CHAPTER-3

Empirical Methodology

CHAPTER-3

3.1 Introduction:

Statistical and econometric tools and techniques used to study the relationship between select macroeconomic variables and stock market in India during the pre and post liberalization periods are described in this chapter. Primarily, to assess the relationship we used very popular and widely used technique called cointegration. But before assessing the relationship we need some pre-testing of the whole set of data especially for the possible structural breaks as the study covers a period from 1966 to 2019, where practically two distinct economic philosophies are experimented by Indian policy planners for the development of the economy .The processing, including pre-testing, for assessment the relationship are described step by step below:

3.2 Structural stability of the time series process:

Literature on structural break in time series is vast, expanding and full of contests but not converging recommendations like known break point, unknown break point, single and multiple break points, gradualism, priors and pre-testing bias, application on linear or nonlinear, finite and infinite time series process, etc. According to Maddala and Kim (2004), more research works need to be done to compare the methods suggested by scholars to determine the structural instability. Some of the recommendations of the studies summarised by him are: (i) it is more important to determine the number and location of the break point(s) than only identifying that three is structural change, (ii) problem of consistent estimation still exists, despite the matter examined nicely by Bai and Perron(1998), (iii) there is no need to search the break points over the entire period of observations under study, because, there are always some prior information about the major break dates, like , war, abrupt or huge changes in economic policies etc, (iv) despite several good works, consensus is not arrived on the use of Bayesian or Classical approach to search the break points and (v) more attention of the scholars is required to the ‗gradual structural changes' as the shift from one regime to the another one is never sudden. Interestingly , may be due to gradualism, Agarwal and Ghosh (2015), found no break in Indian economy during the period 1960 to 2013, especially in per capita GDP, agriculture and manufacturing sector save the service sector in 2003. Some scholars argue that the breaks for ‗small magnitude' have no impact in forecasting. Ignoring rather than modelling small breaks may lead to more accurate forecasts (Stock and Watson, 1996; Pesran and Timmermann, 2005; Boot and Pick, 2017). Against this backdrop, we used all the suggested methodologies, i.e, (i) CUSUM test suggested by Brown,Durbin and Evans,(1975) , (ii) Quandt-Andrews breakpoint test suggested by Quandt(1995,1960), Andrews (1993) and Andrews and Ploberger (1994), and iii) Bai and Perron test (1998,2003) to validate (Bai and Perron,1998) and sub-group the periods.

3.2.1 CUSUM- Test:

Brown, Durbin and Evans (1975) suggests a test of structural break, popularly known as CUSUM(Cumulative Sum)-test, based primarily on the model's ability to predict correctly the observation or observations actually outside the range of the observations used to estimate it. The test is in the right direction to detect the timing of the change which is unknown (Greene, 2003). It is based on the cumulated sum of the recursive residuals and the test statistic recommended is:

$W_t = \sum_{r=K+1}^{t} w_r / \hat{\sigma}$ .... Eq. (1)

where,

$\hat{\sigma}^2 = (T - K - 1)^{-1} = \sum_{r=K+1}^{T} (w_r - \bar{w})^2$, and

$w = (\bar{T} - K)^{-1} \sum_{r=K+1}^{T} w_r$

or, $\hat{\sigma}$ is the standard deviation of the recursive residuals w_t. Under the null hypothesis, w_t has mean zero and variance approximately equal to the number of residuals being summed, because each term has variance 1 and they are independent (Greene, 2003). We have to study the behaviour of w_t by plotting w_t against t to assess the divergence from the ‗zero line'.

Confidence bounds of the cumulated sum are obtained by plotting the two lines that connect the points [K, $\pm$ a $(T - K)^{1/2}$] and [T, $\pm 3a(T-K)^{1/2}$], where ‗a' is the parameter value of which depends on the value of significance level α. The null hypothesis of ―there is no structural change" is rejected if w_t goes outside the area bounded by two critical lines.

3.2.2 Quandt-Andrews Breakpoint test:

According to Brown, Durbin and Evans, 1975, log- likelihood ratio technique suggested by Quandt (1958,1960), is appropriate when the regression relationship changes suddenly from one constant relationship to another constant relationship. The idea behind Quandt-Andrews Breakpoint test is that a single Chow break point test(1960,1983) is performed at every observation between two dates T_1 and T_2.The k statistics from those Chow tests (1960,1983) are then summarised to one test statistic for a test against the null hypothesis of no break points (Diebold and Chan, 1996; Brooks, 2014).

Let ,the unknown break point be , t=r, the first constant regression relationship be specified by $\beta^{(1)}$, σ_1^2 and the second one by $\beta^{(2)}$ and σ_2^2 , then for each r from r=K+1 to

r = T – (K+1), the plot is

$$\lambda_r = \log_{10}\left(\frac{\max\ likeli\ hoodobserv\ ationsgivenbyH\ \ 0}{\max\ likeli\ hoodobservationsgivenbyH\ \ 1}\right)$$

where, H_1 is the alternative hypothesis, i.e. the observations in the time segment (1, …, r) and the segment (r+1,……..,T) come from two different regressions. This form is the standard likelihood ratio statistics for taking decision on the null and alternative hypothesis. However,

$$\lambda_r = \frac{1}{2} r \log \hat{\sigma}_1^{2} + \frac{1}{2} (T - r) \log \hat{\sigma}_2^{2} - \frac{1}{2} T \log \hat{\sigma}^{2} \text{........Eq. (2)}$$

where, $\hat{\sigma}_1^{2}$, $\hat{\sigma}_2^{2}$, and σ^2 are the ratios of the residual sums of squares to number of observations when the regression is fitted to first r observations, the remaining (T – r) observations and the whole set of observations , respectively. When the value of r at which λ_r achieve minimum, at that point shift from one constant relationship to another occurs. The test for the distribution on the minimum of λ_r is given by Andrews (1993).

Andrews (1993) modified the Quandt‘s test and suggested three ‗supremum‘ tests. Let, the null of the structural stability of AR(1) process be:

$$y_t = \rho y_{t-i} + e_t, \quad t=1,\ldots\ldots,T \text{ .... Eq. (3)}$$

(no break or restricted model)

The H_1, one – time break, be:

Model 1: $y_t = \rho_1 y_{t-1} + \epsilon_t$, t= 1,………r .... Eq. (4)

Model 2: $y_t = \rho_1 y_{t-1} + \epsilon_t$, t=(r_{+1},………,T) .... Eq. (5)

(sub–sample or unrestricted model)

Then, the tests are:

$$Sup\ W = max_{\pi} T \left(\frac{\hat{\epsilon}'_t \hat{\epsilon_t} - \hat{\epsilon}'_1 \hat{\epsilon_1} - \hat{\epsilon}'_2 \hat{\epsilon_2}}{\hat{\epsilon}'_1 \hat{\epsilon_1} + \hat{\epsilon}'_2 \hat{\epsilon_2}} \right)$$

$$Sup\ LM = max_{\pi} T \left(\frac{\hat{\epsilon}' \hat{\epsilon} - \hat{\epsilon}'_1 \hat{\epsilon_1} - \hat{\epsilon}'_2 \hat{\epsilon_2}}{\hat{\epsilon}' \hat{\epsilon}} \right)$$

$$Sup\ LR = max_{\pi} T \left(\frac{\hat{\epsilon}'\hat{\epsilon}}{\hat{\epsilon}'_1\hat{\epsilon_1} + \hat{\epsilon}'_2\hat{\epsilon_2}} \right)$$

where , $\hat{\epsilon}$ is the T x 1 vector of OLS residuals from the restricted model, $\hat{\epsilon}_1$ is the r x 1 vector of OLS residuals from the sub-sample 1 model, $\hat{\epsilon}_2$ is the (T - r)*1 vector of OLS residuals from the sub-sample 2 model, and Π= r/T. The standard procedure suggested by Andrews (1993) is to impose Π∈ (0.15,0.85); although , it may vary according to the need of the analyst (Diebold and Chen, 1996). Break date estimate obtained by the OLS for all time series does not depend upon the variation of trimming (Jouini and Boutahar, 2005).

Andrews and Ploberger (1994) developed tests with stronger optimality properties than those of Andrews (1993) (Maddala and Kim, 2004). Using the notations used by Maddala and Kim (2004), let, $T_1<m<T_2$ be the range within which structural break occurs. Let, W^* be the Wald statistics for a break at t=m, then,

Sup W= $\max_{T_1<m<T_2} W^*$,..Eq. (6)

.... The test statistics suggested by Andrews and Ploberger (1994) are:

ExpW = $l_n\left[\frac{1}{T_2-T_1+1} \sum_{m=T_1}^{T_2} exp(w^* \frac{1}{2})\right]$ ....Eq. (7)

AveW= $\frac{1}{T_2-T_1+1} \sum_{m=T_1}^{T_2} W^*$ ....Eq. (8)

They showed that, under some regulatory conditions, the asymptotic distribution of the test statistics are given by the functions of the Wiener process, i.e. like a continuous random walk defined on interval (0,1), and are the same for LM and LR statistics:

Sup W = $\max_{\pi_1<m<\pi_2} Q(\tau)$,..Eq. (9)

ExpW= $l_n\left[\frac{1}{\pi_2-\pi_1} \int_{\pi_1}^{\pi_2} \exp(Q(\tau)/2)\, d\tau \right.$ ....Eq. (10)

Ave W = $\frac{1}{\pi_2-\pi_1} \int_{\pi_1}^{\pi_2} Q(\tau)\, d\tau$ ....Eq. (11)

where, Π_1= T_1/T, Π_2= T_2/T , andQ (τ) = $\frac{(w(\tau)-\tau W(1))'(W(\tau)-\tau W(1))}{\tau(1-\tau)}$

and W is a vector of K independent Wiener process. If τ is known then this is Chi – square (K) (Maddala and Kim, 2004). Hansen (1997) provided the table to consult and compare the probability values.

3.2.3 Bai and Perron Test (1998, 2003):

The Bai and Perron (1998), developed a methodology to identify more than one break point. The methodology in brief, is as follows:

Let us consider the following multiple linear regression with m breaks leading to m+1 periods or regimes.

$$Y_t = x_t'\beta + z_t'\delta_j + u_t, \quad \text{Eq. (12)}$$

where, $t = T_{j-1} + 1, \ldots, T_j$, for $j = 1, \ldots, m + 1$. If $T_0 = 0$ and $T_{m+1} = T$ are denoted then the indices ($T_1, \ldots, T_m$) or the breakpoints are treated as unknowns. For each m-partition ($T_1, \ldots, T_m$) denoted $\{T_j\}$, the associated least-square estimates of β and δ_j are obtained by minimizing the sum of squared residuals:

$$\sum_{i=1}^{m+1} \sum_{i=T_{i-1}-1}^{T_i} \left[y_t - x_t'\beta - z_t'\delta_i\right]^2 \quad \text{Eq. (13)}$$

Let $\hat{\beta}(\{T_j\})$ and $\hat{\delta}(\{T_j\})$ denote the resulting estimates. Substituting them in the objective function and denoting the resulting sum of squared residuals as $S_T(T_1, \ldots, T_m)$, the estimated break-points $(\hat{T}_1, \ldots, \hat{T}_m)$ are such that

$$(\hat{T}_1, \ldots, \hat{T}_m) = \text{argmin}_{(T_1, \ldots, T_m)} S_T(T_1, \ldots, T_m) \quad \text{Eq. (14)}$$

where, the minimization is taken over all possible partitions ($T_1,\ldots,T_m$) such that $T_i - T_{i-1} \geq q$. Here, q is the minimum length assigned to a segment and T_i is the break-point. The procedure considers all possible combination of segments and selects the partition that minimizes the sum of squared residuals. The least-squares estimators of the break-points are the global minima of the sum of squared residuals of the objective function in Eq. (12) (Bai and Perron, 1998).

3.3 Seasonality Test:

Recent philosophy in the time series analysis is to allow the data to speak, hence, maximum modifications of the raw data are suggested to avoid in empirical analysis (Maddala and Kim, 2004; Neuesser, 2016; Brockwell and Davis, 2016).

Removal of the trend and seasonal components are required to get stationary residuals. According to Brockwell and Davis (2016), to achieve this goal it may sometimes be necessary to apply a preliminary transformation to the data. There are several ways in which trend and seasonality can be removed, some involving estimating the components and subtracting them from the data, and others depending on differencing the data, i.e., replacing the original series {X} by {Y, := X, – X-d} for some positive integer d. Whichever method is used, the aim is to produce a stationary series, whose values we shall refer to as residuals. Again, if the magnitude of the fluctuations appears to grow roughly linearly with the level of the series, then the transformed series {In X1, ..., In Xn} will have fluctuations of more constant magnitude.

The magnitude of the fluctuations in the time series under this study appeared to go roughly linearly with the level of the series. Thus, following Brockwell and Davis (2016), we have taken log of the original series at level to transform the series where trend and seasonality are eliminated.

3.4Stationary Process:

Assessment of a time series whether stationary or not is very important as the stationary or otherwise of a series can strongly influence its behaviour and properties (Brooks, 2014). If the mean and variance of a time series change over time, then it is impossible to generate results from regression model for a specific time period to a different or future time period. Hence, for any meaningful statistical inferences, pre-testing the stationarity of the time series is necessary (Balarezo, 2010).

In the literature of time series analysis and econometrics, the concept of weakly stationary or the covariance stationary process is widely discussed and used in empirical finance and economics. The strict form of stationary process which is defined by all the four moments of distribution practically indicates that the distribution of the realised values of the time points in a time series $\{Y_t, t=0, \pm1, \ldots\}$ remains the same as time progresses. It indicates that the probability that a value of the series falls within a particular interval is same now as at any time in the past and future also (Brooks, 2014; Brockwell and Davis, 2016). Simply, if all the moments, namely mean, variance, skewness and kurtosis, of a time series are time invariant, then the series is said to be a strictly stationary process (Balarezo, 2010; Brockwell and Davis, 2016; Commandeur and Koopman, 2007). Some scholars argue that the strict form of stationary process has very little practical application and are rarely observed in practice (Tong, 1990; Greene, 2003; Balarezo, 2010; Brocks, 2014; Herranz, 2017). Moreover, according to Brockwell and Davis (2016) ―it is easy to check" if a time series $\{Y_t, t=0, \pm1, \ldots\}$ is strictly stationary and $EY^2 < \infty$ for all t, then $\{Y_t, t=0, \pm1, \ldots\}$ is also weakly stationary. Again, if a distribution is normal, then weakly stationary also implies strictly stationary because a normal distribution is fully explained by its first two moments (Balarezo, 2010). Thus, in our study, we used the concept of weakly stationary or covariance stationary process to mean ‗stationary process'.

A stochastic process { Y_t } is said to be weakly stationary if its mean and variance are constant over time and the value of the covariance between two time periods depends only on the distance or gap or lag between two time periods and not the actual time at which the covariance is computed (Gujarati, 2003; Greene, 2003). That is,

Mean : $E(Y_t) = \mu$ .... Eq. (15)

Variance : $\mathrm{var}(Y_t) = E(Y_t - \mu)^2 = \sigma^2 < \alpha$ .... Eq. (16)

Covariance : $\mathrm{cov}(Y_{t1}, Y_{t2}) = E[(Y_{t1}-\mu)(Y_{t2}-\mu)] = \gamma_{t1-t2}, \forall t_1, t_2$ .... Eq. (17)

Weakly or covariance stationary will tend to return to its mean and fluctuations around this mean, measured by its variance, will have a broadly constant amplitude (Cuthbertson, Hall and Taylor, 1995; Gujarati, 2003; Brooks, 2014; Balarezo, 2010). In other words, for a autoregressive process of order one (AR(1)) like:

$Y_t = \alpha + \rho Y_{t-1} + u_t$… .... Eq. (18)

is said to be stationary only if the following three conditions are satisfied,

Condition 1: Constant and finite mean: This can be expressed as:

$E(Y_t) = \frac{\alpha}{(1-\rho)}$ .... Eq. (19)

Condition 2: Constant and finite variance : This condition can be expressed as,

$\mathrm{var}(Y_t) = = \frac{\sigma^2}{(1-\rho^2)}$ .... Eq. (20)

Condition 3: Constant and finite covariance: The covariance of the time series,

Eq. (18), for its own lead and lag values, i.e., autocovariance, is constant over time and can be expressed as,

$$\text{Cov}(Y_t, Y_{t-s}) = = \frac{\rho^s.\sigma^2}{1-\rho^2} \quad \text{Eq. (21)}$$

where, ‗s' is the distance between the periods.

The above three conditions that defines stationarity are satisfied only if $|\rho| < 1$. In the case where $|\rho| > 1$, which is, although rare in the economic or finance series, implies that the series is explosive. That is, the future values of the series will not return back to mean values but will, instead, diverge at a faster pace. If $|\rho| = 1$, then we have the case that the mean reverting level is undefined as all the denominations in the above three conditions will turn to zero. In this particular case, i.e., $|\rho| = 1$, the time series is not covariance stationary and exhibits what is called a ‗unit root' (Maddala and Kim, 2004; Brooks, 2014; Balarezo, 2010).

It is worthy to note that, autocovariances are not a particularly useful measure of the relationship between Y_t and its previous values. Scholars suggest to use a more convenient measure by transforming or standardizing, that is, normalizing the autocovariance. The normalization by variance i.e., dividing the autocovariance by variance, we get autocorrelation which is more convenient to use (Maddala and Kim, 2004; Brook, 2014).

3.5 Non-stationary Process:

The opposite of stationary time series process is non–stationary process. The time series process which does not satisfy the stationarity conditions stated in the equations15, 16, and 17, that is, which has a structure of time varying mean or time varying variance or both is coined in the literature of econometrics as non – stationary time series process. In applied finance and economics, we mostly encounter with the non – stationary time series process and the simple example of which is the random walk model (RWM) (Gujarati, 2003; Greene, 2003). A vital feature of these type of series is the ‗persistence of random shock', that is, the impact of a particular shock does not die away and hence have a ‗infinite memory' (Patterson, 2000). Random walk model also falls under the unit root process (Gujarati, 2003). A series follows a unit root process if it is non – stationary and becomes stationary by taking first difference (Brooks, 2014). RWM is a specific case of a more general class of stochastic process known in the literature as integrated process. Series with random walk components are called integrated because they are the integrals of weakly stationary components in the series (Aoki, 1990). If a non – stationary time series becomes stationary after its first difference, then that time series is termed as integrated of order one, and denoted by I(1). In general, if a non – stationary time series has to be differentiated ‗d' times to make it stationary, then the time series is said to be integrated of order ‗d' and denoted by I(d). Thus, if a time series is stationary at level, that is, the series does not require any differencing, it is said to be integrated of order zero, and denoted by I(0) (Gujarati, 2003; Greene, 2003).

Scholars observe that an I(1) series at level form will grow constantly and wander about with no tendency to revert back to a fixed mean. Macroeconomic and financial flows and stocks which relate to population size are mostly I(1) process (Patterson, 2000; Greene, 2003; Gujarati, 2003). Correct identification of the process of a time series whether stationary or non – stationary is heavily researched and debated by the scholars and the issue is so challenging that it seems to ‗never die out'. The basic literature on the issue dates back to the works of Mann and Wald (1943) and Rubin (1950) (Greene 2003). But the huge attention of a large number of scholars is drawn after the seminal paper of Nelson and Plosser (1982) who observed that almost all the macro economic variables in US are I(1). Interestingly, hundreds of studies, both theoretical and empirical on unit roots in finance and economics appeared only in 1980s (Greene, 2003).

3.5.1 Estimation of the Stationary or Non-stationary Process (Unit-root Test):

There are basically, three approaches to identify stationary or non – stationary process i.e. the order of integration of a time series process: (i) graphical inspection or analysis, (ii) observation of the sample autocorrelation and partial autocorrelation functions (Box and Jenkins, 1976) and (iii) formal statistical tests. The ‗eye ball estimation' of the sample time series is the simplest method. It compares the plot of the time series under study with the time series of first and second differences and concludes. It may be an effective procedure for an efficient and experienced scholar but is very subjective one. Under the second method, if the plot of the autocorrelation function (ACF) is slowly decreasing, roughly at a linear rate, and the first value of the partial autocorrelation function (PACF) is very high, then the series under consideration is identified as non – stationary one. The differencing value, i.e., the value of ‗d' then identify the series as I(1) or I(2), etc. This approach is also widely used but suffers from the risk of over – differencing (Gujarati, 2003). Hence, scholars suggest that it is better to approach the identification problem formally using appropriate statistical tests (Maddala and Kim, 2004; Altrova and Fedorova, 2016).

The area of research of the ‗unit-root test' is vast and one can find several formal statistical tests are recommended and used in empirical studies to identify the ‗correct' order of integration. Amongst the battery of tests, the most popular and widely used tests are Dickey Fuller's test (DF tests) (Dickey and Fuller, 1979), Augmented Dickey Fuller test (ADF – test) (Said and Dickey, 1984), Phillips–Perron test (PP–test) (Phillipsand Perron, 1988), KPSS test (Kwiatkowski, Phillips, Schmidt and Sin, 1992) and a bit less frequently used but strongly recommended ADF – GLS test (Elliot, Rothenberg and Stock, 1996) and NGP test (Ng and Perron, 1995,2001). It is worthy to note here that, almost all the tests suffer from ‗lowpower' and virtually a uniformly powerful test does not exist. The choice of an appropriate one largely depends upon the intuition of the analyst which certainly is a subjective judgement

(Schwert, 1989; Dufour and King, 1991; Altrova and Fedorova, 2016). Some scholars strongly argue to reject the use of DF, ADF, PP and KPSS tests and recommended the modified tests like ADF – GLS and NGP tests (Dufour and King, 1991; Maddala and Kim, 2004). Altrova and Fedorova (2016), for a time series of long length (i.e. at least 500 observations) and positive values of the autoregressive parameter AR(1), recommended to use ADF, ADF – GLS and NGP tests. Moreover, Greene (2003), strongly suggests the ‗practioners' to use ADF-test.

In this thesis, to avoid the risk of identifying incorrect order of integration of the time series under the study and also keeping in mind the recommendations of the eminent scholars (Maddala and Kim, 2004; Altrova and Fedorova, 2016), we used ADF test, ADF – GLS test and NGP test.

3.5.2 Augmented Dickey Fuller Test (ADF-test):

The augmented Dickey Fuller test is the modified test procedure of the original Dickey – Fuller test (Dickey and Fuller, 1979). It is observed that many financial time series have more complicated dynamic structure than the simple AR (1) model. Hence, an augmented component was introduced to the basic DF test (Dickey and Fuller, 1979) to accommodate general autoregressive moving average i.e., ARMA (p,q) models with unknown orders. The augmented components are incorporated to soak the disturbances and make the residuals free from autocorrelation. The new or modified test is based on estimating the test regression:

$$y_t = D_t + \varphi y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta y_{t-i} + \varepsilon_t \quad \text{Eq. (22)}$$

where, D_t is a vector of deterministic terms, may be constant, trend, etc, and p is the lagged difference terms or optimum lag order. Δy_{t-i} are used to approximate the ARMA structure of the errors and the value of p is ‗set‘ in such a way that error $\varepsilon_t$is serially uncorrelated. Here, the null hypothesis is Y_t is I(1)‘ against the alternative hypothesis I(0).

An alternative formulation of the test is:

$$\Delta y_t = D_t + \pi y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta y_{t-i} + \varepsilon_t \text{……………………………………………………} \text{Eq. (23)}$$

where, $\pi = \varphi - 1$. Under the null hypothesis Δy_t is I(0) which means to test $\pi = 0$. This Eq. (23) is used in practice for testing the significance of the coefficient Y_{t-1}.

The ADF t-statistics and normalised bias statistics are based on the least square estimate of Eq. (22) and are:

$$ADF_t = t_{\varphi=1} = \frac{\hat{\emptyset}-1}{SE(\emptyset)} \text{………………………………………………………………………………} \text{Eq. (24)}$$

$$ADF_n = \frac{T(\hat{\emptyset}-1)}{1-\gamma 1-\cdots-\gamma p} \text{………………………………………………………………………………} \text{Eq. (25)}$$

The limiting distribution of the test statistic is identical with the distribution of DF test (Dickey and Fuller, 1979) and for $T \rightarrow \infty$is tabulated by Dickey (1976) and MacKinnon (1991).

3.5.3 Augmented Dickey Fuller–Generalised Least Square (ADF–GLS) test:

This test is also known as ERS-test (Elliot, Rothenberg and Stock, 1996) and is recommended by modifying the ADF test. Before the unit root test, ERS test uses the de-trending transformation that primarily removes trend from time series. The constant in the model:

$$y_t = \beta_0 + \emptyset y_{t-1} + \varepsilon t \text{………………………………………………………………………………} \text{Eq. (26)}$$

after estimated by using generalised least square method with the transformation (Altrova and Feorova, 2016):

$\breve{y} = y_1, \breve{y} = y_t - \rho y_{t-1}$,   t=2,……., T

$x_1 = 1, x_t = 1 - \rho$, $t = 2, \ldots, T$ ....Eq. (27)

where, $\rho = 1 + \bar{C} / T$ and $\bar{C} = (-)7$. The constant is estimated based on the equation:

$\breve{y} = \beta_0 x_t + e_t$ … ....Eq. (28)

The parameter is estimated with the help of least square method and is used to remove constant from the time series y_t. The trend, especially the linear trend, is estimated by generalised least square method by transforming the Eq. (28) by:

$Z_1 = 1, Z_t = t - \rho\,(t-1)$,

where, $\rho = 1 + \bar{C} / T$ and $\bar{C} = (-)13.5$. The estimates of the parameters are computed by:

$\breve{y} = \beta_0 x_t + \beta_1 z_t + \varepsilon t$..Eq. (29)

and the estimated parameters $\breve{\beta}_0$ and $\breve{\beta}_1$ are then used to remove constant and a trend from the series y_t :

$y_t^* = y_t - (\breve{\beta}_0 + \breve{\beta}_t)$.. Eq. (30)

after transformation, ADF test is applied on the transformed time series y^*_t and the test statistic is obtained by the equation:

$\Delta y_t^* = \beta_0 + \emptyset y^*_{t-1} + \sum_{i=1}^{p} \gamma i \, \Delta y^*_{t-I} + \varepsilon t$… ....Eq. (31)

The critical values are provided by Elliot, Rothenberg and Stock (1996). It is worthwhile to note here that, the value of $\bar{C}$ for constant = (-)7 and for linear trend (-) 13.5 was deduced for $\propto = 0.05$ by Elliot, Rothenberg and Stock (1996) based on power envelop, and for given T, the tests are optimal at the 50% power.

3.5.4 Ng – Perron test (NGP test):

Ng and Perron (1995, 2001) recommended the test by modifying Phillips – Perron test (Phillips and Perron 1988) with the detrended data y^*_t, Eq. (30), obtained from the ADF – GLS test. Their basic equation is (Maddala and Kim, 2004):

$$y_t = D_t + \emptyset y_{t-1} + \sum_{i=1}^{p} \gamma i \Delta y_{t-I} + \varepsilon t \ldots \ldots \text{Eq. (32)}$$

and the test statistic z of the P P–test (Phillip and Perron, 1988) are modified to the form below (Altrova and Fedorova, 2016):

$$\overline{M}\overline{Z}_{\emptyset} = (T^{-1} y^*_t - S^2_{AR}) (2T^{-2} \sum_{t=1}^{T} y^*_{t-1})^{-1} \ldots \ldots \text{Eq. (33)}$$

$$\overline{M}\overline{S}\overline{B} = (T^{-2} \sum_{1=1}^{T} \frac{y*t-I}{S^2_{AR}})^{1/2} \ldots \ldots \text{Eq. (34)}$$

Here, MSB is the modified Sargan – Bhargava test (Sargan and Bhargava, 1983) and for I(0) series, MSB is related to the P P- test by (Maddala and Kim, 2004):

$$\overline{M}\overline{Z}_T = \overline{M}\overline{Z}_{\emptyset} . \overline{M}\overline{S}\overline{B} \ldots \ldots \text{Eq. (35)}$$

Here, $S^2_{AR} = (\sum_{t=p+1}^{T} \varepsilon_t^2) ((T - K)(1- \sum_{i=1}^{P} \beta_i)^2)^{-1}$, $\emptyset$ = auto regressive parameter, y^*_t is the transformed time series and p is the optimum lag order.

3.6 Optimum lag-order:

The unit-root tests discussed above are, for their power, very sensitive to the lag orders. Scholars recommend several ways to select the optimum lag order but most of them are inappropriate save some modified information criteria (Maddala and Kim, 2003).

The optimum lag order used in this study is derived by ‗information criteria' like:

3.6.1 Akaike Information Criterion:

Akaike Information Criterion which is defined as (Akaike, 1974):

$$AIC = e^{2k/n}\ (\Sigma\ \hat{u}_i^2)/n \quad \text{.......... Eq. (36)}$$

where $\hat{u}_i$ is the estimated error terms from the regression equation involving the time series.The minimum of the criterion serves as the guide to select the optimum lag range.

3.6.2 Schwartz Information Criterion:

Schwartz Information Criterion which is defined as (Schwarz, 1978):

$$SIC = n^{k/n}\ (\Sigma\ \hat{u}_i^2)/n \quad \text{.......... Eq. (37)}$$

where $\hat{u}_i$ is the estimated error terms from the regression equation involving two macrovariables. Like the AIC, here also, the minimum of the criterion serves as the guide to select the optimum lag range.

3.6.3 Final Prediction Error Criterion (FPE):

Final Prediction Error Criterion (FPE) which is generally described in a Error Correction Model as (Hsiao, 1979, 1981):

$$\text{FPE } \Delta y_t(n,m) = \frac{(N+n+m+1)\Sigma(\Delta y_t - \Delta\hat{y}_t)^2}{(N-n-m-1)/N} \quad \text{.......... Eq. (38)}$$

where, n and m are the number of lags in the dependent (explained) and independent (unexplained) variables respectively; N is the number of observations. To obtain the optimum lag, one has to determine the minimum FPE Δy_t by running the regression equations involving the dependent and independent variables under question (Wong et. al. 2004).

3.6.4 Hannan-Quinn Information Criterion (HQC):

Hannan-Quinn Information Criterion (HQC)which is defined as (Hannan and Quinn, 1979):

$$HQC = -2\, l(\hat{a}) + 2k \log \log n \qquad \text{Eq. (39)}$$

where, $l(\hat{a})$ represents the minimum log likelihood as a function of the vector of parameter estimates($\hat{a}$), and k denotes the number of ―independently adjusted parameters within the model".

It is noteworthy here that, none of these criteria is necessarily statistically superior to others (Gujarati, 2003; Maddala and Kim, 2003). Scholars like, Diebold (2001), Stock (1994), however, recommends SIC to be applied. But cautionary approach guided us to explore all the criteria stated above for selecting the optimum lag order.

3.7 Measurement of the Relationship:

It is observed that, almost always non stationary and trending variables are involved in the empirical studies under finance and macroeconomics (Nelson and Plosser, 1982; Phylaktis and Ravazzolos, 2002; Greene, 2003). To avoid spurious and non-standard conclusions, scholars suggest that the best way to handle such series is to use differencing and other transformation like seasonal adjustments etc, and reduce the variables to stationary processes and then analyse the resulting series as VARs (suggested first by Sims, 1980) or the methodology recommended by Box- Jenkins (1970) (Maddala and Kim, 2004).

VAR methodology is very popular for some of its merits but a good number of scholars heavily criticized and pointed to several pre-conditions (which cannot be fulfilled in the real world situation) and disadvantages of VAR methodology like , (i) a very high number of parameters to be estimated , e.g., if we consider k- variables and q- lags for each variable then $k+qk^2$ parameters, which might be a huge number, is to be estimated (Schlegel, 1985; Brooks, 2014), (ii) a good command over the economic or financial theory and proper institutional knowledge is required to solve the identification problem which cannot be solved by VAR (Stock and Watson, 2001), (iii) the results are mostly sensitive to the choice of the optimum lag order and the ordering of the variables in estimation(Mida, 2013), (iv) the problems of multicolinearity often arise due to the fact that a lot of economic time series are correlated with their own past values, hence, increasing the number of variables and lags might cause multicollinearity and in practice, the unrestricted VAR model gives very ‗erratic' estimates because of the multicolinarity of the explanatory variables (Maddala and Kim, 2004; Schegel, 1985), (v) VARs are a-theoretical (Brooks, 2014), (vi) only stationary variables are estimated in VAR, and (vii) a measurement errors or misspecification of the model will induce unexplained information left in the disturbance term resulting interpretation and inferences more difficult (Hendry, 1995).

Again, Box–Jenkins (1970) methodology is one of the most popular methodologies for analysis of time- series data. It is widely used by the scholars as they can estimate any type of series, stationary or non- stationary, with or without seasonality through this methodology. Furthermore, it is very popular as it has well documented computer programmes (Maddala and Kim, 2004). Despite its popularity, it has some serious limitations also. Differencing and transformations of the data at level force to the analyst to lose some important signals or information in forecasting. Autocorrelation of residuals can produce ‗common factors', especially, in over fitted models, which makes estimation difficult and the statistical tests will ill behave (Pelgrin, 2011). Scholars also point to several other limitations of the methodology for using in finance and economics like, i) it requires a long- series, ii) useful for short – run but not long-run analysis, iii) usually financial time series data, contains asymmetries, sudden breaks at irregular intervals and periods of high and low volatility which is impossible to be captured adequately by the ARMA process as that violates the assumption of ‗ constant variance‘ of the ARMA methodology (Petrica et al. 2016) , and (v) in general, financial time series reveals more complex structures than those by ARMA process, hence, against more complex approaches, it, at the best, can be considered as a ‗first starting point‘(Rachev et al. 2007; Bellgard and Goldsmith, 1999; Kon, 1984; Maddala and Kim, 2004).

The recent research and growing literature in macro econometrics has shown that ‗there are more interesting and appropriate ways‘, to analyse financial time series and trending variables (Green, 2003). Cointegration and its related technique – ‗error correction‘ is now widely researched and used as they are concerned with methods of estimation that preserve the information about both form of covariation (Green, 2003; Engels and Granger,1987; Hamilton, 1994; Watson, 1994).

In this study, we attempt to measure the cross-regime relationship between select macroeconomic variables and Indian stock market with the very popular but robust and widely used technique of cointegration.

3.7.1 Cointegration:

Cointegration is a concept whereby time series have a fixed relationship in the long-run (Brooks, 2014). It was introduced by Granger (1981). In simple language, let us consider two variables y_t and x_t which are integrated at order one or I(1). Then, y_t and x_t are said to be cointegrated if there exists a β such that $y_t - \beta_1 x_t$ is stationary or I(0). That is, estimation of the regression equation $y_t = \beta x_t + u_t$ or $u_t = y_t - \beta x_t$, where u_t is the error term, will not produce spurious results. Because, an important property of non- stationary variables, integrated of order one is that, there can be linear combination of these variables which are stationary or (I(0)). And, y_t and x_t do not drift too far apart from each other in the long-run or can drift together at roughly the same rate and might be stable around a fixed mean over time (Greene, 2003). This implies that those time series are moving stochastically together towards some long-run relationship, and the vector [1 - β] or any multiple of it is a cointegrating vector, that is, the set of parameters that describe the relationship between two or more time series (Engle and Granger, 1987; Greene, 2003; Brooks, 2014). Scholars suggest, in such a case, one can clearly distinguish between a long – run equilibrium relationship among the variables y_t and x_t, that is, the manner in which the two variables drift upward together and the short run dynamics depicting the relationship between deviations of y_t from its long – run trend and deviation of x_t from its long – run trend. In this situation, obviously differencing the data is not required to avoid spurious inferences; rather, it would be counterproductive as it would obscure the long-run relationship between the time series (Greene, 2003).There are several methodologies to test co- integration.

3.7.2 CRDW Test:

The most simple and quick test methodology of cointegration is the Cointegrating Regression Durbin-Watson test (CRDW- Test, Gujarati, 2004). Under this method, in a bi-variate setting, we have to estimate the assumed cointegrating equation like:

$$y_t = \beta_1 + \beta_2 x_t + u_t \qquad \text{Eq. (40)}$$

and estimate the Durbin-Watson test statistics (1950,1957) for the first order auto-correlation. Under the null hypothesis that y_t is a random walk and $\beta_2 = 0$, so there is no co-integration and $\hat{u}_t$ becomes a random walk with theoretical first order autocorrealtion equal to unity. If the computed value of ‗*d*' is more than the critical value, then the Eq. (40) is identified as cointegrating equation. The critical value of ‗*d*' is firstly provided by Sargan and Bhargava (1983). Although, the method is very simple, but it suffers from two major problems – (i) it is very sensitive to the assumption of dependant variable be a true random walk and (ii) the critical values of the test statistic are not consistent if the number of regressors increases over the sample size (Sjo, 2019). The use of the test in empirical studies is, therefore, extremely limited (Sjo, 2019).

3.7.3 Engle-Granger Test:

In the literature of finance and economics, till date, two methods gained popularity and are widely used in the empirical studies (Chen, 2012) and one of them is suggested by Engle and Granger (1987). It is a two-step procedure to test cointegration. In the first step, they suggest to estimate the so called cointegrating regression, basically in bi- variate nature but can be extended to multivariate one:

$$y_t = \beta_1 + \beta_2 x_t + u_t \quad \text{Eq. (41)}$$

where, y_t and x_t are the variables and in the regression the assumption is that, all variables are integrated at order one, i.e. I(1) at level and might cointegrate to form a stationary relationship so that the residual term $\hat{u}_t = y_t - \beta_1 - \beta_2 x_t$ is stationary or I(0).

In the second step, Engle and Granger (1987) suggested to test for a unit root in the residual process of the Eq. (41) above. DF (Dicky and Fuller, 1979) or ADF (Said and Dicky, 1984) tests can be used to ascertain the stationarity. As the estimated u_t are based on the estimated cointegrating parameter,β_2 , the DF (Dicky and Fuller, 1979) and ADF (Said and Dicky, 1984) critical significance values are not appropriate because the unit root test is now applied to a derived variable-- the estimated residual from a regression (Sjo, 2019). The new set of critical values is provided now by Engle and Granger (1987). If the null hypothesis of unit root is rejected, then the residuals are identified as the stationary one and the regression equation as the cointegrating equation. More specifically, the variable y_t is then said to be cointegrated with x_t. Hence, it is posited that, the Eq. (41) represents the assumed economically meaningful or understandable steady state or equilibrium relationship among the variables. Moreover, according to the properties of ‗super converge', the estimated parameters are representing correct estimates of the long-run steady state parameters and the residuals at lag one can be used as an error correction term in the error correction model (Sjo, 2019).

Despite the simplicity, easy to perform and popularity of the procedure, the Engle – Granger (1987) two step cointegration methodology suffers from many limitations like, (i) the estimated standard errors of this ‗procedure' are, in general, useless when the variables are cointegtrated, hence, using standard distribution no inference is possible, (ii) asymptotically, the test procedure is independent of the arrangement of the variables in the regression equation but choosing a variable in the left hand side of the regression equation, the cointegrating vector is normalised around that variable which means that normalization corresponds to some long-run economic relationship. Scholars disagree to this proposition as this is not true in limited samples; normalisation matters and makes an economic sense and plays a vital role in economic interpretation (Ng and Perron, 1995), (iii) as the test involves an ADF test (Said and Dicky, 1984) in the second step, all the limitations of ADF test is valid in this cointegration test also, especially, the selection of lag order for the augmentation is a very critical factor, (iv) the test is based on the assumption of one cointegrating vector and, as such, very useful in the model with two variables. Even so, if two variables cointegrate then adding a third integrated variable will not change the results of the test, because, if the third variable do not belong to the cointegrating vector, ordinary least square (OLS) estimation will simply put its parameters to zero leaving error process unchanged (Sjo, 2019) and (v) the test assumes a common factor in the dynamics of the system but the common factor restriction is a severe restriction since all the short – run dynamics is forced to the residual process(Sjo, 2019). Counting the limitations, Maddala and Kim (2004) strongly opposes to use the test procedure in empirical analysis.

Our study, in assessing the relation between select macroeconomic variables and stock market involves more than two variables, hence, there is a possibility of having more than one cointegrating vector. In simple language, the variables in the model might form several equilibrium relationship governing the joint evolution of all the variables as, in general, for n number of variables we can have up to (n-1) cointegrating vectors. So, we need a test procedure which captures more than one cointegrating relationship in the model. The widely popular and ‗superior' test for cointegration is the test suggested by Johansen and Juselius (1990) and Johansen (1991, 1995).

3.7.4 Johansen and Juselius Methodology:

Johansen and Juselius (1990), Johansen (1991,1995) procedure is essentially a process to estimate cointegrating vector based on the maximum likelihood ratio (LR) statistics λ_{max} and λ_{trace}(Johansen 1991, 1995).

Johansen and Jusellius (1990) method is based on the VAR model estimation (Kozhan 2010). Accordingly, we are to start with a simple VAR representation (Maddala and Kim, 2004):

$A_k(L)y_t = \mu_0 + \varepsilon_t$ .... Eq. (42)

and, following Engle and Ganger (1987),under some regulatory conditions and using difference operator Δ (i.e. Δ = 1-L or L = 1-Δ), we can write the cointegrating process, y_t , in VECM ,below:

$$\Delta y_t = \mu_0 + \sum_{i=1}^{k-1} \Gamma_i \Delta y_{t-i} + \Pi y_{t-1} + \varphi D_t + \varepsilon_t \text{ ......} \quad \text{Eq. (43)}$$

where, μ_0 is unrestricted constant component, D_t may include deterministic regressors like, trend and (seasonal) dummy variables, k equals to number of lags, y_t is a p×1 vector and p is the number of variables. The variables are integrated of order ‗d' i.e. $\{y_t\} \sim I(d)$, Γ and Π are the coefficient-matrices representing short and long term impacts, respectively and ε_t is residual vector assumed to be independent and identically distributed as multi-normal distribution with mean zero and variance Ω. Johansen (1991, 1995) decomposes Π in two matrices α and β, where, α represents the effect of each cointegrating vector on the $\Delta y_{p,t}$ variables, that is, the loading vector and β the cointegrating vector. Both α and β are p×r matrices where, $r < p$ such that $\Pi = \alpha\beta'$.

It is worthy to note here that, the number of conintegrating vectors are equal to the number of stationary relationship in the Π matrix. Again, the rank of Π matrix determines the number of independent rows in Π matrix which further identifies the number of cointegrating vectors. The rank ‗r' of the Π matrix is identified by the number of significant eigenvalues found in the estimated Π of the Eq. (43). Each significant eigenvalue depicts a stationary relation and reduced rank in Π matrix, i.e., 0<r< p indicates that there exists cointegrating relation amongst the variables under the study. Here, rank equal to zero indicates that all the variables are non-stationary and rank equals to ‗p' indicates that all the variables are stationary (Sjo, 2019). Hence, for the presence of cointrgration, the rank ‗r' must lie between zero and number of variables p.

Johansen and Jusellius (1990) suggested two methods or two test statistics to identify the cointegrating equation. The first one is the maximum eigenvalue statistics $(\lambda_{\max})$ which consists of ordering the largest eigenvalue is descending order and testify whether they are significantly different from zero (Abbas et al. 2017). The ‗$\lambda_{\max}$ test' tests the null hypothesis of ‗r' cointegrating vector against the hypothesis that there are r+1cointegrating vectors (Maddala and Kim, 2004). The likelihood ratio (LR) test stastistic is:

$$\lambda_{\max} = -T \ln(1-\hat{\lambda}_{r+1}) \quad \text{Eq. (44)}$$

The second one, named ‗trace test', tests the hypothesis that there at most ‗r' conintegrating vector and the LR test statistic is:

$$\lambda_{trace} = -T \sum_{i=r+1}^{p} \ln\left(1-\hat{\lambda}_i\right) \quad \text{Eq. (45)}$$

It has been found that the trace test performs better that the $\lambda_{\max}$ test as (i) the λ_{trace} statistic takes into account all (p-r) of the smallest eigenvalues, (ii) it can be adjusted for degrees of freedom, and (iii) it tends to have more power than $\lambda_{\max}$ statistic. Hence, in empirical analysis and if there be any contradiction in results of $\lambda_{\max}$ and λ_{trace} tests, along with Johansen and Jusellius (1990), scholars advocate to use λ_{trace} statistic (Kasa, 1992; Lutkepohl and Reimers, 1992; Serletis and King, 1997; Lutkepohl et al. 2001; Sjo, 2019). We, throughout our thesis, used the ‗trace test' statistic to identify cointegrating rank over and above the $\lambda_{\max}$ statistic.

The tests for determining the number of cointegrating vectors are nested. Hence, the tests are performed by starting from the hypothesis of zero cointegrating vectors. In simple words,

$H_{0,1}$: 0, i.e., zero conintegrating vector is tested against the alternative $H_{a,1}$: 1 i.e., at most one cointegrating vector. If $H_{0,1}$ is rejected, then the next test is $H_{0,2}$: 1, i.e. , one cointegrating vector against $H_{a,2}$: 2, i.e., at most two cointegrating vectors, and so on.

The LR test statistic for identifying cointegration varies according to the assumptions made on the deterministic trends, i.e., of constants and trends in the model. Five deterministic trend assumptions are suggested by Johansen (1995) and the models to estimate are below (Sjo, 2019):

1. Assumption I (Model 1): Here, the assumption is- ‗there is no deterministic trend in the data and no intercept or trend in the cointegrating equation.' That is, the constant is assumed to be zero ($\mu_0 = 0$) for both the first differenced equations and cointegrating vectors. The estimated model for the p - dimensional vector of variables y_t is:

$$\Delta y_t = \sum_{i=1}^{k-1} \Gamma_i \Delta y_{t-i} + \alpha\beta' y_{t-1} + \varphi D_t + \varepsilon_t \qquad \text{Eq. (46)}$$

2. Assumption II (Model 2): Here, the assumption is- ‗there is no deterministic trend in the data at level and an intercept but no trend present in the cointegrating space of the variables.' That is, μ_0 is assumed to be zero but allows for constant in cointegrating vector. The model to estimate is :

$$\Delta y_t = \sum_{i=1}^{k-1} \Gamma_i \Delta y_{t-i} + \alpha[\beta', \beta_0][y_{t-1}, 1] + \varphi D_t + \varepsilon_t \qquad \text{Eq. (47)}$$

3. Assumption III (Model 3): Here, the assumption is- ‗there is a linear trend in the data at level and intercept with no trend is present in the cointegrating space of the variables.' In

other words, μ_0 is left unrestricted which actually lead to the inclusion of both deterministic trend in the y_t's and constant in thecointegrating vector. The model to estimate is :

$$\Delta y_t = \mu_0 + \sum_{i=1}^{k-1} \Gamma_i \Delta y_{t-i} + \alpha\beta' y_{t-1} + \varphi D_t + \varepsilon_t \text{ ....} \qquad \text{Eq. (48)}$$

4. Assumption IV (Model 4): Here, the assumption is- ‗there is a linear trend in the data at level and both intercept and trend is present in the cointegrating space of the variables.‘ That is, it is assumed to allow for constants and deterministic trends in the cointegrating vectors. The model to estimate is :

$$\Delta y_t = \mu_0 + \sum_{i=1}^{k-1} \Gamma_i \Delta y_{t-i} + \alpha\left[\beta', \beta_1, \beta_0\right]' \left[y_{t-1}, t, 1\right] + \varphi D_t + \varepsilon_t \qquad \text{Eq. (49)}$$

5. Assumption V (Model 5): Here, it is assumed to allow for a quadratic trend in data (y_t's) and both an intercept and trend in the cointegrating equation.‘ That is, least restricted situation and the model to estimate is :

$$\Delta y_t = \mu_0 + \mu_{1t} + \sum_{i=1}^{k-1} \Gamma_i \Delta y_{t-i} + \alpha\left[\beta', \beta_1, \beta_0\right]' \left[y_{t-1}, t, 1\right] y_{t-1} + \varphi D_t + \varepsilon_t \qquad \text{Eq. (50)}$$

Amongst the all five alternative models to select the appropriate deterministic trend, Model 1 is the most restrictive one and carries a very little or no sense from the viewpoint of economics and finance. Because, in social issues, it is very impractical to assume that all the variables in the cointegrating space have the same mean. Similarly, in the case of Model 5, the assumption of the presence of quadratic trends in levels of the variables is also almost unrealistic in economics and finance. It may provide a good fit but will produce dubious out-of sample forecasts (Lada, 2007; Sjo, 2019). Model 2, 3 and 4 are most relevant in empirical studies and out of the five models, Model 3 is advocated as the basic model and is widely chosen and used in empirical studies (Lada, 2007; Sjo, 2019; Wangbangpo and Sharma, 2002). Johansen (1995) suggested to estimate all the models and to perform reduced rank test across the models to determine the most appropriate one. In our study we have estimated only the Model 2, 3 and 4. The search procedure starts from the most restricted to less restricted model and the deterministic component is then determined where the null hypothesis is not rejected for the first time. The technique of the search procedure is explained below (Table-3.1), considering number variables is six (p=6) and models under search are only 2, 3 and 4.

Table-3.1

Testing Deterministic Trends

Null Hypothesis $H_0 : r$	No Number of common trends (unit roots) $p-r$	Model- 2	Model- 3	Model-4
0	6	$\lambda_{trace2,0} \rightarrow$	$\lambda_{trace3,0} \rightarrow$	$\lambda_{trace4,0}$
1	5	$\lambda_{trace2,1} \rightarrow$	$\lambda_{trace3,1} \rightarrow$	$\lambda_{trace4,1}$
2	4	$\lambda_{trace2,2} \rightarrow$	$\lambda_{trace3,2} \rightarrow$	$\lambda_{trace4,2}$
3	3	$\lambda_{trace2,3} \rightarrow$	$\lambda_{trace3,3} \rightarrow$	$\lambda_{trace4,3}$
4	2	$\lambda_{trace2,4} \rightarrow$	$\lambda_{trace3,4} \rightarrow$	$\lambda_{trace4,4}$
5	1	$\lambda_{trace2,5} \rightarrow$	$\lambda_{trace3,5} \rightarrow$	$\lambda_{trace4,5}$
6	0	$\lambda_{trace2,6} \rightarrow$	$\lambda_{trace3,6} \rightarrow$	$\lambda_{trace4,6}$

Note: 1.r = number of cointegrating vector

2.p = number of variables

3.p-r = number of common trends (unit roots)

4. $\lambda_{trace i,j}$ = estimated trace statistics where i represents Model (i=2,3 and 4)

and j represents the number of the significant eigenvalues (j=1,2,...,p)

5. $\rightarrow$= direction of the search procedure

The search procedure starts from row one of the Table 1, i.e., testing Model 2 for cointegrating rank ‗r' equals to zero. If the $\lambda_{trace2,0}$ is rejected, then we have to test r= 0 for Model 3, and so on. Now, in the event that all three models reject the assumptions of zero cointegrating vector in favour of at least one cointegrating vector, we have to move to the second row where the null hypothesis is r= 1, and so on. When for the first time our estimated eigenvalue cannot reject the null hypothesis, then, we have to end the search procedure because looking at the test statistics beyond this point might be misleading (Sjo, 2019). Let, $\lambda_{trace3,3}$ cannot reject the null, then it points to the fact that there exists three cointegrating vectors and the Model 3 is the better model for describing the system under study. Although, at the top of the issue, it is the logic of economics or finance and the objective of the researcher which finally guides the scholar to select the best model (Maddala and Kim, 2004).

Johansen and Jesulius (1990) observed that the first cointegrating vector corresponding to the highest eigenvalue in most related with the stationary part of the model and are very useful in explaining the empirical results. Moreover, the coefficients of the first cointegrating vector seem to possess the signs consistent with the a priori hypothesis of the researchers (Wangbangpo and Sharma, 2002). Accordingly, after normalising the target variable to ‗one', the long-run relationship amongst the variables corresponding to the highest eigenvalue at optimum lag order with the estimated cointegrating rank is reported in our thesis. In our thesis target variable although varies on the basis of the objectives of the study.

3.7.5 Error Correction Model (ECM):

Error Correction Model (ECM) sometimes considered as the model to assess the speed of adjustment towards equilibrium and preserves the information about the short-run form of covariance (Green, 2003). In essence, the ‗short-run dynamics' or ‗speed of adjustment towards equilibrium' depicts the relationship between deviation of y_t from its long-run trend and x_t from its own long-run trend. They are expressed by the error correction term of the model below (Green, 2003; Johansen, 1991):

$$\Delta y_t = \alpha_0 + \alpha_1 ECT_{t-1} + \sum_{i=1}^{k} \alpha_{2i} \Delta y_{t-i} + \sum_{i=0}^{k} \alpha_{3i} \Delta x_{t-i} + \varepsilon_t \text{ ....Eq. (51)}$$

where, Δ is the first difference operator, y_t and x_t are the variables, k is the number of variables, α_1 is the error correction coefficient and ε_t, as usual, are stationary random process with mean zero and constant variance. Along with the long-term relationship, we reported the ‗adjustment coefficients' with short-run dynamics also.

3.7.6 The Significance of the Variables:

It is observed that Johansen and Jesulius (1990) procedure only identifies the number of stationary vectors amongst the variables under the study and the cointegrating vectors are not unique save the cointegrating space (Sjo, 2019; Neusser, 2016). Despite some suitable transformations of the cointegrating vectors on some basis, it is often very difficult to interpret the vectors on the line of economic and financial theories. Empirical researchers are interested to relate the state and direction of variables with theories, willing to know which variables are affected by the vectors and the significance of β 's in the cointegrating relation. In simple words, whether the variables form the long-run equilibrium relationship/ relationships or some are missing, even marginally, the cointegrating space.

The significance of β 's in the cointegrating relation is estimated by the likelihood ratio test statistic (Neusser, 2016; Boswijk and Doornik, 2003):

$$LR(H_0 / H_r) = T \sum_{j=1}^{r} \ln \frac{1 - \tilde{\lambda}_j}{1 - \hat{\lambda}_j} \text{ ....Eq. (52)}$$

where, λ_j 's are the eigenvalues ($\tilde{\lambda}_j$ = estimate the eigenvalues for the restricted estimator, $\hat{\lambda}_j$ = unrestricted estimates of eigenvalues), r is the cointegrating rank, n's are the number of variables, j's are cointegrating ranks such that j = 1,2,...,r, s denotes sub-space in the cointegrating vectors such that $r \leq s \leq n$. The test statistic is asymptotically distributed as a Chi-square distribution with r(n– s) degrees of freedom (Neusser 2016). Similarly, it is possible to test hypothesis on α 's ; and the joint hypothesis on α 's and β 's with s(n– r) degrees of freedom (Neusser, 2016). We, in our thesis, tested the β 's with the null as ‗do not belong to the cointegrating relationship' by excluding variables one by one from the long-run relationship (Wangbangpo and Sharma, 2002; Morin, 2006).

3.7.7 Validity of the Relationship:

The popular and widely used tests like ‗portmanteau' (Box and Pierce, 1970; Ljung-Box, 1978) or LM tests are performed on the residuals to assess the serial correlation in residuals considering ‗no serial correlation' up to the lag order ‗h' in the residuals as null.

3.8 Casualty Test:

We used Granger (1969,1987,1988) causality test to assess which of the variables under our study have statistically significant influence on the future values of each of the other variables in the system. Considering only two variables only eg, Y_t and X_t, the test involves estimating the equations below:

$$\Delta Y_t = \alpha_0 + aEC_{t-1} + \sum_{i=1}^{k} \alpha_{1i} \Delta Y_{t-i} + \sum_{i=1}^{k} \alpha_{2i} \Delta X_{t-i} + \epsilon_t^{y} \quad \text{.....................................Eq. (53)}$$

$$\Delta X_t = \beta_0 + bEC_{t-1} + \sum_{i=1}^{k} \beta_{1i} \Delta X_{t-i} + \sum_{i=1}^{k} \beta_{2i} \Delta Y_{t-i} + \epsilon_t^{X} \quad \text{.....................................Eq. (54)}$$

where, EC_{t-1} is the error correction term obtained from the co- integrating vector and α_i , β_i (i=1, ……,k), a and b are parameters to be estimated, k is the optimum lag order in the system, ϵ^x , ϵ^y are stationary random process with mean zero and constant variance (Wangbangpo and Sharma, 2002).

According to Granger (1998), a VECM provides two ways to detect causality: (i) the significance of the lagged error term coefficients which are tested by t- statistics, and (ii) the joint significance of the lags of each variable tested by F–statistics. Simply, X_t Granger cause Y_t if _a' is significant or α_{2i} are jointly significant i.e. $H_0 : \alpha_{21} = \alpha_{22} = \dots\dots\dots\dots = \alpha_{2k} = 0$ is rejected. In other words, if Y_t Granger cause X_t, then lags of Y_t should be significant in the equation for X_t. If this is the only outcome not the vice versa, then it is said that there is unidirectional causality running from Y_t to X_t. If the case is vice versa, i.e. both sets of lags and both EC_{t-1} coefficients are significant, then it is said that there exists a bi-directional causality between Y_t and X_t . If Y_t cause X_t and not vice versa, then it is said that the variable Y_t is strongly exogenous. If neither set of lags are statistically significant in the equation of the other variable, it is said that Y_t and X_t are independent (Brooks, 2014).

3.9 Impulse Response Analysis:

In a VAR model, all variables depend on each other and on the reaction of the system to a shock, only limited information can be obtained by the estimates of individual coefficients. Moreover, it is very difficult to directly understand and interpret the dynamic interactions between variables as VAR is composed of many coefficients. Impulse response function can be used here to trace out the responses of the dependent variable to the structural shocks of each of the variables, even, related to economic or financial model over time (Neusser, 2016). In impulse response analysis, if the responses subside quickly towards zero, then it is said that transmission between the variables in VAR are relatively efficient (Roca and Tularam, 2011). In Granger causality (1987, 1988) we can find the reaction qualitatively but it fails to give any quantitative idea about the effects for the periods in future or the sign of the relationship or how long the effects require to take place. In impulse response function (IRF), the dynamic relationships between variables are captured (Brooks, 2014). Let us consider the VAR model below (using notations of Lutkepohl, 2007):

$$Y_t = \mu + \sum_{i=1}^{p} \emptyset_i \; Y_{t-i} + \epsilon_t , \qquad t= 1,2,3 \ldots\ldots\ldots\ldots.., T \ldots\ldots \text{Eq. (55)}$$

where, $y_t = (SNX',FX',YLGB', YTB`, IIP',WPI',)'$ is a six block vector (m x 1) in our study, where SNX is a vector of stock market indices, FX is a vector of foreign exchange rate (Rs/$), YLGB is a vector yield on long term govt bonds, YTB is a vector yield on 91-days Treasury bill, IIP is a vector of industrial production and WPI is a vector of whole sale price index ,$\emptyset_i$ are fixed m x m coefficient matrices, $\epsilon_t =(\epsilon_{lt} ,\ldots\ldots\ldots.., \epsilon_{mt})'$ is a Ganssian white noise process with E $(\epsilon_t) = 0$, $E(\epsilon_t \epsilon'_t)=\Sigma_\epsilon$, $E(\epsilon_t\epsilon_s) = 0$, for $t\neq s$ and Σ_ϵ is the covariance matrix of the error terms.

The infinite moving average process (MA) of the model (Eq. 55) under the stability condition is:

$$y_t = C+\sum_{i=0}^{\propto} A_i \, \epsilon_{t-i} \ldots\ldots \text{Eq. (56)}$$

The coefficient matrices can be obtained from the recursive formula:

$$A_i = \sum_{j=1}^{i} A_{i-j} \, \varphi_j, \qquad i= 1,2,3, \ldots\ldots \text{Eq. (57)}$$

where $A_0 = I_m$ and $\varphi_j =0$ for j>p. The constant term e= $(I_m - A_1 - \ldots\ldots\ldots.. - A_p)^{-1} \mu$, and the μ is the mean of the process.

Impulse response function of a ‗one standard deviation' shock to the i th variable in y_t , on jth variable in y_{t+m} is given by:

$$\Psi_{ji} (n) = e'_j \, A_n \, P_{ei} , \; n= 0,1,2, \ldots\ldots \text{Eq. (58)}$$

where, e_i is a column selection vector with unity as the ith element and zeroes otherwise, P is lower triangular matrix obtained by decomposing the covariance matrix Σ_ϵ using the Cholesky method so that $PP' = \Sigma_\epsilon$ (Serwa and Wdowinski 2016).

Impulse response analysis and its outcome depend on the ordering of the variables under study. Impulse response function (IRF) computed using Cholesky decomposition is highly sensitive to the ordering of the variables and on the covariance matrix Σ_ϵ , too. Pesaran and Shin (1998) and Koop, Pesaran and Potter (1996) suggested a method of estimation known as Generalised Impulse Response Function which, authors claim, can avoid the problems of ordering of the variables. But, according to Kim (2012), Generalised Impulse Response Function (GIRF) treats all the shock of the variables as if they were ordered first in a VAR. Moreover, GIRF generates responses that are larger and more frequently statistically significant than ordinary IRF. Hence, using generalised impulse response function may result in misleading inferences caused by their extreme identification schemes (Kim, 2012). We, in this study, used ordinary impulse response function using Cholesky decomposition.

3.10 Variance Decomposition Analysis:

In examining the dynamics of VAR system, Variance Decompositions give the proportion of the movements in the dependent variables that are due to their ‗own shocks' and against shocks to other variables in the system. A shock to the ‗i' th variable not only affect the variable, it also transmits the shock to all other variable through the dynamic structure of VAR. Variance decomposition computes how much of the h-step ahead forecast error variance of a given variable is explained by shocks to each explanatory variable for h = 1, 2, 3……, the horizon (h) chosen by the analyst.

According to Sims (2011), the forecast error of a variable at time t is the change in the variable that could not have been forecast between t-1 and t. This is primarily because of the realisation of the structural shocks in the system, ε_t . Using the notations of Sims (2011), for a n variable system, the total forecast error variance of a variable i at horizon h is:

$$\Omega_i (h) = \Sigma^h{}_{k=0} \Sigma^n{}_{j=1} C_{i,j} (k)^2 \text{.......... Eq. (59)}$$

where, $C(L) = A (L)^{-1}B$ to be the matrix polynomial of structural moving average coefficients, $A (L)^{-1}$ is the inverted AR process, K = 0,1,2,……..h,h = horizon of the study for shocks and A (L) is a matrix of lag polynomial.

Quantification of the shock in explaining the variation in each variable in the system amounts to the fraction or proportion of the forecast error variance of each variable due to each shock at each horizon (Sims 2011). The forecast error variance of variable i due to shock j at horizon h is:

$$\omega_{ij} (h) = \sum_{k=0}^{h} C_{i,j} (k)^2 \text{.......... Eq. (60)}$$

Then the fraction of the forecast error variance of variable i due to shock j at horizon h is the above Eq. (60) divided by the total error variance in Eq. (59), is:

$$\Phi_{i,j} (h) = \frac{\omega i,j\ (h)}{\Omega i(h)} = \frac{forecast\ errorvarianceofivariable}{totalforecasterrorvariance}$$

$$= \frac{Equation\ 57}{Equation\ 56} = \frac{\sum_{k=0}^{h} Ci,\text{j (k)2}}{\sum_{k=0}^{h}\sum_{j=1}^{n} C_{i,j}\text{(k)2}} \text{.......... Eq. (61)}$$

here, it is assumed that the shocks have unit variance and are uncorrelated.

3.11 Summary:

In measuring the relationship between select macroeconomic variables and stock market in pre and post liberalisation period in India and to test the hypothesis under this study we used, step by step, the mythologies recommended by the scholars, below:

Step 1: We transformed the series by taking natural logarithm at level data to eliminate the seasonality of the time series.

Step 2: To identify the order of integration of the time series, we used ADF (Said and Dickey, 1984), ADF-GLS (Elliot, Rothenberg and Stock, 1996) and Ng-Perron tests (Ng and Perron, 1995, 2001).

Step 3: The optimum leg order of the VAR process is identified by the information criteria, namely, AIC (Akaike, 1974), BIC (Schwarz, 1978), HQC (Hannan, and Quinn, 1979) and FPE (Hsiao, 1979, 1981).

Step 4: To test the first two null hypotheses of this study i.e., Indian stock market and select macro-economic variables in the pre and post liberalization period are not integrated, we attempted to use the methodology recommended by Johansen and Juselius(1990), Johansen(1991,1995). In this process we also attempted to identify the number of cointegrating rank, if any, deterministic trend ,long-run equilibrium relationship, short-run dynamics and also to identify the variables which are in the co integrating space , the structural stability of the model etc., suggested by eminent scholars.

Step 5: We, to test the third null hypothesis, that is, past values of none of the select Indian macroeconomic variables under the study influence the future values of stock market and vice versa, we attempted to apply the test suggested by Granger (1969, 1988).

Step 6: The fourth null hypothesis, that is, all the select Indian macroeconomic variables and Indian stock market would remain rigid over the future periods of time, is tested with the help of impulse response and forecast error variance decomposition analysis.

Step 7: The fifth hypothesis, that is, the link and relation between select macroeconomic variables and stock market would remain unchanged across the regimes in India, is tested by comparing the results of the analysis of pre and post liberalisation period.

The literature on time series analysis is expanding very fast. Thus, the recommended and widely used mythologies used in this thesis are obviously not free from limitations and the inferences of the study should be used with caution.

References

Abbas, G. Bhowmik, R., Koju, L., & Wang, S. (2017). Cointgration and Causality Relationship Between Stock Market, Money Market and Foreign Exchange Market in Pakistan. *Journal of System Science and Information*, *5*(1), 1- 20. DOI: 10.21078/ JSSI – 2017- 001 – 20.

Agarwal, M.,& Ghosh, S. (2015). Structural Change in the Indian Economy. *Working Paper No 465*, Nov., Centre for Development Studies, Trivandrum. http://www.cds.edu.

Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEETransactions on Automatic Control*, *AC-19*, 716–23.

Andrews, D.W.K. (1993). Tests for Parameter Instabilty and Structural Change with Unknown Change Point. *Econometrica*, *61*(4), 821-856.

Andrews, D.W.K., & Ploberger, W. (1994). Optimal tests when a Nuisance parameter is Present only under the Alternative. *Econometrica*, *62*(6), 1383-1414.

Aoki, M. (1990). *State Space Modellinf of Time Series*. Springer, Heidelberg,1990.

Arltova, M., & Fedorova, D. (2016). Selection of Unit Root Test on the Basis of Length of the Time Series and Value of AR(1)Parameter. *Statistika*, *96*(3), 47-64.

Balarezo, J. (2010). *International Diversification using Cointegration and Modern Portfolio Theory*. Master‘s Thesis, Applied Economics and Finance, Copenhagen Business School, scholar.google.co.in – cbs.dk.

Bellgard, C., & Goldschmidt, P. (1999). Forecasting across Frequencies: Linearity and Non – linearity. *University of Western Australia Research Paper* , Proceedings of the International Conference on Advanced Technology, Australia. http://www.imm.ecel.uwa.au/~cbeiigar/.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, *66*, 47–78.

Bai, J., & Perron, P. (2003). Computation and Analysis of Multiple Structural Change Models. *Journal of Applied Econometrics*, *18*, 1–22.

Boot, T., & Pick, A. (2017). A Near Optimal Test for Structural Breaks when Forecasting under Square Error Loss. *TI 2017 – 039/111,Tinberger Institute Discussion Paper*. https:// dx.doi.org/10.2139/ssrn.2954468.

Boswijk, H.P., & Doornik, J.A. (2003). Identifying, Estimating and Testing Restricted Cointegrated System : An overview.https://www.nuff.ox.ac.uk/economics/papers/2003/w10/BoswijkDoornik.pdf.

Box, G.E.P., & Jenkins, G.M. (1976). *Time Series Analysis : Forecasting and Control*. Holden–Day, San Francisco.

Box, G., & Pierce, D.A. (1970). Distribution of Residual Autocorrelation in Autoregressive Integrated Moving Average time Series Models. *Journal of American Statistical Association*, *65*, 1509-26.

Brockwell, P.J., & Davis, R.A. (2016). *Introduction to Time Series and Forecasting* (3^{rd} Ed.). Springer International Publishing, Switzerland.

Brooks, C. (2014). *Introductory Econometrics for Finance* (3^{rd} Ed.). Cambridge University Press Cambridge, U.K.

Brown, R.L., Durbin , J., & Evans, J.M. (1975). Techniques for Testing Constancy of Regression Relationship Over Time. *Journal of the Royal Statistical Society*, Series B(Methodological), *37*(2), 149- 192.

Chen, X. (2012). *Empirical Investigations into Stock Market Integration and Risk Monitoring of the Emerging Chinese Stock Markets*. A Thesis Submitted for the Degree of PhD at the University of St. Andrews. http://research-repository.st-andrews.ac.uk.

Chow,G.C. (1960). Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica*, *28*(3), 591-605.

Chow,G.C. (1983). *Econometric Methods*. McGraw Hill, New York, 1983.

Commandeur, J.J.F., & Koopman, S.J. (2007). *Introduction to State Space Time Series Analysis*. Oxford University Press.

Cuthbertson, K., Hall, S.G., & Taylor, M.P. (1995). *Applied Econometric Techniques*. The University of Michigan Press, Michigan.

Dickey, D.A. (1976). *Estimation and Hypothesis Testing in Non stationary Time Series*. Iuwa State University Ph.D. Thesis.

Dickey, D.A., & Fuller, W.A. (1979). Distribution of Estimators for Autoregressive Time Series with a Unit root. *Journal of American Statistical Association*, *74*, 427 – 431.

Diebold, F. (2001). *Elements of Forecasting* (2nd Ed.). South Western Publishing, 83-89.

Diebold, F.X., & Chen, C. (1996). Testing Structural Stability with Endogenous Break Point: A Size Comparison of Analytic Bootstrap Procedure. *Journal of Econometrics*, *70*(1), 221- 24.

Dolado, J.J., Gonzalo, J., & Marmol, F. (2003). *Cointegration* (Chapter 30) (B.H. Baltagi, Ed.). *A Companion to Theoretical Econometrics*. Blackwell Publishing Ltd, Massachusetts, USA.

Dufour, J.M. & King, M.L. (1991). Optimal Invariant Tests for Autocorrelation Coefficient in Linear Regressions with Stationary and Non – Stationary AR (1) errors. *Journal of Econometrics*, *47*, 115- 143.

Durbin, J., & Watson, G.S. (1950). Testing for Serial Correlation in Least Squares Regression ,I. *Biometrika*, *37*(3 – 4), 409- 428.

Durbin, J., & Watson, G.S. (1951). Testing for Serial Correlation in Least Squares Regression ,II. *Biometrika*, 38, no. 1 – 2, pp. 159 -179.

Elliott, G., Rothenberg, T. J., & Stock, J.H. (1996). Efficient Test for an Autoregressive Unit Root. *Econometrica*, *64*, 813-836.

Engle, R.F., & Granger, C.W. (1987). Co-integration Error Correction: Representation, Estimation and Testing. *Econometrica*, *55*(2), 251- 276.

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, *37*(3), 424-438.

Granger, C. W. J. (1988). Some Recent Developments in a Concept of Causality. *Journal of Econometrics*, *39*, 199-211.

Greene, W.H. (2003). *Econometric Analysis* (5th Ed.). Prentice Hall, New Jersey.

Gujarati, D.N. (2003). *Basic Econometrics* (4th Ed.). McGraw Hill, New York.

Hamilton, J.D. (1994). *Time Series Analysis.* Princeton University Press, Princeton.

Hannan, E. J., & Quinn, B. G. (1979*).* The determination of the order of an Autoregression. *Journal of the Royal Statistical Society*, *B41*, 190-195.

Hansen, B.E. (1997). Approximate Asymptotic P Values for Structural Change Tests. *Journal of business and Economic Statistics*, *15*(1), 60-67.

Hendry, D.F. (1995). Econometrics and Business Cycle Empirics. *The Economic Journal*, *105*, 1622 – 1636.

Herranz, E. (2017). Unit Root Tests. *WIRE Computational Statistics*, *9*(3), e2396.http://wires.wiley.com/ wiley CDA.

Hsiao, C. (1979). Autoregressive Modelling of Canadian Money and Income Data. *Journal of American Statistical Association*, *74*, 553-560.

Hsiao, C. (1981). Autoregressive Modelling of Money Income Causality Detection. *Journal of Monetary Economics*, *7*(1), 85-106.

Johansen,S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregression models. *Econometrica*, *59*, 1551-1580.

Johansen,S. (1995). *Likelihood Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford.

Johansen,S., & Juselius, K. (1990). Maximum Likelihood Estimation and Inference on Cointegration – with Application to the Demand for Money. *Oxford Bulletin of Economics and tatistics*, *52*(2), 169 – 210.

Jouini, J., & Boutahar, M. (2005). Evidence on Structural Changes in U.S.Time Series. *Economic Modelling*, *22*, 391 – 422.

Kasa, K. (1992). Common Stochastic Trends in International Stock Markets. *Journal of Monetary Economics*, *29*, 95-124.

Kim , H. (2009). Generalised Impulse Response Analysis : General or Extreme?. *Auburn Economics Working Paper Series, 2012 -2014*, Dept. of Economics, Auburn University. https://mpra.ub.uni-muenchen.de/53951/8/MPRA_paper_53951.pdf.

Kon, S.J. (1984). Models of Stock Returns : A Comparison. *The Journal of Finance*, *39*(1), 147 – 165.

Koop, G., Pesaran, M.H., & Potter , S.M. (1996). Impulse Response Analysis in Non – Linear Multivariate Models. *Journal of Econometrics*, *74*(1), 119 – 147.

Kozhan, R. (2010). *Financial Econometrics*. Roman Kozhan & Ventus Publishing ApS.

Lada, K. (2007). VEC model – The influence of Export on Economic Growth.http://coin.wne.uw.edu.pl/~lgoczek/pdf/macroeconometrics4.pdf.

Ljung, G.M., & Box, G.E.P. (1978). On A Measure of Lack of Fit in Time Series Models. *Biometrika*, *65*(2), 297- 303.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null of stationarity against the alternative of a unit root: How sure are we that the economic time series have a unit root?. *Journal of Econometrics*, *54*, 159-178.

Lütkepohl, H., & Reimers, H.E. (1992). Impulse Response Analysis of Cointegrated Systems. *Journal of Economic Dynamics and Control*, *16*, 53-78.

Lutkephol, H., Saikkonen, P., & Trenkler, C. (2001). Maximum Eigenvalue versus Trace tests for the cointegrating rank of a Var Process. *The Econometric Journal*, *4*(2), 287 – 310.

Lutkepohl, H. (2007). *New Introduction to Multiple Time Series Analysis*. Springer – Verlag.

Mackinnon, J. G. (1991). *Critical Values for Cointegration Tests* (Engle, R. F., & Granger, C. W. J. eds.). *Long Run Economic Relationships*. Oxford University, 267–276.

MacKinnon, R.I. (1973). *Money and Capital in Economic Development*. The Brookings Institute, Washington DC.

Maddala, G.S. & Kim, In – Moo. (2004). *Unit roots, Cointegration and Structural Change*. Cambridge University Press, UK.

Mann, H.B., & Wald, A. (1943). On Stochastic Limit and Order Relationships. *Annals of Mathematical Statistics*, *14*(3), 217 – 226.

Mida, J. (2013). Forecasting Exchange Rates: A VAR Analysis. Institute of Economic Studies, Charles University,Prague.https://is.cuni.cz.

Morin, N.J. (2006). Likelyhood Ratio Tests on Cointegrating Vectors, Disequilibrium Adjustment Vectors, and their Orthogonal Complements. *Finance and Economics Discussion Series*, Division of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C., https://papers.ssrn.com/sol3/papers.cfm?abstract_id=910908.

Nelson, C.R., & Plosser, C.T. (1982). Trends and Random Walks in Macroeconomic Times Series : Some Evidences and Implications. *Journal of Monetary Economics*, *10*, 139 – 162.

Neusser, K. (2016). *Time Series Econometrics*. Springer International Publishing, Switzerland.

Ng, S., & Perron , P. (2001). Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power. *Econometrica*, *69*(6), 1519 – 1554.

Ng,S., & Perron, P. (1995). Unit Root Tests in ARMA models with Data – Dependent Methods for the Selection of the Truncation Lag. *Journal of the American Statistical Association*, *90*(429), 268- 281.

Patterson, K. (2000). *An Introduction to Applied Econometrics*. St. Martin‘s Press, New York.

Pelgrin, F. (2011). Lecture on Box – Jenkens Methodology. *Department of Mathematics*, IMEA.https://math.unice.fr/~frapetti/CorsoP/Chapitre_5_IMEA_1.pdf.

Pesaran, M.H., & Shin, Y. (1998). Generalised Impulse Response Analysis in Linear Multivariate Models. *Economic Letters*, *58*(1), 17- 29.

Pesaran, M.H., & Timmermann, A. (2005). Small Sample Properties of Forecasts from Autoregressive Models under Structural Breaks. *Journal of Econometrics*, *129(*1), 183 – 217.

Petrica, A.C., Stancu, S., & Tindeche, A. (2016), Limitation of ARIMA models in Financial and Monetary Economics. *Theoretical and Applied Economics*, *4*(609), 19 – 42.

Phillips, R. C. B., & P Perron. (1988). Testing for a Unit Root in Time Series Regression. *Biometrika¸75*(2), 335-346.

Phylaktis, K., & Ravazzolo, F. (2003). Stock Market Linkages In Emerging Markets: Implications for International portfolio diversification. *Journal of International Financial Markets Institutions and Money*, *15*(2), 91-106.

Quandt, R. E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes. *Journal of the American Statistical Association*, *53*, 873-880.

Quandt, R.E.(1960). Tests of the Hypothesis that Linear Regression System Obeys Two Separate Regimes. *Journal of the American Statistical Association*, *55*(290), 324-330.

Rachev, S.T., Mittrik, S., Fabozzi, F.J., Fucardi, S.M., & Jasie,T. (2007). *Financial Econometrics: From Basics to Advanced Modelling Techniques*. John Wiley & Sons.Inc. Hoboken, New Jersey.

Ray, H., & Sarkar, J. (2014). Macroeconomic Link to Indian CapitalMarket: A Post-Liberalization Evidence. *Modern Economy*, *5*, 272-288.

Roca, E., & Tularam, A. (2011). Which Way Does Water Flow? An econometric Analysis of the Global Price Integration of Water Stocks. *Applied Economics*, *44*(23), 1-10.

Rubin,H. (1950). *Note on Random Coefficients* (Koopermans,T.C. Ed.). Cowles Commission Monograph 10, 419 – 422, Wiley.

Said, E.S., & Dickey, D.A. (1984). Testing for a Unit Root in Autoregressive Moving Average Models with Unknown Orders. *Biometrika*, *71*(3), 599 – 607.

Sargan, J.D., & Bhargava, A.S. (1983). Testing Residuals from Least Square Regression for Being Generated by the Gaussian Random Walk. *Econometrica*, *51*(1), 153 – 174.

Schlegel, G. (1985). Vector Autoregressive Forecasts of Recession and Recovery:is less more?. *Economic Review*, *26*,1 – 12.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*, 461–64.

Schwert, G.N. (1989). Tests for Unit Roots: A Monte Carlo Investigation. *Journal of Business and Economic Statistics*, *7*, 147 – 160.

Serlatis, A., & King, M. (1997). Common Stochastic Trends and Convergence of European Union Stock Markets. *The Manchester School of Economics & Social Studies*, *65*, 44-57.

Serwa, D., & Wdowinski, P. (2016). Macro – Financial Linkage in the Polish Economy: Combined Impulse – Response Functions in SVAR models. *NBP Working Paper No. 246*, Economic Institute, Warsaw. https://www.nbp.pl/publikacje/materialy_i_studia/246_en.pdf.

Sims , E. (2013). Notes on Time Series. University of Notre Dame, USA.https://www3.nd.edu/~esims1/time_series_notes_sp13.pdf.

Sims, C.A. (1980). Macroeconomics and Reality. *Econometrica*, *48*, 1- 48.

Sjo, B. (2019). Testing for Unit Roots and Cointegration- A Guide. https:// www.researchgate.net/publication/265151092.

Stock, J.M. (1994). *Unit Roots, Structural Break and Trends* (Engle R.F., & McFadden D. Eds.). *Handbook of Econometrics*, *1*(4), Chapter 46, 2739-2841. Elsevier ,North Holland.

Stock, J.H., & Watson, M.W. (1996). Evidence on Structural Instability in Macroeconomic Time Series Relations. *Journal of Business and Economic Statistics*, *14*(1), 11- 30.

Stock, J.H., & Watson, M.W. (2001). Vector Autoregression.https://faculty.washington.edu/ezivot/econ584/stck_watson_var.pdf.

Tong, H. (1990). *Non – Linear Time Series : A Dynamic Systems Approach*. Oxford University Press, Oxford.

Watson, M.W. (1994). *Vector Autoregression and Cointegrtion* (Engle R.F., & McFadden D. Eds.). *Handbook of Econometrics*, *1*(4), Chapter 47,2843-2915. Elsevier, North Holland.

Wongbangpo,P., & Sharma,S.C. (2002). Stock Market and Macroeconomic Fundamental Dynamic Interactions: ASEAN-5 Countries. *Journal of Asian Economics*, *13*, 27-51.

Wong, W.K., Agarwal, A., & Du, J. (2004). Financial Integration for Indian Stock Market: A Fractional Co-Integration Approach. *Finance India*, *18*, 1581-604.