into Short Run and Long run period.

##### 5.3.1.1 FDII and CAB in Long Run

$H_{01}$: FDII does not Granger cause CAB in the Long Run.

$H_{A1}$: FDII Granger cause CAB in the Long Run.

$H_{02}$: CAB does not Granger cause FDII in the Long Run.
H\textsubscript{A2}: CAB Granger cause FDII in the Long Run.

H\textsubscript{03}: There is no significant impact of FDII on CAB in the Long Run.

H\textsubscript{A3}: There is significant impact of FDII on CAB in the Long Run.

5.3.1.2 FDII and CAB in the Short run

H\textsubscript{04}: FDII does not Granger cause CAB in the Short Run.

H\textsubscript{A4}: FDII Granger cause CAB in the Short Run

H\textsubscript{05}: CAB does not Granger cause FDII in the Short Run

H\textsubscript{A5}: CAB Granger cause FDII in the Short Run

5.3.2 FDII and KAB

Theoretically, it is known fact that FDII is a part of Capital Account of Balance of Payments and thus there is a positive correlation between FDII and KAB. However, the KAB is an overall financial account and is affected by many macro factors both internal to the Capital Account and external to the Capital Account. It is thus required to re-examine whether FDII in India is causing KAB and if yes than by how much. This is important to determine in the historical view of increasing KAB.

5.3.2.1 FDII and KAB in Long Run

H\textsubscript{06}: FDII does not Granger cause KAB in the Long Run.

H\textsubscript{A6}: FDII Granger cause KAB in the Long Run.

H\textsubscript{07}: KAB does not Granger cause FDII in the Long Run.

H\textsubscript{A7}: KAB Granger cause FDII in the Long Run.

H\textsubscript{08}: There is no significant impact of FDII on KAB in the Long Run.

H\textsubscript{A8}: There is significant impact of FDII on KAB in the Long Run.

5.3.2.2 FDII and KAB in the Short Run

H\textsubscript{09}: FDII does not Granger cause KAB in the Short Run.
Chapter 5  

Research Methodology

$H_{A9}$: FDII Granger cause KAB in the Short Run.

$H_{010}$: KAB does not Granger cause FDII in the Short Run.

$H_{A10}$: KAB Granger cause FDII in the Short Run.

$H_{011}$: There is no significant impact of FDII on KAB in the Short Run.

$H_{A11}$: There is significant impact of FDII on KAB in the Short Run.

5.4 Econometric Methodology

This section will cover all the econometric models and estimation methods used in the study to reach to the conclusion. The econometric methodology diagrammatically and briefly was discussed in Chapter 1 (see section 1.6). Still, it is to recall here that the study is dealing with Time Series data and therefore all the typical tools of Time Series analysis would be applied. Here, it should also be noted that Applied Time Series Econometrics would guide the methodology as it may differ in practice than the theoretical approach. A Time Series is any set of data ordered by time. The fact that Time Series data are ordered by time implies some of their special properties and also some specific approaches to analysis. The time ordering enables the estimation of models built upon one variable only – so called Univariate Time Series Models. Before moving to the distinct topics of Univariate Analysis and Multivariate Analysis, few important properties and issues related to Time Series data must be discussed.

5.4.1 White Noise

When a Time Series is estimated and the predictive model is correctly specified then the remaining inestimable part of the Time Series, that is, the errors and residuals must be White Noise. Technically, a series of identically and independently distributed (iid) random variables with 0 mean is white noise.

5.4.2 Stationarity

Stationarity is the most crucial property of Time Series Econometrics. A stationary Time Series shows a diminishing effect of a shock that occurs in time $t$ and finally disappears in time $t + s$ as $s$ tends to infinity. This feature is also known as mean reversion or covariance stationarity.
5.4.3 Differencing

It is the common approach applied to achieve stationary time series. The highest order of differencing in practice is two. The transformation equation used for differencing is as follows:

First Order Difference \( \Delta y_t = y_t - y_{t-1} \)

Second Order Difference \( \Delta^2 y_t = \Delta y_t - \Delta y_{t-1} \)

5.4.4 Deterministic Trend

The presence of a deterministic trend in the Time Series implies that the value of \( y_t \) increases in each period by a deterministic value.

The above are the important properties encountered in the study. These will be specifically discussed as and when required. The econometric methodology for the purpose of study is divided into two approaches: Univariate Analysis and Multivariate Analysis and both are discussed in coming section.

5.5 Univariate Analysis

In econometrics, analysis of single variable on the basis of internal dynamics of the Time Series data is known as Univariate Analysis. It is the basic argument of time series econometrics, that a particular time series can be generated with its past values on the basis of an Auto Regressive Moving Average (ARMA) model having error term as the exogenous variable. ARMA model specification is as follows:

a. Autoregressive process of the order p, AR(p) is described as

\[
y_t = a_0 + \sum_{i=1}^{p} a_i y_{t-i} + \varepsilon_t
\]

b. Moving Average process of the order q, MA(q), is described as

\[
y_t = \sum_{i=0}^{q} \beta_i \varepsilon_{t-i}
\]

c. Autoregressive Moving Average process of the orders p and q, ARMA(p,q), is described as
Chapter 5

Research Methodology

\[ y_t = a_0 + \sum_{i=1}^{p} a_i y_{t-i} + \sum_{i=0}^{q} \beta_i \varepsilon_{t-i} \]

The ARMA modelling is the basic approach for univariate analysis and this is not possible if the data in level is non-stationary. Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) are used to identify the stationarity conditions in ARMA modelling. The two summarised conditions are as follows:

a. For AR(1): \( y_t = a_0 + a_1 y_{t-1} + \varepsilon_t \)

Where \( a_1 < 1 \) (a sufficient and necessary condition for the stationarity of such a process) the ACF is defined as \( P_s = a_1^s \)

b. For MA(1) process \( y_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} \) (such a process is always stationary) the ACF is defined as \( P_0 = 1 \) and \( P_1 = \beta_1 / (1 + \beta_1^2) \) where \( P_s = 0 \) for any \( s > 1 \).

5.5.1 Unit Root Tests

Theoretically, Unit Root Tests test only for the presence of unit roots in the series. However, because a unit root process lies on the edge between a stationary and non-stationary time series, it is used to check whether the time series is stationary or not. A stationary time series has following properties:

1. a constant and finite long run mean \( \mu = \text{E}(y_t) \) to which it tends to revert after any disturbance (mean reversion).
2. a constant and finite variance \( \sigma^2 = \text{var}(y_t) = \text{E}(y_t - \mu)^2 \)
3. a constant and finite covariance that can vary only with \( s \) but not \( t \)

\( \gamma_s = \text{cov}(y_t, y_{t-s}) = \text{E}[(y_t - \mu)(y_{t-s} - \mu)] \)

4. the expected length of times between crossing the long run mean value \( \mu \) is finite the theoretical auto correlation \( P_s \) decreases steadily with \( s \) and their sum is finite.

On the other hand, a unit root time series has the following properties:

1. No mean reversion, even if it has a constant long run mean (e.g. a random walk)
2. Its variance goes to infinity as \( t \) goes to infinity.
3. Theoretical auto correlations $P_s$ are not independent of $t$ and converge to 1 for all $s$ as $t$ goes to infinity.

4. Except Cointegration analysis, not much can be investigated with non-stationary Time Series.

5. Correlogram cannot be used in isolation of Unit Root Test.

In this study, two Unit Root test would be simultaneously used to identify the order of Integration for stationary Time Series. It means that if order of Integration is 1, that is, I(1) the series would be stationary at first order differencing. However, if the series is found to be integrated of order 0 (I=0), it means that the series is stationary in level data. Natural Log of the series automatically removes the non-stationarity due to exponential function in a Time series. This method of taking log can be used as a primary method. The two Unit Root Tests would be Augmented Dicky Fuller Test (Dicky & Fuller, 1981) which is a modified version of Dickey Fuller Test (Dicky & Fuller, 1979) and KPSS unit root test that owes its name to the initials of Kwiatowski, Phillips, Schmidt, and Shin (1992). Both tests and their models are discussed below:

5.5.1.1 Augmented Dicky Fuller (ADF) Test

The ADF is the modified version of DF test where the latter used models that estimated the Time Series with simple AR(1) process. Any series having a higher order AR(p) or ARMA (p,q) of more than, would not give residuals as White Noise and Autocorrelation would be present among the observations in the series. This limitation has been solved in the ADF test. Three models for ADF test are available which are as follows:

Model A: Check for Stationarity (Neither Intercept nor Trend)

$$\Delta y_t = \gamma y_{t-1} + \sum P_i \Delta y_{t-i} + \epsilon_t$$

Model B: Check for Level Stationarity (Only Intercept in the equation)

$$\Delta y_t = \mu + \gamma y_{t-1} + \sum P_i \Delta y_{t-i} + \epsilon_t$$

Model C: Check for Trend Stationarity (Intercept and Trend in the equation)

$$\Delta y_t = \mu + \beta t + \gamma y_{t-1} + \sum P_i \Delta y_{t-i} + \epsilon_t$$

Where in all cases $H_0: \gamma = 0$ of a unit root time series

$H_A: \gamma < 0$ of a stationary time series
The t statistic given by $t_{DF} = \hat{y} \div SE(\hat{y})$ enable us to test the hypothesis. If $t_{DF}$ (in absolute terms) is lower than the appropriate absolute critical value, then the $H_0$ of the presence of a unit root is accepted against $H_A$ of a stationary time series. If $t_{DF}$ (in absolute terms) is greater than the appropriate absolute critical value, then $H_0$ of the presence of a unit root is rejected. Another question needs to be answered; that is which model of Unit Root to be selected for checking the order of Integration or Stationarity. In this regard, there are two methods, one is the practitioner’s intuition by observing the pattern of graph and the other being checking for deterministic trend. In the first method, Model A, B or C is selected by a visual inspection of the data (using economic intuition). If the time series steadily increases, then Model C is employed and trend stationarity is checked otherwise Model B is used. If the Time Series oscillates around 0 then Model A of neither trend nor intercept is used. The second method is much more objective than first one. In second method, the series is tested for unit root by selecting Model C of trend and intercept. If the trend value is significant (decided on the basis of probability value), Model C is continued for further analysis. If the trend is found to be insignificant (prob. value is more than 5%) then Model B (only intercept) is used. However, Model A is used only by observing the graph showing oscillation trend.

5.5.1.2 KPSS Test

The main difference in this test is transposition of the hypothesis. In this test the null hypothesis is of “stationary series” unlike “non stationary series” in ADF. KPSS (Kwiatkowski, Phillips, Schmidt & Shin, 1992) test assumes that a Time Series is tested for Trend Stationarity and can be represented in the following equation:

$$T = \beta t + \gamma_t + \epsilon_t$$

In the above equation of KPSS, first term ($\beta t$) shows deterministic trend, second term ($\gamma_t$) shows random trend and the third term ($\epsilon_t$) represents error trend. The equation may be re-written as $y_t = \beta_t + \gamma_t + \epsilon_t$, where $\gamma_t = \gamma_{t-1} + \mu_t$ and $\mu_t$ is normal iid (independently and identically distributed). In KPSS using OLS one of the following model is estimated:

Model A: Check for Level Stationarity (Only Intercept in the equation)

$$y_t = \alpha_0 + \epsilon_t$$
Model B: Check for Trend Stationarity (Intercept and Trend in the equation)

\[ y_t = a_0 + \beta_t + \varepsilon_t \]

The hypothesis are:

- H_0: \( \sigma^2_{\mu} = 0 \) of a stationary time series
- H_A: \( \sigma^2_{\mu} \neq 0 \) of a unit root/ non stationary series

If the KPSS test statistic exceeds the appropriate critical value. Then H_0 of a stationary Time Series is rejected in favour of H_A of a unit root series.

5.5.1.3 Combining ADF and KPSS

The ideal method of Unit Root testing is to combine both the ADF and KPSS test results. If a Time Series is found stationary with the ADF test, then it will be most likely found stationary also when using the KPSS test and it is accepted as stationary, indeed. If a Time Series is found non stationary using the KPSS test, then most likely it will be found non stationary also with the ADF test. However, it can happen quite often that a Time Series that was found non stationary using the ADF test will be marked as stationary with the KPSS test. In such cases the strength of the evidence is checked by the researcher for the stationarity in KPSS and non stationarity in ADF and decide accordingly. Table 5.2 below summarizes the combined result of ADF and KPSS.

Table 5.2

<table>
<thead>
<tr>
<th>Combined Decision for ADF and KPSS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ADF Result</strong></td>
</tr>
<tr>
<td>Stationary</td>
</tr>
<tr>
<td>Non stationary</td>
</tr>
<tr>
<td>Non stationary</td>
</tr>
<tr>
<td>Stationary</td>
</tr>
</tbody>
</table>

Source: Prepared by the research scholar
In case there is contradiction in the results of ADF and KPSS, additional tests would be conducted to take a final decision. In this regard, Dickey Fuller-GLS and PP Test will be additionally used that are discussed in next section.

5.5.1.4 Dickey-Fuller Test with GLS

With a case of having a constant in ADF Regression, ERS (1996) proposed a simple modification of the ADF tests in which the data are detrended so that explanatory variables are “taken out” of the data prior to running the test regression. ERS defined a quasi-difference of \( y_t \) that depends on the value \( a \) representing the specific point alternative against which we wish to test the null:

\[
d(y_t \mid a) = \begin{cases} 
  y_t & \text{if } t = 1 \\
  y_t - ay_{t-1} & \text{if } t > 1
\end{cases}
\]

Next, consider an OLS regression of the quasi-differenced data \( d(y_t \mid a) \) on the quasi-differenced \( d(x_t \mid a) \):

\[
d(y_t \mid a) = d(x_t \mid a)'\delta(a) + \eta_t
\]

Where \( x_t \) contains either a constant, or a constant and trend, and let \( \delta(a) \) be the OLS estimates from this regression. All that is needed now is a value for \( a \). ERS recommends the use of \( a = \bar{a} \), where:

\[
\bar{a} = \begin{cases} 
  1 - 7/T & \text{if } x_t = \{1\} \\
  1 - 13.5/T & \text{if } x_t = \{1, t\}
\end{cases}
\]

It is here important to define the GLS detrended data, \( y_t^d \) using the estimates associated with the \( \bar{a} \):

\[
y_t^d = y_t - x_t'\delta(\bar{a})
\]

Then the DF-GLS test involves estimating the standard ADF test equation after substituting the GLS detrended \( y_t^d \) for the original \( y_t \):

\[
\Delta y_t^d = \alpha y_{t-1}^d + \beta_1 \Delta y_{t-1}^d + \cdots + \beta_p \Delta y_{t-p}^d + \nu_t
\]

Note that since the \( y_t^d \) are detrended, there is no need to include the \( x_t \) in the DF-GLS test equation. As with the ADF test, generally it is considered the t-ratio for \( \bar{a} \) from this test equation. While the DF-GLS t-ratio follows a Dickey-Fuller distribution in
the constant only case, the asymptotic distribution differs when you include both a constant and trend. Thus, the EViews lower tail critical values use the MacKinnon simulations for the no constant case, but are interpolated from the ERS simulated values for the constant and trend case.

5.5.1.5 The Phillips-Perron (PP) Test

Phillips and Perron (1988) propose an alternative (nonparametric) method of controlling for serial correlation when testing for a unit root. The PP method estimates the non-augmented DF test equation and modifies the t-ratio of the \( \alpha \) coefficient so that serial correlation does not affect the asymptotic distribution of the test statistic. The PP test is based on the statistic:

\[
\tilde{\tau}_\alpha = t_\alpha \left( \frac{y_0}{f_0} \right)^2 - \frac{T}{2} \frac{(f_0 - y_0)(se(\widehat{\alpha}))}{se^2(\beta)}
\]

Where \( \hat{\alpha} \) is the estimate, and \( t_\alpha \) the t-ratio of \( \alpha \), \( se(\widehat{\alpha}) \) is coefficient standard error, and \( s \) is the standard error of the test regression. In addition, \( y_0 \) is a consistent estimate of the error variance. The remaining term, \( f_0 \), is an estimator of the residual spectrum at frequency zero. There are two choices available when performing the PP test. First, whether to include a constant, a constant and a linear time trend, or neither, in the test regression. Second, is to choose a method for estimating \( f_0 \). The asymptotic distribution of the PP modified t-ratio is the same as that of the ADF statistic. EViews reports MacKinnon lower-tail critical and \( p \)-values for this test.

5.5.1.6 Unit Root for Structural Change/Breakpoint Unit Root Test

One of the feature of Time Series is the structural changes in the data for the sample period. Particularly, this happens in a data collected over a long period of time. As the frequency of data increases, the possibility of the structural changes also increases. Few of the causes of structural changes in the Time Series data are change in policy, natural disaster, political change of power, changes in the fundamentals of the economic system, international monetary agreements like Bretton Woods etc. From the point of view of the Applied Time Series analysis, the structural break is not always known and they may be predicated under Unit Root testing with structural changes. The structural changes approach of Unit Root testing can be traced to the study of Perron (1989) in which he identified that 14 macroeconomic variables that
were identified as non-stationary by the researchers were actually stationary by considering two shocks as structural changes. The shocks included the Great crash of 1929 and the first oil price shock of 1973. The Perron’s Test allowed testing for the broken trend stationarity of a time series. Zivot and Andrews (1992) argued for the weak results of Perron Unit Root test for arbitrarily choosing the trend break dates. Thus, they transformed the Perron’s broken trend stationarity into an unconditional broken trend stationarity test. To perform Zivot and Andrews test OLS estimation of one of the equations is done:

Model A:  
\[ y_t = \mu + \beta t + \theta DU_t(\lambda) + \alpha y_{t-1} + \sum_{i=1}^{k} p_i \Delta y_{t-i} + \varepsilon_t \]

Model B:  
\[ y_t = \mu + \beta t + \gamma DU_t(\lambda) + \alpha y_{t-1} + \sum_{i=1}^{k} p_i \Delta y_{t-i} + \varepsilon_t \]

Model C:  
\[ y_t = \mu + \beta t + \theta DU_t(\lambda) + \gamma DU_t(\lambda) + \alpha y_{t-1} + \sum_{i=1}^{k} p_i \Delta y_{t-i} + \varepsilon_t \]

In all three models,  
\[ H_0: \alpha = 1 \]
\[ H_A: \alpha < 1 \]

If t statistic is lower than the appropriate critical value, then  \( H_0 \) of a unit root is rejected. Vogelsang Test (1997) as a test identifying single structural change is also very important. This test enables a decision whether there is a single break in the trend function of a given time series and if so, it also enables an estimation of the break date. It focuses only on structural changes and ignores the Unit Root issue though, the critical values will differ for stationarity and non-stationarity Time Series. The potential weakness identified by researchers are that it allows for a single break only in the trend function. For multiple structural change there are several test such as Bai and Perron (1998), Bai and Perron (2003), Gilman and Nakov (2004) etc. EViews9 has been used in the study and it supports the computation of modified Dickey-Fuller tests which allows for levels and trends that differs across a single break date. The framework follows the work of Perron (1989), Perron and Vogelsang (1992), Vogelsang and Perron (1998), Banerjee, Lumsdaine and Stock (1992). These test will be used to identify the structural breaks and short run analysis would be conducted between two selected breaks.
5.5.1.7 Diagnostics of Residuals

Residuals from a correctly specified model that captures the data generating process well should be White Noise. It means that they contain no further information that would help with estimation. White noise has all auto correlation equal to zero.

5.5.1.8 Information Criteria

The parsimonious model is selected that satisfactorily captures the dynamics of the data. The following information criterion will be used in the study:

Akaike Information Criterion (Akaike, 1978)

$$AIC = T \ln \text{SSR} + 2n$$

Schwarz Bayes Information Criterion (Schwarz, 1978)

$$SBIC = T \ln \text{SSR} + n \ln T$$

Hannan-Quinn Information criterion (Hannan & Quinn, 1979)

$$HQIC = T \ln \text{SSR} + 2n (\ln(\ln T))$$

Where, SSR is the sum of the residuals squares;

N is the number of explanatory variables (n = p+q+1 if a constant term is included);

T is the number of usable observations.

In order to select the best model the value of the information criteria is to be minimized.

5.6 Multivariate Analysis

Analysis of more than one variable is considered under Multivariate Analysis. In other words when more than one variable is divided into the category of endogenous and exogenous variables, the approach comes under the category of Multivariate Analysis. When the number of such variables in a particular equation is two, it comes under the category of bivariate analysis. First, the higher order topics such as Vector Auto Regression, Granger causality and Cointegration would be discussed followed by Quantile Regression and Two Stage Least Squares (TSLS) Regression.
5.6.1 Vector Auto Regression Model

Sims (1980) has talked a lot and contributed a lot towards the development of present day Vector Auto Regression (VAR) models. The basic VAR model consists of modelling between two variables. Suppose, two time series $y_{1t}$ and $y_{2t}$ are available and both variables interact with each other. In such a case interaction is assumed without the prior knowledge of which of them is exogenous. The following equations defines the model:

\[ y_{1t} = b_1 + \alpha_{12} y_{2t} + \sum_{i=1}^{p} \left( \gamma_{11}^{(i)} y_{1t-i} + \gamma_{12}^{(i)} y_{2t-i} \right) + \mu_{1t} \]

\[ y_{2t} = b_2 + \alpha_{21} y_{1t} + \sum_{i=1}^{p} \left( \gamma_{21}^{(i)} y_{1t-i} + \gamma_{22}^{(i)} y_{2t-i} \right) + \mu_{2t} \]

Where $y_{1t}$ and $y_{2t}$ are assumed to be stationary and $u_{1t}$ and $u_{2t}$ are assumed to be mutually uncorrelated white noise disturbances with variance $\sigma_1^2$ and $\sigma_2^2$. The matrix form of the model is as follows:

\[ y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix}, A = \begin{bmatrix} 1 & -\alpha_{12} \\ -\alpha_{21} & 1 \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \prod_i = \begin{bmatrix} \gamma_{11}^{(i)} & \gamma_{12}^{(i)} \\ \gamma_{21}^{(i)} & \gamma_{22}^{(i)} \end{bmatrix}, \text{ and } u_t = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}. \]

\[ Ay_t = b + \sum_{i=1}^{p} \prod_i y_{t-i} + u_t \]

The above model is the prominent and popular structural VAR model structured by the $p^{th}$ order of two variables. The model can be extended to more than two variables also. In such cases the N ray matrix will change and will show a different pattern. Thus, when the structural VAR is increased to N interrelated variables, the size of vectors becomes $N * 1$ and the size of matrices from the default $2 * 2$ to $N * N$. The main feature of structural VAR is that its nature is theoretical and in applications it can be estimated due to the presence of specification correlation between $y_{2t}$ and $u_{1t}$ on one hand and $y_{1t}$ and $u_{2t}$ on the other hand. Due to this reason, the standard practice in Applied Econometrics is to transform the structural VAR into the reduced form of VAR by multiplying the structural VAR equation with polynomial $A^{-1}$ to obtain matrix associated with $y_t$ which may be termed as identity matrix.
\[ y_t = \mu + \sum_{i=1}^{p} \Pi^i y_{t-i} + \varepsilon_t, \]

Where \( \mu = A^{-1} b = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad \Pi^i = A^{-1} \begin{bmatrix} \pi_{11}^{(i)} & \pi_{12}^{(i)} \\ \pi_{21}^{(i)} & \pi_{22}^{(i)} \end{bmatrix}, \) and \( \varepsilon_t = A^{-1} u_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \)

The above equation is estimated by econometricians on the basis of it being reduced form by OLS according to equation. On the right hand side of the equation, lagged values of \( y_{1t} \) and \( y_{2t} \) appear along with those that are uncorrelated with \( \varepsilon_{1t} \) and \( \varepsilon_{2t} \). It is to be noted that in case the original structural VAR consists of \( N \) equations for \( N \) variables the reduced form VAR would also consists of \( N \) equations for \( N \) variables.

The study is based on bivariate analysis which is the basic part of Multivariate Analysis. The important and necessary features of the Bivariate VAR are as follows:

1. The Bivariate Model is a replica of simultaneous equation used in mathematical equation. But there is a fundamental difference. The difference being that each equation under the Bivariate VAR contains only its own lagged values and the lagged values of the other variables in the system and no current values of the two variables are included on the right hand side of the set of equations.

2. In most cases, to attain the objective of parsimonious model the number of lagged terms in each equation remains same.

3. In a Bivariate VAR, there can be only two maximum cases of Cointegration. Either there can be found no cointegration or one cointegration. The reason being that for \( n \) variable VAR system Cointegration is defined by \( (n-1) \). If for a Bivariate VAR the Cointegration result is more than one, there may be main two reasons for the same. The first may be due to model misspecification and the second may be due to wrong lag selection.

### 5.6.1.1 Stability and Stationarity of VAR Models

Here it is important to recall that in an AR \((p)\) model with the equation \( y_t = \mu + \sum_{i=1}^{p} a_i y_{t-i} + \varepsilon_t \) or \( A(L)y_t = \mu + \varepsilon_t \) where \( A(L) = 1 - \sum_{i=1}^{p} a_i L^i \), the necessary and sufficient conditions for the stability of the equations (or model) and the stationarity of the generated time series requires that the following condition be fulfilled with respect to the characteristic roots:
This means that the AR characteristic must lie within the unit circle. In simple terms it refers that all the characteristic roots must be less than one in absolute value. In terms of inverse characteristic equations the condition is \( A(L) = 0 \). With respect to a first order VAR system, it is stable if all the eigenvalues \( \alpha_1, \alpha_2, \ldots, \alpha_N \) of the matrix \( \prod 1 \) lie within the unit circle (less than one in absolute terms). The equation representing eigenvalue is as follows:

\[
\alpha l - \prod 1 = 0 \text{ (absolute terms)}
\]

Where the equation indicates the determinant of a matrix M.

The issue of stationarity is an important topic under VAR models. The mainstream view is that VAR models should be applied to stationary time series only. However, in Applied Time Series analysis a strong opinion exists that VAR models can be estimated even with non-stationary data. The main argument in favour of non-differencing the data is that vital information regarding the observations and the internal dynamics of the data are lost while differencing. In this regard, Kocendy & Cerny (2014) suggests that if there is found Cointegration of order 1 than VAR model should be built upon non-differenced data. Further, they have stated that in case of 1 Cointegration, if the data is differenced there will be specification error in the model. On the other hand, the Toda and Yamamoto (1995) approach argues for setting up VAR in level data. This is in order to maintain the asymptotic chi square null distribution that is the requirement to reach to causality particularly in case of macroeconomic variables.

5.6.1.2 Estimation of a VAR Model

The method of estimating a VAR model is through application of OLS to each equation separately. For the same, order of the VAR must be identified that is the number of lags \( p \). Technically, the number of lags \( p \) and the number of regressors are same in all VAR equations. It means that for each equation the lags remain same and therefore it is known as Unrestricted VAR. This condition when not followed leads to a new VAR model known as Near VAR. A Near VAR is identified by the condition that all the equations does not have the same number of lags. Thus, in Near VAR certain restrictions are put on certain equations of VAR. In other words, a restricted
VAR is known as Near VAR. In order to choose the order p, minimization of AIC, HQIC or SBIC will work as a simple method.

### 5.6.2 Granger Causality

Granger Causality is most prominent causality method used in Applied Econometrics and can be traced to Granger (1969). The main reason for its prominence is that it enables causality testing between variables in the VAR models. A simple $X$ Granger cause $Y$ if $Y$ can be better predicted using the histories of both $X$ and $Y$ than it can by using the history of $Y$ alone. In order to test the null hypothesis of $x_t$ not Granger causing $y_t$ the following unrestricted and restricted model specifications are used:

**Unrestricted:**
\[ y_t = \alpha + \sum_{i=1}^{p} \alpha_i x_{t-i} + \sum_{i=1}^{p} \delta_i y_{t-i} + \epsilon_t \]

**Restricted:**
\[ y_t = \alpha + \sum_{i=1}^{p} \alpha_i y_{t-i} + \epsilon_t \]

Under the null hypothesis, the lagged values of $x_t$ are assumed to have no explanatory power on the current values of $y_t$. Thus mathematically,

\[ H_0: \delta_1 = \delta_2 = \delta_3 = 0 \]

Using a simple F-test each pair of variables $x_t$ and $y_t$, two Granger Causality tests can be performed. One can be with the null hypothesis of $x_t$ not Granger causing $y_t$ and second with the null of $y_t$ not Granger causing $x_t$. Each of the hypothesis is checked separately on the basis of respective prob. value. The results of Granger Causality can have three outcomes:

1. If $y_t$ Granger causes $x_t$ and $x_t$ also Granger causes $y_t$, then a feedback causality or bidirectional causality is present between the two variables.
2. If one variable Granger causes the other but not the other way round, then one variable is weakly exogenous to other and only unidirectional causality is present.
3. If no Granger causality is detected, then the two variables do not interact with each variables.

There are two types of approaches to perform Granger causality. First approach that is the mainstream approach is to first make the series stationary and then apply the Granger causality test. Remember that it includes the VAR modelling step. Another
view is of Toda and Yamamoto (1995) where the Granger causality can be performed with the data in levels whether stationary or non stationary. However, knowing the order of integration is necessary in the approach of Toda and Yamamoto.

5.6.2.1 Toda and Yamamoto Approach to Granger Causality

Until few years back this approach to Granger Causality was not used in application. But in last few years overwhelming evidence of the strong estimating strength of this approach has appeared in the literature. In this approach there is no need to stationary the data series but only to know the order of Integration using ADF and KPSS test (for cross check). Cointegration is also used in this approach to cross check the results. The stepwise procedure for checking Granger Causality through T-Y approach is as follows:

Step 1: Testing each Time Series individually with ADF and KPSS for identifying the order of Integration. Both the Unit Root test are simultaneously used for avoiding any pitfall and for cross check.

Step 2: Finding the maximum order of Integration using Step 1 for the two series in Bivariate Analysis. Let the maximum order of integration be m. So if one series is I(1) and the other is I(0), the maximum order of Integration (m) is 1.

Step 3: This is the unique step of T-Y approach. Setting up the VAR in levels of data irrespective of the order of Integration. Remember that the data of the two series must not be differenced or detrended.

Step 4: Determining the appropriate maximum lag length for the variables in the VAR model using the minimized information criteria.

Step 5: Check and correct for any serial correlation in the VAR model.

Step 6: If the two series has same order of Integration it is important to check for Cointegration. For this Johansen’s methodology based on VAR should be used for reliable results.

Step 7: There will be no further role of step 6 in analysis as it is to ultimately cross check the causality results.
Step 8: The VAR model must be re-estimated with an additional lag $m$ (maximum order of integration found in step 1).

Step 9: Test the hypothesis that the coefficients of the first $p$ lagged values of $X$ are 0 in the $Y$ equation and vice versa using a standard Wald test. The coefficients of the extra lags $m$ should not be included while performing the Wald test. The Wald test statistic will be asymptotically chi square distributed with $p$ degree of freedom under null.

Step 10: Rejection of null hypothesis denotes presence of Granger causality.

Step 11: The results in step 10 must be cross checked with the result of Cointegration in Step 6. If two series are found Cointegrated in step 6 the results must suggest Granger Causality. If not then there is conflict in the results and this may happen when the sample size is too small to satisfy the asymptotics that the Cointegration and Causality tests rely on. If two series are not cointegrated in step 6, there is no point available for cross check. However, it does not mean that there cannot be causality.

5.6.3 Quantile Regression

While generally, Classical Linear Regression is used to identify the significant impact of one variable (independent) on another variable (dependent), the nature of the series Current Account Balance (CAB) and Capital Account Balance (KAB) do not allow to use Classical Linear Regression. The reason being that CAB and KAB both consider positive and negative values and this will not be fit for conditional means on which Linear regression is based. Thus, another regression based on another central tendency, that is, Median would be employed for the study known as Quantile Regression.

There is an increasing approach in methods of modelling other aspect of the conditional distribution (other than conditional mean). The recent approach in economics is to use Quantile Regression which models the quantiles of the dependent variable given a set of conditioning variables. The model was originally proposed by Koenkar and Bassett (1978) and it provides estimates of linear relationship between regressors $X$ and a specified quantile of the dependent variable $Y$. When it is use as Least Absolute Deviations (LAD) estimator, the results corresponds to fitting the conditional median of the response variable. The benefit of Quantile Regression is
that it allows a complete description of the conditional distribution to describe how 
the median or $n^{th}$ percentile of the response variable is affected by regressors. It is to 
be added that Quantile Regression approach does not require strong distributional 
asumptions and thus it offers robust method of modelling these relationships.

5.6.3.1 Model of Quantile Regression

Assume a random variable $Y$ with probability distribution function as

$$F(y) = \text{Prob}(Y \leq y)$$

So that for $0 < \tau < 1$, the $\tau^{th}$ quantile of $Y$ may be defined as the smallest $y$ 
satisfying $F(y) \geq \tau$:

$$Q(\tau) = \inf\{y: F(y) \geq \tau\}$$

Given a set of $n$ observations on $Y$, the traditional empirical distribution function is 
given by:

$$F_n(y) = \sum_{k} 1(Y_t \leq y)$$

Where $1(z)$ is an indicator function that takes the value 1 if the argument $z$ is true and 
0 otherwise. The linked empirical quantile is given by:

$$Q_n(\tau) = \inf\{y: F_n(y) \geq \tau\}$$

Or equivalently, in the way of a simple optimization problem:

$$Q_n(\tau) = \arg\min_{\xi} \{\sum_{t} P_{\tau}(Y_t - \xi)\}$$

Where $P_{\tau}(\mu) = \mu(\tau - 1 (\mu < 0))$ is the check function which weights positive and 
negative values asymmetrically. The extension of this simple formulation is allowed 
for regressors $X$ as it is assumed to be a linear specification for the conditional 
quantile of the response variable $Y$ given values for the $P$ vector of explanatory 
variable $X$:

$$Q(\tau|X_i, \beta(\tau)) = X_i\beta(\tau)$$
Where $\beta(\tau)$ is the vector of coefficients associated with the $\tau^{th}$ quantile. The conditional quantile regression estimator is given by the following equation:

$$\beta_n(\tau) = \arg\min_{\beta(\tau)} \{ \sum_i P_i (Y_i - X_i^T \beta(\tau)) \}$$

### 5.6.3.2 Slope Equality Testing

Koenker and Bassett (1982) propose testing for slope equality across quantiles as a robust test of heteroskedasticity. The null hypothesis is given by:

$$H_0: \beta_1(\tau_1) = \beta_1(\tau_2) = \cdots = \beta_1(\tau_k)$$

which imposes $(p-1)(K-1)$ restrictions on the coefficients. We may form the corresponding Wald statistic, which is distributed as a $\chi^2_{(p-1)(K-1)}$. The null hypothesis is “The slopes of the quantiles are equal”. The null hypothesis is rejected if the prob. value of Wald Test is less than 5% indicating that coefficients differ across quantile values and that the conditional quantiles are not identical. On the other hand, if the null hypothesis is accepted it means vice versa.

### 5.6.3.3 Quantile Process Estimates (Process Coefficients)

Under this function, coefficients for 10 quantiles would be processed by the developed model. Even four quantiles can be selected but ten will be able to explain the impact thoroughly. Each coefficient will be shown with the probability value that will help to judge whether there is any significant impact of the independent variable or not on a specific quantile.

### 5.6.4 Two Stage Least Squares Regression (TSLS)

A special case of Regression where there is option to consider an instrumental variable. The instrumental variable is a term used for estimation when correlation between the explanatory variables and the error term is suspected may be due to omitted variables, measurement error etc. This type of regression is based on a two stage analysis. The first stage comprises of finding portions of the endogenous and exogenous variables which can be linked to the instruments. Under this first stage, an OLS regression is estimated for each variable in the model on the set of instruments. Subsequently, the second stage is a regression of the original equation with all the variables replaced by the fitted values from the first step regression. The Two Stage
Least Squares estimates are coefficients from the second stage. The Linear TSLS function is represented as:

$$\psi(\beta) = (Y - X\beta)'Z'(Z'Z)^{-1} Z'(Y - X\beta)$$

Where $Z$ is the matrix of instruments

$Y$ is dependent variable

$X$ is explanatory variable

The coefficients are represented by equation:

$$b_{TSLS} = (X'Z (Z'Z)^{-1} Z)^{-1} X'Z (Z'Z)^{-1} Z'Y$$

The standard estimated covariance matrix of these coefficients is computed using the following equation:

$$\sum_{TSLS} = s^2 (X'Z (Z'Z)^{-1} Z)^{-1}$$

Where $s^2$ is the estimated residual variance. The criteria for accepting or rejecting null hypothesis remains same as for Linear Regression. The validity of the model is additionally checked through Residual Histogram Plots (Jarque-Bera Statistics), Heteroskedasticity Tests and LM Serial Auto Correlation Tests.