The Random-Walk Hypothesis on the Indian Stock Market

Ankita Mishra,* Vinod Mishra+ and Russell Smyth✦

Abstract
This study tests the random walk hypothesis for the Indian stock market. Using 19 years of monthly data on six indices from the National Stock Exchange (NSE) and the Bombay Stock Exchange (BSE), this study applies three different unit root tests with two structural breaks to analyse the random walk hypothesis. We find that unit root tests that allow for two structural breaks alone are not able to reject the unit root null; however, a recently developed unit root test that simultaneously accounts for heteroskedasticity and structural breaks, finds that the stock indices are mean reverting. Our results point to the importance of addressing heteroskedasticity when testing for a random walk with high frequency financial data.

Keywords: India, Unit root, Structural Break, Stock Market, Random Walk
JEL codes: G14, C22

* School of Economics, Finance and Marketing, RMIT University.
+ Department of Economics, Monash University.
✦ Department of Economics, Monash University.
* Address for correspondence: Russell Smyth, Department of Economics, Monash University 3800 Australia. Email: russell.smyth@monash.edu. Telephone: +613 9905 1560. We thank Paresh Narayan for providing the GAUSS codes for the Narayan and Popp (2010) and Narayan and Liu (2013) unit root tests.

© 2014 Ankita Mishra, Vinod Mishra and Russell Smyth
All rights reserved. No part of this paper may be reproduced in any form, or stored in a retrieval system, without the prior written permission of the author.
1 Introduction

There has been a considerable amount of research in the academic literature related to the predictability of stock prices, but there are still no definite agreed upon conclusions. Starting with the influential article by Fama (1970) on various forms of market efficiency, a great deal of research has focussed on the question “are financial markets efficient?”

The reason why so much research has focussed on the question of market efficiency is because of the important implications conveyed by the efficient market hypothesis. If markets are efficient, then prices fully reflect all the information present in the market and, hence, there is no scope for making profits using either technical analysis or fundamental analysis of financial markets. If we assume that financial markets are efficient, then an expert will not do any better than a layman in predicting price moments. This idea was more strongly put by Malkiel (2003) who indicated that in an efficient market an investor cannot earn returns greater than those that could be obtained by holding a randomly selected portfolio of individual stocks (with a comparable level of risk).

Apart from the implications related to making profits, the efficient market hypothesis has many implications for individual investors and policy makers. The basic role of a stock market is to efficiently allocate resources in the economy by converting savings into useful investments. If markets are efficient and participants have full information about all companies (reflected in the price of the corresponding stock), then the stock market will allocate the investment to the most efficient outcome and individual investors can invest in the market without much uncertainty about their investment. However, if markets are not efficient, and prices do not reflect all information present in the market, then the efficiency of this resource allocation mechanism becomes questionable. This, in turn, suggests a bigger role for regulatory mechanisms to protect the investments of individual investors.

There are many empirical studies which have looked at the issue of market efficiency for different stock markets and the recent literature on the efficient market hypothesis has given increasing attention to emerging markets. Some studies of the efficient market hypothesis in emerging stock markets are Chaudhuri and Wu (2003; 2004), Gough and Malik (2005), Cooray and Wickremasinghe (2008), Lean and Smyth (2007) and Phengpis (2006). Lim and Brooks (2011) provide a comprehensive review the relevant literature.
There are two reasons why the recent literature has focussed on emerging markets. First, emerging markets are not characterised by a very well developed information disbursement mechanism. Hence, any news, after its release, may reach different groups of investors at different points in time. This lead-lag relationship between the news and its reception may temporarily make some investors better informed than others, thus creating possibilities for one group of investors to make supra-normal profits. Second, the emerging markets are often not characterised by a well-developed institutional infrastructure to regulate financial markets. A sound institutional arrangement is a necessary requirement for an efficient market; if not always, then at least during the initial development phase of the market.

One of the most commonly adopted approaches to test for the random-walk hypothesis is testing for the presence of a unit root in stock prices. The reasoning behind this approach is that the presence of a unit root suggests that shocks to prices are permanent i.e. any movement in prices permanently changes the price path. This implies that price movements are due to random shocks, which cannot be predicted, thus making it impossible to predict future price movements based on information about past prices. However, if a unit root is not present, this implies that prices are stationary and will revert to their natural mean over time, thus making it possible to forecast future price movements using past data.

The Dickey and Fuller (1979) unit root test has traditionally been used to test for the presence of a unit root. But later studies by Perron (1989) and Zivot and Andrews (1992), demonstrated that the Dickey-Fuller test fails to take into account a structural break in the series and commits a type two error by identifying a stationary series (with a structural break in slope or intercept) as being non-stationary. Since the idea was first suggested by Perron (1989), several tests have been developed to test for a unit root in the presence of structural breaks. These include tests for the presence of a unit root with one or two endogenous structural breaks proposed by Zivot and Andrews (1992), Lumsdaine and Papell (1997), Lee and Strazicich (2003a; 2003b), Narayan and Popp (2010) and Narayan and Liu (2013).

The objective of this study is to examine the efficient market hypothesis in the Indian stock market. To do so, we employ three unit root tests with two structural breaks. These tests are the Lee and Strazicich (2003b) lagrange multiplier (LM) unit root test as well as the recent tests proposed by Narayan and Popp (2010) and Narayan & Liu (2013).
This study extends the existing literature on testing the efficient market hypothesis in emerging stock markets in two ways. First, we contribute to the burgeoning literature pertaining to the efficiency of the Indian stock market. Several recent studies have used unit root tests to examine the efficient market hypothesis in the Indian stock market (see eg. Ahmed et al., 2006; Ali et al., 2013; Alimov et al., 2004; Gupta & Basu, 2007; Jayakumar et al., 2012; Jayakumar & Sulthan, 2013; Kumar & Maheswaran, 2013; Kumar & Singh, 2013; Mahajan & Luthra, 2013; Mehla & Goyal, 2012; Srivastava, 2010). However, these studies typically employ the Dickey-Fuller and/or Phillips-Perron unit root tests without structural breaks (see eg. Ahmed et al., 2006; Ali et al., 2013; Alimov, et al. 2004; Gupta & Basu, 2007; Jayakumar et al., 2012; Jayakumar & Sulthan, 2013; Kumar & Singh, 2013; Mahajan & Luthra, 2013; Mehla & Goyal, 2012; Srivastava, 2010). Moreover, some of the studies use short spans of data (see eg. Ali et al., 2013; Jayakumar & Sam, 2012). The failure to account for structural breaks and the use of short spans of data impair the reliability of existing findings. We advance the literature on the efficiency of the Indian stock market by employing unit root tests, which allow for up to two breaks, have better size and power and, in the case of the Narayan and Liu (2013) test, address the presence of heteroskedasticity.

The second way in which we extend the literature is to contribute to studies of the efficient market hypothesis in stock markets in emerging markets more generally. Existing studies related to efficiency of emerging markets have reached mixed results; for example, Chaudhuri and Wu (2003) found that the random-walk hypothesis was rejected for ten out of the 14 emerging markets considered. Phengpis (2006), using the same dataset as Chaudhuri and Wu (2003), showed that their results weaken when a different methodology for testing the random-walk hypothesis is adopted. Chaudhuri and Wu (2003) followed the Zivot and Andrews (1992) method for determining a structural break whereas Phengpis (2006) used the Lee and Strazicich (2003a) approach. Narayan and Smyth (2005) used data on stock prices for 22 OECD countries, and employed the Zivot and Andrews (1992) sequential trend break test and the Im, Pesaran et al. (2003) panel unit root test. They found that OECD prices follow a random walk, in spite of the presence of significant structural breaks in the data. Narayan and Smyth (2007) examined G7 stock price data using the Lumsdaine and Papell (1997) and Lee and Strazicich (2003a; 2003b) tests and found that the random-walk hypothesis is supported for all the G7 countries except for Japan. A general consensus is that stock prices in developed markets are found to exhibit a random walk, whereas there are no definite conclusions regarding stock prices in emerging markets (Nurunnabi, 2012).
The tests that we use have been shown to have better power and size than those used in most of the existing literature. The LM unit root test with two breaks developed by Lee and Strazicich (2003b) represents a methodological improvement over the Dickey-Fuller-type endogenous two break unit root test proposed by Lumsdaine and Papell (1997), which has the limitation that the critical values are derived while assuming no breaks under the null hypothesis. Narayan and Popp (2013) show that the Narayan and Popp (2010) test has better size and higher power, and identifies the breaks more accurately, than either the Lumsdaine and Papell (1997) or Lee and Strazicich (2003b) tests. The recent generalized autoregressive conditional heteroskedasticity (GARCH) unit root with two structural breaks, proposed by Narayan and Liu (2013), has the advantage that it models heteroskedasticity and structural breaks simultaneously. To the best of our knowledge, the only application of the Narayan and Liu (2013) test to the efficient market hypothesis in stocks is the application in the Narayan and Liu (2013) paper itself, which is to stocks on the New York Stock Exchange.1

The rest of this paper is organised as follows. The next section discusses testing for a unit root in the context of the efficient market hypothesis. The third section provides an overview of the three unit root tests that we employ. The fourth section introduces the dataset, outlines some stylised facts about the indices and presents findings from some preliminary analysis. The empirical results are presented in the fifth section. Our key finding is that the Lee and Strazicich (2003b) and Narayan and Popp (2010) tests are unable to reject the unit root null in the stock price indices, but the Narayan and Liu (2013) test rejects the unit root for each of the indices. This result points to the importance of addressing heteroskedasticity in high frequency financial data when testing for a unit root. The final section summarises the conclusions, and significance, of the study in the context of the efficient market hypothesis.

2 Different Concepts of Market Efficiency

Fama (1970) described three forms of market efficiency subject to three different information sets. A market where future prices cannot be predicted using past historical price data exhibits weak-form market efficiency. Under the semi-strong form of market efficiency prices instantaneously adjust to other relevant information that is publicly available. The strong form takes the theory of market efficiency to the ultimate extreme and suggests that,
even if some investors have monopolistic access to any information relevant for price formation, that this will not help to predict future prices. The definition of the strong form of market efficiency is ambiguous and, in general, it is not possible to empirically test this hypothesis. Hence, the existing literature has primarily focused on testing the weak and semi-strong forms of market efficiency using stock price indices.

A random walk, as suggested by Malkiel (2003), “is a term loosely used in the finance literature to characterise a price series where all subsequent price changes represent random departures from previous prices”. The idea behind the random-walk model is to suggest that all information present in the market is immediately reflected in the price, so today’s news affects only today’s prices, and so on. The news, by its definition, is unpredictable, thus making price changes unpredictable and random. So the broad implication of the random-walk hypothesis is that an uninformed investor purchasing a diversified portfolio will obtain, on average, a rate of return as generous as achieved by an expert.

One approach to test the random-walk hypothesis is to test for the presence of a unit root in stock prices. The basic random-walk hypothesis is that stock prices contain a unit root such that shocks to the stock market are permanent and every random shock moves the price process to a new path. The alternative hypothesis is that shocks are transitory and the price process returns to its mean value over time. If prices are mean reverting, then future movements can be predicted from past values, thus rejecting the random-walk hypothesis.

3. Unit Root Tests in the Presence of Structural Breaks

The first tests for examining the presence of a unit root in a time series were proposed by Dickey and Fuller (1979) and Phillips and Perron (1988). However, the results of applied studies based on these tests, such as Nelson and Plosser (1982), were questioned by Perron (1989) and Zivot and Andrews (1992) on the basis that these tests had low power to reject the unit root null in the presence of structural breaks in the data. Subsequent studies have focussed on refining the methodology of unit root testing in the presence of structural breaks. While there is a whole plethora of tests for a unit root in the presence of one or two structural breaks, in the current study we focus on three of the most recent and robust tests. In the following subsections, we briefly describe the methodology of each of these tests.
Lee and Strazicich (2003b) LM unit root test with two structural breaks

Lee and Strazicich (2003b) suggested using a minimum LM test for testing the presence of a unit root with two structural breaks. The minimum LM test can be specified using the following data-generating process (DGP) (using the notation in Lee and Strazicich, 2003b):

\[ y_t = \delta Z_t + X_t, \quad X_t = \beta X_{t-1} + \varepsilon_t \]

such that \( Z_t \) is a matrix containing exogenous variables and \( \varepsilon_t \sim iid \ N(0, \sigma^2) \). The null of a unit root is given by \( \beta = 1 \). If \( Z_t = [1, t] \) the DGP is reduced to the minimum LM test proposed by Schmidt and Phillips (1992), in which the series in the data under question is characterised by an intercept and a trend, but no structural break. The model with two structural breaks in the intercept (Model AA) is given by the following specification of \( Z_t \):

\[ Z_t = [1, t, D_{1t}, D_{2t}] \]

where \( D_{1t} = 1 \) for \( t \geq T_{B1} + 1 \) and 0 otherwise, \( D_{2t} = 1 \) for \( t \geq T_{B2} + 1 \) and 0 otherwise, while \( T_{B1} \) and \( T_{B2} \) are the breaks in the intercept. The null and alternative hypotheses are given as:

\[ H_0 : y_t = \mu_0 + d_1 B_{1t} + d_2 B_{2t} + y_{t-1} + \nu_{1t} \]
\[ H_A : y_t = \mu_1 + \gamma t + d_1 D_{1t} + d_2 D_{2t} + \nu_{2t} \]

such that \( \nu_{1t} \) and \( \nu_{2t} \) are stationary error terms. It is clear from the above expressions that both the null and the alternative hypothesis include breaks. The model with two breaks in the intercept as well as trend (Model CC) is given by the following specification of \( Z_t \):

\[ Z_t = [1, t, D_{1t}, D_{2t}, DT_{1t}, DT_{2t}] \]

where \( DT_{1t} = t - T_{B1} \) for \( t \geq T_{B1} \) and 0 otherwise, \( DT_{2t} = t - T_{B2} \) for \( t \geq T_{B2} \) and 0 otherwise. The expressions for the null and alternative hypothesis in this case are given as:

\[ H_0 : y_t = \mu_0 + d_1 B_{1t} + d_2 B_{2t} + d_3 D_{1t} + d_4 D_{2t} + y_{t-1} + \nu_{1t} \]
\[ H_A : y_t = \mu_1 + \gamma t + d_1 D_{1t} + d_2 D_{2t} + d_3 DT_{1t} + d_4 DT_{2t} + \nu_{2t} \]

The LM test statistic is estimated by the following regression:
\[ \Delta y_t = \delta' \Delta Z_t + \phi \tilde{S}_{t-1} + u_t \]

where \( \tilde{S}_t = y_t - \tilde{\psi} - Z_t \tilde{\delta} \), \( t = 2, \ldots, T \); \( \tilde{\delta} \) is the vector of coefficients of regression of \( \Delta y_t \) on \( \Delta Z_t \). \( \tilde{\psi} = y_1 - Z_1 \tilde{\delta} \) such that \( y_1 \) and \( Z_1 \) are the first observations of \( y_t \) and \( Z_t \), respectively. The null of a unit root is given by \( \phi = 0 \), and the LM test statistics are given by:

\[ \tilde{\rho} = T \tilde{\phi} \]

\[ \tilde{\tau} = t - \text{statistics testing the null hypothesis } \phi = 0 \]

The location of break points is determined endogenously by conducting a grid search to locate the minimum t-statistics, given by the following rule:

\[ LM_\rho = \inf_{\lambda} \tilde{\rho}(\lambda) \]

\[ LM_\tau = \inf_{\lambda} \tilde{\tau}(\lambda) \]

The grid search is carried out over a trimming region, given by \([kT, (1-k)T]\) to eliminate the break points. There is no general rule for deciding \( k \). We took the value \( k = 0.05 \), trimming 5 per cent of data points at each end of the series. The critical values for the test are tabulated in Lee and Strazicich (2003b). It is to be noted that the critical values for the model with breaks in intercept and trend are dependent on the location of the breaks (i.e. \( \lambda_1 \) and \( \lambda_2 \)).

**Narayan and Popp (2010) unit root test with two structural breaks**

Narayan and Popp (2010) proposed a new test for a unit root that allows for at most two structural breaks in the level and trend of the data series. They argued that Augmented Dickey-Fuller (ADF)-type unit roots tests which do not allow for a break under the null hypothesis or model the break as an innovation outlier (IO) suffer from severe spurious rejections in finite samples when a break is present under the null hypothesis. Narayan and Popp (2010) proposed a new ADF-type unit root test for the case of IOs where the problem of spurious rejection could be avoided by formulating a DGP as an unobserved components model, in which breaks are allowed to occur under both the null and alternative hypotheses.

The DGP of a time series \( y_t \) has two components; a deterministic component (\( d_t \)) and a stochastic component (\( u_t \)). Narayan and Popp (2010) used two different specifications for the
deterministic component; one allows for two breaks in the level, denoted as model 1 (M1) and the other allows for two breaks in the level as well as slope of the deterministic trend component, denoted as model 2 (M2). The test equations for the two models are:

\[ y_{t}^{M1} = \rho y_{t-1} + \alpha_{i} + \beta^{*}t + \theta_{1}D(T_{B})_{1,t} + \theta_{2}D(T_{B})_{2,t} + \delta_{1}DU_{1,t-1} + \delta_{2}DU'_{2,t-1} + \sum_{j=1}^{k} \beta_{j} \Delta y_{t-j} + e_{t} \]

With \( \alpha_{i} = \Psi^{*}(1)^{-1}[(1 - \rho)\alpha + \rho \beta] + \Psi^{*}(1)^{-1}(1 - \rho)\beta, \Psi^{*}(1)^{-1} \) being the mean lag, \( \beta^{*} = \Psi^{*}(1)^{-1}(1 - \rho)\beta, \phi = \rho - 1, \delta_{i} = -\phi \theta_{i} \) and \( D(T_{B})_{1,t} = 1(t = T_{B,i} + 1), i = 1,2. \)

\[ y_{t}^{M2} = \rho y_{t-1} + \alpha^{*} + \beta^{*}t + \kappa_{1}D(T_{B})_{1,t} + \kappa_{2}D(T_{B})_{2,t} + \delta_{1}^{*}DU_{1,t-1} + \delta_{2}^{*}DU'_{2,t-1} + \gamma_{1}^{*}DT_{1,t-1} + \gamma_{2}^{*}DT_{2,t-1} + \sum_{j=1}^{k} \beta_{j} \Delta y_{t-j} + e_{t} \]

Where, \( \kappa_{i} = (\theta_{i} + \gamma_{i}), \delta_{i}^{*} = (\gamma_{i} - \phi \theta_{i}) \), and \( \gamma_{i}^{*} = -\phi y_{i}, i = 1,2. \)

Since break dates \( (T_{B,i}) \) are unknown and have to be estimated, Narayan and Popp (2010) used a sequential procedure\(^2\) along the lines proposed by Kapetanios (2005), to derive their estimates. The unit root null hypothesis of \( \rho = 1 \) is tested against the alternative hypothesis of \( \rho < 1 \). The critical values of the test statistics and the results of Monte-Carlo simulations for the size and power of the test are given in Narayan and Popp (2010). In the current study we test for both M1 and M2 specifications; thereby, allowing for the possibility of a break in the intercept only and a break in both the intercept and trend of the series.

**Narayan and Liu (2013) GARCH unit root test**

Besides structural breaks, high frequency time series data is also characterized by heteroskedasticity. The earlier studies on structural break unit root tests are based on standard linear models that are inappropriate for modelling unit roots in the presence of heteroskedasticity. Narayan and Liu (2013) relax the assumption of independent and identically distributed errors and propose a GARCH(1,1) unit root model that accommodates two endogenous structural breaks in the intercept in the presence of heteroskedastic errors. The test considers a GARCH (1, 1) unit root model of the following form:

\(^2\) For details on the sequential procedure approach, refer to Narayan and Popp (2010)
\[ y_t = \alpha_0 + \pi y_{t-1} + D_1 B_{1t} + D_2 B_{2t} + \varepsilon_t \]

Here, \( B_{it} = 1 \) for \( t > T_{Bi} \) otherwise \( B_{it} = 0 \), \( T_{Bi} \) are structural break points, where \( i = 1,2 \). \( D_1 \) and \( D_2 \) are break dummy coefficients. \( \varepsilon_t \) follows the first order GARCH (1,1) model of the form:

\[ \varepsilon_t = h_t \eta_t, h_t = \kappa + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \]

Here, \( \kappa > 0, \alpha \geq 0 \) and \( \beta \geq 0 \), and \( \eta_t \) is a sequence of independently and identically distributed random variables with zero mean and unit variance. To estimate these equations Narayan and Liu (2013) used joint maximum likelihood (ML) estimation. Since break dates \( (T_{Bi}) \) are unknown and have to be substituted by their estimates, a sequential procedure is used for deriving the estimates of break dates. The unit root hypothesis is tested with the ML \( t \)-ratio for \( \pi \) with a heteroskedastic-consistent covariance matrix.

4 Data and Preliminary Analysis

The data for this study consists of six indices from the two main stock exchanges of India. India has around 20 stock exchanges located in various cities. However, most of the trading is concentrated in the National Stock Exchange (NSE) and the Bombay Stock Exchange (BSE), which are both located in Mumbai, the financial capital of India.

The NSE was established in 1995 and according to the latest figures (as of July 2013) it has 1,685 securities listed on its capital market segment, which are available for trade. NSE is the largest stock exchange in India, accounting for roughly 83 per cent of the entire market turnover in India, with BSE the second largest, accounting for around 13 per cent with the remaining stock exchanges accounting for the other 4 per cent of market turnover. In the current analysis we use monthly data for the period January 1995 – December 2013 (19 years) for the four stock Indices on NSE. The four indices used were:

- **NSE Nifty**: NSE Nifty is the main index of the NSE. It is a well diversified 50 stock index accounting for 22 sectors of the Indian economy. This index is also used for a variety of purposes such as benchmarking fund portfolios, index based derivatives and index funds. The Nifty Index represents about 69 per cent of the free float market capitalization of the stocks listed on the NSE.

- **NSE Nifty Junior**: Nifty Junior represents the next rung of liquid securities after Nifty. NSE Nifty and Nifty Junior constitute the 100 most liquid stocks on the NSE.
The stocks in Nifty Junior are filtered for liquidity, so they are the most liquid of the stocks excluded from the Nifty. The CNX Nifty and the CNX Nifty Junior are synchronized so that the two indices will always be disjoint sets; i.e. a stock will never appear in both indices at the same time. The Nifty Junior Index represents about 13 per cent of the free float market capitalization of stocks listed on the NSE.

- **NSE Defty**: NSE Defty is the NSE Nifty expressed in terms of US Dollars. This index was developed so that institutional investors and offshore fund enterprises, which have an equity exposure in India, would have an instrument for measuring returns on their equity investment in dollar terms. Movement in the Defty index captures movement in both the foreign exchange market and securities market.

- **NSE CNX 500**: NSE CNX 500 is a broad based benchmark index of the Indian capital market. It is a 500 stock index and represents about 97 per cent of the free float market capitalization of the stocks listed on the NSE. The 500 constituents of this index can be disaggregated into 72 industries.

The data for the above indices were downloaded from the NSE website (http://www.nseindia.com/) for the period January 1995 – December 2013; however, the data for NSE Nifty Junior was only available for October 1995 - December 2013.

Even though BSE is the second biggest stock exchange in India in terms of market turnover, it is probably the largest stock exchange in the world in terms of number of listed companies (approximately 8,500). It is also among the oldest stock exchanges in Southeast Asia with 137 years of history. We used the following two indices from BSE in our analysis:

- **BSE SENSEX**: BSE SENSEX is the main market index of BSE. It consists of the 30 most liquid stocks traded on BSE. BSE SENSEX is widely reported in both domestic and international markets through print as well as electronic media.

- **BSE CNX 500**: BSE CNX 500 is a broad based 500 stocks index on BSE, which represents nearly 93 per cent of the total market capitalization on BSE. BSE CNX 500 covers 20 major industries of the economy.

The data for the BSE indices was downloaded from the BSE website (http://www.bseindia.com). The time series plot for all the indices is presented in Figure 1.
Over the course of the last two decades the Indian stock market has fluctuated a lot, but long-term growth has been phenomenal. Some stylised facts observed in the data are summarised in Table 1 and Table 2. The Indian stock market experienced an overall positive return during the near two decades under study. Table 1 presents the ten-year and eight year return (calculated as $R_t = \frac{P_t-P_{t-1}}{P_{t-1}} \times 100$) between 1995 and 2005 (R12) and between 2005 and 2013 (R23) respectively. The last column of Table 1 gives the overall return on the selected indices for the whole period under study (R13). We see that, on the whole, the market has grown and is characterised by a positive value of return on all the indices; however, we notice that in the second half of the sample period returns are higher than in the first half.

Apart from looking at the long-term returns from the Indian stock market (such as 10 year returns and 19 year returns), we also looked at short-term monthly returns in the Indian stock market over the sample period. Table 2 presents the summary statistics for the monthly returns over the sample period. We notice that the average monthly returns are positive for all the indices over the sample period. However, the average values do not decisively suggest that the Indian stock market was profitable in the short term, as we see that all monthly index returns are characterised by a large standard deviation, suggesting large volatility in returns. This conclusion is also confirmed by the large spread between the minimum and maximum values, given in columns five and six respectively. The next to last column of Table 2 gives the values of skewness for all the indices. We see that out of the six indices considered, five exhibit positive skewness whereas one, BSE CNX 500, displays negative skewness. Intuitively, this implies that the probability of occurrence of negative monthly returns is higher on BSE CNX 500 compared to other indices used in the study. Moreover, we also notice that NSE Nifty and BSE SENSEX, the two main indices, have more or less a symmetrical distribution with skewness values of 0.06 and 0.08 respectively.

The last column of Table 2 presents the result of the ARCH LM test. In order to conduct this test, we first filter the data using an AR(12) model, then use the residuals to run an ARCH LM test. The null hypothesis of no arch effect is rejected at the 5 per cent level for all the indices, indicating the presence of significant time varying volatility in the monthly returns.
5. The Empirical Results

As a starting point for testing the random-walk hypothesis we conducted traditional unit root tests, which do not take into account any structural breaks. Specifically, we applied the ADF and Phillips-Perron unit root tests and the KPSS (Kwiatkowski et al., 1992) stationarity test, with and without a trend. All the tests were conducted using the natural logarithm of the stock indices. The results are presented in Table 3. In all cases the tests suggest that there is a unit root in the data. There are, however, two problems with these tests. First, Kim and Schmidt (1993) show that the ADF test is biased in the presence of conditional heteroskedasticity. Second, none of these tests accommodate possible structural breaks. The period studied contains several dramatic changes associated with economic and political factors which may have had some effect on the Indian stock market. Some of the important domestic events which could have caused a structural break in the Indian stock market are as follows:

1. India's decision to detonate nuclear devices (Pokharan-II) in May 1998 resulted in comprehensive economic and technology-related sanctions by a number of countries. Most of the sanctions were lifted within five years of Pokharan-II, but the instantaneous impact of these sanctions was a big shock to the Indian economy.

2. Between May 1999 and July 1999, an armed conflict took place between India and Pakistan in the Kargil District of Kashmir. As this conflict intensified, it attracted much international concern because of the nuclear capabilities of both countries. In the aftermath of the war, the Indian stock market rose by over 1500 points.

3. On 1 March 2001, BSE SENSEX fell by 176 points, shocking Indian investors. This sudden crash\(^3\) caused a panic among investors. As a consequence of the crash eight people committed suicide and hundreds of investors were driven to the brink of bankruptcy, making it one of the biggest stock market scams in Indian history.

---

\(^3\) The main reason for the crash was that the stockbroker Ketan Parekh (KP) and his allies bought large stakes in technology stocks and kept on carrying forward the payments, using securities as collateral. With the crash of technology stocks in the US, the price of technology stocks dropped drastically and KP failed to pay back his loans (or provide more collateral) and defaulted on his loans made for stock purchases. The system that allowed the roll-over of security purchase settlements was popularly known as the *badla* system.
4. As a follow-up to the Ketan Parekh (KP)-induced stock market scam and crash in March 2001, the Securities & Exchange Board of India (SEBI) passed a directive to replace the *badla*\(^4\) system with a system which introduced daily settlements, single stock options and real-time electronic payment in the Indian stock market.

5. On 27 February 2002, there were communal riots in the Indian state of Gujarat. Officially 793 Muslims and 253 Hindus died as a result of the violence and the Indian economy was adversely affected.

6. On 17 May 2004, just after the declaration of the 2004 general election results, the stock markets crashed (activating automatic circuit breakers in both NSE and BSE) in anticipation that the new coalition (Congress and Left parties) coming into power would not continue with the policy of economic liberalisation pursued by previous governments.

7. In July 2005, Mumbai experienced a storm with record rainfall for 24 hours bringing the whole city and all business activity, transport and trading to a complete halt. Roughly 1,000 people were killed in the ensuing floods and landslides.

8. On 11 July 2006, a series of bombs exploded in Mumbai’s local trains during rush hour, killing around 180 people and disrupting life and business activity in the city. Islamic militants were later found to be responsible for the bombings.

9. A major terrorist attack occurred in Mumbai in November 2008, when a series of gunmen led a coordinated attack on the main tourist and business areas of Mumbai, killing at least 200 people and resulting in loss of investor confidence.

10. Manmohan Singh’s led Congress government was re-elected in May 2009 with an almost absolute majority. Manmohan Singh is considered to be the architect of India's liberalisation policies that started in 1991. His re-election in 2009 was seen as an endorsement for the continuation of financial reforms and liberalisation.

---

\(^4\) The *badla* system combined features of forward and margin trading. It was legal, but minimally regulated and this risky system allowed investors to trade stocks with little cash. The investors were allowed to settle the trade up to five days later, and even pay a fee to delay settlement still longer. This system had features of derivatives trading, but instead of exchange being an intermediary between buyer and seller, the stock brokers acted as intermediaries.
11. During the period 2011 - 2012, India witnessed a surge in high profile corruption cases, such a mismanagement associated with the organisation of the Commonwealth Games, bribes associated with allocation of the 2G spectrum and corruption related to allocation of coal-mining licences (colloquially referred to as ‘coalgate’ in the popular press). This resulted in protests demanding tougher laws against government corruption, led by social activist Anna Hazare and his India against corruption team.

Apart from these important domestic events, the period under consideration was also characterised by many global economic and political events that had an impact on global financial markets, including those in India. Some of the important global events were the Asian financial crisis in 1997, the IT bubble burst and crash of technology stocks in 2001, the September 11, 2001 terrorist attacks and the ensuing Afghanistan War in 2001-02, the Iraq war which begin in 2003 and the Global financial crisis in 2007 and 2008.

The results for the two-break LM unit root test are presented in Table 4. Overall, there is strong evidence of a random walk. For the test with break in intercept and trend, the unit root null is only rejected for the NSE Nifty and BSE SENSEX and only at the 10 per cent level. For the test with break in intercept only, the unit root null is not rejected for any series.

The results for the Narayan and Popp (2010) unit root test with two breaks are presented in Table 5. There is strong evidence of a random walk. Both the tests with the break in intercept only and break in intercept and trend fail to reject the unit root null for any of the indices.

On the basis of the traditional unit root tests and the Lee and Strazicich (2003b) and Narayan and Popp (2010) unit root tests with two structural breaks, the six Indian stock indices are characterized by a random walk at the 5 per cent level. This conclusion holds, irrespective of whether one allows for a break in the intercept only or a break in the intercept and trend.

The results for the Narayan and Liu (2013) GARCH unit root test with two breaks in the intercept are presented in Table 6. In contrast to the findings for the Lee and Strazicich (2003b) and Narayan and Popp (2010) unit root tests, there is evidence of mean reversion in all six indices at the 5 per cent level. This finding is consistent with the argument in Narayan and Liu (2013) that the GARCH unit root test with two structural breaks is superior in terms of rejecting the unit root null compared with the Lee and Strazicich (2003b) and Narayan and Popp (2010) unit root tests because Narayan and Liu (2013) take account of both structural breaks and heteroskedasticity. The finding is also consistent with the result in Narayan and

Given that the Narayan and Liu (2013) unit root test suggests a different result than the Lee and Strazicich (2003b) and Narayan and Popp (2010) unit root tests, the question arises as to which test is more accurate. The Narayan and Popp (2010) test chooses the break date more accurately than the Lee and Strazicich (2003b) test and thus avoids the gain in false power from choosing incorrect break dates (or loss in power) to reject the null. The Narayan and Liu (2013) test uses the Narayan and Popp (2010) approach to choosing the structural breaks, but has the advantage that it is more suitable if the data are heteroskedastic. The results of the ARCH LM test suggest that the data are heteroskedastic (see final column of Table 2). Hence, we conclude that the results of the Narayan and Liu (2013) test can be preferred to those of the Lee and Strazicich (2003b) and Lee and Strazicich (2003b) tests.

-----------------------------

Insert Tables 4-6 here

-----------------------------

Turning briefly to the location of the structural breaks, most fall into one of five periods. These are 1998-1999, 2001, 2003, 2006 and 2008-2009. The Narayan and Liu (2013) test suggests that the first break in all six series occurs in May to June 2003 and the second break occurs in July to September 2006. This appears to correspond closely to the plot of the time series for the indices. This is illustrated in Figure 2 using the example of the NIFTY. In the period May to December 2006 Indian financial markets were in a very volatile state (for instance, the BSE SENSEX plunged by 1,100 points during intra-day trading on 22 May 2006, leading to suspension of trading) due to investor panic regarding the fundamentals of the Indian economy. The finance minister of India, the Reserve Bank of India and the SEBI all reassured investors that nothing was wrong with the fundamentals of the economy and advised retail investors to remain in the market and not sell stocks. Nevertheless, the markets remained volatile for the whole period and major indices finished in the red.

-----------------------------

Insert Figure 2 here

-----------------------------
More generally, the identified breaks across the three tests likely reflect a combination of domestic and international shocks to the economy. The breaks in 1998-1999 are associated with the aftermath of the Asian financial crisis and conflict in Kashmir. The breaks in 2001 coincide with the IT bubble burst and 9/11 terrorist attacks as well as the KP scam and the subsequent institutional reforms to the Indian stock exchanges. The breaks in 2003 coincide with the start of the Iraq war and heightening of the US global war on terror. The breaks in 2006 and 2008-2009 occur at the same time as terrorist attacks in Mumbai. The 2008-2009 breaks are also likely to be related to the re-election of Manmohan Singh and the Global financial crisis.

6 Conclusions

In the current study we have tested the random-walk hypothesis for the Indian stock market. Our contribution to the literature on the efficient market hypothesis in stock price indices is twofold. First, we contribute to the growing literature testing the efficient market hypothesis using Indian stock price data. We extend this literature by using recent unit root tests that allow for structural breaks and, in the case of the Narayan and Liu (2013) test, heteroskedasticity in the data. Second, we contribute to the literature employing unit root testing to examine the efficient market hypothesis in high frequency financial data in emerging markets more generally. Most of the existing literature has focused on the low power of conventional unit root tests to reject the unit root null in the presence of structural breaks. The fact that conventional unit root tests are biased when applied to high frequency data in the presence of heteroskedasticity has been largely overlooked. Our results suggest that it is not only important to accommodate structural breaks, but point to the need to consider heteroskedasticity when testing for a random walk with high frequency financial data. When we do this, we find that the Indian stock indices are mean reverting.
References


## Tables and Figures

### Table 1: Returns on selected indices during the sample period

<table>
<thead>
<tr>
<th>Index</th>
<th>Price1 (Jan 1995)</th>
<th>Price2 (Jan 2005)</th>
<th>Price3 (Dec 2013)</th>
<th>R12 (Return in first 10 years)</th>
<th>R23 (Return in last 9 years)</th>
<th>R13 (Overall returns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSE Nifty</td>
<td>1019.20</td>
<td>2115.00</td>
<td>6217.85</td>
<td>107.52%</td>
<td>193.99%</td>
<td>510.07%</td>
</tr>
<tr>
<td>NSE Defty</td>
<td>1125.59</td>
<td>1685.45</td>
<td>3462.36</td>
<td>49.74%</td>
<td>105.43%</td>
<td>207.60%</td>
</tr>
<tr>
<td>NSE CNX 500</td>
<td>972.10</td>
<td>1834.75</td>
<td>4804.40</td>
<td>88.74%</td>
<td>161.86%</td>
<td>394.23%</td>
</tr>
<tr>
<td>NSE Nifty Junior</td>
<td>1167.18</td>
<td>4508.75</td>
<td>12453.65</td>
<td>286.29%</td>
<td>176.21%</td>
<td>966.99%</td>
</tr>
<tr>
<td>BSE SENSEX</td>
<td>3932.09</td>
<td>6679.20</td>
<td>20898.01</td>
<td>69.86%</td>
<td>212.88%</td>
<td>431.47%</td>
</tr>
<tr>
<td>BSE CNX 500</td>
<td>1000.00</td>
<td>2825.95</td>
<td>7828.34</td>
<td>182.59%</td>
<td>177.02%</td>
<td>682.83%</td>
</tr>
</tbody>
</table>

Notes: Data for NSE Nifty Junior were available from October 1995 – December 2013 and the data from BSE CNX 500 were available from February 1999 – December 2013. The data for all the other indices were for the period January 1995 – December 2013.
Table 2: Descriptive statistics of monthly returns for selected indices during the sample period

<table>
<thead>
<tr>
<th>Index</th>
<th>No. of Observations</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
<th>Skewness</th>
<th>ARCH(12) LM Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSE Nifty</td>
<td>227</td>
<td>1.10</td>
<td>7.83</td>
<td>-22.96</td>
<td>30.40</td>
<td>0.06</td>
<td>22.62***</td>
</tr>
<tr>
<td>NSE Defty</td>
<td>227</td>
<td>0.88</td>
<td>8.83</td>
<td>-26.21</td>
<td>39.13</td>
<td>0.23</td>
<td>26.45***</td>
</tr>
<tr>
<td>NSE CNX 500</td>
<td>227</td>
<td>1.07</td>
<td>8.60</td>
<td>-24.06</td>
<td>37.22</td>
<td>0.10</td>
<td>29.46***</td>
</tr>
<tr>
<td>NSE Nifty Junior</td>
<td>218</td>
<td>1.62</td>
<td>10.35</td>
<td>-30.52</td>
<td>45.71</td>
<td>0.18</td>
<td>25.93**</td>
</tr>
<tr>
<td>BSE SENSEX</td>
<td>227</td>
<td>1.04</td>
<td>7.83</td>
<td>-20.82</td>
<td>30.14</td>
<td>0.08</td>
<td>29.57***</td>
</tr>
<tr>
<td>BSE CNX 500</td>
<td>178</td>
<td>1.56</td>
<td>8.88</td>
<td>-24.14</td>
<td>28.96</td>
<td>-0.26</td>
<td>22.70**</td>
</tr>
</tbody>
</table>

Notes: As per Table 1.
Table 3: Results of traditional unit root tests

<table>
<thead>
<tr>
<th>Index</th>
<th>Intercept ADF Test</th>
<th>Intercept PP Test</th>
<th>Intercept KPSS Test</th>
<th>Intercept ADF Test</th>
<th>Intercept PP Test</th>
<th>Intercept KPSS Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSE Nifty</td>
<td>-0.710</td>
<td>-0.500</td>
<td>1.85***</td>
<td>-2.353</td>
<td>-2.694</td>
<td>0.232***</td>
</tr>
<tr>
<td>NSE Defty</td>
<td>-0.995</td>
<td>-0.880</td>
<td>1.67***</td>
<td>-2.134</td>
<td>-2.496</td>
<td>0.247***</td>
</tr>
<tr>
<td>NSE CNX 500</td>
<td>-0.769</td>
<td>-0.466</td>
<td>1.82***</td>
<td>-2.541</td>
<td>-2.947</td>
<td>0.211**</td>
</tr>
<tr>
<td>NSE Nifty Junior</td>
<td>-1.278</td>
<td>-0.923</td>
<td>1.89***</td>
<td>-3.088</td>
<td>-2.665</td>
<td>0.104</td>
</tr>
<tr>
<td>BSE SENSEX</td>
<td>-0.658</td>
<td>-0.404</td>
<td>1.8***</td>
<td>-2.317</td>
<td>-2.573</td>
<td>0.24***</td>
</tr>
<tr>
<td>BSE CNX 500</td>
<td>-0.543</td>
<td>-1.164</td>
<td>1.57***</td>
<td>-2.500</td>
<td>-2.082</td>
<td>0.18*</td>
</tr>
</tbody>
</table>

Notes: As per Table 1.
Table 4: Results for Lee and Strazicich (2003) LM unit root test with two structural breaks

<table>
<thead>
<tr>
<th>Index</th>
<th>Test statistic</th>
<th>Break in Intercept</th>
<th>Break in Intercept and Trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSE Nifty</td>
<td>-2.396</td>
<td>Mar 06</td>
<td>Apr 03</td>
</tr>
<tr>
<td>NSE Defty</td>
<td>-1.905</td>
<td>Dec 03</td>
<td>Jul 08</td>
</tr>
<tr>
<td>NSE CNX 500</td>
<td>-2.038</td>
<td>Jun 98</td>
<td>May 09</td>
</tr>
<tr>
<td>NSE Nifty Junior</td>
<td>-3.354</td>
<td>Dec 03</td>
<td>Jul 08</td>
</tr>
<tr>
<td>BSE SENSEX</td>
<td>2.1961</td>
<td>Oct 07</td>
<td>Mar 03</td>
</tr>
<tr>
<td>BSE CNX 500</td>
<td>-2.551</td>
<td>Jul 08</td>
<td>Apr 03</td>
</tr>
</tbody>
</table>

Critical values for $S_{t-1}$

Model AA (Break in Intercept only)

<table>
<thead>
<tr>
<th>$\lambda_2$</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1$</td>
<td>1%</td>
<td>5%</td>
<td>10%</td>
</tr>
<tr>
<td>0.2</td>
<td>-6.16</td>
<td>-5.59</td>
<td>-5.27</td>
</tr>
<tr>
<td>0.4</td>
<td>-6.45</td>
<td>-5.67</td>
<td>-5.31</td>
</tr>
</tbody>
</table>

Model CC (Break in Intercept and Trend)

<table>
<thead>
<tr>
<th>$\lambda_2$</th>
<th>0.4</th>
<th>0.6</th>
<th>0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1$</td>
<td>1%</td>
<td>5%</td>
<td>10%</td>
</tr>
<tr>
<td>0.2</td>
<td>-6.32</td>
<td>-5.73</td>
<td>-5.32</td>
</tr>
<tr>
<td>0.6</td>
<td>-6.45</td>
<td>-5.67</td>
<td>-5.31</td>
</tr>
</tbody>
</table>

Notes: TB1 and TB2 are the dates of the structural breaks. $\lambda_i$ denotes the location of the breaks. The LM unit root test for model AA is invariant to the location of the breaks; however, this invariance does not hold for model CC, for which the null distribution of the LM test depends on the relative location of the breaks. * (**) *** denotes statistical significance at the 10%, 5% and 1% levels respectively. Data for NSE Nifty Junior were available from October 1995 – December 2013 and the data from BSE CNX 500 were available from February 1999 – December 2013. The data for all the other indices were for the period January 1995 – December 2013.
Table 5 Results for Narayan and Popp (2010) unit root test with two structural breaks.

<table>
<thead>
<tr>
<th>Index</th>
<th>Break in Intercept</th>
<th>Break in Intercept and Trend</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Test statistic</td>
<td>TB1</td>
</tr>
<tr>
<td>NSE Nifty</td>
<td>-2.173</td>
<td>Oct 08</td>
</tr>
<tr>
<td>NSE Defty</td>
<td>-1.839</td>
<td>Oct 08</td>
</tr>
<tr>
<td>NSE CNX 500</td>
<td>-2.242</td>
<td>Mar 01</td>
</tr>
<tr>
<td>NSE Nifty Junior</td>
<td>-2.003</td>
<td>Dec 99</td>
</tr>
<tr>
<td>BSE SENSEX</td>
<td>-2.373</td>
<td>Mar 01</td>
</tr>
<tr>
<td>BSE CNX 500</td>
<td>-1.868</td>
<td>Mar 01</td>
</tr>
</tbody>
</table>

Critical values for unit root test

<table>
<thead>
<tr>
<th>Model</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model M1 (Break in Intercept only)</td>
<td>-4.731</td>
<td>-4.136</td>
<td>-3.825</td>
</tr>
<tr>
<td>Model M2 (Break in Intercept and Trend)</td>
<td>-5.318</td>
<td>-4.741</td>
<td>-4.430</td>
</tr>
</tbody>
</table>

Notes: Data for NSE Nifty Junior were available from October 1995 – December 2013 and the data from BSE CNX 500 were available from February 1999 – December 2013. The data for all the other indices were for the period January 1995 – December 2013. The critical values are taken from Narayan and Popp (2010).
Table 6: Results for Narayan and Liu (2013) GARCH unit root test with two structural breaks in the intercept.

<table>
<thead>
<tr>
<th>Index</th>
<th>Test statistic</th>
<th>TB1</th>
<th>TB2</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSE Nifty</td>
<td>-4.03**</td>
<td>Jun 03</td>
<td>Jul 06</td>
</tr>
<tr>
<td>NSE Defty</td>
<td>-4.69**</td>
<td>May 03</td>
<td>Sep 06</td>
</tr>
<tr>
<td>NSE CNX 500</td>
<td>-5.34**</td>
<td>Jun 03</td>
<td>Sep 06</td>
</tr>
<tr>
<td>NSE Nifty Junior</td>
<td>-5.23**</td>
<td>Jun 03</td>
<td>Sep 06</td>
</tr>
<tr>
<td>BSE SENSEX</td>
<td>-3.78**</td>
<td>Jun 03</td>
<td>Jul 06</td>
</tr>
<tr>
<td>BSE CNX 500</td>
<td>-4.94**</td>
<td>Jun 03</td>
<td>Sep 06</td>
</tr>
</tbody>
</table>

Notes: Data for NSE Nifty Junior were available from October 1995 – December 2013 and the data from BSE CNX 500 were available from February 1999 – December 2013. The data for all the other indices were for the period January 1995 – December 2013. The 5% critical value for the unit root test statistics is -3.76, obtained from Narayan and Liu (2013) [Table 3 for N = 250 and GARCH parameters [α,β] chosen as [0.05, 0.90]]. Narayan and Liu (2013) provide critical values for 5% level of significance only. ** indicates rejection of the null hypothesis of a unit root at the 5% level of significance.
Figure 1: Time series plot of all the indices used.

Figure 2: Breaks in intercept in the NIFTY series using the Narayan and Liu (2013) method.