Can Indian indices bear the news of Indian Capital Market?

Dr. Abdul Aziz Ansari

Professor and Head of the Department of Commerce and Business Studies, Jamia Millia Islamia – A central university

and

Samiran Jana

Programme Manager of International Management Institute, New Delhi

Abstract

An index bears the news of a stock market or a country where it exists. So it is a barometer of a country's economy. It is also getting used as information source in many researches. But this could be properly happen if Indian market follow random walk hypothesis. This implies that price manipulation at Indian capital market will not be possible. But since its liberalisation and gloabalisation, Indian capital markets have experienced various kinds of price rigging. Sensex at Bomaby Stock Exchange and S&P CNX Nifty at National Stock Exchange are the main two indices at Indian stock market. Authors have analysed five years daily return (from 1st January 2004 to 31st December 2008) of S&P CNX Nifty and Sensex to check the price behaviour of these two indices. Various statistical tests do not give any definite conclusion about randomness of their return.

Key words: Indices, Run test, Serial correlation test, Augmented Dickey-Fuller test, Variance Ratio test and noise trader's risk.

Every stock price moves for two possible reasons: news about the company (e.g. a product launch, or the closure of a factory, etc.) or news about the country (e.g. explosion of nuclear bombs, or a budget announcement, etc.). The job of an index is to purely capture the second part, the movements of the stock market as a whole (i.e. news about the country). Definition of index says that by proper averaging a certain number very liquid stocks from various key economic sectors of a country, news about the company can be erased. Therefore this diversification helps to mitigate the unsystematic risk of individual stocks. Index also gets revised from time to time to remove illiquid stocks from it. Therefore at a certain point in time the value of an index should properly reflect the news of that stock market, where it exists and by comparing one's own portfolio with the index, an investor can get ideas about his own position. That is why index is a benchmark of fund manager's performance. It is also applied as information source in many economic researches. News is random hence the associate price will also be random. This is called random walk hypothesis. The probability distribution of random variables becomes normal when the amount of time separating them becomes large. Again canceling out of the effect of any information on contained stocks of an index is only possible if it is getting reflected purely and perfectly on these. Therefore the purpose of an index will be properly served if all contained stocks exist in a market, which follow random walk hypothesis. Theoretically it will not be possible to predict or manipulate the return of an index or its contained stocks.

Indian Market:-

Liberalisation was brought to Indian capital market in the beginning of 1990s. Security Exchange Board of India was given statutory power to regulate the capital market. Many new amendments were prescribed to control the price manipulation. The National Stock Exchange, set up in 1994, and the Over the Counter Exchange of India, set up in 1992. The National Securities Clearing Corporation (NSCC) and the National Securities Depository Limited (NSDL) were set up in April 1995 and November 1996 respectively for improved clearing and settlement of dematerialized trading. The Securities Contracts (Regulation) Act, 1956 was amended in 1995-96 for introduction of options trading. Moreover, rolling settlement was introduced in January 1998 for the dematerialized segment of all companies. Technology brought radical changes in the trading mechanism. With automation and geographical spread, Indian stock market participation has increased in many folds. But in 1992 Harshad Mehta scam was one of the first black spot on libaralised and globalised Indian capital market. The consequence was so serious that Bombay Stock Exchange was closed for one month. M. S. Shoes case in 1995 rigged up prices of shares in a huge way and eventually leaded a crash. Once again the market was closed for three days. In December 1995 Bombay Stock Exchange stopped trading of Reliance share for its unfair and unexpected growth. The cases of C. R. Bansali, J. C. Prakash and Anand Rathi in 1997, 1998 and 2001 respectively prove that price manipulation is possible at Indian market. Very recent the case of Ketan Parekh and Satyam are also in this row. Ketan Parekh found out the loopholes of Indian market and hammered to investors. Satyam hided its unhealthy business situation from investors for long time and kept its share price in a very high position.

S&P CNX Nifty and Sensex are the two indices of two main stock markets of India i.e. National Stock Exchange (NSE) and Bombay Stock exchange (BSE) respectively and both of these were containing the stock of Reliance and Satyam since long time. Website of these two stock exchanges describe that these two indices bear the news of both the stock market and also serve as a barometer of Indian economy. This will be only possible if Indian market follow random walk hypotheses. But this will come in front of a question mark if price manipulation happens at Indian capital market. Through various statistical tests price behaviour of S&P CNX Nifty and Sensex has been checked and reasons in that regard has been identified in the current study.

Literature Review

Fama (1965) found that returns of American indices were negatively skewed: more observations were in the left hand (negative) tail than in the right hand tail. In addition the tails were flatter and the peak around the mean was higher than predicted by the normal distribution. This distribution is called "leptokurtic". Turner and Weigel (1990) and Peters (1996) got the same result by using daily S&P index returns from 1928 through 1990 and from January 1928 to December 1989 respectively. Niarchos and Alexarkis (2003) saw that the distributions of the index of banking and industrial sectors returns of Greek Stock Exchange are found to be skewed slightly to the left. Panas (2001), did a study on Athens Stock Market with 13 Greek stocks and saw that skewness and kurtosis were positive. Bera-Jarque (BJ) statistics also reject the null hypothesis of normality. Hasan (2004) used autocorrelation function, unit root test, variance Ratio test and BDS test on Bangladesh all share price index of Dhaka Stock Exchange. He considered data from January 1990 to December 2000 and concluded that price of Dhaka Stock Exchange is not random.

An efficient stock market must ensure rapid information access so that it can instantaneously process the information to reflect on security prices. This information transmission mechanism ensures that the stock returns across all days of the week are equal. No market participant can earn any extra normal profits. Fields (1934) examined Dow Jones Industrial Average (DJIA) for the period 1915-1930 and found that weekend prices were more than rest of the week days for 52 percent of the time, while they were lower only 36 percent of the time. French (1980) examined the calendar time hypotheses that the return generating process was continuous and the expected return for Mondays were three times than expected return for other days of the weeks. Indian market has also experienced anomalies. Using Kruskal-Wallise test on BSE sensitive index between June 1988 and January 1990 Chaudhuri (1991 a) saw that average returns on Mondays were negative and highest returns were on Fridays. With the same test Broca (1992) saw in BSE sensex between 1st April 1984 and 31st December 1989 that all Wednesday's returns were negative and Friday's returns were positive. Arumugam (1999) saw that there were significant negative returns on Mondays in bull phase and significant positive returns on same days in bear phase. Friday returns were positively significant only for the bear phase not for bull phase. He used closing prices of the Bombay Stock Exchange Sensitive Index of 30 scrips from April 4, 1979 to March 31 1997. All these researches prove that identical mean return is not possible for Indian stock investors.

Several authors have used run test and serial correlation test to check the randomness of data of Indian Capital Market. Ray (1976) studied seven daily index series and conducted runs test, serial correlation test and spectral analysis for the period from January 1966 to July 1972 and found that the random walk model held only for iron and steel, and cement industries. Bombay, London and New York stock exchanges were compared by Sharma and Kennedy (1977). They used the runs test and spectral analysis techniques on the monthly data from 1963-73 and found that BSE stocks obey random-walk model. Gupta (1985) used 'week-end closing prices of 39 shares from January 1971 to March 1976 and tested the log random walk model, and found support for the weak form of efficiency using serial correlation and runs tests. Chaudhuri (1991b) using log random walk model, on daily price quotations of 93 actively traded shares for the period January 1988 to April 1990 found that 70 shares were significantly autocorrelated for 1 day lag and 17 shares showed the same behaviour for a lag of 2 days or more. Madhusoodanan (1998) applied varianace ratio test on both Indian individual stocks, BSE sensitive index and S&P CNX Nifty index and saw that data of Indian market was persistent. He concluded that short run prediction will be possible. Singh and Kumar (2009) used variance ratio test on S&P CNX nifty. They used 5 minutes, hourly and daily data from Jaunary 2008 to December 2008 and saw that the variance ratio show mean reversion tendency.

Data:

Daily closing price data of Sensex and S&P CNX Nifty from 1st January 2003 to 31st December 2008 has been used for analysis. Data have been collected from the website of Bomaby Stock Exchange and National Stock Exchange of India. Changes of log normal of the data or return have been taken for calculation. Economic time series are often analyzed after computing their logarithms or the changes in their logarithms. One reason for this is that many economic series, exhibit growth by a certain percentage per year on average; if so, the logarithm of the series tends to grow by a certain percentage per year on average; and the logarithm of the series generally grows approximately linearly. Another reason is that the standard deviation of many economic time series is approximately proportional to its level, that is, the standard deviation is well expressed as a percentage of the level of the series; if so, the standard deviation of the logarithm of the series is approximately constant. In either case, it is useful to transform the series so that changes in the transformed series are proportional (or percentage) changes in the original series, and this is achieved by taking the logarithm of the series.

Test of normality

Test of normality has been done through Jarque-Bera (JB) Test. The statistics is given by:-

$$JB = \frac{T-K}{6}\left[S^2 + \frac{(K-3)^2}{4}\right]$$

Where T is the number of observations, S is a measure of skewness, defined as :

$$S = \frac{\frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})^3}{\sigma^3}$$

Skewness measures the amount of asymmetry in a distribution. If the skewnesss equals zero, the distribution is symmetric; the large the absolute size of the skewness statistics, the more asymmetric is the distribution. A large positive value indicates a long right tail and a large negative value indicates a long left tail. The skewness of a normal distribution which is a symmetrical distribution is zero. K is a measure of kurtosis, defined as:

$$K = \frac{\frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})^4}{\sigma^4}$$

The kurtosis of a random variable is a measure of the thickness of the tail of its distribution relative to those of a normal distribution. A normal random variable has a kurtosis of 3; a kurtosis above 3 indicates fat tails or leptokurtic; that is the distribution has more probability mass in the tails than the normal distribution, Under the null hypothesis of normality, the Jarque-Bera statistic is distributed χ^2, with 2 degrees of freedom if JB> $\chi^2_{(2)}$ then we reject the null hypothesis of normality.

If the computed p value of the JB statistics in an application is sufficiently low, which will happen if the value of the statistics is very different from zero, one can reject the hypothesis that the residuals are normally distributed. But if the p value is reasonably high, which will happen if the value of the statistics is close to zero, we do not reject the normality assumption.

Table 1 Result of the data of S&P CNX Nifty and Sensex

	Nifty	Sensex
Mean	0.000350	0.000392
Median	0.001393	0.001428
Maximum	0.079691	0.079311
Minimum	-0.130539	-0.118092
Standard Deviation	0.018725	0.018408
Skewness	-0.793465	-0.562272
Kurtosis	8.727671	7.648065
Jarque-Bera	1838.349	1190.147
Probability	0.000000	0.000000

Table 1, shows that in both the cases the skewness are not zero. It is negative. Therefore more data are in the left hand side of the distribution. Kurtosis is more then 3. Therefore it is leptokurtic and tail is fat. Probability of JB statistics is zero for both the indices. Therefore we can reject the null hypothesis of normal distribution of data in both the cases. Randomness of the data has been checked through run test and serial correlation test.

Run test

Mean: $E(k)=\frac{2n_1n_2}{n_1+n_2}+1$ $\qquad \sigma_k^2=\frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}$

If the hypothesis of randomness is sustainable, we should expect k, the number of runs obtained in a problem, to lie between $E(k)\pm Z\sigma_k$. If number of obtained runs is R then

$$|Z|=\frac{R-E(k)}{\sigma_k}$$

$n=$ total number of observations $=n_1+n_2$

$n_1=$ number of + ve symbols (i.e., + residuals)

$n_2=$ number of – ve symbols (i.e., - residuals)

$k=$ number of runs Gujarati (1995)

The null hypothesis i.e. randomness hypothesis will be rejected at 1, 5 and 10 percent level of significance in favor of alternative hypothesis (non-randomness hypothesis) if observed value of $|Z|$ is greater then 2.58, 1.96 and 1.65 respectively and vice versa.

Calculated Z value is less then the critical Z value in case of S&P CNX Nifty even at ten percent significance level (table 2). So data of S&P CNX nifty is random. But in case of sensex the calculated Z value is significant at five percent significance level. So run test is indicating that data of sensex is not random.

Table 2 Run test result:

Indices	$\lvert Z \rvert$ values	Expected number of runs	Actual number of runs	Standard Error
Sensex	2.498**	618.630	574	17.469
S&P CNX Nifty	1.473	618.653	593	17.413

** implies the significance at 5% level.

Nonparametric tests ignore a certain amount of information. They are unable to capture huge fluctuation in number. The estimate of an interval at the ninety five percent confidence level using a nonparametric test may be twice as large as the estimate using a parametric test. So parametric test has been used to get proper information. One of the parametric test is Serial Correlation Test.

Serial correlation test

Serial correlation coefficients provide a measure of relationship between value of a random variable (X_t) in time t and its value k periods earlier, where $k = 1, 2, 3, \ldots\ldots\ldots\ldots\ldots, n$. In the present study we have considered time lags of 12 days. The autocorrelation matrix is estimated by:

$$\rho_k = c_k / c_0$$

Wherein,

$$c_k = \frac{1}{n}\sum_{t=1}^{n-k}(X_t - \mu)(X_{t+k} - \mu)$$

$k = 1, 2, 3, \ldots\ldots\ldots\ldots\ldots$, n

$$\mu = \frac{1}{n}\sum_{t=1}^{n} X_t$$

Wherein,

c_0 = Variance of (X_t), and

n = number of observation.

The test of significance of ρ_k (against the null hypothesis of population ρ_k being equal to zero) can be carried out by computing z statistics given by

$$|Z| = \rho_k \sqrt{n}$$

If the estimated $|Z| \geq 1.96$, we can draw the inference that ρ_k is significantly different from zero at 5 per cent level, and, hence, price changes are not serially independent (or random).

Instead of testing the significance of any individual autocorrelation matrix, a joint hypothesis has been tested that all ρ_k upto certain lags are simultaneously equal to zero. This is obtained by using Q statistics developed by Box and Pierce (1970), and is defined as:

$$Q = n\sum_{k=1}^{m} \rho_k^{\ 2}$$

Wherein,

ρ_k is autocorrelation coefficient at lag k,

n = sample size   m = lag length

The Q statistics is often used as a test of whether a time series is white noise. In large samples, it is approximately distributed as a chi-square distribution with m degree of freedom. In case the computed Q exceeds its critical value on the chi-square distribution at a given significance level (0.01, 0.05), null hypothesis is rejected, that is all (true) ρ_k are nonzero; atleast some of them must be nonzero. For a large sample Lunge – Box statistics follows chi-square distribution with m degree of freedom:

Lung-Box statistic $= n(n+2)\sum_{k=1}^{m}\left(\rho_k^{\,2} / n-k\right) \sim \chi^2$ m

Table 3 Serial Correlation test result

	Sensex			S&P CNX Nifty		
Lags	Autocorrela-tions	Z values	Q Values	Autocorrela-tions	Z Values	Q Values
K=1	0.072	2.534**	6.4378 **	0.069	2.445**	5.9941 **
K=2	-0.066	-2.33**	11.900*	-0.064	-2.251**	11.078*
K=3	-0.004	-0.889	11.916*	-0.004	-0.144	11.099 **
K=4	-0.009	-1.751	12.028 **	0.000	0.002	11.099 **
K=5	-0.025	-0.889	12.821 **	-0.022	-0.749	11.732 **
K=6	-0.049	-1.751	15.895 **	-0.053	-1.887	15.305 **
K=7	0.008	0.283	15.975 **	0.013	0.459	15.517 **
K=8	0.061	2.145**	20.600*	0.054	1.896	19.131 **
K=9	-0.011	-0.379	20.743 **	-0.012	-0.436	19.322 **
K=10	0.019	0.682	21.210 **	0.020	0.713	19.831 **
K=11	-0.022	-0.793	21.840 **	-0.027	-0.960	20.754 **
K=12	0.012	0.413	22.012 **	0.005	0.171	20.784

* implies that the value is significant at 1 percent significance level and ** implies that the value is significant at 5 percent significance level. K is the lag length.

Calculated Z are significant at some lags of Sensex and S&P CNX Nifty (table 3). Same way calculated Q statistics is greater then critical values of χ^2 at various lags for both the indices. Some of them are significant at one percent confidence level. This implies that some ρ_k s are not zero. This test implies that data of both the indices are not random. Serial correlation test gives more perfect and concluding result then the run test. Serial correlation zero would imply that price changes in consecutive time periods are uncorrelated with each other viewed as a rejection of hypothesis that investors can learn about future price changes from past ones. This has happened in all lags except 1 and 2 in both the indices. Serial correlation in lag 1 is positive and statistical significant. This suggests that returns in this period are more likely to be positive (negative) if the prior period returns were positive (negative). Similarly, negative and statistically significant serial correlation in lag 2 could be evidence of price reversals and would be consistent with a market where positive returns were to follow negative returns and vice versa.

If data of indices are random and normally distributed then the probability distribution will not change overtime. This implies that data should be stationary. The random walk model is represented by a first order autoregressive process, i.e. AR(1):

$$\log P_t = \rho \log P_{t-1} + \varepsilon_t$$

Where P_t signifies the natural log of the price index series as a proxy of share prices: ε_t is a classical error-term. The random walk model hypothesis that $\rho = 1$. The autoregressive process is stationary if and only if $|\rho| < 1$. Stationarity has been checked using Augmented Dickey-Fuller unit root test. The results (table 4 and 5) indicate that, regardless of the lag length, the null hypothesis of a unit root can not be rejected for the logarithm variable, P. In contrast, similar results applied to the continuously compounded stock return series as measured by the logarithmic first difference of stock prices indicate that the null hypothesis of unit root is rejected at least at the 5% level. This implies that data series are non-stationary in level but not in the first difference form.

Table 4 Result of Augmented Dickey-Fuller Test on Sensex

	Level		First difference	
	t_μ	t_λ	t_μ	t_λ
K=0	-1.273	0.06632	-32.862*	-32.899*
K=1	-1.261	-0.181384	-25.818*	-25.870*
K=2	-1.237	0.07031	-20.663*	-20.725*
K=3	-1.267	0.0690	-18.069*	-18.139*
K=4	-1.257	0.1232	-16.460*	-16.532*
K=5	-1.206	0.1879	-15.627*	-15.711*
K=6	-1.1961	0.3543	-14.083*	-14.174*
K=7	-1.2103	0.3298	-12.386*	-12.474*
K=8	-1.2024	0.15323	-11.898*	-11.991*
K=9	-1.185	0.2192	-10.927*	-11.029*
K=10	-1.2298	0.1550	-10.743*	-10.860*
K=11	-1.2528	0.2862	-10.076*	-10.187*
K=12	-1.2208	0.2081	-9.134*	-9.254*

Table 5 Result of Augmented Dickey-Fuller Test on S&P CNX Nifty

	Level		First difference	
	t_μ	t_λ	t_μ	t_λ
K=0	-1.231	-0.303	-32.944*	-32.966*
K=1	-1.235	-.05440	-25.785*	-25.816*
K=2	-1.1967	-0.3139	-20.67118*	-20.711*
	Level		First difference	
K=3	-1.222	-0.3104	-17.8920*	-17.938*
K=4	-1.228	-0.2862	-16.3108*	-16.3575*
K=5	-1.1742	-0.2412	-15.557*	-15.6119*
K=6	-1.1550	-0.0732	-13.951*	-14.0127*
K=7	-1.1802	-0.1095	-12.382*	-12.442*
K=8	-1.1822	-0.2622	-11.883*	-11.946*

K=9	-1.1641	-0.2049	-10.9171*	-10.987*
K=10	-1.2084	-0.2670	-10.789*	-10.872*
K=11	-1.2272	-0.1160	-10.190*	-10.270*
K=12	-1.1996	-0.1734	-9.205*	-9.292*

t_μ and t_λ are the t statistics based on Augmented Dickey-Fuller regression with allowance for a constant and constant and trend respectively. K signifies the lag length of ADF regression. * implies the significance level at 5%.

Variance ratio test has a clear edge over ADF unit root test, because ADF test is one sided test. The interval estimated by ADF test is much higher than normal, this is called size limitation of ADF test and the power of ADF test depends on the time span not on sample. Persistency in the data has been checked through variance ratio test. It is given as:

$$R\ (q) = (1/q)\left[VarP_{t-q} - P_t\right] / Var(P_{t-1} - P_t)$$

$R\ = 1$ would suggest a random walk and $R\ > 1$ implies persistency and $R\ < 1$ is for mean reversion.

Table 6: Result of variance test ratio

	S&P CNX Nifty		Sensex	
Lag	log P	Δ log P or retrun	log P	Δ log P or return
K=1	1.000	1.000	1.000	1.000
K=2	1.071	0.572	1.073	0.501
K=3	1.052	0.360	1.054	0.384
K=4	1.042	0.269	1.044	0.271
K=5	1.036	0.220	1.035	0.218
K=6	1.026	0.189	1.020	0.185
K=7	1.003	0.152	0.996	0.163
K=8	0.989	0.130	0.980	0.135
K=9	0.991	0.122	0.980	0.113
K=10	0.991	0.106	0.980	0.110
K=11	0.994	0.090	0.983	0.097
K=12	0.993	0.090	0.983	0.093

Result in table 6 shows in case of Sensex and S&P CNX Nifty till 6^{th} and 7^{th} lag respectively are persistent and after that it is mean reversion. Though Log P are randomly distributed because variance ratios are not significantly different from 1 but it is not true in the case of Δ log P or return. This implies that short run prediction is possible in both the indices.

Concluding Remarks

Above tests results shows that data of Sensex and S&P CNX Nifty are not random. Based on free float market capitalization weight these two indices were made. This methodology has been followed to prepare most of the world indices. On any given day, there would be good stock-specific news for a few companies and bad stock-specific news for others. Theory says that free float market capitalization weighting methodology helps to cancel out stock specific news.

Human psychology says that Investor's behaviour may play a opposite role is this place. According to past experiences, investors show cognitive biases on 'best' and 'worst' stocks. Investors will get negative shock if 'best' stocks are giving lower return than investor's expectations and positive shock comes if the 'worst' stocks are giving higher return. Psychology says that negative shocks have higher effects than positive shocks. Therefore investors will buy or sell more 'best' stocks than 'worst' ones. Optimism is another psychological bias and makes investors underestimate the probabilities of bad outcomes. The traders, whose mental process gets affected with biases, are termed as noise traders. Theory says that those, whose decision process is free from mental biases, are called rational traders. These noise traders will give more weight to such information, which have no relevance. This will be more common when information availability is not enough at a stock market. This is very common phenomenon at India. Hence noise traders will wait or follow each other's mistakes. Evil market makers take this opportunity and start price rigging. Classical finance theory says that rational traders will come in this time and start arbitrage and erase the effect of noise traders from ultimate price and random walk will exist at stock market. Financial psychology says that these mental biases also affect the decision process of rational traders because they are also human being. As a result arbitrage will not be always possible and market will not follow random walk.

Therefore free float market capitalization weight will not be helpful to mitigate fully the effect of news on content stocks of Indian indices and the purpose of making index at Indian market will not be served properly.

This study has considered only the daily closing price of S&P CNX Nifty and Sensex. Weekly, monthly and yearly return of these indices may follow randomwalk hypothesis. So future research on these areas may give clearer ideas.