CHAPTER 3
DATA AND METHODOLOGY

3.1 Nature and Period of Data Set

The study is based on secondary data set. In this study Rupee/Dollar exchange rate, M_1 money supply in India has been used. We have used the historical dataset in this study. It is quarterly by nature. The period covers 1975 (I) to 2006 (IV). Specified variables are e_t and m_t where

e_t = Rupee/Dollar exchange rate,

m_t = M_1 money supply in India ('00 Billions of Rupees)

3.2 Source of Data Set

The time series data for Rupee/Dollar Exchange rate and M_1 money supply in India have been used. These datasets have been collected from different Year Books of *International Financial Statistic (IFS),* published by the IMF.

3.3 Rationale behind the Period Chosen

The period since 1970s in economic history is made colorful by the growth of unprecedented events and subsequent developments of important innovative ideas. The most important events in the history of international economics during 1970s is the breakdown of fixed exchange rate system and the adoption of flexible exchange rate system instead. It is significant because it spelled the end of the Bretton wood system as incorporated in the IMF chapter. Since then several changes occurred in the realm of international economics. Attention of economists was diverted from balance of payments to exchange rate system. Economists got pre-occupied with the determination of exchange rate and explanation of its variations.

However, in India the flexible exchange rate system was initiated from mid 1970s. Since then the relevance of exchange rate determination became more akin. The value of rupee in terms of major international currency underwent spectacular depreciations. Several models explaining the variation of exchange rate were developed. These models were mainly developed and verified in the context on developed countries.

3.4 Rationale behind the Use of IFS Data Set

In this study we have opted for the IFS data-sets though the data on the relevant variables are available from the Reserve Bank of India Bulletins This is done because of the fact that the Reserve Bank of India collects detailed data which it does not release in India. It does, however, make these data available to World Bank and the International Monetary Fund, so that the foreigners with access to these institutions can easily get hold of them. It is the common place for Indian researchers to realize that for the 'correct' data on the Indian economy, whether it is the state of the external and internal debt, and actual payments made on total defense amount, one has to turn to foreign publication which use data made available by the Indian Government. This has motivated our choice of the IFS as the source of data set for our present study. [Ref. *The Front Line, July 20, August 1991, Vol. 8 No. 15, pp11*].

3.5 Rationale Behind The Use of M_1 Money Supply

The Reserve Bank of India controls the money supply by altering **base money** which consists of currency plus commercial bank reserves held against deposits. As **base money** changes, Increases in base money tend to result in an expansion of the money supply whereas decreases in base money tend to contract the money supply for MABP and MAER purposes, it is useful to divide base money into domestic and international components.

The domestic component of base money is called **domestic credit**, where the reminder is made up of **international resources** (money items that can be used to settle international dept, primarily foreign exchange) at home affect base money and then money supply. For instance, if an Indian exporter receives payments in foreign currency, this payment will be presented to an Indian Commercial Bank to be converted into reserves and deposited in the exporter's account. If the commercial bank has no use for the foreign currency, the bank will exchange the foreign currency for dollars with the RBI. The RBI creates new money to buy the foreign currency by increasing the commercial bank's reserve deposit with the RBI. Thus, the RBI is accumulating international reserves, and this reserve accumulation brings about an expansion of base money. In the case of an excess supply of money at home, either domestic credit falls to reduce base money, or else international reserves will fall in order to lower base money to the desired level. In our study the effects of money supply on exchange rate one being examined. Consequently, the study involves the use of base money to capture money supply in Indian Economy over the period of study.

3.6 Methodological Issues: Stationarity Test

The study involves the use of time series data. These series are not deterministic variables. They share some stochastic properties. Charles. R. Nelson and C.I. Plosser (1982) hold that macro-economic time series usually behave like random walks. These series are not *'trend reverting.'* Consequently, these variables do not tend to revert back to a long-run trend after a shock. In such case, standard regression with non-stationary data makes all regression coefficients misleading and, therefore, leads to spurious relationships with erroneous conclusions. It, therefore, becomes pertinent to enquire into the nature of the stochastic processes of the macroeconomic time series like government expenditures and government revenues in countries concerned

Stationarity of the time series data on Rupee/Dollar exchange rate and money supply has been enquired through *Unit-Root tests* like *Augmented Dickey Fuller (ADF)* and *Philips-Perron (PP) unit root tests.*

The results of unit-root tests are very sensitive with the included assumptions about the time series, such as trend or intercept or both trend and intercept etc. Therefore, to confirm about the nature of the underlying series, we can plot them graphically and can find the particular nature (at levels and after differencing).

3.6.1 Test of Stationarity: Augmented Dickey-Fuller (ADF) Test

For testing the presence of *Unit Roots* in the time series dataset, Dickey and Fuller have developed this test. The methodology of *ADF test* is explained below:

For the time series E_t , the *ADF unit root test* can be done with the following three assumptions:

(i) the macro-economic variable (E_t) can follow random walk with a drift and having a stochastic trend,

(ii) E_t can follow random walk with drift but without having any stochastic trend,

(iii) E_t can follow random walk without any drift and stochastic trend,

Now, for the first assumption, the relevant equation for the ADF test is:

$$\Delta E_t = \alpha_1 + \beta\, t + \gamma_1 E_{t-1} + \delta_{1i} \sum_{i=1}^{m} \Delta E_{t-1} + \varepsilon_{1t} \tag{3.1}$$

For the second assumption, the ADF test equation is:

$$\Delta E_t = \alpha_2 + \gamma_2 E_{t-1} + \delta_{2i} \sum_{i=1}^{m} \Delta E_{t-1} + \varepsilon_{2t} \tag{3.2}$$

But if we take the third assumption, the estimable equation will take the form:

$$\Delta E_t = \gamma_3 E_{t-1} + \delta_{3i} \sum_{i=1}^{m} \Delta E_{t-1} + \varepsilon_{3t} \tag{3.3}$$

where, $\varepsilon_{1t} \sim$ i i d N $(0, \sigma^2)$, and $\Delta E_{t-1} = (E_{t-1} - E_{t-2})$, $\Delta E_{t-2} = (E_{t-2} - E_{t-3})$, etc.

The number of lagged difference terms (m) in the test equations (3.1) – (3.3) is often determined empirically. It should be included in such a way that ε_{1t}, ε_{2t} and ε_{3t} becomes serially uncorrelated.

3.6.2 Test of Stationarity: Phillips-Perron Unit Root Test

The second motivation for alternative unit root test statistic is to allow for disturbance processes, 'ε_t', which are not ~ N i i d $(0, \sigma_\varepsilon^2)$. Philips (1987), Philips and Perron, (1988), generalized the DF test to situations where the 'ε_t' are serially correlated, other than by augmenting the initial regression with lagged dependent variables as in the *ADF procedure*. The regression equation for the *Philips-Perron (PP) test* defines an AR (1) process:

$$\Delta E_t = \alpha + \beta E_t + \varepsilon_t$$

The correction is nonparametric since we use an estimate of the spectrum of 'ε_t'at frequency zero that is robust to heteroskedasticity and autocorrelation of unknown form. Here we use the *Newey-West heteroskedasticity autocorrelation* consistent estimate:

$$\omega^2 = \gamma_0 + 2\sum_{j=1}^{q}\left(1 - \frac{j}{q+1}\right)\gamma_j, \qquad \gamma_j = \frac{1}{T}\sum \hat{\varepsilon}_t \hat{\varepsilon}_{t-j}$$

where, q is the function lag. The *Philips-Perron* t-statistic is computed as:

$$t_{pp} = \frac{\sqrt{\gamma_0}\, t_0}{\omega} - \frac{\left(\omega^2 - \gamma_0\right) T s_b}{2\omega\hat{\sigma}}$$

The asymptotic distribution of the PP t-statistic is the same as the ADF 't'-statistic. As with the ADF test, here also we have to specify whether to include a constant, a constant and linear trend, or neither in the test regression. For the PP test, we also have to specify the truncation lag q for the *Newey-West correction,* that is, the number of periods of serial correlation to include. The truncation lag, q, can be calculated as:

$$\mathrm{Q} = \mathrm{floor}[4(T/100)^{2/9}]$$

This is based solely on the number of observations used in the test regression. We may, of course, specify any integer.

3.6.3 Test of Stationarity: The Correlogram Analysis

The random walk process for the time series implies non-stationarity in the sense that the series has been growing over time so that mean and other moments of the series are time dependent. The confirmation of the existence of non-stationarity in any series requires that the plots of *Autocorrelation Function* (ACF) and *Partial Autocorrelation Function* (PACF) be examined. Graphical presentation of *Autocorrelation Functions* (ACF) and *Partial Autocorrelation Functions* (PACF) of different lags with corresponding Q-statistic is called *Correlogram.* Therefore, *Correlogram* test is a simple test of stationarity, which is based on the *Autocorrelation Function* (ACF) and *Partial Autocorrelation Functions* (PACF).

The *Correlograms* of the stationary and the non-stationary time series have some distinguishing features. If in any *Correlogram* the autocorrelation coefficients start from a very high level and decline very slowly towards zero as the lags lengthen, we can conclude that the time series in question is non-stationary and it follows a *random walk process*. So in case of non-stationary time series we see that solid spikes of the initial lags are very large and these decline very slowly. On the other hand, in case of stationary series the *Correlogram* has only few solid spikes at lower lags. Thus stationarity of the variables can be confirmed by examining the *Correlograms* of the time series concerned.

3.7 Cointegration: Meaning and Relevance

Random walks process attains stationarity after differencing. If a test fails to reject the hypothesis of a random walk, one can difference the series before using them in regression. Since many economic times series follow random walks, variables are subject to first differencing before using them in a regression.

However, differencing the data has a cost. The cost arises from the fact that differencing may result is a loss of information about the long-run relationship between variables concerned. This occurs because the models, estimated with differenced data, do not have a long-run solution. Moreover, the level of a variable and its first difference will typically be very different in terms of mean, variance etc. The theory of *Cointegration*, therefore, may be used as a diagnostic for linear regression.

Engle and Granger (1987) hold that non-stationary random walk time series data can still be used for the study of long-run equilibrium relationship among the variables concerned, provided that the variables are *Cointegrated.* In many cases two or more variables follow random walk processes but the linear combinations of these variables are found to be stationary. If this be the case, then these variables are called *Cointegrated* variables. *Cointegration* provides a method for eliminating the cost of differencing by retaining terms in levels but only in linear combination, which are stationary.

The justification of the *Cointegration* study is that in equilibrium relationship of a set of variables, the individual variables cannot move independently. Therefore, equilibrium relationship among a set of non-stationary variables implies that their stochastic trends must be linked. This linkage among the stochastic trends indicates that the variables are *cointegrated.*

Again, when the trends of *Cointegrating* variables are linked, the dynamic paths of the variables are also linked. The dynamic path of the relationship must bear some relation to the current deviation from the equilibrium relationship. The connection between the change in a variable and the deviation from equilibrium needs to be examined in detail.

3.7.1 Methods of Cointegration

Cointegration among the macro-economic time series is studied by applying different methods, namely,

(i) *Engle-Granger Two-Step Method,* and

(ii) *Johansen Method.*

3.7.2 Engle-Granger Method of Cointegration

Engle-Granger Method of Cointegration study involves two steps of estimation as given below

(i) Let there be three non-stationary variables, e_t, and m_t, Then in the first step, one variable is regressed on the others such that:

$$e_t = \alpha_1 + \beta_1 m_t + u_{1t} \tag{3.4}$$

$$m_t = \alpha_2 + \beta_2 e_t + u_{2t} \tag{3.5}$$

From the equation (3.4) and (3.5) the residuals are obtained such that

$$u_{1t} = e_t - \alpha_1 - \beta_1 m_t \tag{3.6}$$

$$u_{2t} = m_t - \alpha_2 - \beta_2 e_t \tag{3.7}$$

where, u_{1t} and u_{2t} are white noise error terms.

(ii) In the second step stationarity of the residuals $\hat{u}_{1t}$ and $\hat{u}_{2t}$ is examined. If e_t, f_t and m_t are *Cointegrated,* any linear combination of these variables would generate stationary residuals. For the purpose of testing stationarity, *Augmented Dickey Fuller (ADF)* Test, *Phillips-Perron Test* etc. may be applied. Again, the stationarity of the residuals can be confirmed through the examination of their *Correlograms.*

(iii) If residuals ($\hat{u}_{1t}$ and $\hat{u}_{2t}$) exhibit *random walk*, the time series are subject to filtering like differencing.

(iv) The *cointegration equation* has to be re-estimated through the use of filtered (first differenced) data. For example, if E_t, F_t and M_t be the filtered series, then the estimable equations are

$$E_t = \gamma_1 + \delta_1 M_t + v_{1t} \quad (3.8)$$

$$M_t = \gamma_2 + \delta_2 E_t + v_{2t} \quad (3.9)$$

(vi) Residuals of the re-equation are

$$v_{1t} = E_t - \gamma_1 - \delta_1 F_t - \mu_1 M_t \quad (3.10)$$

$$v_{2t} = M_t - \gamma_2 - \delta_2 E_t \quad (3.11)$$

The residuals again are subject to *ADF, PP* test for ascertaining if the residuals are white noise. Again the stationarity of residuals can be confirmed through Correlogram.

These procedures have to be repeated until the residuals of the estimated equation are free from random walk. Thus the order of *Cointegration* will be determined.

3.7.3 Johansen Cointegration Test

Both the Johansen (1988) and the Stock and Watson (1988) methodologies rely heavily on the relationship between the rank of the matrix and the characteristic roots. The *Johansen cointegration test equation* is presented below:

$$\Delta x_t = \beta + \pi x_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta x_{t-i} + \varepsilon_t \quad (3.12)$$

where, β is the vector of constants, x_t is the m dimensional vector of variables, (i.e., E_t, R_t in our analysis), p is the number of lags, ε_t is the error vector, which is multivariate normal and independent across observations. $\pi = -(1 - \sum_{i=1}^{p} A_i)$ and $\pi_i = -\sum_{j=i+1}^{p} A_j$.

Here, the rank of the matrix π is equal to the number of independent co-integrating vectors. Specifically,

If π =0, the matrix is null and is the usual *VAR model* in first differences.

If π is of rank n, the vector process is *stationary*.

If π=1, there is a single *cointegrating vector*.

If 1<π< n, there are multiple cointegrating vector.

To show this, suppose we obtained the matrix π and ordered the n characteristic roots $\lambda_1, \lambda_2, \ldots\ldots, \lambda_k$ such that $\lambda_1 \rangle \lambda_2 \rangle \ldots\ldots \lambda_n$. If the variables in x_t are not *Cointegrated*, the rank of π is zero and all of these characteristic roots will be zero. Since $Log(1) = 0$ each of the expressions $Log(1 - \lambda_i)$ will equal zero if the variables are not *Cointegrated*. Similarly, if the rank of π is unity, $0 \langle \lambda_1 \langle 1$, so the expression $Log(1 - \lambda_1)$ will be negative and the other $\lambda_i = 0$, so that $Log(1 - \lambda_2) = Log(1 - \lambda_3) = \cdots = Log(1 - \lambda_n) = 0$

Here the number of distinct *cointegrating* vectors can be determined by checking the significance of the characteristic roots of π. Now, the test for the number of characteristic roots that are significantly different from unity can be obtained using the following two test statistic:

(a) *The Trace Statistic*,

(b) *The Max-Eigen Statistic*.

The *Trace Statistic* can be calculated in terms of the following expression:

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{n} Log(1 - \hat{\lambda}_i) \tag{3.13}$$

On the other hand, the *Max-Eigen Statistic* can be calculated as

$$\lambda_{\max}(r, r+1) = -T Log(1 - \hat{\lambda}_{r+1}) \tag{3.14}$$

where, $\hat{\lambda}_i$ = the estimated values of the characteristic roots (i.e., Eigenvalues) obtained from the estimated π matrix, T = the number of unstable observations.

The *'Trace Statistic'* is used to test the null hypothesis that the number of distinct *cointegrating vectors* is less than or equal to 'r' against the general alternative. The *'Max-Eigen Statistic'* test the null hypothesis that the number of *cointegrating* vector is 'r' against the alternative of (r +1) co-integrating vectors. The critical values of the λ_{trace} and the $\lambda_{\max}$ statistic are calculated using the *Monte Carlo* approach.

3.8 The Vector Error Correction Mechanism

The time path of any variables is influenced by the extent of its deviations from the long-run equilibrium level. The *Vector Error Correction (VEC)* specification restricts the long-run behavior of the endogenous variables to converge to their *cointegrating relationships* while allowing for a wide range of short-run dynamics. The cointegration term is known as the *error correction* term since the deviation from the long-run equilibrium is corrected gradually through a series of partial short-run adjustments. Therefore, *VEC* modeling gives important information about the short-run relationship between the *cointegrated* variables.

3.9 Vector Error Correction (VEC) Model

The *Vector Error Correction (VEC)* may be applied to analyze the short-run dynamics between two variables. Let the variables be Rupee/Dollar exchange rate (e_t) and money supply (m_t). Then the relevant VEC equations are:

$$\Delta E_t = \alpha_1 + \rho_1 Z_{1t-1} + \beta_i \sum_{i=1}^{n} \Delta E_{t-i} + \gamma_i \sum_{i=1}^{n} \Delta M_{t-i} + \omega_t \qquad (3.15)$$

$$\Delta M_t = \alpha_2 + \rho_2 Z_{2t-1} + \beta_2 \sum_{j=1}^{n} \Delta M_{t-j} + \gamma_j \sum_{j=1}^{n} \Delta E_{t-j} + \vartheta_t \qquad (3.16)$$

where, ΔE_{t-i} = First Differenced Series of E_t at time t-i; i = 1, 2, 3, ---, n.

ΔM_{t-i} = First Differenced Series of M_t at time t-i; i = 1, 2, 3, ---, n.

The lag length (n), in estimation, are determined through AIC and SIC, and

$\omega_t \sim$ i i d N $(0, \sigma_w^2)$, and

$\upsilon_t \sim$ i i d N $(0, \sigma_\vartheta^2)$

Z_{1t-1} and Z_{2t-1} are error correction terms.

The focus of the vector error correction analysis is on the lagged Z_{1t-1} and Z_{2t-1} terms. These lagged terms are the residuals from the previously estimated *cointegrating* equations. In the present case the residuals from two lag specifications of the *cointegrating* equations have been used in the vector error correction estimates. These Z_{1t-1} and Z_{2t-1} terms provide an explanation of short-run deviations from the long-run equilibrium for the test equations above. Lagging these terms means that disturbance of the last period impacts upon the current time period. In general, finding a statistically insignificant coefficient of the Z_{it-1}, Z_{2t-1} terms imply that the system under investigation is in the short- run equilibrium. If the coefficient of Z_{it-1} and Z_{2t-1} terms are found to be statistically significant, then the system is in the state of the short-run disequilibrium. In such a case the sign of Z_{it-1} and Z_{2t-1} terms give an indication of the causality direction between the two test variables.

3.10 Vector Autoregression Model (VAR)

Christopher Sims has developed the ***Vector Autoregressive Model (VAR),*** where true simultaneity among a set of variables necessitates that these are to be treated on equal footing. There should not be any *a prior* distinction between endogenous and exogenous variables. Economic theories contain behavioral, structural, and/or reduced form relationships that can be incorporated into a *VAR analysis*. The *VAR model* is designed for forecasting of the interrelated variables by analyzing the dynamic impact of the random disturbances of the system.

The model of Vector Autoregression (VAR) for the two variables E_t and M_t may be presented through the following equations

$$E_t = \alpha_1 + \sum_{i=1}^{k} \beta_{1i} E_{t-i} + \sum_{i=1}^{k} \gamma_{1i} M_t + u_{1t} \tag{3.17}$$

$$M_t = \alpha_2 + \sum_{i=1}^{k} \beta_{2i} M_{t-i} + \sum_{i=1}^{k} \gamma_{2i} E_{t-i} + u_{2t} \tag{3.18}$$

where, u_{it}, u_{2t} are the stochastic error terms, called *impulse* or *innovations* or *shocks.*

In this VAR model

(i) The variables (like E_t and M_t) must be stationary, and

(ii) u_{1t} and u_{2t} must be white-noise terms.

3.11 Selection of Lag Length for the VAR Estimation

In the estimation of any *VAR model*, the selection of maximum lag length (k) is important since inclusion of too many lagged terms consumes degrees of freedom and problem of multicollinearity may arise. On the other hand, inclusion of few lags may lead to selection errors.

Akaike Information Criterion (AIC), Schwartz Bayesian Criterion (SBC), Hannan-Quinn Criterion (HQ), Sequential Modified LR test statistic, etc are generally used for the determination of optimum lag length. Another popular method has been developed by Enders (1995). Under this method, model has to be estimated with high lags (specific number depends on the nature of the underlying time series i.e., whether it is annual or quarterly time series etc). Then lags are to be reduced by one and carried out the estimation, given that estimated t-static for the coefficient of the discarded lag term involved is insignificant. Finally, the lag corresponding to the estimated model with maximum number of significant coefficients has to be undertaken for analysis.

AIC and *SBC* corresponding to any lag length may be calculated as follows

AIC = T In (sum of squired residuals) + 2n

SBC = T In (sum of squired residuals) + n In (T)

where, n = number of parameters estimated

T = no of usable observations

A distributed lag econometric model is called good fit and, therefore, valid inferences can be drawn from such model if AIC and SBC values obtained from the model are negative and as small as possible. In other words, fit of the model improves, the AIC and SBC approach $-\infty$.

Of the two criteria, the SBC has superior large sample properties. It is asymptotically consistent. On the other hand, AIC is biased towards selecting an over parameterized model. However, in small samples, the AIC can work better than SBC.

3.12 Impulse Response Functions: Relevance

A shock to the 'i'-th variable not only directly affects the i-th variable itself but the shock is also transmitted to all of the other endogenous variables through the dynamic (lag) structure of the *VAR*. An *Impulse Response Function* traces the effect of a one time shock to one of the innovations on current and future values of the endogenous variables. In other words, *Impulse Response Function* traces the response of a variable through time to an unanticipated change in itself or other interrelated variables. Therefore, the *Impulse Response Function* may be used in *VAR* system to describe *the dynamic behaviors of the whole system with respect to shocks in the residual of the time series.*

In a linear model, the *Impulse Responses* are not history dependent and the magnitude of the shock does not alter the time profile of the responses. However, the interpretation of the *Impulse Response Functions* for a nonlinear model is not straightforward. In the nonlinear model *Impulse Responses* are history dependent. The effect of a particular shock on the time path of the system depends on the magnitudes of the current and the subsequent shocks. In this case sign of the shocks is important.

3.13 Mathematical Exposition of Impulse Response Functions

Let the *Vector Autoregression (VAR) model* consist of following equations:

$$E_t = \alpha_1 + \sum_{i=1}^{k} \beta_{1i} E_{t-i} + \sum_{i=1}^{k} \gamma_{1i} M_{t-i} + e_{1t} \tag{3.19}$$

$$M_t = \alpha_2 + \sum_{i=1}^{k} \beta_{2i} M_{t-i} + \sum_{i=1}^{k} \gamma_{2i} E_{t-i} + e_{2t} \tag{3.20}$$

where, e_{it}, e_{2t} are the stochastic error terms, called *impulse* or *innovations* or *shocks*.

Let, i = 1. In this case, these equations can be written in the matrix form as:

$$\begin{bmatrix} E_t \\ M_t \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} \beta_1 & \gamma_1 \\ \beta_2 & \gamma_2 \end{bmatrix} \begin{bmatrix} E_{t-1} \\ M_{t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \tag{3.21}$$

Now expressions of each of E_t and R_t can also be expressed in terms of their errors,

$$\begin{bmatrix} E_t \\ M_t \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \beta_{11} & \gamma_{21} \\ \beta_{21} & \gamma_{22} \end{bmatrix}^i \begin{bmatrix} e_{1t-i} \\ e_{2t-i} \end{bmatrix} \tag{3.22}$$

Consequently, the VAR model is converted to *Vector Moving Average (VMA) model.*

Equation (3.22) expresses E_t and M_t in terms of the e_{it} and e_{2t} sequences. The vector errors can be written as

$$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \frac{1}{1 - b_{12} b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{Et} \\ \varepsilon_{Rt} \end{bmatrix} \tag{3.23}$$

Combining Equation (3.19) and (3.23), we have

$$\begin{bmatrix} E_t \\ R_t \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \frac{1}{1 - b_{12} b_{21}} \sum_{i=0}^{\infty} \begin{bmatrix} \beta_{11} & \gamma_{12} \\ \beta_{21} & \gamma_{22} \end{bmatrix}^i \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{Et-i} \\ \varepsilon_{rt-i} \end{bmatrix} \tag{3.24}$$

Now let, $\phi_i = \dfrac{A_1^i}{1 - b_{12} b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}$

Then equation (3.24) can be written as

$$\begin{bmatrix} E_t \\ M_t \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{Et-1} \\ \varepsilon_{Rt-1} \end{bmatrix} \tag{3.25}$$

or, more compactly, $\quad x_t = \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i} \quad$ (3.26)

The cumulative effects of one unit impulses in ε_{Et} and/or ε_{Rt} can be obtained by the appropriate summation of the coefficient of the *Impulse Response Function*. For example, suppose we are interested in finding the effect of ε_{Rt} on the value of E_{t+n}. This effect will be $\phi_{12}(n)$. Thus after n periods, the cumulative sum of the effect of ε_{Rt} on the E_t sequence is $\sum_{i=0}^{n}\phi_{12}(i)$. Now n approaching infinity yields the long-run multiplier. Since the E_t and M_t sequences are assumed to be stationary, then for all j and k, $\sum_{i=0}^{\infty}\phi_{jk}^{2}(i)$ is finite

The four sets of coefficients $\phi_{11}(i), \phi_{12}(i)$, $\phi_{21}(i)$ and $\phi_{22}(i)$ are called the ***IMPULSE RESPONSE FUNCTIONS.*** Plots of the coefficients of $\phi_{jk}(i)$ against 'i' provides the visual representation of the behavior of the E_t and M_t series in response to various shocks.

3.14 Variance Decomposition

Impulse Response Functions trace the effects of a shock to one endogenous variable on to the other variables in the *VAR*, while *Variance Decomposition* separates the variations in an endogenous variable into some component shocks. Thus *Variance Decomposition* provides information about the relative importance of each random innovation in the matter of affecting the variables in the *VAR*.

Unrestricted VARs are not particularly useful for *short-term forecasts* since they are over parameterized. The *Forecast Error Variance Decomposition* reflects the proportion of forecast error variance of a variable which is explained by an unanticipated change in itself as opposed to that proportion attributable to change in other interrelated variables. In other words, *the Forecast Error Variance Decomposition* tells us the proportion of the movement in a sequence due to its own shocks versus shocks of other variables.

3.15 Long-Run Causal Relationship: Granger Causality Test

Following Weiner, Granger (1969) has proposed the terminology that "y causes x if, given all values of x, past values of y help predict x. Sim extends Granger's terminology of causality and asserts that *"the projection of y on the entire x equals the projection of y on current and past x's if and only if y fails to Granger cause x".*

Let, $\begin{bmatrix} x_t \\ y_t \end{bmatrix}$ be a bivariate, jointly covariance stationary stochastic process, and

$\begin{bmatrix} x_t \\ y_t \end{bmatrix}$ be a strictly linearly in deterministic process with mean zero.

Under this conditions, the bivariate version of *Wold's Representation Theorem* states that there exists a moving average representation of the (x_t, y_t) process.

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} c^{11}(L) & c^{12}(L) \\ c^{21}(L) & c^{22}(L) \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix} \tag{3.27}$$

where, $c^{11}(L) = \sum_{k=0}^{\infty} c_k^{11} L^k$ are squire summable polynomials in the lag operator L that are one-sided in non-negative powers of L. ε_t and u_t are serially uncorrelated processes with E(u_t, ε_t) = 0 for all t, s.

Here E(ε_t^2) = σ_ε^2 and E(u_t^2) = σ_u^2

The one step ahead predictions are given by

$$x_t - P[x_t / x_{t-1}, \dots\dots\dots, y_{t-1} \dots\dots\dots] = c_0^{11} \varepsilon_t + c_0^{12} u_t$$

$$y_t - P[y_t / y_{t-1}, \dots\dots\dots, y_{t-1} \dots\dots\dots] = c_0^{21} \varepsilon_t + c_0^{22} u_t \tag{3.28}$$

Thus ε_t and u_t are "jointly fundamental for x and y". *Wold's Representation Theorem* establishes that a vector moving average is a general representation for an indeterministic covariance stationary vector process. This idea can be used to explain Sim's assertions.

Let (x_t, y_t) process have an autoregressive representation. Then we have a sequence of projections.

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = F_1^\eta \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \dots\dots\dots + F_n^\eta \begin{bmatrix} x_{t-n} \\ y_{t-n} \end{bmatrix} + \begin{bmatrix} a^\eta x_t \\ a^\eta y_t \end{bmatrix} \tag{3.29}$$

where $F_1^{\eta}, F_2^{\eta}, \ldots\ldots F_n^{\eta}$ are 2 x 2 matrics of least square coefficients and we have the orthogonality conditions

$$E\begin{bmatrix} x_{t-j} \\ y_{t-j} \end{bmatrix}\begin{bmatrix} a_{xt}^{\eta} & a_{yt}^{\eta} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{3.30}$$

for j = 1...... n.

As $n \to \infty$, the F_j^{η} converges to F_j for each j.

Then the autoregressive representation for (x_t, y_t) can be written as

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \sum_{j=1}^{\infty} F_j \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} a_{xt} \\ a_{yt} \end{bmatrix}$$

$$= \mathrm{F(L)}\begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} a_{xt} \\ a_{yt} \end{bmatrix}, \quad \mathrm{F(L)} = \sum_{i=1}^{\infty} F_j L^{j-1} \tag{3.31}$$

Where, the random variables (a_{xt}, a_{yt}) obey the least square orthogonal

conditions $\begin{bmatrix} x_{t-j} \\ y_{t-j} \end{bmatrix}\begin{bmatrix} a_{xt} & a_{yt} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (3.32)

for all $j \geq 1$.

The random variables (a_{xt}, a_{yt}) are the one-step ahead errors in predicting (x_t, y_t) for all past values of x and y.

Now the particular representation for (x_t, y_t) process be

$$\mathrm{A(L)}\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix} \tag{3.33}$$

Multiplying (3.33) by A_0^{-1} we get

$$A_0^{-1}A(L)\begin{bmatrix} x_t \\ y_t \end{bmatrix} = A_0^{-1}\begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix} \qquad (3.34)$$

or, $\begin{bmatrix} x_t \\ y_t \end{bmatrix} = A_0^{-1}\left[A_1 L + A_2 L^2 + \cdots\cdots\right]\begin{bmatrix} x_t \\ y_t \end{bmatrix} + A_0^{-1}\begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix}$

$$= A_0^{-1}H(L)\begin{bmatrix} x_t \\ y_t \end{bmatrix} + A_0^{-1}\begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix} \qquad (3.35)$$

where, H(L) = $A_1 L + A_2 L^2 + \cdots\cdots$

The linear least squares projection of the $\begin{bmatrix} x_t \\ y_t \end{bmatrix}$ process based on all lagged x's and all lagged y's from (3.35) is

$$P_{t-1}\begin{bmatrix} x_t \\ y_t \end{bmatrix} = A_0^{-1}H(L)\begin{bmatrix} x_t \\ y_t \end{bmatrix} = F(L)\begin{bmatrix} x_t \\ y_t \end{bmatrix} \qquad (3.36)$$

since by construction $P_{t-1}\begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix} = 0$

Thus x prediction errors and y prediction errors are contemporaneously correlated so long as A_0 is not diagonal.

As $A_0^{-1}A(L)$ is lower triangular, then (3.33) can be inverted to yield the vector moving average representation.

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = c(L)\begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix} \qquad (3.37)$$

where, A(L) $^{-1}$ = C(L) = $c_0 + c_1 L + c_2 L^2 + \cdots\cdots$ and C_j being 2 x 2 matrix and c(L) is lower triangular.

Conversely, a moving average representation of lower triangular form (3.37) with ε_t and u_t being serially uncorrelated process with E($\varepsilon_t u_t$) = 0 for all t and s, then A(L) gives the representation

$$C(L)^{-1}\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \varepsilon_t \\ y_t \end{bmatrix}$$

$$\text{Or, } A(L)\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \varepsilon_t \\ y_t \end{bmatrix}$$

where A(L) is lower triangular and one sided on the present and past when

$c(L)^{-1}$ exists.

So Sim asserts, 'Let (x_t, y_t) be a jointly covariance stationary, strictly indeterministic process with zero mean. Then {y_t} fails to Granger cause {x_t}if and only if there exists a vector moving average representation".

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} c^{11}(L) & 0 \\ c^{21}(L) & c^{22}(L) \end{bmatrix}\begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix} \qquad (3.38)$$

where ε_t and u_t are serially uncorrelated process with zero mean and E(ε_t, u_t) = 0 for all t

So, Sim's (1972) causality idea is as follows:

"Y_t can be expressed as a distributed lag of current and past x's with a disturbance process that is orthogonal to past, present and future x's if and only if y does not Granger cause x".