Price Discovery and Information Sharing between Futures and Spot Market: Evidence from India

Anshul Jain*

Research Fellow, Management Development Institute (MDI), Gurgaon, Haryana 122001, India. Email: fpm11anshul_j@mdi.ac.in, jainanshul@gmail.com; Phone: +91-124-4560010

Pratap Chandra Biswal

Associate Professor, Management Development Institute (MDI), Gurgaon, Haryana 122001, India. Email: pcbiswal@mdi.ac.in; Phone: +91-124-4560379

*Corresponding author

Abstract: This study examines the price discovery and information sharing relationships between the S&P CNX NIFTY index and its futures. An Autoregressive Distributed Lag (ARDL) Bounds Test for multivariate cointegration shows that the spot index and its futures are cointegrated and the relationship is robust to choice of dependent variables. Augmented VAR approach of Toda & Yamamoto was used to test for direction of causality in prices. Results indicate a unidirectional causality from spot to futures market, indicating higher efficiency in the spot market. Gonzalo-Granger information share measure indicates the highest share of the NIFTY index followed by the next month futures. This study provides further support to the investor structure issue highlighted in Bohl, Salm and Schuppli (2011) and Yang, Yang and Zhou (2012).

Keywords: Nifty; Price Discovery; ARDL Bounds Test; Johansen Juselius Cointegration Test; Toda Yamamoto Augmented VAR; Gonzalo-Granger Information Share

Price Discovery and Information Sharing between Futures and Spot Market: Evidence from India

1. INTRODUCTION

Price Discovery is the process through which asset markets reach equilibrium price levels. Increased information efficiency of a market allows for faster price discovery. This process in the stock market is aided by the presence of derivative markets, which allow for information to flow through another channel. Futures are one of the most commonly traded derivatives which support the spot markets in discovering the equilibrium price. As futures have inherent leverage and can be easily shorted, these markets tend to have higher liquidity than the underlying cash markets. Higher liquidity implies greater participation by various groups of investors and traders and hence faster information absorption. The addition of options market further improves the liquidity in the derivatives market which allows it to play a greater role in the price discovery process.

The futures market and its underlying cash market are linked by the Cost of Carry model which is based on the principle of no arbitrage. The Cost of Carry model proposes that the futures price is the sum of the spot price and the cost of carrying the underlying asset net of dividend over the time period of the futures life. The underlying assumption is that both markets receive and absorb information at the same time and that no arbitrage is possible by taking offsetting positions in these markets. This is represented by the following equation:

$$F_T = S_t * e^{[(r-d)(T-t)]} \quad (1)$$

Where F_T is the price of the futures contract expiring at time T, S_t is the price of the underlying on the date of valuation, r is the rate of borrowing, d is the dividend yield and T-t is the time to expiry of the futures contract. The borrowing rate and dividend yield are specified in continuously compounded form. This relationship allows for a difference in prices between the futures and spot market, which is known as the cost of carry or basis.

Theoretically, in the absence of taxes, costs and other market frictions, futures price is an unbiased estimator of the spot price at a fixed future date. Hence future prices should lead the spot prices in price discovery. But the presence of market frictions and the differences between the investor perceptions in the cash and futures markets can lead to a breakdown of this proposed relationship.

Research in the area of price discovery and information sharing amongst futures and cash markets is important from two main standpoints. Firstly, it has implications for market efficiency. Presence of arbitrage opportunities indicates an inefficient market. Secondly, the fundamental reason behind the introduction of derivatives markets is to increase liquidity and price discovery in the underlying cash markets through the trading linkages between these markets. This hypothesis assumes the importance of futures markets in absorbing and disseminating information to the cash market.

Plethora of research has happened in the last three decades examining the relationships between the futures market and their underlying cash market. Kawaller, Koch and Koch (1987) found an intraday bi-directional causality between the S&P 500 cash market and its futures market. The futures market was found to lead the spot market over a period of 1 minute and vice versa over a longer period of 20 to 45 minutes. Examining the Major Market Index (MMI), Stoll and Whaley (1990) found that the futures lead the spot by 5 minutes. Ghosh (1993) used cointegration to assess the price discovery process in the S&P500 spot and futures market and found evidence to support the superiority of futures in price discovery. Bidirectional causality was found between S&P500 spot and futures returns by Wahab and Lashgiri (1993). Tse (1995) used a VECM method to examine the Nikkei 225 spot and futures market for a period of 5 years and found that the futures market leads the spot market, but not vice versa. FTSE 100 futures were found to lead the FTSE 100 index by Brooks, Rew and Ritson (2001). So and Tse (2004) used Hang Seng index, futures and tracker data for three years and measured the individual contributions to price discovery. They found that almost three fourth of the price discovery was due to the futures market and rest by the cash market. The tracker did not contribute to the price discovery. All these results have been based on developed markets in which institutional traders are the major investor group. This allows for a bias in these studies as they do not look at emerging markets like China and India. The methodology used in these studies to test for cointegration has been the Johansen Cointegration Test, which necessitates the series to be integrated of the same order, amongst other restrictions, and hence is dependent on the power of the unit root tests being used. More recent methodologies like the Autoregressive Distributed Lag (ARDL) Bounds Test for cointegration, as developed by Pesaran and Shin (1998) and Pesaran *et al* (2001) have not been utilised. The ARDL test is more general in its assumptions as compared to the Johansen Cointegration test and hence is more robust in its results.

In the area of emerging market research, Bohl, Salm and Schuppli (2011) show that in markets with presumably uninformed private investors, the superiority of the futures market for price discovery might break down. In fact, in such cases, the spot market will lead the futures market in price discovery. Yang, Yang and Zhou (2012) found evidence to support the Bohl, Salm and Schuppli (2011) results. They studied China's future market and found the contrarian result of the cash market leading the futures market.

Karmakar (2009) looked at the price discovery and volatility spill over between the Nifty Index and its near month future in India. Using Johansen Cointegration Tests and the VECM approach to causality, a bidirectional relationship between the spot and futures market was established. But in doing so, the role of the other two futures which trade with the Nifty as the underlying was ignored. These two futures are the next month and far month futures, which are useful for hedging purposes as they allow traders to take a longer term view of beyond a month, while reducing rollover costs and basis risk. Indian institutional investors like mutual funds, banks and insurance companies can only use the derivatives market for hedging purposes[1]. Hence testing the price discovery and information share relationships, in the Indian market, taking the full set of the three futures contracts is important to the understanding of emerging markets as utilising just one futures contract in testing price discovery is essentially a restricted model.

This study examines the price discovery and information sharing functions in the Indian stock market utilising data from the S&P CNX Nifty Index and its three rolling futures. We use the Autoregressive Distributed Lag (ARDL) Bounds testing approach developed by Pesaran and Shin (1998) and Pesaran *et al* (2001) to test for cointegration amongst the spot index and its futures. Direction of causality is then tested using the Augmented VAR approach developed by Toda &Yamamoto (1995) to ascertain the direction of information flow and relative importance of markets in price discovery. Johansen Cointegration test is conducted to find the number of cointegrating vectors in the data. Using the number of cointegrating vectors, information share of the index and futures in price discovery is quantified using the Gonzalo & Granger (1995) and De Jong (2002) methodology.

The rest of the study is organised as follows: Section 2 provides a brief overview of the S&P CNX NIFTY index and futures along with their descriptive statistics. Section 3 discusses the various methodologies used in this study. Section 4 discusses the empirical results and Section 5 concludes.

2. S&P CNX NIFTY AND NIFTY FUTURES

National Stock Exchange of India (NSE) is the largest stock exchange in India in terms of volume of trading, both in the Cash as well as Futures markets. S&P CNX Nifty is the benchmark index of NSE and is composed of 50 blue chip stocks across 22 sectors of the Indian Economy. These stocks are chosen based on a methodology which gives importance to several factors, prime amongst them being the liquidity in these stocks and the resulting impact factors. Hence these 50 stocks are amongst the most liquid and widely traded stocks on the NSE. As of March 31st, 2012 the Nifty Index represents 65.57% of the free float market capitalisation on the NSE market, hence is representative of the cash market as a whole. Futures on the Nifty index were introduced on June 12th, 2000 and have been trading successfully with increasing volumes ever since. The exchange lists three futures on a rolling basis, one each for the near month expiry, next month expiry and third month expiry.

This study utilises daily closing prices for the Nifty index and its three rolling futures from 1st January 2001 to 30th December 2011, a period of eleven years with 2870 observations. Four price series are constructed, SPOT for the index closing and three for the rolling futures closing (FUT1, FUT2 and FUT3). FUT1 denotes the daily closing price series for the near month expiry, FUT2 for the next month and FUT3 for the far month.

Table I provides the descriptive statistics for these price series. These series are converted into new price series by taking logarithms of the existing series. These new series are denoted by attaching “L” before the name of the original series. Returns series are generated using log difference and are denoted by “DL” before the name of the original series.

Table I

Summary Statistics of Level, Log and Return Series of Spot and Rolling Futures Contracts

Panel A: Summary Statistics of Levels				
	SPOT	**FUT1**	**FUT2**	**FUT3**
Mean	3141.915	3140.737	3141.041	3142.194
Median	2968.050	2960.175	2949.000	2943.275
Maximum	6312.450	6333.450	6361.050	6380.750
Minimum	854.200	855.400	860.300	865.150
Std. Dev.	1707.844	1710.900	1713.950	1715.860
Skewness	0.171	0.176	0.182	0.186
Kurtosis	1.549	1.554	1.558	1.560
Jarque-Bera	265.647	264.852	264.567	264.443
Probability	0.000	0.000	0.000	0.000
Observations	2870	2870	2870	2870
Panel B: Summary Statistics of Logs				
	LSPOT	**LFUT1**	**LFUT2**	**LFUT3**
Mean	7.8731	7.8721	7.8720	7.8723
Median	7.9957	7.9930	7.9892	7.9873
Maximum	8.7503	8.7536	8.7579	8.7610
Minimum	6.7502	6.7516	6.7573	6.7629
Std. Dev.	0.6307	0.6315	0.6317	0.6315
Skewness	-0.2881	-0.2848	-0.2797	-0.2761
Kurtosis	1.5808	1.5803	1.5777	1.5763
Jarque-Bera	280.5564	279.8428	279.3326	278.8512
Probability	0.000	0.000	0.000	0.000
Observations	2870	2870	2870	2870
Panel C: Summary Statistics of Returns				
	DLSPOT	**DLFUT1**	**DLFUT2**	**DLFUT3**
Mean	0.0005	0.0005	0.0005	0.0005
Median	0.0012	0.0009	0.0009	0.0007
Maximum	0.1633	0.1619	0.1562	0.1566
Minimum	-0.1305	-0.1626	-0.1656	-0.1673
Std. Dev.	0.0159	0.0168	0.0168	0.0169
Skewness	-0.2916	-0.4641	-0.5622	-0.6349
Kurtosis	12.1079	13.1002	13.4531	13.4826
Jarque-Bera	9957.1290	12298.0100	13213.1100	13328.5500
Probability	0.000	0.000	0.000	0.000
Observations	2869	2869	2869	2869

Note: Spot refers to the series of daily closing values for the NIFTY Index. FUT1, 2 and 3 refer to series of the daily closing values of the near month, next month and far month futures respectively. L in front of the series name indicates logarithm of the series, DL indicates log differenced form.

The higher standard deviations of the futures returns suggest the higher volatility of futures prices as compared to index prices, which is because futures prices are expectations of spot prices in the future. The higher volatility reflects the higher uncertainty implied in the futures price due to the intervening period of time till the expiry. The mean returns are similar, which along with the higher volatility of futures prices indicates that the risk adjusted returns are better for investors in the cash market as compared to the futures market. Higher skewness and kurtosis of futures returns also indicates the higher volatility of futures prices as compared to the index.

3. EMPIRICAL METHODOLOGIES

3.1 Autoregressive Distributed Lag Bounds Test for Cointegration

The ARDL Bounds test for cointegration was developed by Pesaran and Shin (1998) and Pesaran *et al* (2001). The method does not require pretesting the series for the presence of unit root as long as the series are not I(2) in nature. It can also simultaneously handle series of both I(0) and I(1) nature and hence is more robust than the commonly used Johansen Juselius technique which requires all series to be integrated of the same order. Hence it does not require the testing of unit roots, which can give erroneous results in the presence of structural breaks. Also, unlike the Johansen Juselius technique, the ARDL method is not sensitive to the ordering of the variables, as all variables are tested as a dependent variable for cointegration with the other variables.

The ARDL model is a general dynamic specification. It regresses the differenced form of the dependent variable on lagged forms of the differences and level of the independent variables. The general form of the regression is as follows,

$$\Delta LSPOT_t = a_{0LSPOT} + \sum_{i=1}^{n} b_{iLSPOT} \Delta LSPOT_{t-i} + \sum_{i=1}^{n} c_{iLSPOT} \Delta LFUT1_{t-i}$$
$$+ \sum_{i=1}^{n} d_{iLSPOT} \Delta LFUT2_{t-i} + \sum_{i=1}^{n} e_{iLSPOT} \Delta LFUT3_{t-i} + \beta_{1LSPOT} LSPOT_{t-1}$$
$$+ \beta_{2LSPOT} LFUT1_{t-1} + \beta_{3LSPOT} LFUT2_{t-1} + \beta_{4LSPOT} LFUT3_{t-1} + \varepsilon_{1t} ....(2)$$

$$\Delta LFUT_t = a_{0LFUT} + \sum_{i=1}^{n} b_{iLFUT} \Delta LSPOT_{t-i} + \sum_{i=1}^{n} c_{iLFUT} \Delta LFUT1_{t-i}$$
$$+ \sum_{i=1}^{n} d_{iLFUT} \Delta LFUT2_{t-i} + \sum_{i=1}^{n} e_{iLFUT} \Delta LFUT3_{t-i} + \beta_{1LFUT} LSPOT_{t-1}$$
$$+ \beta_{2LFUT} LFUT1_{t-1} + \beta_{3LFUT} LFUT2_{t-1} + \beta_{4LFUT} LFUT3_{t-1} + \varepsilon_{2t} ....(3)$$

The equation (3) is the generalised equation for all the future series where LFUT can be substituted with LFUT1, LFUT2 and LFUT3 to get the relevant equations for all the futures prices.

The null hypothesis is then tested, which implies no correlation. An F-statistic is computed to check whether the estimated coefficients for the lagged level variables are significantly different from zero. Essentially, for (2), the null hypothesis is,

H0: $\beta_{1LSPOT} = \beta_{2LSPOT} = \beta_{3LSPOT} = \beta_{4LSPOT} = 0$ (4)

Which is tested against, H1: $\beta_{1LSPOT} \neq \beta_{2LSPOT} \neq \beta_{3LSPOT} \neq \beta_{4LSPOT} \neq 0$ (5)

The computed F-statistic is non standard in nature. This F-statistic is then compared to the critical values for large samples of size greater than 1000, given by Pesaran and Shin (1998), Pesaran *et al* (2001) or for small samples of size less than 100, by Narayan (2005). The critical values decrease as the sample size increases; hence there is a sharp deviation in the values for large and small samples.

Two sets of values are provided, assuming that either all variables are I(0) or I(1) in nature and whether a trend term has been used as an independent in the regression. If the computed F-Statistic value falls outside the band of the critical value and hence is greater than both, then the null hypothesis can be rejected and the presence of cointegration confirmed without regard to the order of integration of the variables. If the F-statistic falls below both the critical values, then null hypothesis cannot be rejected and cointegration is not present and if it falls in the band of critical values then the results are inconclusive.

3.2 Toda Yamamoto Version of Granger Non-causality Test using Augmented VAR

Toda & Yamamoto (1995) developed a method of testing for causality amongst a group of variables which is robust in spite of the presence or absence of cointegration. The method does not require the series to be integrated of the same order and can utilise series with different orders of integration also. Hence it is robust to the results of the unit root tests.

The test estimates a Vector Autoregression with the lag length of ($k+d_{max}$) where "k" is the correct order of the VAR and d_{max} is the maximum order of integration of any of the series. A Modified Wald (MWALD) statistic is then computed testing whether the first "k" coefficients of each equation for each lagged variable in the VAR is significantly different from zero or not. This MWALD statistic follows the normal chi-square distribution with degree of freedom equal to the number of excluded lagged variables. In our case, the VAR can be expanded as,

$$\begin{bmatrix} LSPOT_t \\ LFUT1_t \\ LFUT2_t \\ LFUT3_t \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} + \begin{bmatrix} A_{11,1} A_{12,1} A_{13,1} A_{14,1} \\ A_{21,1} A_{22,1} A_{23,1} A_{24,1} \\ A_{31,1} A_{32,1} A_{33,1} A_{34,1} \\ A_{41,1} A_{42,1} A_{43,1} A_{44,1} \end{bmatrix} \begin{bmatrix} LSPOT_{t-1} \\ LFUT1_{t-1} \\ LFUT2_{t-1} \\ LFUT3_{t-1} \end{bmatrix} + \ldots + \begin{bmatrix} A_{11,k} A_{12,k} A_{13,k} A_{14,k} \\ A_{21,k} A_{22,k} A_{23,k} A_{24,k} \\ A_{31,k} A_{32,k} A_{33,k} A_{34,k} \\ A_{41,k} A_{42,k} A_{43,k} A_{44,k} \end{bmatrix}$$

$$\begin{bmatrix} LSPOT_{t-k} \\ LFUT1_{t-k} \\ LFUT2_{t-k} \\ LFUT3_{t-k} \end{bmatrix} + \begin{bmatrix} A_{11,p} A_{12,p} A_{13,p} A_{14,p} \\ A_{21,p} A_{22,p} A_{23,p} A_{24,p} \\ A_{31,p} A_{32,p} A_{33,p} A_{34,p} \\ A_{41,p} A_{42,p} A_{43,p} A_{44,p} \end{bmatrix} + \begin{bmatrix} LSPOT_{t-p} \\ LFUT1_{t-p} \\ LFUT2_{t-p} \\ LFUT3_{t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \end{bmatrix} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

Where “k” is the correct order of integration and “p” is ($k+d_{max}$). For illustration purposes the null hypothesis to test for the effect of LFUT1 prices on LSPOT prices would be,

H_{01}: $A_{12,1} = A_{12,2} \ldots\ldots = A_{12,k} = 0$, which implies that LFUT1 does not granger cause LSPOT

The Toda & Yamamoto (TY) method tests for causality using the level series which are I(1) in nature and not the difference I(0) series. This allows the method to take into account the long run information contained in the level series which is lost upon differencing when using the VECM method.

Unit root tests and cointegration tests suffer from biases arising from the presence of structural breaks, as discussed by Zapata & Rambaldi (1997) and Clarke & Mirza (2006). The TY method does not depend on the presence of unit root or cointegration, hence is more robust to these biases. Toda & Yamamoto (1995) showed that the MWALD statistic is valid till the time the maximum order of integration is less than the correct lag length. A drawback of the MWALD statistic is its inefficiency in small samples, typically of size below 100. As we are using a much larger sample, this drawback does not impact our results.

3.3 Gonzalo Granger Information Share

There are two commonly used models to calculate the information shared by two or more related markets in price discovery. These are the Hasbrouck (1995) and Gonzalo & Granger (1995) methods. The Hasbrouck method focuses on the variance in return innovation and the Gonzalo & Granger method focuses on the permanent and transitory components in the long term relationship amongst the variables. As discussed by Tse (1999) and Huang (2000), in the presence of contemporaneous correlations in disturbances across markets, the Hasbrouck method gives a wide range of information shares for the same market. This correlation amongst the disturbances across markets is very low in the Hasbrouck (1995) study, as high frequency data (1 second resolution) has been used and hence the information share range is narrow. As we have used daily observations and the disturbances across markets show high correlation, we have decided to use the Gonzalo & Granger method.

We estimate the following Vector Error Correction Model (VECM) as discussed in De Jong (2002),

$$\Delta P_t = \gamma Z_t + A_1 \Delta P_{t-1} + ... + \varepsilon_t \tag{7}$$

Where ΔP_t is a n*1 vector of index and futures returns. 'n' is the number of prices being considered. 'γ' is a n*n-1 matrix of the coefficients of the error correction terms Z_t, which are defined as,

$$Z_t = [LFUT1 – LSPOT, LFUT2 – LSPOT, LFUT3 – LSPOT]' \tag{8}$$

All the error correction terms have been measured from the spot prices as the Toda Yamamoto causality test indicate that the spot index prices cause the futures prices.

Using the γ matrix from the VECM estimation, we find γ^* of size 1*n such that γ^* is orthogonal to γ. We then compute $\beta' = (\gamma^{*\prime}\ \iota)^{-1}\ \gamma^{*\prime}$ where ι is a n*1 vector of 1's. This effectively normalises the vector γ^* so that all its terms add up to one. The vector β now contains the information shares of all the futures and spot prices as each of its components corresponds to the information share of a given price.

4. EMPIRICAL RESULTS

As a first step, the price series were tested for the order of integration. This is useful for computing the maximum order of integration for the Toda-Yamamoto causality test. The unit root tests are also helpful to rule out the presence of series which are I(2) in nature. Augmented Dickey Fuller (Dickey & Fuller (1979)) and Phillip Perron (1988) tests were used. The test revealed that all the series in log form are I(1) in nature and no I(2) series are present. Table II presents the results of the same on both Log series as well as the return series.

Table II: Unit Root Tests Results

Panel A: Tests on Level				
ADF TEST	**LSPOT**	**LFUT1**	**LFUT2**	**LFUT3**
With Trend	-1.755 (0.725)	-1.676 (0.761)	-1.707 (0.747)	-1.751 (0.727)
Without Trend	-0.908 (0.786)	-0.911 (0.785)	-0.896 (0.789)	-0.878 (0.795)
PP TEST	**LSPOT**	**LFUT1**	**LFUT2**	**LFUT3**
With Trend	-1.628 (0.781)	-1.711 (0.746)	-1.733 (0.736)	-1.784 (0.712)
Without Trend	-0.902 (0.788)	-0.911 (0.785)	-0.898 (0.789)	-0.882 (0.794)
Panel B: Tests on First Differences				
ADF TEST	**DLSPOT**	**DLFUT1**	**DLFUT2**	**DLFUT3**
With Trend	-49.375*(0.000)	-51.788*(0.000)	-52.016*(0.000)	-52.280*(0.000)
Without Trend	-49.382*(0.000)	-51.795*(0.000)	-52.024*(0.000)	-52.288*(0.000)
PP TEST	**DLSPOT**	**DLFUT1**	**DLFUT2**	**DLFUT3**
With Trend	-49.273*(0.000)	-51.759*(0.000)	-51.995*(0.000)	-52.265*(0.000)
Without Trend	-49.281*(0.000)	-51.767*(0.000)	-52.003*(0.000)	-52.273*(0.000)

Note: Table indicates the value of the respective test statistics. p-values are in parentheses. * indicates value significant at 5% level of significance.

ARDL Bounds Test was conducted to check for cointegration. The optimal lag order was found to be four by the Akaike Information Criteria (AIC), both with and without the time trend[2]. The residuals were checked for serial correlation and its absence verifies the choice of lag length.

The results are reported in Table III. Cointegration is present, in all the trend inclusive cases, when any of the price series is taken as the dependent variable. This is an indicator of the spot and futures market being cointegrated and the presence of a price discovery mechanism which forces the prices to move in tandem with each other and correct to the differences between them. From these results it is not clear which is the leading or lagging price series as cointegration exists for all series being considered as the dependent variable. The absence of cointegration would have implied that the markets do not move together in the long run and that there was no information sharing happening between these markets.

Table III: ARDL Bounds Test for Cointegration results

F-statistics	**Without a time trend**		**With a time trend**	
F_{LSPOT}(LSPOT | LFUT1, LFUT2, LFUT3)	3.162		5.185*	
F_{LFUT1}(LFUT1 | LSPOT, LFUT2, LFUT3)	5.085*		7.070*	
F_{LFUT2}(LFUT2 | LSPOT, LFUT1, LFUT3)	4.812*		7.215*	
F_{LFUT3}(LFUT3 | LSPOT, LFUT1, LFUT2)	9.551*		12.291*	
***F-critical at 5% level**	**I(0)**	**I(1)**	**I(0)**	**I(1)**
	2.86	4.01	3.47	4.57

Note: Critical Values of F-statistics taken from Pesaran, Shin & Smith (2001)

Cointegration implies the presence of causation, either unidirectional or bidirectional. This presence of directional causality implies a direction of information flow between the markets. We have not used the VECM model to test for causality using the error correction term as we have used the Toda Yamamoto causality test which is more robust as it tests for causality using the level variables instead of their differenced form. Hence the long term information in the level prices is not lost in the Toda Yamamoto method but is lost in the VECM approach.

The Toda Yamamoto Augmented VAR test was conducted with an optimal lag (k) of four as determined by Schwarz information criterion (SIC) and Hannan Quinn Information Criterion (HQ)[3]. It was augmented by a lag of one as the maximal order of integration (d_{max}) in the price series was one. The results are presented in Table IV.

Table IV: Toda Yamamoto Causality Test Results

Dependent Variable	LSPOT	LFUT1	LFUT2	LFUT3
LSPOT	-	5.894(0.207)	6.521(0.163)	5.448(0.244)
LFUT1	35.401*(0.000)	-	5.712(0.221)	7.764(0.100)
LFUT2	25.830*(0.000)	7.940(0.093)	-	5.178(0.269)
LFUT3	20.308*(0.000)	5.343(0.253)	5.076(0.279)	-

Note: Table gives MWALD statistic along with the p-values in parentheses. * indicates significance at the 5% level

The results of the TY Causality test indicate that the null hypothesis of LSPOT does not Granger cause LFUT1 is rejected at the 5% critical level. The same result holds for the other futures. These results show the presence of causality from the LSPOT to LFUT1, LFUT2 and LFUT3. The information flow is from the Nifty index to the future prices and not vice versa. These results indicate that the Nifty Index and the cash market are more efficient than the futures market in India as the cash market prices determine the Nifty Index price which plays the major role in price discovery.

To quantify the contribution of the index and futures to price discovery, the Gonzalo & Granger (1995) information share measure was computed. Following the estimation procedure of De Jong (2002) which is valid in the presence of only one common factor, the number of common factors was estimated. The number of common factors is the difference between the number of prices and the number of cointegrating relationships between them.

The Johansen cointegrating test was performed to estimate the number of cointegrating relationships. Results are presented in Table V.

Table V: Johansen Cointegration Test Results

H0	λ_{max}	5% CV	λ_{trace}	5% CV
$r \leq 0$	482.255*	27.584	780.457*	47.856
$r \leq 1$	245.803*	21.131	298.201*	29.797
$r \leq 2$	51.4678*	14.264	52.398*	15.494
$r \leq 3$	0.931	3.841	0.9301	3.841

Note: * indicates significance at 5% level. λ_{max} and λ_{trace} are the maximum eigenvalue and trace eigenvalue statistics respectively

The results of the Johansen Cointegration Test indicate the presence of 3 cointegrating relationships in these 4 price series. This indicates the presence of only one common factor in these price series (4 price series – 3 cointegrating relationships) and hence supports the usage of the De Jong (2002) process for finding the Gonzalo Granger information shares. Estimation of the VECM (7) was done and the vector β computed. Table VI presents the results.

Table VI : Gonzalo Granger Information Sharing Results

	SPOT	FUT1	FUT2	FUT3
β_j (Information Share)	0.477	0.098	0.322	0.103

Note: β_j denotes the percentage share of each price series in information sharing for price discovery

The results indicate that the Nifty index has the highest contribution to price discovery of 47.7%, which is in line with the results from the causality test. In India, restrictions exist on financial institutions, like mutual funds, banks and insurance companies, to trade only in the cash market. For these institutions, use of derivatives market is allowed only for hedging purposes. This makes the cash market more informationally efficient as compared to the futures markets.

Amongst the three futures, second month futures have the next highest contribution of 32.2%. This indicates that informed investors in the futures markets tend to have longer term views and hence they transact more in the next month futures. Portfolio managers from financial institutions would find next month futures a more suitable hedge as compared to near month futures due to the longer horizon of hedging while saving on costs. Though these futures are less liquid than the near month futures, transacting in the next month future saves the investors roll over costs and basis risk associated with roll over. Hence investors with trading horizon of over a month would prefer to transact in the next month futures instead of the near month futures which makes the next month futures more informationally efficient as compared to the near month and far month futures.

5. CONCLUSION

This study looked at the long term relationship between the stock index S&P CNX Nifty and its futures. The purpose was to analyse the price discovery mechanism and direction of information flow along with the relative importance of the two cash and futures markets in determining equilibrium prices. Cointegration was found between the index and its futures indicating a long term relationship and co-movement in a specified range, supporting the Cost of Carry hypothesis. Tests for Causality indicate the unidirectional flow of information from the index to the futures market, indicating the importance of the cash market in price discovery. Gonzalo Granger information test also indicate that the Nifty index has an almost half share in price discovery, with the next month futures coming second at 32%. This indicates the presence of informed investors in the next month futures market. These results also show that the cash market is more informationally efficient as compared to the futures market in India.

Our results are in line with those by Bohl, Salm & Schuppli (2011) and Yang, Yang & Zhou (2012), but are contrary to the established body of research on developed markets. As highlighted by Bohl, Salm and Schuppli (2011), investor structure is an important constituent of the price discovery function and market efficiency. Our results provide support to this argument, as unlike other developed markets, the Indian futures market is still young and many large institutional investors (mutual funds, banks and insurance companies) are restricted to the cash markets by Securities and Exchange Board of India. This difference in investor structure in an emerging market like India can be the reason for our counterintuitive results.

These results have significance for both policy makers and investors. In order to improve the price discovery participation of the futures market, the restrictions on institutional investors have to be relaxed. They should be allowed to take directional risks in the futures markets as they do in the cash markets, with restrictions to curb speculative activities. For investors, the superiority of the second month futures indicates its importance for construction of quantitative trading systems. As it has a higher contribution to price discovery as compared to the first month futures, it contains more information about the direction of the index.

ENDNOTES

1. By Securities and Exchange Board of India (SEBI) Order, Circular MFD/CIR/011/061/2000 dated 1st February 2000.

2. Findings are not provided due to lack of space and are available upon request.

3. Findings are not provided due to lack of space and are available upon request.

BIBLIOGRAPHY

Bohl, M. T., Salm, C. A., & Schuppli, M. (2011). Price discovery and investor structure in stock index futures. *Journal of Futures Markets, 31*, 282-306.

Brooks, C., Rew, A. G., & 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*, *17*(1), 31-44.

Clarke, J. A., & Mirza, S. (2006). A comparison of some common methods for detecting Granger noncausality. *Journal of Statistical Computation and Simulation*, *76*(3), 207-231.

De Jong, F. (2002). Measures of contributions to price discovery: a comparison. *Journal of Financial Markets*, *5*(3), 323-327.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 427-431.

Ghosh, A. (1993). Hedging with stock index futures: Estimation and forecasting with error correction model. *Journal of Futures Markets*, *13*, 743-752.

Gonzalo, J., & Granger, C. (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics*, 27-35.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *Journal of Finance*, 1175-1199.

Huang, R. D. (2002). The quality of ECN and Nasdaq market maker quotes. *The Journal of Finance*, *57*(3), 1285-1319.

Karmakar, M. (2009). Price discoveries and volatility spillovers in S & P CNX nifty futures and its underlying index CNX nifty. *Vikalpa*, *34*(2), 41-55.

Kawaller, I. G., Koch, P. D., & Koch, T. W. (1987). The temporal price relationship between S&P 500 futures and the S&P 500 index. *Journal of Finance*, 1309-1329.

Narayan, P. K. (2005). The saving and investment nexus for China: evidence from cointegration tests. *Applied Economics*, *37*(17), 1979-1990.

Pesaran, H. H., & Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics letters*, *58*(1), 17-29.

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of applied econometrics*, *16*(3), 289-326.

Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, *75*(2), 335-346.

So, R. W., & Tse, Y. (2004). Price discovery in the Hang Seng index markets: index, futures, and the tracker fund. *Journal of Futures Markets*, *24*, 887-907.

Stoll, H. R., & Whaley, R. E. (1990). The dynamics of stock index and stock index futures returns. *Journal of Financial and Quantitative Analysis*, *25*(04), 441-468.

Toda, H. Y., & Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of econometrics*, *66*(1-2), 225-250.

Tse, Y. K. (1995). Lead-lag relationship between spot index and futures price of the nikkei stock average. *Journal of Forecasting*, *14*(7), 553-563. Wiley Online Library.

Tse, Y. (1999). Price discovery and volatility spillovers in the DJIA index and futures markets. *Journal of Futures markets*, *19*, 911-930.

Wahab, M., & Lashgari, M. (1993). Price dynamics and error correction in stock index and stock index futures markets: A cointegration approach. *Journal of Futures Markets*, *13*, 711-742.

Yang, J., Yang, Z., & Zhou, Y. (2012). Intraday price discovery and volatility transmission in stock index and stock index futures markets: Evidence from China. *Journal of Futures Markets*, *32*, 99-121.

Zapata, H. O., & Rambaldi, A. N. (1997). Monte Carlo evidence on cointegration and causation. *Oxford Bulletin of Economics and Statistics*, *59*, 285-298.