CHAPTER 3

RESEARCH METHODOLOGY

This chapter discusses need and research objectives of the study. The focus of this chapter is to highlight the data collection sources. This is followed by discussing the econometrics tools used to analyze the data.

3.1 Need of the Study

The active participation of retail investors in the Indian stock market has been increased from last five years. According to Central Depository Service India Limited Annual Report 2017-18, there is increase in active participation of retail investors, represented with the increase in turnover of shares traded in NSE rising to 140 lakh crore in the financial year 2017-18 from 28 lakh crore in the financial year 2013-14. Now the investors are also interested to make investment in the commodity market. NCDEX data has shown that there is 23% increase in the retail investors investing in the Indian commodity market (Singh, 2011). There is increase in retail investors in the NCDEX, represented in the form of volume traded increased to 2,17,736 thousand tones in the financial year 2016-17 from 1,94,255 thousand tones in the financial year 2015-16. The volume traded in the MCX has been increased from 89,331 thousand tones in the year 2015-16 to 93,078 thousand tones in the financial year 2016-17 (SEBI Annual Report, 2016-17). The investors use commodities for the purpose of risk management as these are less volatile as compare to stock market. Despite the improvements in the performance of Indian commodity market and stock market, the active participation of percentage of retail investors of total population is still very less. SEBI investor survey, 2015 stated that out of total population of India, 1.9 crore investors are investing in stock market and there are only 21 lakh investors in commodity futures. The less participation in these markets is due to the less knowledge of investors to carry out fundamental analysis of financial assets; they tend to go with general market flow and direction. This study will be helpful for these investors to get more information about the financial markets.

This study examines the co-integration between commodity market and stock market to provide better insights regarding the hedging effectiveness of commodities against
the unexpected fluctuations in the stock market. Furthermore the linkage between prices of raw material and their related stock indices will provide relevant information about the optimal substitution tactics between commodities and stocks (Creti et al., 2013). This study will help the policy makers to increase the participation of investors in commodity market and stock market with the help of optimal weights and hedge ratios, calculated on the basis of results of this study. Investors can use these weights and ratios to hedge their risk effectively. This way they will be better equipped to anticipate and prepare for unexpected fluctuations in commodity and stock prices.

3.2 Objectives of the Study
1. To examine the seasonality in mean return and volatility for the Commodity Market and Stock Market
2. To examine the long run co-integration between the Commodity Market and Stock Market
3. To examine the causal relationship between the Commodity Market and Stock Market
4. To examine the return links and volatility transmission between the Commodity Market and Stock Market
5. To examine the dynamic correlation between the Commodity Market and Stock Market

3.3 Research Design and Methodology
Research design is a framework which provides direction to conduct investigation effectively and efficiently

3.3.1 Data Collection
The spot price data related to individual commodities has been collected from the official website of Multi Commodity Exchange (MCX) and National Commodities and Derivatives Exchange (NCDEX), India from the year 2007 to 2017. Trading on NCDEX mainly concentrated on the agricultural commodities and MCX focuses on non-agricultural commodities. The total number of commodities and their related stock indices selected for this study are given in table 3.1. The total number of commodities traded on NCDEX is 23 out of which 16 agricultural commodities have
been selected in this study. The other agricultural commodities which are chana, coriander, castor seed, cotton, kapas, maize and sugar have not been taken into account due to non availability of data.

Table 3.1: Commodities and Related Stock Indices

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>COMMODITY NAME</th>
<th>STOCK INDEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Aluminum</td>
<td>NSE Metal Index</td>
</tr>
<tr>
<td>2</td>
<td>Copper</td>
<td>NSE Metal Index</td>
</tr>
<tr>
<td>3</td>
<td>Lead</td>
<td>NSE Metal Index</td>
</tr>
<tr>
<td>4</td>
<td>Nickel</td>
<td>NSE Metal Index</td>
</tr>
<tr>
<td>5</td>
<td>Zinc</td>
<td>NSE Metal Index</td>
</tr>
<tr>
<td>6</td>
<td>Gold</td>
<td>NIFTY Index</td>
</tr>
<tr>
<td>7</td>
<td>Silver</td>
<td>NIFTY Index</td>
</tr>
<tr>
<td>8</td>
<td>Crude oil</td>
<td>NSE Energy Index</td>
</tr>
<tr>
<td>9</td>
<td>Natural gas</td>
<td>NSE Energy Index</td>
</tr>
<tr>
<td>10</td>
<td>Barley</td>
<td>NSE FMCG Index</td>
</tr>
<tr>
<td>11</td>
<td>Cotton Seed Oil Cake</td>
<td>NSE FMCG Index</td>
</tr>
<tr>
<td>12</td>
<td>Crude Palm Oil</td>
<td>NSE FMCG Index</td>
</tr>
<tr>
<td>13</td>
<td>Guar gum</td>
<td>NSE FMCG Index</td>
</tr>
<tr>
<td>14</td>
<td>Guar Seed</td>
<td>NSE FMCG Index</td>
</tr>
<tr>
<td>15</td>
<td>Gur</td>
<td>NSE FMCG Index</td>
</tr>
<tr>
<td>16</td>
<td>Jeera</td>
<td>NSE FMCG Index</td>
</tr>
<tr>
<td>17</td>
<td>Mustard seed</td>
<td>NSE FMCG Index</td>
</tr>
<tr>
<td>18</td>
<td>Pepper</td>
<td>NSE FMCG Index</td>
</tr>
<tr>
<td>19</td>
<td>RGB Palm oil</td>
<td>NSE FMCG Index</td>
</tr>
<tr>
<td>20</td>
<td>Soya Oil</td>
<td>NSE FMCG Index</td>
</tr>
<tr>
<td>21</td>
<td>Rubber</td>
<td>NSE FMCG Index</td>
</tr>
<tr>
<td>22</td>
<td>Soy bean</td>
<td>NSE FMCG Index</td>
</tr>
<tr>
<td>23</td>
<td>Turmeric</td>
<td>NSE FMCG Index</td>
</tr>
<tr>
<td>24</td>
<td>Wheat</td>
<td>NSE FMCG Index</td>
</tr>
<tr>
<td>25</td>
<td>Yellow Peas</td>
<td>NSE FMCG Index</td>
</tr>
</tbody>
</table>

Source: Official Websites of NSE, MCX and NCDEX

The total number of non-agricultural commodities traded on MCX is 9. The data related to these 9 non-agricultural commodities, covering three sectors which are Precious Metals, Base Metals and Energy, is collected from the official website of
MCX. The data related to selected stock indices has been collected from the official website of National Stock Exchange of India (NSE) from the year 2007 to 2017. The selected stock indices are NIFTY FIFTY, NIFTY Energy, NIFTY Metals and NIFTY FMCG.

3.3.2 Data Analysis Tools and Techniques

**Generalized Auto-Regressive Conditional Heteroskedasticity (GARCH)**

GARCH model is used widely for modeling volatility and seasonality in the financial markets. In this study, GARCH model is applied to model seasonality in commodity market and stock market.

Before applying the GARCH model, it is required to estimate ARCH-LM (Lagrange Mutiple) test to study the presence of ARCH effect in the residuals ($\varepsilon_t$). Then the residuals are squared and regressed on their own lagged return of order one to four.

The estimated equation is given below:

$$\varepsilon_t^2 = c + \sum_{i=1}^{4} \alpha_i \varepsilon_{t-i}^2 + k_t$$  \hspace{1cm} (1)

Here $k_t$ is the error term. The null hypothesis for ARCH-LM test is the absence of ARCH effect in the error term. If the coefficient of ARCH-LM test is statistically significant, it confirms the presence of ARCH effect in the error term.

Engle (1982) introduced the concept of modeling volatility in the financial markets by introducing Autoregressive conditional heteroskedasticity model. This model states that the forecasted conditional variance of the mean equation is changed with the change in previous period’s squared error term. The error terms should be serially uncorrelated. The generalized version of ARCH model is known as GARCH model. Further Bollerslev (1986) extended the ARCH model based on the assumption that conditional variance not only depends upon the past error term, but it is also affected by its own past lagged variance. The mean equation of GARCH model is given below:

$$r_t = c + \varepsilon_t$$  \hspace{1cm} (2)
The variance equation is:

\[ h_t = \mu + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \]  \hspace{1cm} (3)

Here \( c \) is the constant term, \( \varepsilon_t \) is the error term at time \( t \), the terms \( \alpha \) and \( \beta \) represents the ARCH and GARCH terms respectively. If the GARCH term is high, it means the volatility is highly persistent while the higher ARCH term represents the insensitivity of conditional variance to the unexpected market reactions. The sum of \( \alpha \) and \( \beta \) should be close to one, implying the high persistence of shock in the market.

Further in order to measure the seasonality, Dummy Augmented GARCH model is employed. Here eleven dummies are introduced that represents the month of the year in the mean equation to study the monthly seasonality in the mean equation. Now the mean equation is

\[ r_t = c + \sum_{t=1}^{11} \phi_t D_t + \varepsilon_t \]  \hspace{1cm} (4)

The eleven dummies are introduced again but now as exogenous variable in GARCH model in order to captures the monthly seasonality in the conditional variance equation. The variance equation is:

\[ h_t = \mu + \sum_{t=1}^{11} \phi_t D_t + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \]  \hspace{1cm} (5)

Here \( r_t \) is the return of either stock market or commodity market. \( D_t \) is the dummy variable, \( \phi_t \) is the coefficient of dummy variable, where \( t=1,2,3,........11 \) representing the months of the year from January to November respectively. The constant term captures the December effect in this model.

**Unit Root Test**

In order to study the order of integration for each series, various methods are used. These methods include Augmented Dickey Fuller (ADF) Test, Phillip-Perron Test, KPSS Test. The majority of study used ADF test. The null hypothesis for ADF test is
the series has a unit root which implies that the series is non-stationary (Dickey and Fuller, 1979). The regression equation for this test is given below:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \sum_{i=1}^{n} \alpha_i \Delta y_t + \varepsilon_t$$  \hspace{1cm} (6)

Here $y_t$ is the price series of individual commodity or stock index series, $\Delta$ is first difference operator, $i=1,2...n$ represents the number of lags.

**Johansen Co-integration Test**

Further Johansen co-integration test is applied to predict long run relationship between the commodity market and stock market. Two or more markets move jointly in the long run regardless the markets individually drifted, and then the difference between them is constant, known as co-integration and it is also termed as long run equilibrium association (Hall and Henry, 1989). If there is absence of co-integration between these variables, it means they drifted away from one another (Dickey et al., 1994). The johansen co-integration equation is given below:

$$y_t = \mu + \Delta_1 y_{t-1} + \Delta_p y_{t-p} + \varepsilon_t$$  \hspace{1cm} (7)

Here $y_t$ is the price series of individual commodities or stock index, $\varepsilon_t$ is error term

Further two different methods are used to find out the co-integration vector. The first one is trace statistics and the second is Eigen value criteria (Johansen, 1988).

**Toda and Yamamoto Granger Causality Test**

Furthermore in order to study the causal relationship between the commodities and stocks, Granger Causality Test approach proposed by Toda & Yamamoto (1995) has been employed. This approach is relatively more efficient than the other traditional methods used to study the causal relationship. Firstly validity of this method does not depend upon the order of integration of the variables under study. This method can be applied on any order of integration. Secondly, it is not required to find out the co-integrating relationship between the variables before detecting the causal relationship between them. Thirdly the bias associated with the unit root test and co-integrating properties of the variables has been reduced by this method.
Toda and Yamamoto method is based on the idea of applying Vector Autoregressive Model at level \((p= k+d_{\text{max}})\) with correct VAR order \(k\) and \(d\) extra lag, where \(d\) represents the maximum order of integration of time series. At last the wald statistics has been used in order to study the causality between the variables under study. The implementation of Toda and Yamamoto approach of Granger Causality linking both the variables under study as follow:

\[
Y_t = A_0 + A_1 Y_{t-1} + A_2 Y_{t-2} \ldots \ldots A_k Y_{t-k} + \epsilon_t
\]  
(8)

Where \(Y_t = \begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} \text{Com}_{t} \\ \text{SI}_{t} \end{bmatrix} \) and \(\epsilon_t \sim i.i.d N(0, \mu)\). Here com stands for individual commodities and SI stands for stock index.

The following equation represents the augmented level VAR \((k+d_{\text{max}})\) in order to detect causal relationship between the variables.

\[
Y_t = \alpha + A_1 Y_{t-1} + Y_{t-k} + A_{k+1} Y_{t-k+1} + A_p Y_{t-p} + \epsilon_t
\]  
(9)

**VAR-GARCH Model**

In this study, the VAR (1)-GARCH (1,1) model is used to study the spillover effects across stock market and commodity market. Ling and McAleer (2003) introduced this model and subsequently it is used by various researchers (Arouri et al., 2012; Jouini, 2013; Mensi et al., 2013; Bouri et al., 2017). This model is appropriate to find out the return and volatility transmission across financial markets. It provides the benefit of multi-variate analysis of conditional variance of the individual market and volatility spillover across financial markets. Secondly this model provides appropriate results with less computational complications (Arouri et al., 2012). The mean equation for this model is:

\[
y_t = c + \partial y_{t-1} + \epsilon_t
\]  
(10)

\[
\epsilon_t = D_t \varphi \epsilon_t
\]  
(11)

Here \(y_t=(r^s_t, r^c_t)\), \(r^s_t\) and \(r^c_t\) represents the return series of stock market and commodity market respectively.
\( \varepsilon_t = (\varepsilon_s^t | \varepsilon_c^t) \), \( \varepsilon_s^t \) and \( \varepsilon_c^t \) represents the residual term for mean equation of stocks and commodities respectively

\( \phi_t = (\phi_s^t | \phi_c^t) \), refers to independently identical white noise terms

\( D_t = (\sqrt{h_s^t} | \sqrt{h_c^t}) \), \( h_s^t \) and \( h_c^t \) represents the conditional variance for mean equation of stocks and commodities respectively

The variance equation is

\[
\begin{align*}
    h_s^t &= c_s + \alpha_s (\varepsilon_s^{t-1})^2 + \beta_s h_s^{t-1} + \alpha_c(\varepsilon_c^{t-1})^2 + \beta_c h_c^{t-1} \\
    h_c^t &= c_c + \alpha_c(\varepsilon_c^{t-1})^2 + \beta_c h_c^{t-1} + \alpha_s(\varepsilon_s^{t-1})^2 + \beta_s h_s^{t-1}
\end{align*}
\]

(12) (13)

Here \( \alpha \) is the coefficient of ARCH term and \( \beta \) is the coefficient of GARCH term. The above equation represents the volatility spillover across commodity market and stock market. \((\varepsilon_s^{t-1})^2 \) and \((\varepsilon_c^{t-1})^2 \) represents the impact of own one period lagged and cross market lagged return innovations on the current conditional volatility of stock market. \( h_s^{t-1} \) and \( h_c^{t-1} \) represents the impact of own one period lagged conditional variance and cross market lagged conditional variance on the current conditional variance of stock market. the second equation represents the opposite of it.

**Dynamic Conditional Correlation**

To investigate the conditional correlation between stock market and commodity, DCC-GARCH model is applied. This model is introduced by Engle (2001). The mean equation of this model is given below

\[
r_t = \mu + \varepsilon_t
\]

(14)

Here \( r_t \) is return series for commodities and stock indices. The covariance matrix is given below

\[
h_t = D_t R_t D_t
\]

(15)

Where \( D_t = (\sqrt{h_s^t} | \sqrt{h_c^t}) \). It is a diagonal matrix of dynamic conditional standard deviation estimated from the univariate GARCH model.
Here $R_t$ is the conditional correlation matrix of standardized return $\varepsilon_t$

$$R_t = \begin{bmatrix} 1 & q_{12t} \\ q_{21t} & 1 \end{bmatrix}$$

Further the above matrix decomposed

$$R_t = Q_t^{-1}Q_tQ_t^{-1}$$

Here $Q_t$ is the positive definite matrix containing the conditional variance and co-variance of $\varepsilon_t$ and $Q_t^{-1}$ is the inverted diagonal matrix.

The DCC(1,1) model specification is as follows

$$Q_t = \theta + \alpha \varepsilon_{t-1} \varepsilon_{t-1} + \beta Q_{t-1}$$ (16)

Here $\alpha$ is ARCH term and $\beta$ is GARCH term

Finally dynamic conditional correlation is represented as

$$\rho_{12t} = \frac{q_{12t}}{\sqrt{q_{11t}q_{22t}}}$$ (17)

**Optimal Weights and Hedge Ratio**

If an investor is holding stocks of an industry and want to hedge his position against the unexpected fluctuations in the commodity market. Therefore the main motive of investor is to minimize the risk without sacrificing the expected return. The optimal weights of stocks and commodities are given in the table by following the formula given by Kroner and Ng (1998).

The formula is given below:

$$w_{t}^{SC} = \frac{h_t^{c} - h_t^{SC}}{h_t^{c} - 2h_t^{SC} + h_t^{c}}$$ (18)

$$w_{t}^{SC} = \begin{cases} 0, & \text{if } w_{t}^{SC} < 0 \\ w_{t}^{SC}, & \text{if } 1 \leq w_{t}^{SC} \leq 0 \\ 1, & \text{if } w_{t}^{SC} > 1 \end{cases}$$
Here $w_t^{SC}$ is the weight of commodity in 100 rupees portfolio of stocks and commodities. $h_t^s$ and $h_t^c$ are the conditional variance of stocks and commodities respectively. $h_t^{SC}$ is the conditional co-variance between stocks and commodities. The optimal weight of stock sector index is given by $1-w_t^{SC}$.

The investors can also calculate the optimal hedge ratio for their portfolio. The long position in commodities is hedged by taking the short position in the stock market in $\beta_t$ rupees. The formula to calculate hedge ratio, describe by Kroner and Sultan (1993) is given below:

$$\beta_t^{SO} = \frac{h_t^{SO}}{h_t^s}$$

(19)