CHAPTER-6
LEAD-LAG RELATIONSHIP BETWEEN SPOT AND INDEX FUTURES MARKETS IN INDIA

6.1 INTRODUCTION

The introduction of the Nifty index futures contract in June 12, 2000 has offered investors a much greater degree of flexibility in the construction of their investment portfolios and in the timing of transactions associated with such portfolios. With the emergence of such markets worldwide, there is a growing body of literature, primarily concerned with the stock index futures contracts in the United States (especially the S&P 500 index futures contract), examining the pricing relationship between the stock and stock index futures markets (Kawaller, Koch and Koch, 1987; MacKinlay and Ramaswamy, 1988; Stoll and Whaley, 1990; and Chan, 1992). Much of this examination of the pricing relationship has been concerned with identifying the lead-lag relationship between prices in the two markets to try and determine which market, if either reacts to new information first. In this respect the study of the relation between stock market index and index futures prices has attracted the attention of researchers, financial analysts and traders since last two decades.

This chapter addresses to the nature of the pricing relationship between the two markets, arguing that the focus of those studies of the US stock index futures markets are inappropriate and flawed, both in the way they approach the issue conceptually and in the econometric methods they employ to test their proposed models. The estimation methodology employed in this study is the Cointegration and error correction modelling technique. The investigation of the Cointegration and causal relationship between futures and spot prices is very significant especially in an emerging market economy like India. Indian capital market has witnessed significant transformations and structural changes due to implementation of financial sector

---

1Given the much wider availability of finer data bases in the US, many of the studies that examine the pricing relationship use intra-day data over long periods of time. For example, Stoll and Whaley (1990) use prices quoted at five minute intervals from April 1982 to March 1987. Unfortunately, such data over reasonable periods of time is not widely available in the UK and thus we are restricted to using daily data. Nevertheless, this does not diminish the arguments that will follow.
reform measures by the Govt. of India since early 1990s. In this process, index futures trading were launched on June 9, 2000 at BSE and on June 12, 2000 at NSE and India started trading in derivative products. The main objectives behind the introduction of derivatives market were to control the increasing volatility of the asset prices, and to introduce sophisticated risk management tools leading to higher returns by reducing risk and transaction costs as compared to individual financial assets. The introduction of stock index futures has profoundly changed the nature of trading on stock exchanges. Futures market offer investors flexibility in altering the composition of their portfolios and also provide opportunities to hedge the risks involved with holding diversified equity portfolios. As a consequence, significant portion of cash market equity transactions are tied to futures market activity.

Thus, it is desirable that an empirical analysis be conducted to investigate the lead-lag relation between spot and index futures market in India, i.e., whether the daily changes of futures price index constitute information relevant with the trend that will follow the stock market, or changes of spot market constitute predicting tool for trend of prices in the market of futures contracts traded in the National Stock Exchange (NSE) of India.

The rest of this chapter is organized as follows: Second section discusses the nature of lead-lag relationships and how they might arise; Section three focuses on the 'traditional' method of estimating and testing for lead-lag relationships and points out the deficiencies with such an approach. A new and alternative framework and method for addressing the question of lead-lag relationships is proposed in section four. This framework demonstrates that the issue to be examined is one of whether equity markets function effectively. In section five, we link this framework to the issue of market efficiency, suggesting that tests of efficiency should be conducted in the framework of effectively functioning equity markets. In section six, we focus our attention on the behaviour of mispricing, using the framework proposed in this chapter to argue that it is a path independent, stationary, mean reverting stochastic process. Section seven concludes.
6.2 NATURE OF LEAD-LAG RELATIONSHIP

The argument that underlies the analysis of lead-lag relationships between indices and index futures is predicated on the observation that this relationship is indicative first of how well integrated the markets are and second of how quickly the markets reflect the arrival of new (and relevant) information relative to each other. If markets were perfect and investors fully rational with costless and equal access to the same information set then as Zeckhauser and Niederhoffer (1983) point out, it is not unreasonable to assume that stock index futures prices would carry no predictive information and would therefore have no role to play. However, the existence of transactions costs and other imperfections ensure that stock index futures do have a role to play because in this situation, they will convey relevant information about future movements in the stock index.

There are several reasons as to why this may be the case. One intuitive reason is similar to Black's (1975) analogy concerning the role of option contracts in the provision of relevant information for the underlying asset. Futures markets are very liquid with relatively low transactions costs. Moreover, investing in a futures contract requires no capital outlay since the margin can be posted in the form of interest-bearing securities and as such there is no opportunity cost. Thus, suppose an investor acquires new information on the health of the economy, say, that is worth acting upon. The investor has to decide whether to purchase stocks or a stock index futures contract. Purchase of the stocks requires a substantial amount of capital, a substantial amount of time and relatively substantial transactions costs. Purchase of the index futures contract, on the other hand, can be affected immediately with little up-front cash. Therefore, if the investor is willing to trade in futures, the futures transaction is the one to choose. The information will be incorporated in the futures price, driving it upwards. This will widen the differential between the futures and spot price which in turn will attract arbitrageurs. Since arbitrageurs trade simultaneously in cash and futures markets the information will be transmitted from the futures to the cash market. Thus, the futures price will lead the cash price.
Other reasons as to why the futures will lead the cash stem from institutional arrangements such as short-sale restrictions that are present in the cash market but not in the futures market. In this setting, Diamond and Verrecchia (1987) demonstrate that prices will be slower to adjust especially to bad news if traders who have private information are not allowed to short the security/securities. Such constraints are not present in the futures market; hence traders can short the futures contract. This will drive the futures price down, narrowing the differential between spot and futures prices and again attracting arbitrageurs. The futures price will thus lead the cash price. The relationship will, of course, not be as one-sided as it appears from the above discussion. A stock index futures price will tend to react to economy-wide information as opposed to security specific information. Thus, information concerning a specific security or group of securities may cause the cash market to lead the futures market, such that a (potentially complex) feedback relationship exists. This recognition that the futures price should lead the stock price has formed the basis of a great deal of empirical work geared to testing this very proposition (Kawaller, Koch and Koch, 1987; Stoll and Whaley, 1990; and Chan, 1992). However, as we shall see, these studies are flawed and the results that they generate are potentially misleading.

6.3 THEORETICAL UNDERPINNING OF THE LEAD-LAG RELATIONSHIP

Typically, test of lead-lag relationship in the extant literature are similar in spirit to Granger-Sims-type causality tests (Granger, 1969; and Sims, 1972). The model that is usually estimated is of the following form:

\[ \Delta S_t = \alpha_0 + \alpha_i \sum_{i=-k}^{i=+k} \Delta F_{t-i} + u_t \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (6.1) \]

Where \( \Delta S_t \) corresponds to the change in the spot price, \( \Delta F_t \) is the change in futures prices and \( u_t \) is the usual white noise error term. Tests of the lead-lag relationship then consist of testing the significance of the lag and lead coefficients on the futures prices. If the lags are significant and the leads are zero, the futures lead the spot. If the opposite is true then the spot leads the futures. If some of both the lead and lag coefficients are statistically non-zero, then a feedback relationship exists.
There are, however, two important and inter-related criticisms that can be addressed to the 'traditional' method of testing lead-lag relationships. One is concerned with the estimation of, and inference about, models such as (6.1) and the second is concerned with the specification of such models. To formalize matters, first note that whilst theory models suggest that an asymmetric feedback relationship is likely to exist, they give little guidance about the nature and form this asymmetry takes. Thus, models such as (6.1) are inevitably statistical models within which what effectively amounts to Granger-Sims causality tests are undertaken. The method of estimation in this context becomes vitally important if valid inference is to be sustained. This is one of the criticisms that can be leveled at Stoll and Whaley (1990) who estimate (6.1) by Ordinary Least Squares, immediately casting doubt on their results.

Perhaps more important here, however, is the nature of the interaction between spot and futures markets and the effect this has on the specification of models such as (6.1). The reason for such specification problems stems from the fact that in considering the pricing relationship between stock index futures markets and the underlying stock market, two quite distinct and seemingly independent strands have emerged in the literature: those studies that analyze mispricing by comparing the actual futures price with its fair, or theoretically correct, value to determine whether profitable arbitrage opportunities are available (MacKinlay and Ramaswamy, 1988; Yadav and Pope, 1990 and Chung, 1991) and those that analyze the lead-lag relationship between the two markets (Kawaller, Koch and Koch, 1987; Harris, 1989) and Stoll and Whaley, 1990). Most studies tend to focus on either the former or the latter issue, but not both.

This is where the specification problems arise for rather than being apparently independent areas of investigation, the former, that is, mispricing provides some valuable insights into the likely behaviour of lead-lag relationships and indicates that, in addition to those points mentioned above, results from studies of the lead-lag relationship must be viewed with some caution. To demonstrate, consider two commonly used and well known theoretical models showing the relationship between
the stock index futures price and the underlying stock index portfolio. First, we have (Cornell and French, 1983a, b)

\[ F_{t,T} = S_t e^{r(T-t)} - \sum_{k=t+1}^{T} D_k e^{r(T-k)} \]  \hspace{1cm} (6.2)

Where \( F_{t,T} \) is the fair or, equivalently, the theoretically correct stock index futures price quoted at time \( t \) for delivery at time \( T \), \( S_t \) is the value of the underlying stock index (spot portfolio), \( r \) is a riskless interest rate of approximately the same duration as the time to expiration of the futures contract and \( D \) is the daily dividend inflow from the portfolio until maturity of the stock index futures contract. Alternatively, we can consider the following model (MacKinlay and Ramaswamy, 1988):

\[ F_{t,T}^* = S_t e^{(r-d)(T-t)} \]  \hspace{1cm} (6.3)

Where \( F_{t,T}^* \) and \( S_t \) are defined as above, \( r \) is the risk free rate of interest, \( d \) is the yield on dividends from the underlying portfolio and \( (T-t) \) is the time to maturity of the futures contract. The expression \( (r-d)(T-t) \) is generally referred to as the cost of carrying the spot portfolio until maturity.

Now, studies that analyze mispricing and the existence of arbitrage opportunities typically compare the differential between the actual futures price quoted at time \( t \) for delivery at time \( T \), \( F_{t,T} \), with the fair futures price \( F_{t,T}^* \).

However, it is straightforward to demonstrate the role of the simple basis\(^2\) in this analysis. For ease of exposition, we will work with (6.3). The theoretical basis, \( F_{t,T} - F_{t,T}^* \), is compared with transactions costs to determine if arbitrage opportunities are present. If the theoretical basis falls outside of the no arbitrage window determined by transactions costs then dependent on whether the futures contract is undervalued (overvalued) due to, say, bearish (bullish) speculation in the stock index futures market, arbitrageurs will buy (sell) futures and sell (buy) stocks. It is clear that the theoretical basis is very important in the pricing relationship given that

\(^2\) One must be careful in talking about the basis for there are several definitions. Where there may be confusion, we will refer to the futures to cash price differential as the simple basis. When there is no risk of confusion, we will refer to it as the basis. The futures to fair price differential will be referred to as the theoretical basis.
index arbitrage links the two markets and the theoretical basis determines whether arbitrage opportunities are available or not.

To see the importance of the basis itself in the pricing relationship, take natural logs of (5.3) (lower case letters denote variables in natural logarithms):

\[ F_{t,T} = s_t + (r - d)(T - t) \quad (6.4) \]

If the futures market is pricing the stock index futures contract correctly then we have that:

\[ F_{t,T} - F_{t,T}^* = 0 \quad (6.5) \]

Now, to see the importance of the basis in the pricing relationship, substitute (6.4) into (6.5) and rearrange to obtain:

\[ F_{t,T} - s_t = (r - d)(T - t) \quad (6.6) \]

It is clear from (6.6) that the simple basis also has an important role to play in the arbitrage process. From the theoretical viewpoint the basis is crucial given that arbitrage provides an important link between the two markets. From an econometric point of view, the basis also has the rather appealing interpretation as the error correction mechanism which prevents prices in the two markets drifting apart without bound. The importance of the basis cannot be understated.

The traditional method is not suitable to address these specifications arising from the interaction of spot and futures prices. In this backdrop it is imperative to formulate and analyze the lead-lag relationship in a sophisticated manner. Thus, the following section tests the lead-lag relationship using sophisticated econometric techniques.

6.4 EMPIRICAL TESTING OF THE LEAD-LAG RELATIONSHIP

In the study made here the entire estimation procedure has been divided into three interrelated steps: first, unit root test; second, Cointegration test; third, the error correction estimation.

The econometric methodology, first examines the stationarity properties of each time series of consideration. The present study uses Augmented Dickey-Fuller (ADF)
unit root test to examine the stationarity of the data series. It consists of running a regression of the first difference of the series against the series lagged once, lagged difference terms and optionally, a constant and a time trend. This can be expressed as follows:

\[ \Delta R_t = \alpha_0 + \alpha_1 t + \alpha_2 R_{t-1} + \sum_{j=1}^{p} \alpha_j \Delta R_{t-j} + \epsilon_t \] ....................... (6.7)

Here, \( R_t \) is the daily compounded return on index. In this model the additional lagged terms are included to ensure that the errors are uncorrelated. In this ADF procedure, the test for a unit root is conducted on the coefficient of \( Y_{t-1} \) in the regression. If the coefficient is significantly different from zero, then the hypothesis that \( Y_t \) contains a unit root is rejected. Rejection of the null hypothesis implies stationarity. Precisely, the null hypothesis is that the variable \( Y_t \) is a non-stationary series (\( H_0 : \alpha_2 = 0 \)) and is rejected when \( \alpha_2 \) is significantly negative (\( H_a : \alpha_2 < 0 \)). If the calculated value of ADF statistic is higher than McKinnon’s critical values, then the null hypothesis (\( H_0 \)) is not rejected and the series is non-stationary or not integrated of order zero, I(0). Alternatively, rejection of the null hypothesis implies stationarity. Failure to reject the null hypothesis leads to conducting the test on the difference of the series, so further differencing is conducted until stationarity is reached and the null hypothesis is rejected. If the time series (variables) are non-stationary in their levels, they can be integrated with I(1), when their first differences are stationary. Hendry and Juselius (2000) investigated the properties of economic time series that were integrated processes, such as random walks, which contained a unit root in their dynamics. Here we extend the analysis to the multivariate context, and focus on cointegration in systems of equations. We showed in Hendry and Juselius (2000) that when data were non-stationary purely due to unit roots (integrated once, denoted I(1)), they could be brought back to stationarity by the linear transformation of differencing, as in \( x_t - x_{t-1} = \Delta x_t \). For example, if the data generation process (DGP) were the simplest random walk with an independent normal (IN) error having mean zero and constant variance \( \sigma^2_\epsilon \):
\[ x_t = x_{t-1} + \epsilon_t \text{ where, } \epsilon_t \sim IN \left[0, \sigma^2 \right] \]  

Then by subtracting, \( x_{t-1} \) from both sides of the equation (6.7a) delivers \( \Delta x_t \sim IN \left[0, \sigma^2 \right] \), which is certainly stationary. It is natural to enquire if other linear transformations than differencing will also induce stationarity. The answer is ‘possibly’, but unlike differencing, there is no guarantee that the outcome must be I(0). Thus cointegration analysis is designed to find linear combinations of variables that also remove unit roots. Cointegration vectors are of considerable interest when they exist, since they determine I(0) relations that hold between variables which are individually non-stationary. Such relations are often called ‘long-run equilibria’, since it can be proved that they act as ‘attractors’ towards which convergence occurs whenever there are departures there from (see e.g., Granger (1986), and Banerjee, Dolado, Galbraith, and Hendry (1993). Once a unit root has been confirmed for a data series, the next step is to examine whether there exists a long-run equilibrium relationship among variables. This is called Cointegration analysis which is very significant to avoid the risk of spurious regression.

Cointegration analysis is important because if two non-stationary variables are cointegrated, a VAR model in the first difference is misspecified due to the effects of a common trend. If Cointegration relationship is identified, the model should include residuals from the vectors (lagged one period) in the dynamic VECM system. In this stage, Johansen’s Cointegration test is used to identify cointegrating relationship among the variables. The Johansen method applies the maximum likelihood procedure to determine the presence of cointegrated vectors in non-stationary time series. The testing hypothesis is the null of non-cointegration against the alternative of existence of cointegration using the Johansen maximum likelihood procedure. In the Johansen framework, the first step is the estimation of an unrestricted, closed \( p^\text{th} \) order VAR in \( k \) variables. The VAR model as considered in this study is:

\[ R_t = A_1 R_{t-1} + A_2 R_{t-2} + \ldots + A_p R_{t-p} + BX_t + \epsilon_t \]  

Where \( R_t \) is a \( k \)-vector of non-stationary I(1) endogenous variables, \( X_t \) is a \( d \)-vector of exogenous deterministic variables, \( A_1, \ldots, A_p, \text{ and } B \) are matrices of coefficients to
be estimated, and $\varepsilon_t$ is a vector of innovations that may be contemporaneously correlated but are uncorrelated with their own lagged values and uncorrelated with all of the right-hand side variables. Since most economic time series are non-stationary, the above stated VAR model is generally estimated in its first-difference form as:

$$
\Delta R_t = \Pi R_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta R_{t-i} + BX_t + \varepsilon_t \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots
$$

Here, $\Pi = \sum_{i=1}^{p} A_i - I$, and $\Gamma_i = -\sum_{j=i+1}^{p} A_j$ Granger’s representation theorem asserts that if the coefficient matrix $\Pi$ has reduced rank $r < k$, then there exist $k \times r$ matrices $\alpha$ and $\beta$ each with rank $r$ such that $\Pi = \alpha \beta$ and $\beta R_i$ is I(0). $r$ is the number of co-integrating relations (the co-integrating rank) and each column of $\beta$ is the co-integrating vector. $\alpha$ is the matrix of error correction parameters that measure the speed of adjustments in $\Delta R_t$. The Johansen approach to Cointegration test is based on two test statistics, viz., the trace test statistic, and the maximum eigen value test statistic. The trace test statistic can be specified as: $\tau_{trace} = -T \sum_{i=r+1}^{r} \log(1 - \lambda_i)$, where $\lambda_i$ is the $i^{th}$ largest eigen value of matrix $\Pi$ and $T$ is the number of observations. In the trace test, the null hypothesis is that the number of distinct cointegrating vector(s) is less than or equal to the number of cointegration relations ($r$). On the other hand, the maximum Eigen value test examines the null hypothesis of exactly $r$ cointegrating relations against the alternative of $r + 1$ cointegrating relations with the test statistic: $\tau_{max} = -T \log(1 - \lambda_{r+1})$, where $\lambda_{r+1}$ is the $(r+1)^{th}$ largest squared eigenvalue. In the trace test, the null hypothesis of $r = 0$ is tested against the alternative of $r + 1$ cointegrating vectors. It is well known that Johansen’s cointegration test is very sensitive to the choice of lag length. So first a VAR model is fitted to the time series data in order to find an appropriate lag structure. The Akaike Information Criterion (AIC), Schwarz Criterion (SC) and the Likelihood Ratio (LR) test are used to select the number of lags required in the cointegration test. In the event of detection of cointegration between the time series we know that there exists a long-term equilibrium relationship between them so we apply VECM in order to evaluate the short run properties of the
cointegrated series. In case of no cointegration VECM is no longer required and we directly precede to Granger causality tests to establish causal links between variables. The regression equation form for VECM is as follows:

$$\Delta X_t = \alpha_0 + \lambda_1 \varphi_{t-1} + \sum_{i=1}^{m} \alpha_i \Delta X_{t-i} + \sum_{j=1}^{n} \alpha_j \Delta R_{t-j} + \varepsilon_{1t} \tag{6.10}$$

$$\Delta R_t = \beta_0 + \lambda_2 \varphi_{t-1} + \sum_{i=1}^{m} \beta_i \Delta R_{t-i} + \sum_{j=1}^{n} \beta_j \Delta X_{t-j} + \varepsilon_{2t} \tag{6.11}$$

Here, $\varphi_{t-1}$ is the error correction term lagged one period; $\lambda$ is the short-run coefficient of the error correction term ($-1 < \lambda < 0$); and $\varepsilon$ is the white noise. The error correction coefficient ($\lambda$) is very important in this error correction estimation as greater the co-efficient indicates higher speed of adjustment of the model from the short-run to the long-run. On the other hand, the lagged terms of $\Delta X_t$ and $\Delta R_t$ appeared as explanatory variables, indicate short-run cause and effect relationship between the two variables. Thus, if the lagged coefficients of $\Delta X_t$ appear to be significant in the regression of $\Delta R_t$ this will mean that $X$ causes $R$. Similarly, if the lagged coefficients of $\Delta R_t$ appear to be significant in the regression of $\Delta X_t$, this will mean that $R$ causes $X$.

In VECM the cointegration rank shows the number of cointegrating vectors. For instance a rank of two indicates that two linearly independent combinations of the non-stationary variables will be stationary. A negative and significant coefficient of the ECM (i.e. equ.-6.10 in the above equations) indicates that any short-term fluctuations between the independent variables and the dependant variable will give rise to a stable long run relationship between the variables.

At the outset, it is required to determine the order of integration for each of the two series used in the analysis. The Augmented Dickey-Fuller unit root test has been used for this purpose and the results of such test are reported in Table -1. It is clear that the null hypothesis of no unit roots for both the time series are rejected at their first differences since the ADF test statistic values are less than the critical values at 10%, 5% and 1% levels of significances. Thus, the variables are stationary and integrated of same order, i.e., $I(1)$.
Table 6.1: Results of Augmented Dickey-Fuller Unit Root Test

<table>
<thead>
<tr>
<th>Variables in their First Differences with trend and intercept</th>
<th>ADF Statistic</th>
<th>Critical Values</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>X = LNIFTY</td>
<td>-49.33</td>
<td>At 1% : -3.96</td>
<td>Reject Null hypothesis of no unit root</td>
</tr>
<tr>
<td></td>
<td></td>
<td>At 5% : -3.41</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>At 10% : -3.12</td>
<td></td>
</tr>
<tr>
<td>Y = LFUTIDX</td>
<td>-51.12</td>
<td>At 1% : -3.96</td>
<td>Reject Null hypothesis of no unit root</td>
</tr>
<tr>
<td></td>
<td></td>
<td>At 5% : -3.41</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>At 10% : -3.12</td>
<td></td>
</tr>
</tbody>
</table>

In the next step, the Cointegration between the stationary variables has been tested by the Johansen’s Trace and Maximum Eigenvalue tests. The results of these tests are shown in Table-6.2. The Trace test indicates the existence of one cointegrating equation at 5% level of significance. And, the maximum eigenvalue test makes the confirmation of this result. Thus, the two variables of the study have long-run equilibrium relationship between them. But in the short-run there may be deviations from this equilibrium and we have to verify whether such disequilibrium converges to the long-run equilibrium or not. And, Vector Error Correction Model can be used to generate this short-run dynamics.

Table 6.2: Results of Johansen’s Cointegration Test

<table>
<thead>
<tr>
<th>Hypothesized Number of Cointegrating Equations</th>
<th>Eigen Value</th>
<th>Trace Statistics</th>
<th>Critical Value at 5% (p-value)</th>
<th>Maximum Eigen statistics</th>
<th>Critical Value at 5% (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>0.038091</td>
<td>106.3986</td>
<td>15.49471(0.0001)</td>
<td>106.1764</td>
<td>14.26460(0.0001)</td>
</tr>
<tr>
<td>At Most 1</td>
<td>0.000081</td>
<td>0.222223</td>
<td>3.841466(0.6373)</td>
<td>0.222223</td>
<td>3.841466(0.6373)</td>
</tr>
</tbody>
</table>

* denotes rejection of the hypothesis at the 0.05 level

Error correction mechanism provides a means whereby a proportion of the disequilibrium is corrected in the next period. Thus, error correction mechanism is a means to reconcile the short-run and long-run behavior.

The estimation of a Vector Error Correction Model (VECM) requires selection of an appropriate lag length. The number of lags in the model has been determined according to Schwarz Information Criterion (SIC). The lag length that minimizes the SIC is 2. Then an error correction model with the computed t-values of the regression coefficients is estimated and the results are reported in Table-6.3.
Table 6.3: Estimates for VECM Regression

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>$\Delta X_t$</th>
<th>$\Delta Y_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.000508</td>
<td>0.000477</td>
</tr>
<tr>
<td>[t-statistic]</td>
<td>[ 1.61170]</td>
<td>[ 1.41731]</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.1071)</td>
<td>(0.1564)</td>
</tr>
<tr>
<td>$EC_{t-1}$</td>
<td>-0.036232</td>
<td>-0.001726</td>
</tr>
<tr>
<td>[t-statistic]</td>
<td>[-9.83951]</td>
<td>[-0.43884]</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.0000)</td>
<td>(0.6608)</td>
</tr>
<tr>
<td>$\Delta X_{t-1}$</td>
<td>0.062283</td>
<td>-0.007751</td>
</tr>
<tr>
<td>[t-statistic]</td>
<td>[ 3.28524]</td>
<td>[-0.38281]</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.0010)</td>
<td>(0.7019)</td>
</tr>
<tr>
<td>$\Delta X_{t-2}$</td>
<td>-0.043126</td>
<td>0.049592</td>
</tr>
<tr>
<td>[t-statistic]</td>
<td>[-2.27434]</td>
<td>[ 2.44871]</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.0230)</td>
<td>(0.0144)</td>
</tr>
<tr>
<td>$\Delta Y_{t-1}$</td>
<td>-0.032932</td>
<td>0.022920</td>
</tr>
<tr>
<td>[t-statistic]</td>
<td>[-1.79467]</td>
<td>[ 1.16949]</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.0728)</td>
<td>(0.2423)</td>
</tr>
<tr>
<td>$\Delta Y_{t-2}$</td>
<td>-0.029763</td>
<td>-0.036931</td>
</tr>
<tr>
<td>[t-statistic]</td>
<td>[-1.62077]</td>
<td>[-1.88296]</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.1051)</td>
<td>(0.0598)</td>
</tr>
</tbody>
</table>

The estimated coefficient of error-correction term in the $\Delta X_t$ equation is statistically significant and has a negative sign, which confirms that there is not only any problem in the long-run equilibrium relation between the independent and dependent variables in 5% level of significance, but its relative value (-0.0362) shows the rate of convergence to the equilibrium state per year. Precisely, the speed of adjustment of any disequilibrium towards a long-run equilibrium is that about 3.62% of the disequilibrium in Nifty index is corrected each year. Furthermore, the negative and statistically significant value of error correction coefficient indicates the existence of a long-run causality between the variables of the study. And, this causality is unidirectional in our model being running from the index futures market price to spot market price index. In other words, the changes in spot prices can be explained by prices of futures.
The existence of Cointegration implies the existence of Granger causality at least in one direction (Granger, 1988). The long-run causality test from the VECM indicates that causality runs from futures market to spot market, since the coefficient of the error term in $\Delta X_t$ equation is statistically significant and negative based on standard t-test which means that the error correction term contributes in explaining the changes in spot market prices.

The coefficients of the first difference of $\Delta X_t$ lagged two periods in $\Delta Y_t$ equation in Table-6.4 is statistically significant which indicate the presence of short-run causality from spot market to futures market based on VECM estimates. In order to confirm the result of the short-run causality between the $\Delta X_t$ and the $\Delta Y_t$ based on VECM estimates, a standard Granger causality test has been performed based on F-statistics.

Table 6.4: Results of Granger Causality Test

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>F-Statistic</th>
<th>Probability</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta Y$ does not Granger Cause $\Delta X$</td>
<td>1.797</td>
<td>0.109</td>
<td>Accept</td>
</tr>
<tr>
<td>$\Delta X$ does not Granger Cause $\Delta Y$</td>
<td>2.448</td>
<td>0.031</td>
<td>Reject</td>
</tr>
</tbody>
</table>

(Number of lags = 5)

The result in Table-6.4 indicates spot market price index does not Granger cause the futures market is rejected at the 5% level of significance. This result supports the previous result obtained from VECM that there exists short-run causality from spot market to futures market at the 5% level of significance. In addition, it is inferred that futures market prices does not Granger cause the spot market prices is accepted at the 5% level of significance. Based on this causality tests, it can be said that change in the futures prices causes change in spot market prices in the long-run only, but not in the short-run. And, change in the spot market prices causes change in the prices in futures market in the short-run only, but not in the long-run.

The results imply that the derivative contracts on NSE Nifty lead the underlying cash market. Thus, the derivative markets are indicative of futures price movements and this will certainly be helpful to potential investors to design their risk-return portfolio while investing in stocks and derivatives contracts. Thus, investors
will be able to use futures price as a good indicator in predicting spot price. The causal relationship suggests that policy makers should take into consideration the impact of futures market towards cash market when developing a policy for the futures market.

This chapter examines the dynamics of the short- and long-run relationship between the spot and index futures market in India over the sample period June 2000 to May 2011 using the popular time series models, viz., ADF stationarity test, Johansen’s Cointegration test, and Granger causality test in the vector error correction framework.

5.6 TO SUM UP

The future market trading in Indian financial markets was introduced in June 2000 and options index was commenced from June 2001 and subsequently the options and futures on individual securities trading was commenced from July 2001 and November 2001, respectively. The futures trading on stock indexes has grown rapidly since inception and provides important economic functions such as price discovery, portfolio diversification and opportunity for market participants to hedge against the risk of adverse price movements. Hence, the movements of spot market price have been largely influenced by the speculation, hedging and arbitrage activity of futures markets. Thus, understanding the influence of one market on the other and role of each market segment in price discovery is the central question in market microstructure design and has become increasingly important research issue among academicians, regulators and practitioners alike as it provides an idea about the market efficiency, volatility, hedging effectiveness and arbitrage opportunities, if any. Price discovery is the process of revealing information about future spot prices through the future markets.

The essence of the price discovery function hinges on whether new information is reflected first in changes of future prices or changes of spot prices. Hence, there exists lead-lag relationship between spot and futures market by information dissemination. All the information available in the market place is immediately incorporated in the prices of assets in an efficient market. So, new
information disseminating into the market should be reflected immediately in spot and futures prices simultaneously. This will lead to perfect positive contemporaneous co-movement between the prices of those markets and there will be no systematic lagged response and therefore no arbitrage opportunity. This prediction arises directly from the Cost of Carry (COC) model. In addition, if there are economic incentives for traders to use one market over the other, a price discovery process between the two markets is likely to happen. This implies that futures and spot market prices are inter-related and can be traced under different market frictions through price discovery mechanism.

Accordingly, there exist diversified theoretical arguments pertaining to the causal relationship between spot and futures markets by information dissemination and raises the major question that which market price reacts first (lead) whether; Futures prices tend to influence spot prices or; Spot prices tend to lead futures prices or; A bidirectional feedback relationship exists between spot and futures prices.

The main arguments in favour of futures market leads spot market are mainly due to the advantages provided by the futures market includes higher liquidity, lower transaction costs, lower margins, ease leverage positions, rapid execution and greater flexibility for short positions. Such advantages attract larger informed traders and make the futures market to react first when market-wide information or major stock-specific information arrives. Thus, the future prices lead the spot market prices.

On the other hand, the low cost contingent strategies and high degree of leverage benefits in futures market attracts larger speculative traders from a spot market to a more regulated futures market segments. Hence, this ultimately reduces informational asymmetries of the spot market through reducing the amount of noise trading and helps in price discovery, improve the overall market depth, enhance market efficiency and increase market liquidity. This makes spot market to react first when market-wide information or major stock-specific information arrives. Hence, spot market leads the futures market.
Besides, there exists a bidirectional relationship between the futures and spot markets through price discovery process. This may be mainly due to future markets attracts larger informed traders to enjoy the advantages of higher liquidity, lower transaction costs, lower margins and greater flexibility for short positions. Hence, these advantages make futures markets to lead the spot markets around macroeconomic or major stock-specific information releases. Consequently, the spot markets will lead the futures market under the circumstances that these advantages of futures markets attracts larger speculative traders from a spot market and reduces informational asymmetries of the spot market through reducing the amount of noise trading and helps in price discovery, improve the overall market depth, enhance market efficiency and increase market liquidity. This makes spot market to react fast when market-wide information or major stock specific information arrives. Thus, both the spot and futures markets are said to be informationally efficient and reacts more quickly to each other.

Johansen’s Cointegration technique followed by the Vector Error Correction Model (VECM) was employed to examine the lead-lag relationship between S&P CNX Nifty and Nifty Futures. The empirical analysis was conducted for the daily data series from June, 2000 to March 2011. The analysis reveals the bidirectional relationship between spot and futures markets.

The study also provides the evidence of long-run equilibrium relationship between the spot market price index and its futures price. It implies that either of these two historical prices will help to forecast the other, which is the evidence for disapproving market efficiency hypothesis between these two markets. Furthermore, the study provides the evidence of long-run causality running from the index futures market price to spot market price index. In other words, in India’s capital market futures prices lead the spot prices only in the long-run, but not in the short-run. In particular, the findings show that the S & P CNX Nifty based index futures market tends to lead the underlying stock index (i.e., S & P CNX NIFTY) over a long period of time.
The theoretical implication is that if these arbitrage opportunities exist, traders will take advantage of them. But the finding of no short-run causality provides evidence to the finance literature that spot market price index and its future markets do not offer obvious arbitrage opportunities in the short-run and are reasonably efficient as far as the daily data is concerned in India. Another interesting finding is there that in the short-run spot market price tends to lead the futures market price. In other words, this result also indicates that spot price do lead futures price but the lead-lag relationship is relatively weak as compared to the impact of futures price on spot price. And, this may be due to the impact of bad news in the short-run. Since the interests of Indian investors on the futures markets are comparatively low when compared to the spot markets, research contribution to the price discovery mechanisms on these markets can help them to predict in which markets they can invest for higher returns. The present study suggests that depending on the relative proportions of informed to uninformed (noise) traders migrating from the spot market to the futures market, the lead-lag relationship between futures and spot market may differ.

REFERENCES


