CHAPTER-IV
RESEARCH DESIGN
CHAPTER-IV

RESEARCH DESIGN

4.1 Introduction
The impact of introduction of financial derivatives on the stock market volatility has mainly been studied in the developed economies since long and the empirical research studies on the same line have also been studied in some of the emerging economies recently. After the seminal work of Engle (1982), on the Autoregressive Conditional Heteroscedastic (ARCH) model and its generalized form (GARCH) by Bollerslev (1986), much of empirical works have been done on these models. The ARCH, GARCH and the extensions of GARCH models like Exponential GARCH (EGARCH), Threshold GARCH (TGARCH), GJR-GARCH, GARCH-in-Mean, Component GARCH (CGARCH), Asymmetric Component GARCH (ACGARCH), Power GARCH (PGARCH) and Integrated GARCH (IGARCH) have been used to examine the characteristics of stock market volatility like clustering, asymmetry and persistence after the introduction of financial derivatives.

In the international context, studies on the impact of financial derivatives on stock markets’ volatility were done by Kiymaz Halil and Girard Eric, (2009) [studied on the hypothesis; TGARCH is a manifestation of the daily time dependence in the rate of information arrivals to the market for daily stock returns]; Floros Christos (2008)[ studied on capturing financial time series characteristics by employing GARCH, EGARCH, TGARCH, CGARCH, PGARCH, AGARCH]; Alexakis Panayiotis (2007)[studied the variation in return distributions by ARCH and GARCH models] ; Spyrou Spyros I (2005)[ Studied estimation of volatility and volatility persistence by using ARCH and GARCH] ; Cavallos Laura and Bologna Pierluigi (2002) [studied the effect of future trading on the volatility of the spot market for the Italian stock Exchange using GARCH]; Pilar Corredor and Rafael Santamaria (2002) [Studied impact of derivatives on Spanish stock market using GARCH, EGARCH and GJR GARCH] ;Reyes Mario G (1996)[ studied impact of index futures trading on variation in volatility by EGARCH]; Holmes Phi (1996)[ studied the volatility by using Standard deviation, GARCH and IGARCH]; Franses Philip Hans and Dijk Dick Van (1996) [studied to forecast weekly stock market volatility by using GARCH and QGARCH model]; and many others till recently. In recent years, studies in the similar line were also done in the Indian context by Gahlot Ruchika, Datta Saroj and Kapil Sheeba (2010)[ studied on the impact of futures on stock market volatility by using
GARCH model only; Karmarkar Madhusudan (2007) studied asymmetric Volatility and Risk-return Relationship in the Indian Market by using GARCH and GARCH-M models; Sas Ash Narayan and Omkarnath G (2006) studied derivatives trading and Volatility using ARCH, and GARCH model; Thenmozhi M and Thomas Sony M (2004) used GARCH and IGARCH models to study the Impact of Index Derivatives on S&P CNX Volatility; Kaur Harvinder (2004) studied time varying volatility in the Indian Stock Market using GARCH, EGARCH, TGARCH and many others. However these studies suffered from the following limitations in one way or the other:

1. The periods under study were not adequate.
2. Asymmetric effects in volatility were not considered.
3. Spillover effects were not considered.
4. Effects of market wide factors were not considered.

The present study addresses all the drawbacks of the previous studies and represents a holistic approach to study the impact of the introduction of financial derivatives on stock market volatility in India.

4.2 Objective of the study

The primary objective of the study though is to determine the impact of the financial derivatives such as financial futures and options contracts on the underlying stock market indices of the country i.e. BSE SENSITIVE INDEX comprising 30 shares and NSE NIFTY INDEX comprising 50 shares, the specific objectives considered in the study are stated below:

1. To study the pattern of conditional volatility in Sensex and Nifty during pre-derivatives period, post-derivatives period and whole period of study by using ARCH, GARCH and GARCH family of models.
2. To empirically study the characteristics of Indian stock markets which are represented by the BSE SENSITIVE INDEX and NSE NIFTY INDEX during the period of study by using Jarque-Bera test, Augmented Dickey Fuller test, Phillips-Peron test and Box-Jenkins test.
3. To study the stylized facts of financial time series return data such as autoregressive, clustering, asymmetry and persistence nature of conditional volatilities of BSE SENSEX and NSE NIFTY during Pre-derivative, Post-derivative, and Whole period of the study by
employing ARCH, GARCH, TGARCH/ GJR-GARCH, EGRACH, PGARCH, and IGARCH models.

4. To determine the impact of market wide factors on the conditional volatilities of BSE SENSEX and NSE NIFTY during Pre-derivative, Post-derivative and Whole Period by taking BSE-100 and Nifty Junior as surrogate indices (having no derivatives on them) to act as independent variables in the GARCH model.

5. To study the spillover effect on the conditional volatilities of BSE SENSEX and NSE NIFTY during the Pre-derivative, Post-derivative and Whole period by using GARCH model and taking S&P 500 and Nasdaq Composite Index of the US Stock Market as independent variables over the Pre-derivative, Post-derivative and whole period of study.

4.3 Hypotheses of the Study

The following null hypotheses (H₀) and the alternate hypotheses (Hₐ) are formulated in order to test the followings:

10. H₀ : Time series return data of both Sensex and Nifty do not follow normal distribution with skewness and kurtosis as zero and three respectively.
    Hₐ : Time series return data of both Sensex and Nifty follow normal distribution with skewness and kurtosis as zero and three respectively.

11. H₀: The return distributions of Sensex and Nifty are not stationary.
    Hₐ : The return distributions of Sensex and Nifty are stationary.

12. H₀ : Auto-correlation is not present in the ordinary and squared residuals obtained from the OLS Regression of return series with its lagged for both Sensex and Nifty during the pre and post derivatives period.
    Hₐ: Auto-correlation is present in the ordinary and squared residuals obtained from the OLS Regression of return series with time for both Sensex and Nifty during the pre and post derivatives period.

13. H₀: There is no difference in the conditional volatility of Sensex and Nifty during the pre and post derivatives periods.
    Hₐ: There is difference in the conditional volatility of Sensex and Nifty during the pre and post derivatives periods.

14. H₀: There is no clustering in the volatility of Sensex and Nifty during pre and post derivatives period.
H₀: There is clustering in the volatility of Sensex and Nifty during pre and post derivatives period.

15. H₀: There is no asymmetric or leverage effect on the conditional volatility of Sensex and Nifty.
Hₐ: There is asymmetric or leverage effect on the conditional volatility of Sensex and Nifty.

16. H₀: There is no persistence on the conditional volatility of Sensex and Nifty.
Hₐ: There is persistence on the conditional volatility of Sensex and Nifty.

17. H₀: The impact of market wide factors represented by the presence of surrogate indices like BSE-100 and Nifty Junior on the conditional volatility of Sensex and Nifty respectively is not significant.
Hₐ: The impact of market wide factors represented by the presence of surrogate indices like BSE-100 and Nifty Junior on the conditional volatility of Sensex and Nifty respectively is significant.

18. H₀: there is no spillover effect from US stock markets to Indian stock markets.
Hₐ: there is spillover effect from US stock markets to Indian stock markets.

4.4 Research Methodology

The methodology adopted for the study of the impact of financial derivatives on the stock market volatility of India is described under the following heads:

- Sample
- Data
- Period of Study
- Techniques of Data Analysis

4.4.1 Sample

The present study is undertaken in order to determine the impact of the introduction of financial derivatives on the volatility of Indian stock market. Accordingly, BSE Sensitive index comprising 30 shares of Bombay Stock exchange (BSE Sensex) and S&P CNX Nifty comprising 50 shares of National Stock exchange (NSE Nifty) of the country are taken as the sample indices for study purpose as these two indices represent the stock market of the country by means of their magnitude and dimension in market capitalization/volume of transactions/turnover ratios/listing of the companies etc. The two surrogate indices having no derivatives trading on them
like BSE-100 from BSE and Nifty Junior from NSE are also included in the ambit of volatility analysis in order to know the contribution of derivatives on volatility moderation/stabilization in a relative manner. Apart from these four domestic indices from the Indian Stock Markets, international indices like S&P 500 and Nasdaq Composite Index from the United States (US) stock markets are considered in the study in order to determine the spillover effects, if any, of the events in the US markets on Indian stock markets volatility.

4.4.2 Data
For the conduct of the study in the estimation of descriptive statistics, conditional volatility, volatility clustering, volatility persistence etc, daily closing levels are considered for both the sample indices for impact study under consideration i.e. BSE Sensex and NSE Nifty and also for the surrogate indices i.e. BSE-100 and NSE Nifty Junior along with S&P 500 and Nasdaq Composite Index of the US stock market. The total number of observations in the volatility study of BSE Sensex, BSE-100, S&P 500 and Nasdaq Composite Index are respectively 4473 each over the period 1992-2011. But in case of the volatility of NSE, NSE Nifty, Nifty Junior, S&P 500 and Nasdaq Composite Index, the daily data observations are only 3723 over the period of the study from 1995-2011. The sources of the data are:

1. www.bseindia.com
2. www.nseindia.com
3. inфинанси.yahoo.com

4.4.3 Period of the Study
The study considers the two primary stock markets in India i.e., Bombay Stock Exchange (BSE) and National Stock Exchange (NSE) and their respective indices BSE-30 (SENSEX) Index and S&P CNX Nifty Index, which are well recognized and have significant importance in terms of listing, trading and settlement mechanism on daily basis transparently. The period of study covers a total of 19 years from 2.04.1992 to 31.03.2011 with a total daily closing Index level of 4473 observations of BSE SENSEX and BSE-100 each. For NSE NIFTY and NIFTY Junior, the study period covers 16 years from 1.11.1995 to 31.3.2011 with 3723 daily closing observations each.

Though, the NSE was incorporated in November 1992 as per the recommendation of a high power committee headed by J. Pherwani with the objective of establishing nationwide trading facility for equities, debts and hybrid securities; providing fairness, enabling shorter settlement cycles and meeting international securities market standard, but in fact started functioning
vigorously from 1994-95 in terms of turnover, no. of listed companies and market capitalization. Hence, the study period and no. of daily observations for the above two exchanges are not equal.

The total period of study is further sub-divided into two as pre-derivative period and post-derivative period in order to compare the conditional volatility with and without financial derivatives in Sensex and Nifty indices. Financial derivatives like futures and options have been introduced in these above two indices since June, 2000 and June, 2001 respectively. Initially, the traders and the market participants were not well acquainted with the valuation, trading mechanism and settlement procedures of the financial derivatives. As a result, the benefits of derivatives in volatility moderation could not be experienced for at least one year after derivatives introduction in the market. Hence, the study considers from 02.4.1992 to 31.3.2001 as the pre-derivative period instead of 02.04.1992 to June 2000 and therefore 1.7.2001 to 31.3.2011, as the post-derivative period for BSE Sensex. In case of NSE Nifty Index, the pre-derivative and post derivative periods are taken as 1.11.1995 to 31.3.2001 and 1.4.2001 to 31.3.2011 respectively as per the above logic and rationality.

To examine the impact of market wide factors on the volatility of the BSE Sensex and NSE Nifty, two surrogate indices i.e. BSE-100 from BSE and Nifty Junior from NSE having no derivatives trading on them are included. The daily closing levels are collected for these two surrogate indices in line with the Sensex and Nifty. For BSE-100, closing data are collected for the period 2nd April 1992 to 31st March 2001 and 1st April 2001 to 31st March 2011 for pre and post derivative period respectively. However for Nifty Junior the pre derivatives period starts from 1st November 1995 to 31st March 2001 and 1st April 2001 to 31st March 2011 is the post derivative period.

For the study of spillover effects separately on Sensex and Nifty due to the US stock markets represented by S&P 500 and Nasdaq Composite Index, daily closing levels of S&P 500 and Nasdaq Composite Index are collected for the period from 1st April 1992 to 31st March 2011.

4.4.4 Techniques of data analysis.

The present study uses various appropriate techniques/ models for the calculation/testing of the following:

a. Daily Return calculation,
b. Calculation of Descriptive statistics,
c. Normality test of descriptive statistics,
d. Stationarity test of the Data distribution.
e. Testing of the presence of AR terms, MA terms and ARMA terms in the data.
f. Estimation of conditional volatility,
g. Testing of volatility clustering in conditional volatility.
h. Testing of asymmetry in conditional volatility.
i. Testing of persistence in conditional volatility.

4.4.4 (a) Daily Return Calculation.

The daily closing levels of BSE Sensex, BSE-100, NSE Nifty, NSE Nifty Junior, S&P 500 and Nasdaq composite indices collected from the www.bseindia.com, www.nseindia.com and finance.yahoo.com over the period of study from 1992-2011 in case of BSE Sensex, BSE-100, S&P 500 and Nasdaq Composite index and from 1995-2011 for NSE Nifty, Nifty Junior S&P 500 and Nasdaq composite index covering a total of 4473 and 3723 observations respectively are converted to daily returns of such indices. The present study uses the natural logarithmic difference of closing index levels of two successive days for the calculation of daily rate of return. The logarithmic return is symmetric between up and down movement and is calculated as follows:

\[ y_t = \ln \left( \frac{Y_t}{Y_{t-1}} \right) \]

Where,

- \( \ln \): Natural logarithm with base e.
- \( Y_t \): Closing index level of the day ‘t’
- \( Y_{t-1} \): Closing index level for the day ‘t-1’
- \( y_t \): Daily logarithmic return of the index

The daily returns calculated by using the natural logarithm of two successive days’ closing levels for all the indices under consideration over the period of study constitute the time series data distribution for statistical and econometric modeling in order to determine the impact of financial derivatives on the stock market volatility of India.

4.4.4 (b) Calculation of Descriptive statistics

The descriptive statistics like mean, median, mode, maximum, minimum, standard deviation, skewness and kurtosis of daily log return of all the indices (BSE Sensex, BSE-100, NSE Nifty, NSE Nifty Junior, S&P 500 and Nasdaq Composite Index) are computed for the whole period as well as for pre and post derivative periods as defined under the period of study. Normality test of
the descriptive statistics is carried on by using an asymptotic or large size Jarque-Bera (1981) test in order to know whether the return distribution of the indices considered in the study follow normal or non-normal distribution. The formulae of Jarque-Bera (JB) statistics is stated below:

$$JB\,\text{statistics} = T \left[ \frac{S^2}{6} + \frac{(K-3)^2}{24} \right]$$

Where,

- $T =$ No. of Observations,
- $S =$ skewness Coefficient
- $K =$ Kurtosis coefficient.

For a normal distribution, skewness(S) must be zero and kurtosis (K) be equal to three. Therefore, the JB test of normality is the test of the joint null hypotheses if S and K are 0 and 3, respectively. If S is zero and K is three, the JB statistics will be equal to zero. This test statistic asymptotically follows a chi-square distribution with 2 degree of freedom $[\chi^2 (2)]$. The return distribution of the series will be symmetric and mesokurtic (normally distributed about the mean) under the null hypothesis if p (probability value) of JB statistic is less than the table value of p (probability) at a given significance level.

4.4.4 (c) Stationarity Test of Data Distribution:
The very purpose of the present study is to determine the impact of financial derivatives trading on the daily volatility of Indian Stock Markets during the post derivative period. Accordingly, various econometric models in the form of ARCH, GARCH, EGARCH, TGARCH/GJR GARCH, PGARCH, and IGARCH are applied in order to identify the presence of the type of stylized facts like volatility clustering, leverage effects or asymmetry in the volatility and persistence of volatility during the post derivative period. Daily continuous compounded return over the period of study is obtained by taking the natural logarithm of current and previous index levels of both Sensex and Nifty along with other indices for the purpose of comparison.

Financial time series is said to be stationary if the mean, variance and autocovariance of the time series remain the same. A stationary financial time series tends to return to its mean and fluctuates around this mean. Whereas, a nonstationary time series has time varying mean or time varying variance or both. Non-stationary time series has little practical application from the point of view of forecasting as non-stationary time series study the behavior of only for period of time.
Thus it is not possible to generalize a non-stationary time series to other time period. So for the purpose of forecasting a stationary time series should be taken into consideration.

The present study uses the following tests to test the stationarity

(i) Augmented Dickey Fuller Test
(ii) Phillips-Peron Test

(i) The Augmented Dickey-Fuller (ADF) Test

Dickey and Fuller (1976) have developed procedures for testing the stationarity or non-stationarity of a time series data with the absence or presence of unit root respectively. The first difference ($\Delta y_t = y_t - y_{t-1}$) of the dependent variable ($y_t$) is regressed in three different forms with three different null hypotheses in order to allow the various possibilities under the DF test. These possibilities are:

1. $y_t$ is a random walk: $\Delta y_t = \delta y_{t-1} + u_t$
2. $y_t$ is a random walk with drift: $\Delta y_t = \beta_1 + \delta y_{t-1} + u_t$
3. $y_t$ is a random walk with drift around a stochastic trend: $\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + u_t$

Where,

$y_t$ - log return series of index
$\Delta y_t$ = first difference of $y_t = y_t - y_{t-1}$
$\delta = (p-1)$, and also the first difference operator
$\beta_1$ = Drift parameter
$t$ = time or the trend variable
$\beta_2$ = coefficient of the regressor ‘t’
$u_t$ = white noise error term/random shock/random error i.e. $u_t \sim N(0,1)$

The three null hypotheses under the above three different models of Dickey and Fuller are stated as per the following:

1. null hypothesis $H_0: \delta = 0$ (time series is non-stationary)
   alternate hypothesis $H_a: \delta \neq 0$

for model 1 (a) $\Delta y_t = \delta y_{t-1} + u_t$
If $\delta = 0$, there is unit root and the time series is non-stationary. If $\delta \neq 0$ and $\delta < 0$ (if $\delta > 0$, the underlying time series will be explosive), the null hypothesis will be rejected and the alternative hypothesis will be accepted i.e., $y_t$ is stationary time series with ‘0’ mean.

2. null hypothesis $H_0: \delta = 0$
   alternate hypothesis $H_a: \delta \neq 0$

for model 2 (b) $\Delta y_t = \beta_1 + \delta y_{t-1} + u_t$

If $\delta = 0$, $y_t$ is not stationary, if $\delta \neq 0$ and $\delta < 0$, $y_t$ is stationary with a non-zero mean as equal to $\beta_1/1+\delta$

3. null hypothesis $H_0: \delta = 0$
   alternate hypothesis $H_a: \delta \neq 0$

for model 3: $\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + u_t$

If $\delta = 0$, then $\rho = 1$, as $\delta = \rho - 1$, meaning that $y_t$ is having unit root or the financial time series is not stationary. If $\delta \neq 0$ and $\delta < 0$, $y_t$ is stationary around a deterministic trend. The above three different forms of models for testing the presence or absence of stationarity in the time series have assumed that the error term ‘$u_t$’ is uncorrelated. In conducting Dickey Fuller test, it is assumed that error term $u_t$ is uncorrelated. But in case, $u_t$ are correlated, Dickey and Fuller have developed another test, known as the Augmented Dickey Fuller (ADF) test. The third form of DF test i.e. $\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + u_t$ has been augmented by the presence of an extra term called lagged values of the dependent variable ‘$\Delta y_t$’. Hence the ADF model can be stated as:

$$\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + \sum_{i=1}^{m} \alpha_i \Delta y_{t-i} + \epsilon_t$$

Where,

$y_t =$ log return series of index
$\Delta y_t =$ first difference of $y_t = y_t - y_{t-1}$
$\delta =$ coefficient of $y_{t-1}$
$\beta_1 =$ Drift parameter
$t =$ time or the trend variable
$\beta_2 =$ coefficient of the regressor ‘$t$’
\[ \varepsilon_t = \text{a pure white noise error term} \]

In the above model under ADF, the null and the alternative to null hypothesis are taken as:

- \( H_0: \delta = 0 \) (i.e. return series is non-stationary)
- \( H_1: \delta \neq 0 \) (i.e. return series is stationary)

For the acceptance/ rejection of the null hypothesis i.e., \( \delta = 0 \), tau (\( \tau \)) statistic is used. As the t-value of the estimated coefficient of \( y_{t-1} \) does not follow the t-distribution even in large sample, so Dickey-Fuller computed the critical values of the tau (\( \tau \)) statistics on the basis of Monte Carlo simulations (D.A. Dickey and W.A. Fuller, 1979). If the computed absolute value of tau statistics exceeds the DF tau statistics critical value null hypothesis of \( \delta \) being 0 will be rejected. If the computed absolute value of tau (\( \tau \)) does not exceed the critical value, null hypothesis that \( \delta = 0 \) will be accepted.

\[ \text{Computed value of tau (} \tau \text{)} = \frac{\delta}{\sqrt{\frac{1}{n}}} \]

Where, \( \delta = \text{coefficient of } y_{t-1} \)

\( n = \text{no. of observations which is asymptotic in nature. (Large sample size)} \)

\( \frac{1}{n} = \text{variance} \)

\[ \sqrt{\frac{1}{n}} = \text{standard error of the estimate} \]

ii) Phillips- Perron (PP) tests

The DF test is carried on with the assumption that the error terms \( u_t \) are not correlated. They are identically and independently distributed. The above assumption is dropped and in order to take care of the possible correlation in the error terms, ADF test have been developed. In ADF test, lagged values of the regressand have been used as regressors in addition to the regressors of the DF test to take care of the serial correlation in the dependent variable.

Stationarity of the return distribution of the indices over the period is also tested by employing a test called Phillips -Perron (PP).

Phillips-Perron (1988) proposes an alternative (non-parametric) method of controlling for serial correlation while testing for a unit root. The PP method estimates the non-augmented DF test equation and modifies the t-ratio of the \( \delta \) co-efficient so that serial correlation does not affect the
asymptotic distribution of the test statistics. Nonparametric statistical method is used to take care of the serial correlation in the error terms without adding lagged difference terms. PP test is based on the statistics on:

\[
\text{Modified } t_\delta = t_\delta \left( \frac{\gamma_0}{f_0} \right) + \frac{T(f_0 - \gamma_0)(\text{se}(\delta))}{2f_0^{1/2}s}
\]

Where,

- Modified \( t_\delta \) = t-values of the coefficient (\( \delta \)) of \( y_{t-1} \) of ADF equation, to be estimated under PP test
- \( \delta \) = coefficient of \( y_{t-1} \)
- \( t_\delta \) = t-ratio of \( \delta \),
- \( \text{se}(\delta) \) = coefficient's standard error,
- \( s \) = standard error of the test regression,
- \( \gamma_0 \) = consistent estimate of the error variance
- \( f_0 \) = an estimator of residual spectrum at frequency zero.
- \( T \) = No of observations.

The asymptotic distribution of the PP test is like that of the ADF test statistics if the absolute value of the tau statistic (|\( \tau \)|) exceeds the DF tau statistics critical tau value, the null hypothesis that series is nonstationary will be rejected under PP test and the alternative that time series is stationary will be accepted. On the other hand, if the computed |\( \tau \)| does not exceed the critical tau value, the null hypothesis will not be rejected, in which case the time series is nonstationary.

4.4.4 (d) Testing of Auto-Regressive Moving Average and Integration in the Data.

Auto Regressive Integrated Moving Average (ARIMA) model is used as a new generation forecasting tools developed by Box and Jenkins (1976) and is known as Box-Jenkins methodology. The emphasis of this method is to analyze the probabilistic or stochastic properties of the time series data on their own under the philosophy “let the data speak for themselves”. The ARIMA model allows \( y_t \) to be explained by its past or lagged values of \( y_t \) itself and its stochastic error term. Financial time series data are integrated in nature and therefore, are non-stationary in nature which means the time series have unit roots. If a time series is integrated of order one [i.e. it is I(1)], its first difference is stationary i.e. I(0). Similarly, if a time series is integrated of order two i.e. I(2), its second difference will make the series stationary i.e. I(0), that is
stationary. In general if a time series is I (d), after differencing it d times, then an I(0) series or stationary series is obtained. I (1) and I (2) series can wander a long way from their mean value and cross the mean value, while I(0) series should cross the mean frequently. Hence a time series is to be differenced 'd' (where d may be 1, 2, 3 etc) times to make it stationary. After obtaining stationary time series by means of differencing the time series original data for d times, the next step is to get the AR terms as well as MA terms in the differenced series.

1. Auto-Regressive (AR) Model
An autoregressive model is one where the current value of a variable ‘\( y_t \)’ depends on its previous value at different lags. An autoregressive model of order ‘\( p \)’ denoted as AR (p) can be stated as:

\[
y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i} + u_t
\]

(Chris Brooks, Introductory Econometrics for Finance, 2nd, p-215)

Where,
- \( \alpha_0 \) = constant term
- \( \alpha_1, \alpha_2, \ldots, \alpha_p \) = AR coefficients of the lagged values of \( y_t \) respectively for \( y_{t-1}, y_{t-2}, y_{t-3}, \ldots, y_{t-p} \) (i varies from 1, 2, ..., p)
- \( y_t \) = daily log return series of Sensex/ Nifty
- \( y_{t-1}, y_{t-2}, y_{t-3}, \ldots, y_{t-p} \) = lagged log return series of Sensex/Nifty upto period ‘p’
- \( u_t \) = residual

The above model states that the current value series \( y_t \) is dependent on its previous lagged values of order ‘\( p \)’ provided the autocorrelation coefficients i.e. \( \alpha_i \) are statistically significant.

2. Moving Average (MA) model
The Concept of MA model is developed when the current value of a time series depends on the current and previous values of residuals obtained from the above AR model. This model can be stated below:

\[
y_t = \alpha_0 + \sum_{j=1}^{d} \beta_j u_{t-j} + u_t
\]

(Chris Brooks, Introductory Econometrics for Finance, 2nd, p-211)
This model is the qth order moving average denoted by MA(q) and it indicates a linear combination of residuals, so that $y_t$ depends on the current and previous values of a residuals.

Where,

$y_t = \log \text{return series of Sensex / Nifty at time } 't'$

$a_0 =$ Constant term

$\beta_0, \beta_1, ..., \beta_q =$ MA coefficients for the MA terms or residuals ($j$, varies from 1,2,...q)

$u_t =$ current value of residuals

$u_{t-j} =$ previous values of residuals upto lag q

**Autoregressive Moving Average (ARMA) model**

Auto-Regressive Moving Average (ARMA) model is obtained by combing the AR (p) and MA(q) models. ARMA (p, q) model states that the current values of time series data $y_t$ depends linearly on its own previous values plus a combination of current and previous values of residuals. An ARMA (p, q) model follows the following linear approach:

$$y_t = a_0 + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=1}^{q} \beta_j u_{t-j} + u_t$$

The above equation states that the current value of the series depends linearly on its own previous values upto lags p plus a combination of current and previous values of residual ($u_t$) upto lag q.

Where,

$y_t =$ log return series of Sensex/Nifty at time 't'

$a_0 =$ Constant term

$\alpha_1, \alpha_2, ..., \alpha_p =$ AR coefficients of the lagged values of $y_t$ respectively for $y_{t-1}, y_{t-2}, y_{t-3}, ..., y_{t-p}$ (i varies 1,2,...p)

$\beta_0, \beta_1, ..., \beta_q =$ MA coefficients for the MA terms or residuals ($j$ varies from 1,2,...q)

$u_t =$ current value of residual

$u_{t-1}, u_{t-2}, ..., u_{t-q} =$ lagged values of residuals ‘$u_t$’.

**Box-Jenkins (BJ) Methodology**

Box-Jenkins (1976) methodology is to be employed to study whether the return series of Sensex and Nifty follows a purely AR process or a purely MA process or ARMA process or ARIMA
process. The lag lengths of p, d and q as applicable for respective model are obtained by using BJ methodology. It primarily consists of three following steps:

1. Identification of tentative AR/MA/ ARMA and ARIMA order by visual inspection of Autocorrelation (AC) and Partial Autocorrelation (PAC) of the return series of Sensex and Nifty through Correlogram. Graphically plotting the values of AC and PAC against different lags is known as Correlogram.

2. Estimation involves the followings steps:
   a. Estimation of the statistical significance of the values of the parameters (co-efficients) of the tentative AR/MA/ARMA and ARIMA model.
   b. Estimation of Akaike’s Information Criteria (AIC) and Schwarz’s Bayesian Information Criteria (SBIC).
   c. Estimation of stationarity and Invertibility of AR and MA terms.

3. Diagnostic Checking involves the following steps:
   a) Diagnostic Checking of no autocorrelation in the ordinary residual, obtained from Ordinary Least Square (OLS) regression by specifying appropriate order of AR/MA/ARMA
   b) Diagnostic Checking of autocorrelation in the squared residual, obtained from Ordinary Least Square (OLS) regression by specifying appropriate order of AR/MA/ARMA

Methodology suggested by Box-Jenks follows a repeated process. The above stated steps will be repeated till an appropriate (parsimonious) model is selected. A parsimonious model describes all the features of the data using as few parameters as possible. A graphical presentation of Box-Jenkins Methodology is stated below

```
Identification of tentative order AR/MA/ARMA model by visual inspection of values of autocorrelation (ac) and partial autocorrelation function (pac) of the return series of Sensex and Nifty

Parameter estimation of the chosen model

Diagnostic checking to observe out no autocorrelation in the ordinary residuals and autocorrelation in the squared residuals, obtained from OLS
```
If the diagnostic checking suggests autocorrelation in the squared residuals then the variance of the residuals is not constant and there is a presence of heteroscedasticity. Heteroscedasticity is modeled by using ARCH and GARCH model.

If no autocorrelation is found in the squared residuals then step 1 through 3 would be repeated to find out autocorrelation in the squared residuals.

Each one of the above discussed step under Box-Jenkins methodology is elaborated as followings:

1. **Identification**- Under this step visual inspection of the values of autocorrelation(ac) and partial autocorrelation(pac) of the return series of Sensex and Nifty is done to study the appropriate values of p, d, and q. This step involves in determining the order of the model (AR/MA/ARMA/ARIMA) required to capture the dynamic features of the data by graphically, plotting the values of AC and PAC against different lags. Graphically plotting the values of AC and PAC against different lags is also known as Correlogram.

   The visual inspection of order of AR, MA and ARMA are done on the basis of the following parameters with reference to ACF and PACF test

   1.1 An autoregressive process has:
      - a geometrically decaying acf
      - number of non-zero points of pacf is AR order

   1.2 A moving average process has:
      - number of non-zero points of acf is MA order
      - a geometrically decaying pacf

   1.3 A combination autoregressive moving average process has:
      - A geometrically decaying acf.
      - A geometrically decaying pacf

**Autocorrelation Function (ACF) Test**

If there is significant linear dependence between \( y_t \) and \( y_{t-i} \) is of interest, the concept of correlation is generalized to autocorrelation (AC). The correlation between \( y_t \) and \( y_{t-k} \) is called the lag-k autocorrelation of \( y_t \) and is commonly denoted by \( \rho_k \). The AC at lag k, denoted by \( \rho_k \) is formulated as:
\[ \rho_k = \frac{\text{Cov}(y_t, y_{t-k})}{\sqrt{\text{var}(y_t)} \sqrt{\text{var}(y_{t-k})}} = \frac{\text{Cov}(y_t, y_{t-k})}{\text{var}(y_t)} = \frac{y_k}{\gamma_0} \]

Where the property \( \text{Var}(y_t) = \text{Var}(y_{t-k}) \) for a weakly stationary series is used and from the definition, \( \rho_0 = 1 \), \( \rho_k = \rho_{-k} \), and \(-1 \leq \rho_k \leq 1\).

**Partial Autocorrelation Function (PACF) Test**

Partial Autocorrelation Function denoted by \( \rho_{kk} \), measures the correlation between an observation \( k \) periods ago and the current observations at immediate lags (i.e. all lags < \( k \)) - i.e. the correlation between \( y_t \) and \( y_{t+k} \), after removing the effects of \( y_{t+k+1}, y_{t+k+2}, \ldots y_{t+k} \). For example, the pacf for lag 3 would measure the correlation between \( y_t \) and \( y_{t-3} \) after controlling for the effects of \( y_{t-1} \) and \( y_{t-2} \). At lag 1, the autocorrelation and partial autocorrelation coefficients are equal, since there are no intermediate lag effects to eliminate.

In AR (1) process \( y_t \) and \( y_{t+2} \) are correlated even though \( y_{t+2} \) does not directly appear in the model. The correlation between \( y_t \) and \( y_{t+} \) denoted by \( \rho_2 \) is equal to the correlation between \( y_t \) and \( y_{t+1} \) (\( \rho_1 \)) multiplied by the correlation between \( y_{t+1} \) and \( y_{t+2} \) (\( \rho_2 \)) so that \( \rho_2 = (\rho_1)^2 \). In AR process all such indirect correlations are present in the ACF, whereas the partial autocorrelation controls the effects of such indirect correlations.

Mathematically,

- Partial autocorrelation at lag 1 = Autocorrelation at lag 1 (\( \rho_1 = \rho_1 \))
- PAC at lag 2 (\( \rho_{22} \)) = \((\rho_2 - \rho_1^2)/(1 - \rho_1^2)\)

For computational ease the PAC at more than lag 2 are computed by the EVIEWS -7.1 software package.

The order determination through visual inspection of the values of ACF and PACF may not always be possible. However, a possible model could be identified through the visual inspection of the values of ACF and PACF. At identification stage, a tentative AR/MA/ARMA model could be determined, although it is hard to precisely determine the appropriate order of AR/MA/ARMA (Introductory Econometrics for Finance, Chris Brooks, 2nd Ed, p-235). So, to precisely determine the order of the model the second step suggested by Box-Jenkins is followed.
Estimation

The statistical significance of the value of the parameters are estimated under this step by using an Ordinary Least Square (OLS) regression on the tentative model identified under identification step. Statistically significant values of the parameters suggest the model selected is appropriate (parsimonious).

If the values of the parameter are statistically significant then the Akaike’s Information Criteria (AIC) and Schwraz’s Bayesian Information Criteria (SBIC) are checked in the regression equation. Smallest or even negative AIC and SBIC suggest that the model and lags estimated are appropriate. Appropriateness of the model is also determined with reference to Akaike’s Information Criteria(AIC) and Schwarz’s Information Criteria (SBIC). If the AIC and SBIC of AR/MA/ARMA model is small or negative then the model selected is also said to be a parsimonious model. A parsimonious model ideally should have smallest or even negative AIC and SBIC. A parsimonious model describes all of the features of data of interest using as few parameters as possible. The formula used to calculate the AIC and SBIC

Akaike’s Information Criteria (AIC)

\[ AIC = T \ln(\text{sum of squared residuals}) + 2k \]

Schwarz’s Information Criteria (SBIC)

\[ SBIC = T \ln(\text{sum of squared residuals}) + k \ln(T) \]

Where,

\[ k = \text{number parameter to be estimated (p+q+ possible constant term)} \]

\[ T = \text{total number of observations} \]

\[ \ln = \text{natural logarithm} \]

Lastly, under the estimation stage the inverse of the AR and MA roots of the characteristics equation are also checked to observe to the process implied by the model is stationary and invertible. Invertibility condition is mathematically the same as the stationary condition but refers to Moving Average (MA) rather than AR.

3. Diagnostic Checking

This step involves the process of checking the no autocorrelation in ordinary residuals and autocorrelation in squared residuals. Residuals and squared residuals are obtained from the OLS regression in the estimation stage.

The values of the autocorrelation (ac) and partial autocorrelation (pac) of residuals and squared residuals are plotted to test the statistically significance with reference to the Q-statistics developed by Ljung-Box (1978) and the values of the probability.
If the values of Q-statistics do not exceed the chi-square values at different lags then the null hypothesis of no autocorrelation is not rejected. Similarly if the value of the probability is more than the level of significance then the null hypothesis is not rejected.

If there is no autocorrelation and autocorrelation are found in the ordinary residuals and squared residuals respectively then model selected is considered as an appropriate model, otherwise steps 1 through 3 should be repeated. Thus Box-Jenkins follows an iterative process, in all the lags selected, then the diagnostic checking of squared residuals are done.

The calculation procedure of ACF and PACF test remain same as discussed under identification stage for return series replacing \( u_t \) and \( u_t^2 \) in place of \( y_t \).

**Q-Statistics**

Q-statistics developed by Ljung-Box (LB) (1978) is used in the present to study to test the no autocorrelation in the ordinary residuals and autocorrelation in the squared residuals. If computed Ljung-Box Q-statistics exceeds the Q value from the chi-square distribution at the chosen level of significance, **one can reject the null hypothesis**, of no autocorrelation or accept the alternate hypothesis of autocorrelation at different lags.

**Ljung-Box (LB) Q Statistics**

\[
Q(m) = T(T+2) \sum_{k=1}^{m} \left( \frac{\hat{\rho}_k^2}{k-1} \right) - \chi^2_m
\]

Where,

- \( Q_m = \) Ljung-Box Q statistics
- \( T = \) no of observation
- \( m = \) maximum lag length
- \( k = \) no of lags = 0, 1, 2...m

\( \hat{\rho}_k^2 = \) sample autocorrelation coefficient of lag k
\( \chi^2_m = \) chi-square distribution with m degree of freedom.

Value of the probability given in the correlogram of residuals and squared residuals are also considered for accepting or rejecting the null hypothesis of 'no autocorrelation' at a certain level of significance.
Autocorrelation in the squared residuals suggests that the variance of the squared residuals is not constant and there is heteroscedasticity. Presence of heteroscedasticity violates the assumption of Classical Linear Regression Model (CLRM). CLRM assumes that the variance of the residuals is constant and hence presence of heteroscedasticity is a problem. Rather than removing the heteroscedasticity, it is modeled under financial time series as ARCH model developed by Engle (1982).

4.4.4 (e) Volatility Modeling Building

The volatility model for the return series consists the following four steps:

1. AR/MA/ARMA model is to test the autocorrelation in the squared residuals by employing Box-Jenkins methodology as stated above.
2. Squared Residuals from the AR/MA/ARMA model constructed in step 1 are used to test the ARCH effects.
3. Volatility model(s) is/are to be specified if ARCH effects are statistically significant and the AR/MA/ARMA model and volatility equation are jointly to be specified.
4. Fitted model is checked and to be refined, if necessary.

4.4.4 (f) Study of ARCH Effect in the Data

The residuals series (ut) arising from the AR/MA/ARMA model should be checked for serial correlation by employing Box-Jenkins methodology. If the residual series are serially uncorrelated, then the squared series of residuals (u^2_t) is to be checked for autocorrelation. Autocorrelation in the squared residuals suggest that the variance of the squared residuals is not constant or heteroscedasticity. Heteroscedasticity is a problem under Classical Linear Regression Model (CLRM) and rather removing it, heteroscedasticity is modeled to observe ARCH (Autoregressive conditional heteroscedasticity) effects in the financial time series.

The ARCH effect in the data is to be tested by applying the followings:

a. Ljung-Box statistics Q (m) test is applied to the Squared residuals series and also by referring the p-value (Probability value) with the null hypothesis is that the first m lags of ACF of the u^2_t series are zero.

b. Lagrange Multiplier (LM) test of Engle (1982) is applied which the test is equivalent to the usual F statistic for testing \( \alpha_i = 0 \) (i=1,......,q) in the following linear regression of u^2_t

121
The generalized form

\[ u_t = \alpha_0 + \alpha_1 u_{t-1} + \ldots + \alpha_q u_{t-q} + \varepsilon_t. \]

Or the generalized form

\[ u_t^2 = \alpha_0 + (\sum_{i=1}^{q} \alpha_i u_{t-i}^2) + \varepsilon_t. \]

Where,

- \( \varepsilon_t \) = error term arising in the lagged values of the residuals \( (u_t) \),
- \( q \) = the lag length,
- \( T \) = no. of observations
- \( u_{t}^2 \) = squared residual at time ‘t’
- \( u_{t-1}^2 \) = lagged squared residuals

If there are no ARCH or GARCH effects, the estimated values of \( \alpha_i \) through \( \alpha_q \) should be zero. Hence the regression will have little explanatory power so that the coefficient of determination (i.e. the usual R²) will be quite low. Using a sample of T residuals, under the hypothesis of no ARCH effects, the test statistics TR² converge to \( \chi^2 \) distribution with \( q \) degree of freedom.

If \( TR^2 \) is sufficiently large, null hypothesis of no ARCH effects is rejected. For small Sample F-test is used to reject or accept the null hypothesis by comparing the computed F value with the F-table with \( q \) degree of freedom.

The null hypothesis of ‘No ARCH’ effect can also be accepted or rejected by ‘p-value’ approach or Probability value approach under hypothesis testing. Under p-value approach of testing hypothesis, the null hypothesis is accepted or rejected on the basis following:

1. If the p-value is greater than or equal to the observed level of significance, null hypothesis cannot be rejected
2. If p-value is less than observed level of significance, the null hypothesis can be rejected

4.4.4(g) Conditional Volatility Estimation by ARCH and GARCH models

In this study GARCH framework is used in order to examine any possible effect of derivative trading on the volatility of the Sensex and Nifty and the impact of financial derivatives on the volatility of Sensex and Nifty is studied for the post derivative period in comparison to the pre-
derivative periods. To examine whether the volatility has increased/decreased in period under consideration, ARCH and GARCH models are employed.

The GARCH model has been developed by Bollerslev (1986) from the ARCH model previously introduced by Engle (1982). In the ARCH model, the conditional variance at time \( t \) depends on the size of the squared error at \( t-1 \), thus allowing the conditional variance to change over time.

**ARCH Model**

The objective for which ARCH model is to be employed in the present study is:

i. To model the conditional variance of residuals (errors) to find out the presence of clustering effect in the data.

ARCH model is non-linear model and does not follow the assumption of constant variance of the residuals as it is taken under the Classical Linear Regression Model (CLRM). However, if the residuals variance at different point of time are not constant, this would be known as heteroscedastic \([\text{var}(u_t) = \sigma_t^2]\). In case of financial time series the variance of the residuals does not remain constant over time and hence makes sense to consider a model that does not assume constant variance. For studying the financial time series where the variance of errors is not constant over time, Autoregressive Conditional Heteroscedastic (ARCH) model is proposed by Engle (1982). Another important feature of many series of financial asset returns that provides a motivation for the ARCH class models, is known as ‘volatility clustering’ or ‘volatility pooling’.

Volatility clustering describes the tendency of large changes in asset prices (of either sign) to follow large changes or small changes (of either sign) to follow small changes. When many large positive and large negative returns are observed during a period of time, it could be stated that ‘volatility is autocorrelated’.

How could this phenomena, which is common to many series of financial asset returns, be parameterized (modeled). One approach is to use an ARCH model and to understand how the model works, a definition of the conditional variance of a random variable, \( u_t \), is required. The distinction between the conditional and unconditional variance of a random variable is exactly the same as that of the conditional mean and unconditional mean. The conditional variance of \( u_t \) may be denoted \( \sigma_t^2 \), which is written as,

\[
\sigma_t^2 = \text{var}(u_t | u_{t-1}, u_{t-2}, \ldots) = E[(u_t - E(u_t))^2 | u_{t-1}, u_{t-2 }, \ldots]
\]

And usually it assumed that \( E(u_t) = 0 \), so
The above equation states that the conditional variance of zero mean normally distributed random variable \( u_t \) is equal to the conditional expected value of the square of \( u_t \) and under ARCH model, the 'autocorrelation in volatility' is modeled by allowing the conditional variance of the error term \( \sigma^2_t \) to depend on the immediately previous value of the squared residuals. Hence

\[
\sigma^2_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_q u_{t-q}^2 .... \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdot
Generalized ARCH (GARCH) model

The following are the objectives for which GARCH model is employed in the present study:

i. To study the impact of introduction of derivatives on the volatility of Sensex and Nifty during the post derivative period.

ii. To study the phenomenon like "volatility clustering" for both the indices under consideration.

iii. To study the impact of market wide factors on the volatility of Sensex and Nifty.

iv. To study the impact worldwide factors on the Indian stock markets represented by BSE Sensex and NSE Nifty.

Bollerslev (1986) extended Engle's original work by developing a technique that allows the conditional variance to be an ARMA process. One of the most appealing features of the GARCH framework, which makes this model so widely used in financial literature, is that it captures one of the well known empirical regularities of asset returns, the volatility clustering. GARCH model is more parsimonious model in comparison to ARCH model and avoids overfitting. The GARCH specification is as follows:

\[
\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta \sigma_{t-1}^2 \quad \text{GARCH (1,1)}
\]

\[
\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \quad \text{GARCH (p, q)}
\]

The important point in GARCH model is that, the conditional variance of \( u_t \) is given by \( E_{t-1} u_t^2 = \sigma_t^2 \), thus the conditional variance of \( u_t \) is the ARMA process given by the expression \( \sigma_t^2 \) in GARCH.

In GARCH (1,1) and GARCH (q, p),

- \( \alpha_0, \alpha_1 \) and \( \beta_1 \) are the coefficients of the regression
- \( \alpha_0 \) is the measure of long term constant volatility i.e. unconditional variance estimation.
- \( \alpha_1 + \beta_1 \) represents persistence- tendency of an index being affected by the previous days’ volatility
- \( \alpha_1 \) is the coefficient of the squared error term of the previous day, describes volatility due to one day old news.
\( \beta_1 \) is the coefficient of the lagged conditional variance, describes the volatility due to news which are old by more than one day.

GARCH (1, 1) model can be extended to a GARCH (q, p) formulation, where the current conditional variance is parameterized to depend upon q lags of the squared error and p lags of the conditional variance.

**Unconditional variance under GARCH**

The conditional variance is changing, but the unconditional variance of \( u_t \) is constant and given by

\[
\text{Var}(u_t) = \frac{\alpha_0}{1 - (\alpha_1 + \beta_1)}
\]

so long as \( \alpha_1 + \beta_1 < 1 \). For \( \alpha_1 + \beta_1 \geq 1 \) the unconditional variance of \( u_t \) is not defined, and this would be termed ‘non-stationary’ in variance. \( \alpha_1 + \beta_1 \) usually observed very close to zero, which signifies that the volatility of asset returns is highly persistent. The effect of any shock in volatility dies out at a rate of \((1 - \alpha_1 - \beta_1)\) and if \((\alpha_1 + \beta_1) \geq 1\) the effect of shock will never die out. The volatility will be defined only if \((\alpha_1 + \beta_1) < 1\). Therefore this condition is imposed while estimating the GARCH model.

GARCH model considers both ARCH parameter and GARCH parameter. A large ARCH parameter and a smaller GARCH parameter would imply that volatility in the return series is highly responsive to recent ‘news’ or ‘shock’ and therefore the volatility pattern could be ‘spiky’ (recent information is more important than old ones) and less ‘persistence’ (information decay is very fast). A smaller ARCH parameter and Large GARCH parameter imply, just the opposite. But GARCH model enforces symmetric response of positive and negative shocks.

**Dummy Variable in GARCH(1,1) model**

To study whether the introduction of derivatives altered the volatility of the Sensex and Nifty, a dummy variable is to be introduced in the conditional variance equation of GARCH model that measure volatility. Dummy variable, \( D \), which will take the value zero (0) for pre- derivative period and one (1) for post introduction of derivatives and thus the conditional variance equation becomes:

\[
\sigma^2_t = \alpha_0 + \alpha_1 u^2_{t-1} + \beta \sigma^2_{t-1} + \gamma D
\]

126
The sign of the coefficient of the Dummy variable is important. A negative (positive) value implies fall (rise) in the spot market volatility with the introduction of derivatives. If $\gamma$, the coefficient of the dummy variable, is statistically significant, then it can be said that existence of derivative trading has had the impact on spot market volatility. $\gamma$ indicates the direction of changes in the spot market volatility. If the coefficient is negative, it can be said that the volatility has been reduced post introduction of derivatives and vice versa if the coefficient is positive.

4.4.4(h) Asymmetric GARCH models
One of the primary restrictions of GARCH models is that they enforce a symmetric response of volatility to positive and negative shocks. This arises since conditional variance in GARCH model is a function of the magnitudes of the lagged residuals and not their sign (as squaring the lagged error the sign is lost). However, it has been argued that a negative shock to financial time series is likely to cause volatility to rise by more than a positive shock of the same magnitude. In the case of equity returns, such asymmetries are typically attributed to leverage effects.

To study the leverage effect in the Indian stock markets the asymmetric models are to be employed in this study. The asymmetric models which are included in this study are-

Threshold GARCH (TGARCH)/ GJR-GARCH
TGARCH model is to be employed in the study to study the asymmetric nature of the various types of shocks in the volatility of Sensex and Nifty during the post derivatives period. An interesting feature of asset prices is that “bad” news seems to have a more pronounced effect on volatility than “good” news. The tendency for volatility to decline when returns rise and to rise when returns fall is often called the Leverage effect. Glosten, Jaganathan, Runkle (1994) showed how to allow the effects of good and bad news to have different effects on volatility. In a sense, $u_{t-1} = 0$ is a threshold such that shocks greater than the threshold have different effects than shocks below the threshold and to address this in the volatility of Sensex and Nifty TGARCH model is identified.

TGARCH (1,1)

$$\sigma_i^2 = \alpha_0 + \alpha_i u_{i-1}^2 + \beta \sigma_{i-1}^2 + \gamma u_{i-1}^2 I_{i-1}$$

TGARCH (q, p)

$$\sigma_i^2 = \alpha_0 + \sum_{i=1}^{q} (\alpha_i + \gamma I_{i-1}) u_{i-1}^2 + \sum_{j=1}^{p} \beta_j \sigma_{i-j}^2$$

127
Where, \( I_{t-1} \) is an indicator for negative \( u_{t-1} \), that is,

\[
I_{t-1} = \begin{cases} 
1 & \text{if } u_{t-1} < 0 \\
0 & \text{if } u_{t-1} \geq 0
\end{cases}
\]

\( \alpha, \beta, \) & \( \gamma \) are non-negative parameters satisfying conditions similar to those of GARCH models.

Form the above model, it is seen that a positive \( u_{t-1} \) contributes \( \alpha_i u_{t-1}^2 \) to \( \sigma_t^2 \), whereas a negative \( u_{t-1} \) has a larger impact \((\alpha_i + \gamma)u_{t-1}^2\) with \( \gamma > 0 \).

In this model, good news \((u_t > 0)\) and bad news \((u_t < 0)\) have differential effects on the conditional variance. Good news has an impact of \( \alpha \), while bad news has an impact of \( \alpha + \gamma \). \( \gamma > 0 \) then the leverage effect exists and bad news increases volatility, while if \( \gamma \neq 0 \) the news impact is asymmetric. For leverage effect, we would see \( \gamma > 0 \) and the condition for non-negative will be \( \alpha_0 > 0, \alpha_i > 0, \beta \geq 0 \) and \( \alpha_i + \gamma \geq 0 \). Model is still admissible, even if \( \gamma < 0 \), provided that \( \alpha_i + \gamma \geq 0 \).

**Exponential GARCH (EGARCH) Model**

To capture the asymmetric effect of news another model was proposed by Nelson (1991), known as Exponential GARCH (EGARCH), which do not require non negativity constraints.

The model is stated below

\[
\ln(\sigma_t^2) = \alpha_0 + \alpha_1 \left[ \frac{|u_{t-1}|}{\sqrt{\sigma_{t-1}^2}} - \sqrt{\frac{2}{\pi}} \right] + \gamma \frac{u_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \beta \ln(\sigma_{t-1}^2) \quad \text{EGARCH (1, 1)}
\]

\[
\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{q} \alpha_i \frac{|u_{t-1}|}{\sqrt{\sigma_{t-1}^2}} + \gamma \frac{u_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \sum_{j=1}^{p} \beta_j \ln(\sigma_{t-j}^2) \quad \text{EGARCH (q, p)}
\]

The above two models are the EGARCH (1,1) and EGARCH (p, q) and the models have three interesting properties:

1. The equation for the conditional variance is in log-linear form. Regardless of magnitude of \( \ln(\sigma_t^2) \), the implied value of \( \sigma_t^2 \) can never be negative. Hence, it is permissible for the coefficients to be negative.

2. Instead of using the value of \( u_{t-1}^2 \), EGARCH model uses the level of standardized value of \( u_{t-1} \) [i.e. \( \frac{u_{t-1}}{\sqrt{\sigma_{t-1}^2}} \)]. Nelson argues that this standardization allows for a more natural interpretation of the size and persistence of shocks. After all, the standardized value of \( u_{t-1} \) is a unit-free measure.
3. The EGARCH model allows for leverage effects. If $\frac{u_{t-1}}{\sqrt{\sigma^2_{t-1}}}$ is positive, the effect of the shock on the log of the conditional variance is $\alpha_1 + \gamma_1$. If $\frac{u_{t-1}}{\sqrt{\sigma^2_{t-1}}}$ is negative, the effect of the shock on the log of the conditional variance is $-\alpha_1 + \gamma_1$.

The logarithm form of the conditional variance implies that the leverage effect is exponential (so the variance is non-negative). The presence of leverage effects can be tested by the hypothesis that $\gamma < 0$. If $\gamma \neq 0$, then the impact is asymmetric. The model has several advantages over the pure GARCH specification since the log ($\sigma^2_t$) in the equation is modeled, then even if the parameter are negative, $\sigma^2_t$ will be positive. There is thus no need to artificially impose non-negativity constraints on the model parameters. Secondly, asymmetries are allowed for under the EGARCH formulation, since if the relationship between volatility and returns is negative, $\gamma$, will be negative.

**Power GARCH (PGARCH)**

PGARCH model is to be employed in this study with the objective to study the asymmetric effects on volatility of Sensex and Nifty. The PGARCH model nests the GARCH model of Bollerslev (1986), which features a conditional variance equation, as well as the model of Taylor (1986), which features a conditional standard deviation equation. PGARCH model also nests familiar asymmetric volatility models such as TGARCH and EGARCH. Hence parameter restrictions on PGARCH model produce a number of nested GARCH models, symmetric and asymmetric including the well-known GARCH, TGARCH and EGARCH models.

PGARCH is proposed by Ding et al. (1993). Standard deviation GARCH model is introduced, where standard deviation is modeled rather than the variance. In Power GARCH model, the power parameter $d$ of the standard deviation can be estimated rather than imposed and the optional $\gamma$ parameters are added to capture asymmetry.

$$
\sigma^d_t = \alpha_0 + \alpha_1(|u_{t-1}|-\gamma u_{t-1})^d + \beta_1 \sigma^d_{t-1} \text{ .............PGARCH(1,1)}
$$

$$
\sigma^d_t = \alpha_0 + \sum_{i=1}^{\infty} \alpha_i(|u_{t-i}|-\gamma u_{t-i})^d + \sum_{j=1}^{\infty} \beta_j \sigma^d_{t-j} \text{ .............PGARCH(q,p)}
$$

Where $u_t$ is the error term, $\gamma$ is the asymmetric term and $d$ is power transformation that can be estimated from the data or restricted to take values such as 1 or 2.
4.4.4 (i) Integrated GARCH (IGARCH) Model

Many financial time series data possess the characteristics of persistence i.e. the states of continuing to exist for a long period of time without interruption. For example, one shock in one instant of time may influence the variance exponentially over a long period time. IGARCH model proposed by Engle and Bollerslev (1986) takes care of the shock on the conditional volatility which does not die out asymptotically. IGARCH model is like GARCH model, but ARCH and GARCH coefficients do not obey the stationary conditions.

The volatility is said to be persistence if
\[ \sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j = 1, \]

In above model, there would be "a unit root in variance" and is termed as 'Integrated GARCH or IGARCH'. Thus, IGARCH models are unit-root GARCH models, similar to ARIMA models. The volatility is said to be persistence if the summation of \( \alpha_1 \) and \( \beta_1 \) equal to one (i.e. \( \alpha_1 + \beta_1 = 1 \)).

Persistence in the conditional volatility is tested through IGARCH model. Nelson (1990) argued that constraining \( \alpha_1 + \beta_1 \) to equal unity can yield a very parsimonious representation of the distribution of an asset's return. In some respects, this constraint forces the conditional variance to act like a process with a unit root, known as Integrated GARCH (IGARCH).

The relative magnitude and significance of ARCH and GARCH term corresponds to the dynamic of the volatility patterns in the return series. A smaller GARCH and larger ARCH term terms imply that the return series is highly responsive to the most recent news, whereas larger GARCH and smaller ARCH term imply that the return series is highly responsive to the old news. GARCH model enforces symmetric response of positive and negative shocks. It has been found that volatility is more when the market is under bearish trend than to the bullish trend.

4.4.4 (j) Modeling for Surrogate Indices

BSE-100 and Nifty Junior are the two surrogate indices of BSE and NSE respectively and on these indices derivatives trading is not allowed. So the volatility in these two indices is due to the influence of only market-wide factors and not due to derivatives trading. Hence it is important to examine that whether Sensex and Nifty are also influenced by the market factors or only by derivatives trading as derivatives trading is allowed on these two indices. To examine the impact of market-wide factors on Sensex and Nifty, the returns of BSE-100 and Nifty Junior are taken as proxy variables representing the market factors and the following regression model is

...
constructed considering BSE-100 and Nifty Junior as exogenous variables and Sensex and Nifty as the endogenous variables.

To study the impact of BSE-100 on BSE Sensex and Nifty Junior on Nifty the following techniques and methodologies are adopted:

a. The following regression equation is formed by considering the time series return data of BSE-100 and Nifty Junior as the independent variable and the time series return data of Sensex and Nifty as dependent variable:

\[ y_t = \lambda_0 + \lambda_1 y_{st} + u_t \]

where,

- \( y_t \) = daily log return series of Sensex/Nifty
- \( y_{st} \) = daily log return of surrogate indices i.e. BSE-100/Nifty Junior
- \( \lambda_0 \) = constant
- \( \lambda_1 \) = coefficient of the \( y_{st} \)
- \( u_t \) = residuals/error term

b. The coefficient of \( y_{st} \) i.e. \( \lambda_1 \) and its significance is important to study the impact of BSE-100 and Nifty Junior on the BSE Sensex and NSE Nifty. If the coefficient (\( \lambda_1 \)) is high and statistically significant than there is an impact of market factors on the BSE Sensex and NSE Nifty otherwise not.

c. Further to study the impact of market factors on the conditional variance of BSE Sensex and NSE Nifty the variance of the error term ‘\( u_t \)’ is modeled, as the error term is expected to follow normal distribution with mean zero and variance as \( \sigma_t^2 \), \( N \sim (0, \sigma_t^2) \) where \( \sigma_t^2 \) is the conditional variance.

d. The conditional variance thus can be modeled by the application of GARCH (1,1).

4.4.4 (k) Modeling of World-wide factor (Spillover effect)

In order to isolate the unique impact of the introduction of financial derivatives on the volatility of Sensex and Nifty in the post derivative period, the predictability associated with world returns is removed by alternatively including two US stock indices in the mean equation of the GARCH model. S&P 500 and Nasdaq Composite Index of US are alternatively regressed in the mean equation of the GARCH model with the return series of Sensex and Nifty respectively. The following equation is formed to control the impact of world wide factor in the mean equation of the GARCH model for Sensex and Nifty.
To study the spillover effect from the US stock markets represented by S&P 500 and NASDAQ Composite Index on Indian stock markets represented by BSE Sensex and NSE Nifty the following techniques and methodology adopted:

a. The following regression equation is formed by considering the one day lagged time series return data of S&P 500 and NASDAQ Composite Index as independent variable and the time series return data of Sensex and Nifty as dependent variable:

\[ y_t = \lambda_0 + \lambda_1 y_{w,t-1} + u_t \]

Where,

- \(y_t\) = log return series of Sensex/ Nifty
- \(y_{w,t-1}\) = one day lagged log return series of world indices represented by S&P500/ Nasdaq Composite Index
- \(\lambda_0\) = constant
- \(\lambda_1\) = coefficient of the \(y_{w,t-1}\)
- \(u_t\) = error term

In the equation one day lagged return for S&P 500 and NASDAQ Composite index are considered due to the difference in trading timings between the US stock market and Indian stock market.

b. The coefficient of \(y_{w,t-1}\) i.e. \(\lambda_1\) and its significance is important to study the impact of spillover effect from S&P 500 and NASDAQ Composite Index to BSE Sensex and NSE Nifty. If the coefficient \((\lambda_1)\) is high and statistically significant than there is an impact of market factors on the BSE Sensex and NSE Nifty otherwise not.

c. Further to study the impact of world factors on the conditional variance of BSE Sensex and NSE Nifty the variance of the error term \(u_t\) is modeled, as the error term is expected to follow normal distribution with mean zero and variance as \(\sigma_t^2\), \(N \sim (0, \sigma_t^2)\) where \(\sigma_t^2\) is the conditional variance.

d. The conditional variance thus can be modeled by the application of GARCH (1,1).

4.5 Conclusion

Research on financial time series depends on the econometric model for analysis, interpretation and forecasting. Econometrics models provide both simple as well as complicated models with great derivation. Few econometrics models are also included in the present studies with appropriate simplification with reference to the objective of the study.


135


**Book Reference**