Information, Price Discovery and Causality in the Indian Stock Index Futures Market

Pratap Chandra Pati* and Purna Chandra Padhan**

This study examines the price discovery process and lead-lag relationship between NSE S&P CNX Nifty stock index futures and its underlying spot index, using daily data from January 1, 2004 to December 31, 2008. It investigates the long-term and short-term dynamics of prices between spot and futures market, using Johansen-Juselius cointegration test, Vector Error Correction Model (VECM), impulse response functions, and variance decomposition. In addition to it, the recently developed Granger non-causality tests of Toda and Yamamoto (1995) and Dolado and Lütkepohl (1996) have also been applied to examine the causal relationship between spot and futures markets. The obtained results support the existence of a long-run relationship between spot and futures prices. Further, VECM indicates short-run unidirectional causality from futures to spot market. In addition, the study finds unidirectional Granger causality from futures market to spot market through Toda-Yamamoto-Dolado-Lütkepohl (TYDL) causality test. The shape of the impulse response graphs shows that spot market has a larger response to shocks in the futures index than the futures responses to spot innovations. The results of variance decomposition indicate that the futures market shocks dominate over spot market in explaining the variation in spot market. However, disturbance originating from spot market contributes very less percentage variability to futures market. To conclude, futures price leads spot price and performs the price discovery function. The obtained results have important implications for traders, regulatory bodies and practitioners.

Introduction

The role of futures markets in providing an efficient price discovery mechanism has been an area of extensive empirical research. The existence of price discovery and market efficiency is the centerpiece of market microstructure design and of utmost importance for the practitioners, regulators and academicians. Price discovery refers to the process through which financial markets converge and reach the efficient equilibrium price. The essence of the price discovery function of futures market hinges on whether new information is reflected first in changed futures price or in changed spot price. In a perfectly efficient and frictionless market, futures price should move concurrently with the underlying spot price without any lead or lag in price movements across markets. Futures price should be an unbiased estimator of the future spot price at the expiration date. However, due to market friction such as transaction cost or market microstructure effects, futures market processes information faster than the spot market. Futures price leads spot

* Academic Associate, Indian School of Business (ISB), Hyderabad, India; and is the corresponding author. E-mail: pratap_pati@isb.edu

** Faculty, Xaviers Labour Relations Institute (XLRI), Jamshedpur, India. E-mail: pcpadhan@xlri.ac.in

price, indicating that the futures market performs the price discovery function. It illustrates how rapidly one market incorporates information relative to the other, and also indicates the efficiency of their functioning as well as the degree of integration between the two markets.

The study of dynamic price relationship between equity spot and futures is useful for the traders, regulatory bodies, practitioners, and academicians for several reasons, such as price discovery, hedging and arbitrage opportunities. Since the most notable objectives of security market design are optimal price discovery and market efficiency, it is essential to determine the nature and location of price discovery. The implementation of arbitrage trading strategies must take into account the lead-lag relationship between cash and futures markets. It provides valuable information to the traders regarding the prospective direction of price movement in the cash market, which in turn directs the trader's future actions. Mispricing of futures provides especially short-run riskless profit-making opportunity to the arbitrageurs. Therefore, their actions will result in correcting the magnitude of basis. But if mispricing is governed by factors other than market risk, it may spoil the market sentiment and will contribute to information asymmetry. This in turn may damage the trader's confidence. Efficient price discovery in the futures market has critical implications, especially for the hedgers who want to reduce the risk element contained in the cash market portfolio by taking reverse positions in the futures market. Since, both markets are related to each other through the cost-of-carry phenomenon, movement of basis may play a critical role.

Considerable empirical studies have been conducted on price discovery and lead-lag relationship between spot and futures prices for developed countries like the US, the UK, Australia, Japan, and New Zealand. While in the context of India, hardly existence of any such study is noticed. However, two diametrically opposing views have emerged. On the one hand, Thenmozhi (2002) puts forward that futures market leads the spot market, whereas Mukherjee and Mishra (2006) argue that spot market plays a major role in price discovery and leads the futures market. The present study investigates the long-term and short-term dynamic relationship of prices between spot and futures markets, using Johansen (1988) and Johansen and Juselius (1990) cointegration test, Vector Error Correction Model (VECM), Granger non-causality tests by Toda and Yamamota (1995) and Dolado and Lütkephol (1996) (popularly known as the TYDL model), impulse response function, and variance decomposition.

The paper reviews the theoretical and existing empirical evidence on the relationship between spot and futures prices. Subsequently, it presents the nature and preliminary analysis of data and describes the methodology and the model specifications used. Then, it provides empirical results and findings, and finally offers the conclusions.

Conceptual Framework and Literature Review

Theoretically, the spot and futures markets are linked by the no-arbitrage Cost-of-Carry (COC) model. The COC model is a continuous-time representation of the relationship between the 'fair value' of a futures contract and the 'fair value' of the underlying spot index, plus cost of carrying spot index for the duration of the contract. Mathematically, the COC model is represented by the following equation:

$$F_t = S_t \, e^{[(r-d)\,(T-t)]} \qquad \ldots(1)$$

where F_t is the index futures price at time t; S_t is the spot index price at time t; r is the continuously compounded cost of carrying the cash index basket from time t; d is the dividend yield on the index; and T is the expiration date of the futures index contract. $T - t$ is the time remaining for expiration of the futures contract, while $r - d$ is the time value of wealth held in a portfolio which matches the stock index, net of the flow of dividends from the index. Equation (1) demonstrates the contemporaneous difference between futures price (F_t) and spot price (S_t). The difference between futures and spot prices is called 'basis' and represents the net cost of carry. An assumption of the COC model is that the two markets are perfectly efficient, frictionless and act as perfect substitutes. Accordingly, arbitrage opportunities should not exist because new information arrives simultaneously at both the markets and is reflected immediately in both futures and spot prices.

Substantial number of studies are found in the literature on these issues of spot and futures prices. Kawaller *et al.* (1987) examined the minute-by-minute data on S&P 500 futures and spot over the period 1984-1985 and found that futures price led the spot index price by some 20-45 minutes. Stoll and Whaley (1990) analyzed the 5 minutes intraday returns of the S&P 500 and MMI index futures for the period 1982-1987 and concluded that S&P 500 and MMI futures led the spot indices by about 5 minutes. Similarly, Cheung and Ng (1990) examined the 5 minutes returns of S&P 500 spot and futures for the period 1983-1987 and found that futures returns led the spot returns by 15 minutes. Using daily data on S&P 500 for the period 1988-1992, Wahab and Lashgari (1993) studied the causality between spot and futures prices and found bidirectional causality between spot and futures returns. Tse (1995) analyzed the daily data on Nikkei 225 futures for the period 1988-1993 by applying VECM and found that futures returns led the spot returns, but not vice versa. Pizzi *et al.* (1998) analyzed the 1 minute returns of S&P 500 for the period January 1987-March 1987 using cointegration approach and found bidirectional causality between spot and futures. Futures price led the spot price by 20 minutes and spot market led the futures market by 3-4 minutes. Abhyankar (1998) studied the lead-lag relationship, using 5 minutes returns of FTSE 100 for the four near month futures contracts traded in 1992 and found that the futures returns led the spot returns by 5-15 minutes. Frino *et al.* (2000) examined the 1 minute returns of the All Ordinaries Index (AOI) and Share Price Index (SPI) futures for 1995-1996 and found that futures led spot by 18 minutes, while spot led futures by 4 minutes. Brooks *et al.* (2001) examined the lead-lag relation between the FTSE 100 stock index futures and the FTSE 100 index using 10 minutes observations from June 1996-1997. The results indicate that futures market led the spot market and this was attributed to the faster flow of information into futures market mainly due to lower transaction costs. Roope and Zurbruegg (2002) considered transaction data for four months in 1999 on Taiwan MSCI futures in Singapore, Taifex futures in Taiwan, and two underlying securities. They concluded that price discovery was dominated by Taiwan MSCI futures in Singapore. So and Tse (2004) used transaction data for 1999-2002 to compare the relative contribution of the Hang Seng index, Hang Seng index futures and Hang Seng tracker fund to price discovery. They found that 75-80% of the price discovery was by the futures market, and the remaining 20-25% by the spot market, and nothing by the tracker fund. Zhong *et al.* (2004) investigated the lead-lag relationship in returns between Mexican IPC index futures contracts and the spot market using daily data for 1999-2002. They found that futures returns led the spot returns.

The literature on the lead-lag relationship indicates that futures market is the major source of market-wide information and the futures lead the spot market. Very little evidence of spot index leading the futures market is noticed. It has been argued that persistence in the lead-lag relationship between index futures and spot index prices can be traced to market imperfections, transaction costs, liquidity differences between the markets, no synchronous trading effects, the automation of market, and short-selling restrictions.

Data and Preliminary Analysis

The data used in this study consists of daily closing prices of S&P CNX Nifty stock index and its index futures traded at the National Stock Exchange (NSE) from January 1, 2004 to December 31, 2008. The Mumbai-based NSE is a leading electronic limit order-driven market without market makers. S&P CNX Nifty Index (Nifty) is the benchmark stock index of the Indian capital market. Nifty is a value-weighted index of 50 largest and the most liquid stocks from about 22 sectors in India. The sample period is characterized by high liquidity and volume in Indian derivatives market. Since most trading activities take place in near month contract, only near month contract data are examined. Closing prices for spot and futures data are obtained from the official web page of the NSE (www.nseindia.com). The daily continuous compound return is defined as the first logarithmic difference of closing prices on consecutive trading days, i.e., $d\ln S_t = (\ln S_t - \ln S_{t-1})$ and $d\ln F_t = (\ln F_t - \ln F_{t-1})$.

Table 1: Descriptive Statistics of Nifty and Its Index Futures Returns

	$d\ln S_t$	$d\ln F_t$
Mean	0.0001	0.0013
Median	0.0006	0.00054
Std. Dev.	0.0081	0.0089
Skewness	–0.7951	–0.8849
Kurtosis	8.7479	9.8646
Jarque-Bera	1845.1270 (0.000)	2607.019 (0.000)
LB-Q(8)	48.453 (0.000)	20.101 (0.000)
LB^2-Q(8)	315.746 (0.000)	405.816 (0.000)
ARCH-LM (8)	139.471 (0.000)	113.137 (0.000)

Note: *p*-values are given in parentheses. LB-Q(*k*) is the portmanteau statistics. Ljung-Box Q-test is used for testing the joint significance of autocorrelation up to order *k*.

Table 1 reports the descriptive statistics of the returns of the S&P CNX Nifty spot and futures index.

The average daily returns from the two markets are almost equal over the sample period. The futures volatility is relatively greater than the cash market volatility as revealed from the standard deviation. This is to be expected as the futures market is regarded as a source of price stability in the spot market, since it absorbs the brunt of price adjustments. The negative skewness coefficients indicate that frequency distribution of futures and spot returns series are negatively skewed or have longer tail to the left. The unconditional distribution of both spot and futures returns exhibit fat tails and excessive peak at the mean than the corresponding normal distributions. The non-normality is also confirmed by Jarque-Bera test where the null hypothesis is that the given series is normally distributed. Here the Jarque-Bera statistic is highly statistically significant for both spot and futures returns series, and hence we reject the null hypothesis of normality. The significant autocorrelation in the squared returns series indicates the presence of volatility clustering. The presence of significant LM(5) test statistic indicates the evidences of ARCH effect and time-varying volatility.

Table 2 presents the results of Augmented Dickey-Fuller (ADF) (Dickey and Fuller, 1979) and Phillips-Perron (PP) (Phillips and Perron, 1988) unit root tests applied on the log-levels and log-first differences of daily cash and futures price series to test the existence of unit roots and identify the order of integration of each variable.

Table 2: Results of Unit Root Tests

Variables	ADF		PP	
	Intercept	Intercept and Trend	Intercept	Intercept and Trend
Levels				
ln S_t	–1.225	–1.654	–1.225	–1.693
ln F_t	–1.221	–1.615	–1.220	–1.609
First Differences				
dln S_t	–16.207	–16.257	–32.807	–32.832
dln F_t	–16.412	–16.457	–35.045	–35.074

Note: The MacKinnon critical values for ADF and PP tests are: –3.438, –2.864, –2.568 with only intercept and –3.970, –3.415, –3.129 with intercept and trend at 1%, 5%, and 10% significance levels, respectively.

The null hypothesis of a unit root in the closing prices could not be rejected at the 5% significance level, suggesting that the logarithm of the closing prices of the spot and futures are non-stationary. But the first differences of logarithm of the closing prices yielded larger ADF statistics that rejected the null hypothesis. The results indicate that there is unit root in both the logarithmic spot and futures prices series. Hence, these log-level series can be described as non-stationary processes, but are stationary at their first differences. Thus, the time series plot of logarithm of the closing price series of Nifty spot and futures are observed to be trending upwards over time and exhibit non-stationarity.

Figure 1 shows that the two series tend to move together due to the cost-of-carry relationship and arbitrage effects, while Figure 2 shows that the logarithmic first differences of spot and futures appear to be oscillating around a constant value, indicating stationary behavior.

Figure 1: Time Series Plot of Nifty Spot and Futures Index: January 1, 2004 to December 31, 2008

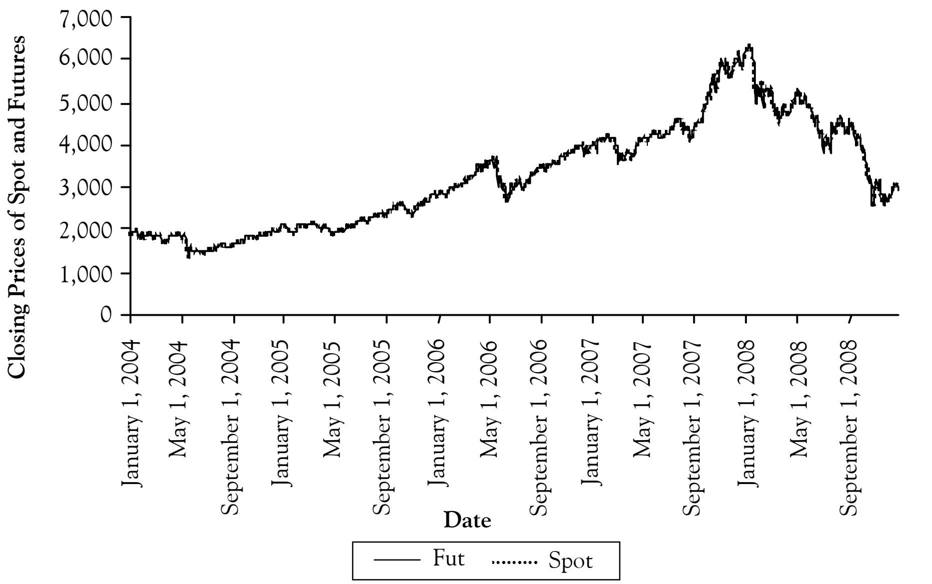
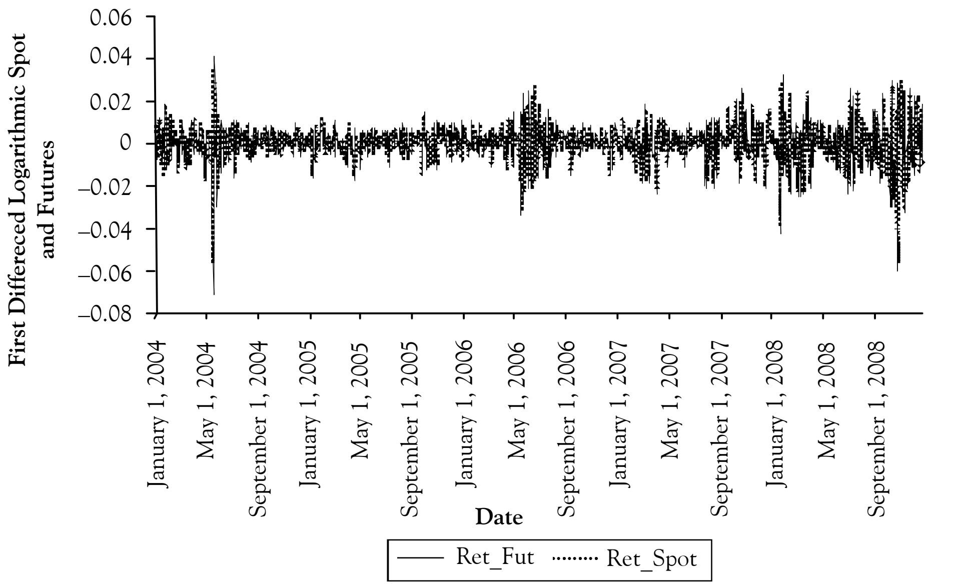


Figure 2: Time Series Plot of Differenced Logarithmic Nifty Spot and Futures: January 1, 2004 to December 31, 2008



Model Specification

The empirical analysis of data reveals that the log of both spot and futures price series is non-stationary at levels, but stationary at their first differences. The high degree of positive correlation between the spot and the futures suggests that both markets are interrelated. To investigate the dynamic linkages between spot and futures markets, the present study applies cointegration technique of Johansen-Juselius, Vector Error Correction Model (VECM), Toda-Yamamota Dolado-Lütkepohl (TYDL) causality test, impulse response function, and variance decomposition.

Johansen's cointegration test has been applied to test the long-run relationship between spot and futures prices, which is investigated by estimating the following VECM:

$$\Delta Y_t = A_0 + \Pi Y_{t-k} + \sum_{i=1}^{k-1} \Gamma_i \Delta Y_{t-i} + \varepsilon_t \qquad ...(2)$$

where $\Pi = \sum_{j=1}^{k} A_i - I$ and $\Gamma_i = \sum_{j=1}^{i} A_j - I$

Y_t = [ln S_t ln F_t]′ is a (2 × 1) vector of non-stationary log-spot and log-futures prices; ΔY_t = [dln S_t dln F_t]′ is a (2 × 1) column vector first differenced series; $\varepsilon_t = [\varepsilon_{s,t} \ \varepsilon_{f,t}]'$ is a (2 × 1) column vector of white noise Gaussian error; A_0 is a (2 × 1) column vector of constants; and A_i is a (2 × 2) matrix of coefficients.

The existence of cointegrating relations among the variables can be examined through the Π matrix. Mathematically, the Π matrix can be rewritten as $\Pi = \alpha\beta'$, where α and β are $n \times r$ matrices of rank r. Here, β represents the matrix of cointegrating parameters and α is the matrix of the speed of adjustment parameters. The test for cointegration between the Ys is calculated by looking at the rank of the Π matrix via eigenvalues. Johansen developed two likelihood ratio tests for testing the number of cointegrating vectors (r)—the trace test and the maximum eigenvalue test. The trace statistic (λ_{trace}) tests the null hypothesis of $r = 0$ (i.e., no cointegration) against the alternative of $r > 0$, i.e., there is one or more cointegrating vectors. The maximal eigenvalue test statistic (λ_{max}) examines the null hypothesis that the number of cointegrating vectors are less than or equal to r against the alternative of $r + 1$.

According to Granger (1986) representation theorem, if the variables are cointegrated, then there must be an associated VECM. The VECM examines the short-run dynamics between the integrated variables by including the lagged equilibrium error term obtained from cointegration equation and lagged values of the first differences of each variable. The VECM model is given by Equations (3) and (4).

$$\mathrm{d}\ln S_t = \alpha_1 + \alpha_s ect_{t-1} + \sum_{i=1}^{k} \alpha_{11}(i)\mathrm{d}\ln S_{t-i} + \sum_{i=1}^{k} \alpha_{12}(i)\mathrm{d}\ln F_{t-i} + \varepsilon_{s,t} \qquad ...(3)$$

$$d\ln F_t = \alpha_2 + \alpha_f ect_{t-1} + \sum_{i=1}^{k} \alpha_{21}(i) d\ln S_{t-i} + \sum_{i=1}^{k} \alpha_{22}(i) d\ln F_{t-i} + \varepsilon_{f,t} \quad \ldots(4)$$

ect_{t-1} is the equilibrium error correction term derived from the cointegrating equation. Coefficients on the equilibrium errors α_s and α_f are the speed of adjustment coefficients and quantify the speed at which the short-run deviation from the long-run equilibrium is adjusted or corrected in the next period.

The study has also applied the recently developed TYDL causality test (after Toda and Yamamoto, 1995; and Dolado and Lütkepohl, 1996) for examining the causal relationship between spot and futures markets. The TYDL model avoids the problems associated with Granger or Sims tests by ignoring any possible non-stationarity or cointegration between series while testing for Granger causality. By fitting a standard VAR in the levels of the variables (rather than first differences, as is the case with the Granger and Sims causality tests), it minimizes the risks associated with possibly wrongly identifying the orders of integration of the series, or the presence of cointegration, and minimizes the distortion of the tests' sizes as a result of pre-testing. The TYDL procedure basically involves the estimation of a standard VAR at levels with a total of $p = [k + d_{max}]$ lags, where k is the optimal lag length in the original VAR system, and d_{max} is the maximal order of integration of the variables in the VAR system. It is not necessary to pre-test the variables for integration and cointegration properties, provided the maximal order of integration of the process does not exceed the true lag length of the VAR model. The TYDL procedure uses a modified Wald (MWald) test for zero restrictions on the parameter of the VAR(k) model. This test has an asymptotic chi-squared distribution with k degrees of freedom in the limit.

To shed more light on the dynamic linkages between spot and futures markets, innovation accounting analysis, such as impulse response function and variance decomposition pioneered by Sims (1980), is conducted in a VAR framework. Impulse Response Function (IRF) traces the effect of one market's shock (through innovation or news) on other markets. It measures how a one standard deviation shock or innovation to a variable affects the current and future values of other endogenous variables through the dynamic (lag) structure. A shock in spot market ($\varepsilon_{s,t}$) will not only directly affect the spot returns, but will also be transmitted to the futures market over time, since dln S_t appears as an explanatory variable in the futures market regression equation. Similarly, a change of one standard deviation in $\varepsilon_{f,t}$ of the dln F_t equation will have an impact on dln S_t. The ordering of the variables is an important issue in IRF. By contrast, variance decomposition analysis provides the proportion of the n-periods ahead forecast error variance of a variable, which can be attributed to its own shock and the shocks to the other variables in the VAR system. In the short run, most of the variation is due to own shock. But as the lagged variables' effect starts kicking in, the percentage of the effect of other shocks increases over time. If $\varepsilon_{s,t}$ can explain none of the forecast error variance of the sequence dln F_t at all forecast horizons, then dln F_t is exogenous. If $\varepsilon_{s,\ t}$ can explain most of the forecast error variance of the sequence dln F_t at all forecast horizons, then dln F_t is endogenous.

Empirical Analysis

The estimation procedure of Johansen and Juselius (1990) cointegration test is based on maximum likelihood estimation with a VAR model. However, prior to the application of VAR model, the selection of lag length is important. The AIC, SIC, HQ, FPE and LR statistics can be applied to determine the VAR order (i.e., lag length, k). The resulting lag structures are reported in Table 3.

Table 3: VAR Lag Order Selection Criteria

Lag	log L	LR	FPE	AIC	SIC	HQ
0	6,401.88	NA	1.11e – 07	–10.33	–10.33	10.33
1	10,721.48	8618.26	1.04e – 10	–17.31	–17.28	–17.30
2	10,819.25	194.75	8.94e – 11	–17.46	–17.40	–17.44*
3	10,824.09	9.62	8.93e – 11*	–17.46*	–17.42*	–17.44
4	10,827.23	6.24	8.94e – 11	–17.46	–17.38	–17.43
5	10,828.62	2.75	8.98e – 11	–17.45	–17.36	–17.42
6	10,833.02	8.66	8.97e – 11	–17.45	–17.35	–17.41
7	10,835.59	5.12	8.99e – 11	–17.45	–17.33	–17.40
8	10,841.27	11.19	8.97e – 11	–17.45	–17.31	–17.40

Note: * indicates lag order selected by the criterion; FPE: Final Prediction Error; AIC: Akaike Information Criterion; SIC: Schwarz Information Criterion; and HQ: Hannan-Quinn Information Criterion.

We started with the upper bound for $k = 8$ (approximately one week) and successively tested the coefficient of the largest lag. Using individual significance level, a lag length of $k = 3$ is selected based on these criteria. The VAR estimated with three lags satisfies the stability test and normality test, and no serial correlation is found among the residuals in the VAR model. The additional results can be obtained upon request.

The results of Johansen-Juselius cointegration test based on λ trace and λ maximal eigenvalue tests are reported in Table 4.

The value of λ_{trace} statistic under the null hypothesis of $r = 0$ at 63.628 is greater than the 5% significance level. However, for all other values of r, the trace statistic is less than the critical value, allowing us to reject the null hypothesis of more than one cointegrating vectors. Hence, the trace statistics, together with their associated critical values, suggest the presence of at least one cointegrating vector, implying that long-run equilibrium relationship can be detected between spot and futures prices. Similarly, the alternative

Table 4: Johansen-Juselius Maximum Likelihood Cointegration Test				
λ Trace Test				
H_0	H_1	Eigenvalues	λ Trace Statistic	5% Critical Value
$r = 0$	$r > 0$	0.048	63.628**	15.487
$r \leq 1$	$r > 1$	0.001	1.481	3.841
λ Maximal Eigenvalue Test				
H_0	H_1	Eigenvalues	λ Max Statistic	5% Critical Value
$r = 0$	$r = 1$	0.048	62.146**	14.264
$r = 1$	$r = 2$	0.001	1.481	3.841

Note: Linear deterministic trend with intercept and trend in CE assumption is assumed. ** denotes rejection of the hypothesis at 0.05 level.

measure used to identify the number of cointegrating vectors is λ_{max} statistic, formulates the null hypothesis of no cointegrating vector ($r = 0$), which is rejected at 5% significance level. The λ_{max} test is performed again to test the null hypothesis of one cointegrating vector ($r = 1$) against the alternative of two cointegrating vectors. The maximal eigenvalue test statistic indicates that the null hypothesis cannot be rejected as the calculated maximal eigenvalue is less than the critical value at 5% level. Therefore, the overall results indicate the existence of one cointegrating vector between futures and spot index. Hence, in the long run, both the variables move in the same direction.

Having identified a cointegrating vector between spot and futures prices, the VECM is estimated. The results are reported in Table 5.

Table 5 indicates that both the error correction terms (ect_{t-1}) are statistically significant and non-zero, indicating that the model is out of equilibrium. Since the coefficients of error correction terms are positive both the series come forward (move upward) to return to the equilibrium. Of course, the moment in each period depends upon the sign of the error factor (ect_{t-1}) and its magnitude. Since the error correction term in the

Table 5: Estimates of Vector Error Correction Model

	dln S_t	dln F_t
ect_{t-1}	1.224* (3.839)	4.137* (6.641)
dln S_{t-1}	–2.283* (–4.683)	–1.801 (–1.242)
dln S_{t-2}	–1.623* (–4.449)	–1.031 (–1.225)
dln S_{t-3}	–1.039* (–4.966)	–0.171 (–1.205)
dln F_{t-1}	1.6115* (3.356)	2.093* (4.054)
dln F_{t-2}	1.107* (3.110)	1.991* (3.633)
dln F_{t-3}	0.762* (3.805)	0.881* (4.089)

Note: *t*-statistics are given in parentheses. * represents statistically significant at 5% significance level.

futures equation is greater (4.137) in magnitude than the one in the cash equation (1.224), it indicates how quickly the futures market restores equilibrium. This indicates that futures prices react more rapidly in restoring equilibrium than spot, indicating that futures lead spot in price discovery. However, the coefficients of lags of changes in futures are significant in the cash equation, while the coefficients of lags of changes in the spot are insignificant in the futures equation. This implies that there is a unidirectional Granger causal relationship from the futures to spot market, and therefore, futures prices lead spot prices.

Table 6 reports the results of the TYDL causality test based on augmented VAR ($k + d_{max} = 3 + 1 = 4$) model. The unit root tests show that maximal order of integration is one, i.e., ($d_{max} = 1$). The optimal lag length of three is selected ($k = 3$) following the AIC, SIC and FPE criteria.

Table 6: TYDL Causality Test

Null Hypothesis	MWald	*p*-Value
ln S_t does not Granger cause ln F_t	1.467	0.209
ln F_t does not Granger cause ln S_t	4.939**	0.000

From the above results, the null hypothesis that spot price does not Granger cause futures price cannot be rejected, as MWald statistic is statistically not significant at 5% significance level. But the second null hypothesis that futures price does not Granger cause spot price could be rejected. So, the study finds unidirectional causality from futures market to spot through TYDL causality test.

To provide a more detailed insight into the findings of VAR model, impulse response function and variance decomposition are estimated. Figure 3 illustrates the estimated impulse response functions for seven days ahead time horizons.

The graphs of impulse response functions depicted in Figure 3 have been plotted for seven periods ahead forecasting horizon. It is evident from the shape of the impulse response graphs that spot market has a larger response to one standard deviation shocks to the futures index than the futures responses to spot innovations. Each variable responds positively to shocks to itself and to the other variable in the system, and there is a contemporaneous effect only from futures to spot. Initially, the response of spot price to shocks to futures prices is very quick and lasts for at least one week and then it starts decaying. On the other hand, response of futures price to shocks to spot prices is weak and low initially, and it fluctuates more or less at the same low level. The response of spot prices is higher than the response of futures prices. Thus, it can be concluded from this analysis that futures price leads spot price.

The forecast error variance decomposition provides an alternative way to look at the finding of the impulse response analysis. It enables in innovating the extent to which a variable helps in explaining the other variables. The estimates of the variance decomposition are reported in Table 7 for seven-day time horizons. The reported figures

Figure 3: Impulse Response Functions

Response to Cholesky One SD Innovations

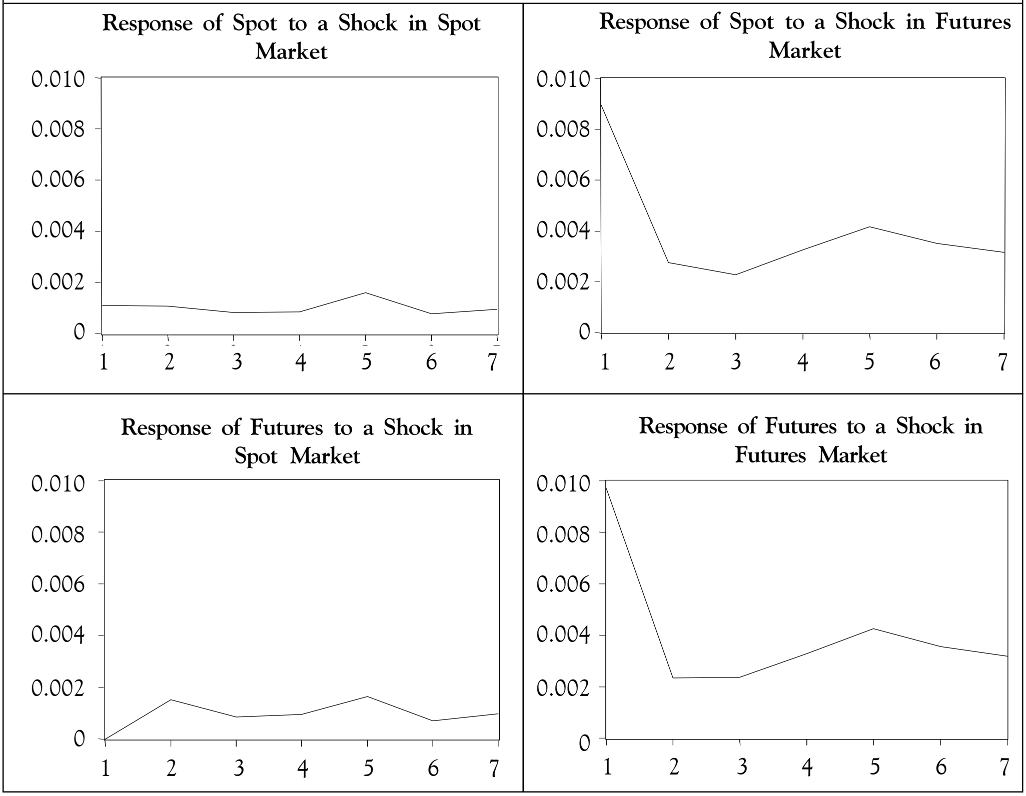


Table 7: Forecast Error Variance Decomposition

Panel A: Forecast Error Variance Decomposition of Spot Return			Panel B: Forecast Error Variance Decomposition of Futures Return		
Period	Futures Return	Spot Return	Period	Futures Return	Spot Return
1	98.511	1.4888	1	100.00	0.0000
2	97.385	2.6145	2	97.690	2.3098
3	96.832	3.1675	3	97.164	2.8352
4	96.478	3.5216	4	96.676	3.3238
5	95.000	4.9991	5	95.250	4.7490
6	95.037	4.9623	6	95.342	4.6579
7	94.806	5.1932	7	95.072	4.9271

indicate the percentage of movement in each variable that can be attributed to its own shock and the shocks to the other variables in the system.

Panel A of Table 7 shows the forecast error variance decomposition of spot return. Initially a very low percentage (1.49%) of its total forecast error variance is explained by itself. After that it increases but at a very slow pace. However, in the first period of the forecasting horizon, 98.511% variation in the forecast error of spot market is explained by the futures market. After that it starts decaying at a slower pace. This indicates that much percentage of the forecast error variance in spot price is due to futures prices. In other words, the futures prices lead to spot prices. Similarly, Panel B of Table 7 reports the forecast error variance decomposition of futures returns. Throughout the forecast period, it explains a high level of forecast error variance of itself. At the initial period, it explains 100% variation in its forecast error, but after that it shows a decreasing trend. However, only a small percentage changes in forecast error of futures market is explained by the spot market, though over the period of time it shows an increasing trend, but the rate of increase is very low. This indicates that futures prices lead the spot prices.

Conclusion

This paper explores the price discovery role and lead-lag relationship between stock index futures and its underlying spot index in the S&P CNX Nifty. From the Johansen-Juselius test, it can be concluded that there is only one cointegrating vector suggesting the long-run comovement of both the spot and futures prices. The coefficient of error correction terms are positive and significant, indicating that both current spot and futures prices adjust to the long-run differences between itself and other variables. The error correction term in the futures equation is greater in magnitude than the one in the spot equation, therefore implying that futures lead spot in price discovery. However, the study finds unidirectional Granger causality from futures market to spot through Toda and Yamamoto (1995) and Dolado and Lütkepohl (1996) causality test. The shape of the impulse response graphs show that spot market has larger response to shocks to the futures index than the futures responses to spot innovations. The results of variance decomposition indicate that the futures market shocks dominate over spot market in explaining variations in spot market. However, disturbance originating from spot market contributes very less percent of variability in futures market. To conclude, futures price leads spot price and performs the price discovery function.

Implications: The results of the study have economic significance and important implications for the traders, regulatory bodies and practitioners. The implementation of arbitrage trading strategies must take into account the lead-lag relationship between cash and futures market. It provides important information to the traders regarding the prospective direction of price movement in the spot market, which in turn directs the trader's future actions. Mispricing of futures provides short-run riskless profit-making opportunity to the arbitrageurs and their actions will result in correcting the magnitude of basis, but if mispricing is governed by factors other than market risk, it may spoil the market sentiment and will contribute to information asymmetry, which in turn may damage the trader's confidence.❖

Acknowledgment: *The authors are thankful to an anonymous referee for his/her extremely helpful suggestions and comments on an earlier version of the paper.*

References

1. Abhyankar A (1998), "Linear and Nonlinear Granger Causality: Evidence from the UK Stock Index Futures Market", *Journal of Futures Markets*, Vol. 18, No. 5, pp. 519-540.
2. Brooks C, Rew A G and Ritson S (2001), "A Trading Strategy Based on the Lead-Lag Relationship Between the Spot Index and Futures Contract for the FTSE 100", *International Journal of Forecasting*, Vol. 17, No. 1, pp. 31-44.
3. Cheung Y W and Ng L K (1990), "The Dynamics of S&P 500 Index and S&P 500 Futures Intraday Price Volatilities", *Review of Futures Markets*, Vol. 9, No. 2, pp. 458-486.
4. Dickey D A and Fuller W A (1979), "Distribution of the Estimators for Autoregression Time Series with a Unit Root", *Journal of the American Statistical Association*, Vol. 74, No. 366, pp. 427-431.
5. Dolado J J and Lütkepohl H (1996), "Making Wald Test Work for Cointegrated VAR Systems", *Econometrics Reviews*, Vol. 15, No. 4, pp. 369-386.
6. Frino A, Walter T and West A (2000), "The Lead-Lag Relationship Between Equities and Stock Index Futures Markets Around Information Releases", *Journal of Futures Markets*, Vol. 20, No. 5, pp. 467-487.
7. Granger C W J (1986), "Developments in the Study of Co-integrated Economic Variables", *Oxford Bulletin of Economics and Statistics*, Vol. 48, No. 3, pp. 213-228.
8. Johansen S (1988), "Statistical Analysis of Co-integration Vectors", *Journal of Economic Dynamics and Control*, Vol. 12, Nos. 2-3, pp. 231-254.
9. Johansen S and Juselius K (1990), "Maximum Likelihood Estimation and Inference on Co-integration with Application to the Demand for Money", *Oxford Bulletin of Economics and Statistics*, Vol. 52, No. 2, pp. 169-209.
10. Kawaller I G, Koch P D and Koch T W (1987), "The Temporal Price Relationship Between S&P 500 Futures and the S&P 500 Index", *Journal of Finance*, Vol. 42, No. 5, pp. 1309-1329.
11. Mukherjee K N and Mishra R K (2006), "Lead-Lag Relationship Between Equities and Stock Index Futures Market and Its Variation Around Information Release: Empirical Evidence from India", www.nseindia.com
12. Phillips P C B and Perron P (1988), "Testing for a Unit Root in Time Series Regression", *Biometrika*, Vol. 75, No. 2, pp. 335-346.
13. Pizzi M A, Economopoulos A J and O'Neill H M (1998), "An Examination of the Relationship Between Stock Index Cash and Futures Markets: A Co-integration Approach", *Journal of Futures Markets*, Vol. 18, No. 3, pp. 297-305.

14. Roope M and Zurbruegg R (2002), "The Intraday Price Discovery Process Between the Singapore Exchange and Taiwan Futures Exchange", *Journal of Futures Markets*, Vol. 22, No. 3, pp. 219-240.

15. Sims C (1980), "Macroeconomics and Reality", *Econometrica*, Vol. 48, No. 1, pp. 1-48.

16. So W R and Tse Y (2004), "Price Discovery in the Hang Seng Index Markets: Index, Futures and the Tracker Fund", *Journal of Futures Markets*, Vol. 24, No. 9, pp. 887-907.

17. Stoll H R and Whaley R E (1990), "The Dynamics of Stock Index and Stock Index Futures Returns", *Journal of Financial and Quantitative Analysis*, Vol. 25, No. 4, pp. 441-468.

18. Thenmozhi M (2002), "Futures Trading, Information and Spot Price Volatility of NSE-50 Index Futures Contract", NSE Research Paper, NSE India.

19. Toda H Y and Yamamoto T (1995), "Statistical Inference in Vector Autoregressions with Possibly Integrated Processes", *Journal of Econometrics*, Vol. 66, No. 1, pp. 225-250.

20. Tse Y K (1995), "Lead-Lag Relationship Between Spot Index and Futures Price of Nikkei Stock Average", *Journal of Forecasting*, Vol. 14, No. 7, pp. 553-563.

21. Wahab M and Lashgari M (1993), "Price Dynamics and Error Correction in Stock Index and Stock Index Futures Markets: A Co-integration Approach", *Journal of Futures Market*, Vol. 13, No. 7, pp. 711-742.

22. Zhong M, Darratt A F and Otero R (2004), "Price Discovery and Volatility Spillovers in Index Futures Markets: Some Evidence from Mexico", *Journal of Banking and Finance*, Vol. 28, No. 12, pp. 3037-3054.

Reference # 37J-2009-09/12-01-01