CHAPTER 3
MODEL SPECIFICATION AND ECONOMETRICS METHODOLOGY

3.1 Introduction

Before estimating the empirical models in the next chapter, we will discuss a brief about the methodology. The study will develop the model based on the theoretical framework, as given the previous chapter. The research utilizes data from different macroeconomic indicators and BD and CAD for BRICS countries from 1990 to 2018. The methodological research was conducted to analyse the interaction between the variables. A serious attempt has been made by the author to discuss all the objectives for each country in their respective chapters (chapter 4 to 8) while keeping up an overall progression of the perspective all through the thesis. To achieve the objectives of the study, various econometric methods have been employed keeping into consideration the objective of the study. In defining the variables, considerable attention has been given to the units/denominations to prevent any fallacious or deceptive effects.

The first section 3.2 will give model specification. Section 3.3 belongs to methodological of the study. The methodological section has four sub-sections. The first sub-section gives Augmented Dickey-Fuller (ADF) (1981) and Phillips-Perron (PP) (1988) unit root in sections 3.3.1. The cointegration estimation of autoregressive distributed lag (ARDL) system provide the next segment of 3.3.2. Another 3.3.3 sub-section would include the methodology of causality estimation. This will assist us in understanding the causality trajectory between the variables. Final sub-section 3.3.4 will discuss impulse response function. This technique will give us input and output relationship with the respect one positive shock.

3.2 Model Specification

The model for the study is calculated, according to the theoretical literature and methodology of previous analytical studies, the Indian current account deficits are focused on budget deficits, inflation, rate of interest and exchange rate. The relationship between twin deficits can be defined in an implied form to provide this equation:
In order to explore the relationship between BD and CAD in the macroeconomic framework (twin deficits theorem) for BRICS, this segment introduces a tractable open economy by including current account deficit, budget deficit, inflation, interest rate, exchange rate, supply of money and tax revenue based on past literature review and theoretical framework. There may be an issue of concurrence between the current account deficit and real exchange rate. However, all the variables are incorporated in order to catch the transmission mechanism of the twin deficit as described by Kim and Roubini (2008), Miller & Russek (1989) and Barro’s (1974). Based on the open economy model of Mundell/Fleming with greater global capital mobility, the association between the CAD and BD can happen directly through higher absorption capacity or indirectly by monetary shocks. The below equation (1) represents the twin deficit model:

$$BD_t = \alpha_0 + \alpha_1 CAD_t + \alpha_2 INF_t + \alpha_3 INT_t + \alpha_4 REER_t + \alpha_5 MS_t + e_t \qquad (1)$$

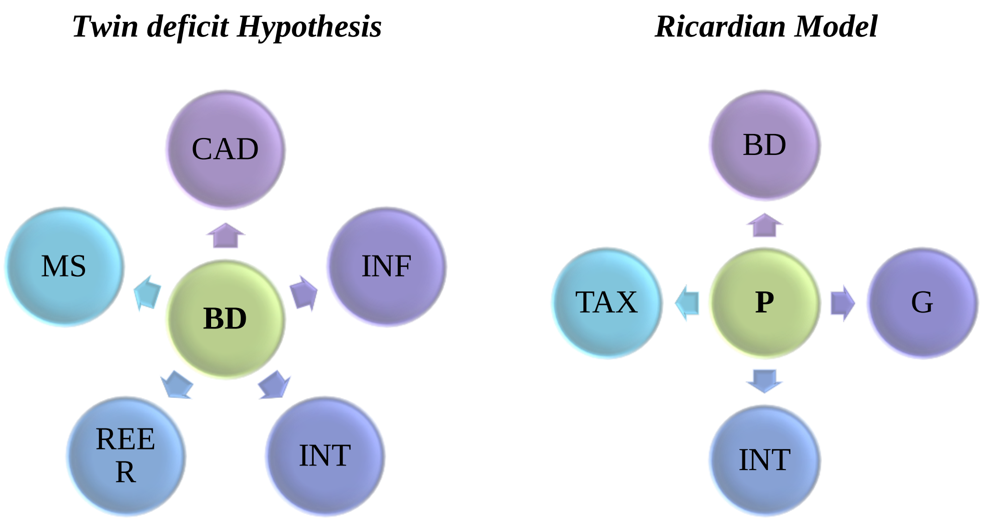
Where BD_t is the budget deficit, CAD_t is the current account deficit, INF_t is inflation, REER is the real effective exchange rate, MS_t is money supply, INT_t is the interest rate and e_t is a white random process. Based on the macroeconomic theory, estimation of α_1; α_2; α_3; and α_5 are supposed to be positive. This means that budget deficit, inflation, interest rate and money supply, may deteriorate current account balance. However, the impact of α_4 real effective exchange rate may have a positive or negative relationship because the exchange rate is characterized as per US dollar. The depreciation in the exchange rate will raise the value of the foreign currency, it will increase the demand for domestic money and α_4 will have a positive relationship. However, the depreciation in the exchange rate, people will hold more foreign currency as compared to domestic currency and α_4 will be negative.

The model 2 will estimate Ricardian Equivalence hypothesis based on Bernheim (1987) consumption equation. Most of the empirical literature estimate Euler equation or reduced form of the consumption equation. But, in this study we will apply the reduced form of the consumption function as given below:

$$P_t = \beta_0 + \beta_1 Y_t + \beta_2 G + \beta_3 BD_t + \beta_4 TAX_t + \beta_5 INT_t + X_t\beta + \varepsilon_t \qquad (2)$$

Where P is the private consumption, G is the government expenditure, BD is the budget deficit, Tax is the tax revenue and INT is the interest rate. We applied ARDL bound testing approach for long-run relationship. Based on equation (1 and 2) we have estimated two models the first one is twin deficit hypothesis and the second one is Ricardian equivalence hypothesis as given below in figure (3.1).

Figure 3.1: Model estimation twin deficit hypothesis (TDH) and Ricardian theorem



Note: Current account deficit (CAD), budget deficit (BD), inflation (INF), interest rate (INT), real effective exchange rate (REER), money supply (MS), Government expenditure (G) and private consumption (P).

3.3 Econometric Methodology

Different econometric methods have been used to estimate the linkage between the variables. Therefore, in accordance with other analytical studies on the twin-deficit theory, the study tests the long and short-term relationship and the course of the causality between the two deficits. We will render the variables stationary before estimating the model. Because empirical literature has argued that calculating time-series with unit root will show questionable results, the stationarity test can help to assess whether or not the time-series are stationary.

The co-integration approach will also analyse the possible long-term and short-term connections between model variables and the coefficient degree of significance. The Granger causality test, which is the main objective of this work, is performed in comparison to the generally utilised bivariate framework within the multivariate framework. This helps to determine the course of causality and input among the variables.

3.3.1 Unit root test

Until estimating empirical models, it is necessary to check and estimate the time series properties of the results. The test of the stationary data is important as the F-test and t-value are null if the sequence is not stationary. Several function root checks were done for the integration and root class of the variables. Unit root tests for Augmented Dickey-Fuller (ADF) in 1981 and Phillips-Perron in 1981 were conducted (PP 1988). The ADF test is based on the ARMA framework where you are AR defined (1). However, the Phillips-Perron (PP) root tests vary fundamentally from the ADF tests in the manner in which similarity and heteroscedasticity are treated. Specifically, while parametric self-regression in the test hypothesis is used by ADF tests to approximate the structure of ARMA, serial regression correlation is not used by the PP tests. Augment Dickey-Fuller (ADF) and Phillips-Perron (PP) claim that I(1) is a time series and I is an alternative hypothesis I(0).

In the other side, a non-stationary series is one in which $/\rho/> 1$. Any stochastic shock then cannot return to the right average stage. A non-stationary series is, therefore, a random phase in which the absolute value is equal to 1 (that is, unity). Such a vector could then be named "unit root" Nkang et al (2006). The ADF and PP test was used to verify the variables stationarity.

$$y_t \ = \ \alpha D_t \ \ + \gamma Y_{t-1} + \sum_{i=1}^{k} \beta i \, \Delta Y_{t-k} + \ \varepsilon_t . \ldots \ldots \ldots . . (3)$$

Where Yt is the time sequence, Δ is the first difference operator, T is the linear pattern and α is a constant. The null hypothesis that the unit root remains is β.is 0.

3.3.2 Autoregressive distributed lag model (ARDL)

There are different types of cointegration methodology like Johansen Juselius (1990), Johansen (1991), Engel and Granger (1987) and Gregory and Hansen (1996). The Johansen process, therefore, is more commonly used, but it also has drawbacks such as low ability in limited samples and includes order (1) variables. The research uses the Autoregressive-Distributed Lag (ARDL) model to analyse the twin-deficit relationship in the macroeconomic system and Ricardian equivalence centered on the consumption function of Bernheim (1987). Pesaran et al., (2001) sponsored the (ARDL) boundary checking method has more value than others. Thirdly, we have overlapping short-term and long-lasting relations between variables. The low and upper limit importance is suggested by the co-integration technique of ARDL. We accept that co-integration between the variables occurs where the F-statistics are larger than the upper limit. If F data are lower than the upper limit, we embrace zero non-cointegration hypotheses (Pesaran et al., 2001). This study is based on the above attributes, with simple incorporation order and small sample size, as the most efficient approach. There are four phases to test an ARDL model. The first step explores the long-term partnership by utilising the Bound Test method (Pesaran and Pesaran 1997; Pesaran et al., 2001). In the second and third steps, we measure the long-term and short-term coefficients. Finally, in the context of (CUSUM) and (CUSUMSQ), the model provides stability performance. The ARDL model can be written by the following equation.

$$\partial(\mathrm{L},\mathrm{p})y_t = a_0 + \sum_{i=1}^{k} n_i\,(L,q_i)x_{it} + \lambda' w_t + \varepsilon_t \qquad (4)$$

Where

$$\partial(\mathrm{L},\mathrm{p}) = 1 - \partial_1\mathrm{L} - \partial_2 L^2 - \ldots - \partial_p L^p \qquad (5)$$

$$n_i(L,q_i) = n_{i0} + n_{i1}L + n_{i2}L^2 + \cdots + n_{iq}L^{qi} \qquad (6)$$

Where the dependent variable is $y_{t,,}$ constant is $a_{0,}$ L defines the lags, w_t is deterministic trend of vectors. The long-run estimates of the ARDL model is as:

$$\emptyset_i = \frac{n_i\,(1,q_i)}{\partial(1,\mathrm{p})} = \frac{n_{i0} + n_{i1} + \cdots + n_{iq}}{1 - \partial_1 - \partial_2 - \cdots - \partial_p} \qquad (7)$$

Where q_i is the estimator of long-run coefficient in the ARDL model

The ECM value of the ARDL model is derived by the first difference of lagged values. In ARDL approach the first approach gives long-run relationship and the second approach gives long-run, short-run and ECM value.

$$\Delta \mathrm{y}_t = \Delta \mathrm{a}_0 - \partial(1,\mathrm{p})ECM_{t-1} + \sum_{i=1}^{k} n_{i0}\Delta x_{it} + \boldsymbol{\lambda}'^{\Delta x_t} - \sum_{j=1}^{p-1} \partial\, j\Delta y_{t-1} - \sum_{i=1}^{k}\sum_{j=1}^{qi-1} n_{ij}\,\Delta_{1,t-j} + \varepsilon_t \qquad (8)$$

We assume x_t is not co-integrated and ε_t is the error term. The first model we estimate is Ricardian proposition based on private consumption model and the second model we estimate is to twin deficit hypothesis. The F-Statistics is contrasted with upper and lower limits. The ARDL equation model (1 and 2) can be written as follows:

$$\Delta \mathrm{BD}_i = \alpha_0 + \sum_{i=1}^{n} \alpha_{1i} \Delta BD_{t-i} + \sum_{i=0}^{n} \alpha_{2i} \Delta CAD_{t-i} + \sum_{i=0}^{n} \alpha_{3i} \Delta INF_{t-i} + \sum_{i=0}^{n} \alpha_{4i} \Delta REER_{t-i}$$
$$+ \sum_{i=0}^{n} \alpha_{5i} \Delta INT_{t-i} + \sum_{i=0}^{n} \alpha_{6i} \Delta MS_{t-i} + BD_{t-i} + \beta_2 CAD_{t-i} + \beta_3 INF_{t-i}$$
$$+ \beta_4 REER_{t-i} + \beta_5 INT_{t-i} + \beta_6 MS_{t-i} + \varphi EC_{t-1}$$
$$+ \varepsilon_t \qquad (9)$$

3.3.3 Granger causality

As discussed in the theoretical background on the basis of national income accounting, budget deficit and current account deficit either have bidirectional, unidirectional or neutral relationship. The study attempts to test the authenticity of a Ricardian theorem and Keynesian proposition for BRICS. However, the other macroeconomic variables which influence the BD and CAD are also taken in the model like; the impact of exchange rate depreciation can cause the current account deficit. The increase in interest rate will cause an inflow of funds and deteriorating current account balance (CAB), a decrease in tax revenue or tax rate will cause the budget deficit. The increase in money supply can bring inflation with more demand for goods and services which will further deteriorate CAB. An increase in growth rate can have a positive impact on the CAB; by increasing exports (see; Hoffmaister and Roldos, 1997).

Granger causality approach is to find out the link between the variables. Granger (1969, p. 430) causality test includes the estimation of the regression equations as pursues: if y_t contains past information that aides in the forecast of x_t, meaning y_t causes x_t. To estimate the causality, the equations for the model can be specified as:

$$x_t = \gamma_1 + \sum_{j=1}^{p} \alpha_{1j} x_{t-j} + \sum_{j=1}^{p} \beta_{1j} y_{t-j} + \varepsilon_{1t} \qquad (10)$$

$$y_t = \gamma_2 + \sum_{j=1}^{p} \alpha_{2j} x_{t-j} + \sum_{j=1}^{p} \beta_{2j} y_{t-j} + \varepsilon_{2t} \qquad (11)$$

k is the maximum lag and d is the order of integration and ε_{1t} is the error term. The null hypothsis for equation (10) will be $H_o = \beta_{1j} = \beta_{1j+1} = \beta_{1k} = 0$ which means no causlaity from y_t to x_t and vice versa.

This association between causality and monotony drove Granger to express direction of causality in a parametric structure, based on traditional time series data. It is important to check stationarity and lag structure before applying Granger causality. Causality analysis is sensitive to lag selection, we applied the (AIC) for optimum lag length. The auto regression model based on the equation (10 and 11) can be written in the below form for estimating the relationship between the variables:

$$\Delta BD_t = \alpha_1 + \Sigma\beta_1\Delta CAD_{t-i} + \Sigma\theta_1\Delta INF_{t-i} + \Sigma\gamma_1\Delta REER_{t-i} + \Sigma\delta_1\Delta INT_{t-i} + \Sigma\lambda_1\Delta MS_{t-i} + \Sigma\boldsymbol{\partial}_1\Delta TAX_{t-i} + \varepsilon_t \quad (12)$$

$$\Delta CAD_t = \alpha_2 + \Sigma\beta_2\Delta BD_{t-i} + \Sigma\theta_2\Delta INF_{t-i} + \Sigma\gamma_2\Delta REER_{t-i} + \Sigma\delta_2\Delta INT_{t-i} + \Sigma\lambda_2\Delta MS_{t-i} + \Sigma\boldsymbol{\partial}_2\Delta TAX_{t-i} + \varepsilon_t \quad (13)$$

$$\Delta INF_t = \alpha_3 + \Sigma\beta_3\Delta BD_{t-i} + \Sigma\theta_3\Delta CAD_{t-i} + \Sigma\gamma_3\Delta REER_{t-i} + \Sigma\delta_3\Delta INT_{t-i} + \Sigma\lambda_3\Delta MS_{t-i} + \Sigma\boldsymbol{\partial}_3\Delta TAX_{t-i} + \varepsilon_t \quad (14)$$

$$\Delta REER_t = \alpha_4 + \Sigma\beta_4\Delta BD_{t-i} + \Sigma\theta_4\Delta CAD_{t-i} + \Sigma\gamma_4\Delta INF_{t-i} + \Sigma\delta_4\Delta INT_{t-i} + \Sigma\lambda_4\Delta MS_{t-i} + \Sigma\boldsymbol{\partial}_4\Delta TAX_{t-i} + \varepsilon_t \quad (15)$$

$$\Delta INT_t = \alpha_5 + \Sigma\beta_5\Delta BD_{t-i} + \Sigma\theta_5\Delta CAD_{t-i} + \Sigma\gamma_5\Delta INF_{t-i} + \Sigma\delta_5\Delta REER_{t-i} + \Sigma\lambda_5\Delta MS_{t-i} + \Sigma\boldsymbol{\partial}_5\Delta TAX_{t-i} + \varepsilon_t \quad (16)$$

$$\Delta MS_t = \alpha_6 + \Sigma\beta_6\Delta BD_{t-i} + \Sigma\theta_6\Delta CAD_{t-i} + \Sigma\gamma_6\Delta INF_{t-i} + \Sigma\delta_6\Delta REER_{t-i} + \Sigma\lambda_6\Delta INT_{t-i} + \Sigma\boldsymbol{\partial}_6\Delta TAX_{t-i} + \varepsilon_t \quad (17)$$

$$\Delta TAX_t = \alpha_7 + \Sigma\beta_7\Delta BD_{t-i} + \Sigma\theta_7\Delta CAD_{t-i} + \Sigma\gamma_7\Delta INF_{t-i} + \Sigma\delta_7\Delta REER_{t-i} + \Sigma\lambda_7\Delta INT_{t-i} + \Sigma\boldsymbol{\partial}_7\Delta MS_{t-i} + \varepsilon_t \quad (18)$$

G-causality gives four possible outcomes, unidirectional causality from BD to CAD, unidirectional causality from CAD to BD, bidirectional causality from BD to independent variables and no-causality among the variables. Note here that association itself doesn't really suggest a development in the forecast. Relationship is a proportion of coupling quality, which can start from both causation and reliance on normal causes. Granger causality is a proportion of coupling, with directionality. Thus, it depends on forecast errors instead of linear relationships among the variables.

3.3.4 Impulse response function

Finally, a novel attempt is made to investigate the time way or (input and output behavior of the system) of these components and their responses to shocks from the selected macroeconomic variables. Based, on the Granger causality outcomes, policy makers cannot predict the future policy based on the present results. Secondly, these results can be clarified with sample tests that may give more explanation on the dynamic properties of this relationship Masih and Masih (1995).

This approach includes calculating abrupt shifts in time t in one variable X (the impulse) and estimating its impact on the other variable Y in time t, t+1, t+2, etc... (the answers). The IRF explains how the dependent variable reacts to the error shocks of the VAR model. In other terms, in one of the developments, the IRF detects the influence of a particular shock on existing and future values of endogenous variables. The basic structure for the IRF will be:

$$Y_t = \alpha + \varepsilon_t + \Theta 1\, \varepsilon_{t-1} + \Theta 2 \varepsilon_{t-2} + \ldots + \Theta i \varepsilon_{t-i} \qquad (19)$$

Where yt is a function of dependent variables, ε is a function of shock for all VAR models and Θi is a vector parameter, which measures the dependency variable's responses to developments in all the VAR model variables.

For two variables (Yt and Xt), however, the IRF form will be:

$$Yt = \alpha 1 + \varepsilon Y,t + \eta 1\varepsilon\ Y,t\text{-}1 + \eta 2\varepsilon Y,t\text{-}2 + \ldots + \eta i\varepsilon Y,t\text{-}I \qquad (20)$$

$$Xt = \alpha 2 + \varepsilon X,t + \varphi 1\varepsilon\ X,t\text{-}1 + \varphi 2\varepsilon X,t\text{-}2 + \ldots + \varphi i\varepsilon X,t\text{-}I \qquad (21)$$

Equations 20 and 21 describe how the predictor variables, Yt or Xt, reacts to past developments that occurred in the VAR model's dependent variable (εX's and εY's). The quantities of responses are, however, presented by the coefficients (φ's and π's).

In this study we use generalize impulse response functions (GIR) which investigates the time impacts of a one-time shock to every factor as given below.

$$\Theta \frac{g}{i} = \emptyset_i \sigma_{jj} - \frac{1}{2}\Sigma, \qquad (22)$$

Both structural and symmetrical impulse functions are compelled either by finding the correct order of factors or by the distinguishing proof of the evaluated structural parameters. Koop et al. (1996) propose an alternate sort of impulse function, called generalised impulse responses (GIR). The functions are independent of the order of variables since they combine the impacts of different shocks out of the responses. σ_{jj} is the variance of the jjth variable.