The present study attempts to examine the existence of long run diversification benefits of investment allocation in Indian Stock Market with that of the Developed World by using daily data of closing stock prices for the period 1997 to 2011 and applying a modified GPH test, along with a battery of other tests such as KPSS, GPH and the Rescaled Range. The study revealed conflicting results: the powerful GPH test indicated long memory, while the earlier generation of tests pointed towards its absence. The results indicate the existence of fractional cointegration of Indian Stock Market with the Developed world and hence absence of diversification benefits for the selected markets in the long run.

Keywords: Long Memory, Fractional Co-integration, Modified GPH, KPSS, RS.

JEL Classification: G15.
1. Introduction

The Stock Market is an important part of a financial system that meets its long-term fund requirements. In doing so, it also acts as a barometer for the economy. Movements in stock markets are indicative of the domestic economic conditions as well as the level of confidence that investors – both domestic and foreign – place in an economy. An important phenomenon witnessed during the decades after liberalization of the Indian economy in 1991 was the dramatic increase in its international capital inflows. The Indian economy saw substantial Foreign Direct Investment (FDI) inflows and portfolio investment being made by Foreign Indirect Investors (FII). India was termed as a promising emerging market, given the high levels of growth taking place during the past two decades. This in turn points towards a greater integration of the Indian stock market with that of the developed world which in turn would have effects on the exchange rate, trade and the monetary and the fiscal policies of the country. Thus, it becomes necessary to understand the linkages between the Indian stock market and that of the developed world. With this background, our study will look into the cointegration between the Indian stock market with the developed world.

The Efficient Markets Hypothesis (EMH) by Fama (1965) holds that an efficient market is where security prices adjust rapidly to new information and hence the current prices of securities would reflect all available information about a security. Thus, while a market can over-adjust or under-adjust, it would not be biased and an investor would not be able to predict the occurrences over a period of time. Thus, since market prices would reflect all available information, the differences between prices at two different points in time cannot be predicted i.e., a random event. However, practitioners are skeptical about the concept and believe that the
Indian markets are increasingly becoming integrated with the global economy. However, there are only a few studies that examine Long memory effects for the same. The paper revisits the question of whether or not actual stock market prices exhibit long-range dependence in the context of India and with that of the stock markets of major developed markets, during the post-liberalization phase of the Indian economy by using Long memory approach.

A number of important reasons for the investors, investment fund managers, policy makers and academicians make study worthy by understanding of the international linkages of stock market. This study might be helpful in understanding the benefits that arises from the international portfolio diversification, stock market efficiency, capital budgeting, formulation of macro-economic policy and financial regulation. First, the interdependence of stock markets is helpful for the international investors in understanding the global stock market environment and to anticipate the risk and rewards due to international diversification. Greater integration among the world stock markets reduces the potential gain arises due to international diversification. Second, evidence of high interdependence among the world stock markets might lead us to reject the efficient markets hypothesis. Third, to study the world stock market integration is important from the macroeconomic and monetary policies perspective also as stock market co-movements might lead to market contagion because investors incorporate into their trading decisions information about price changes in other markets. Fourth, examining the world stock market integration has fundamental implication in international capital budgeting. For example, in case of segmented capital markets, the cost of capital for a project will depend upon the source of finance, whereas this cost will be irrespective of the location where it is raised in an integrated world market. Last but not least, increasing integration of stock markets of the world can be
related with numerous benefits such as the development of financial markets and securities market design, and effective price discovery, higher savings, investment and economic progress.

India has engaged in various bilateral and/or multilateral trade and economic cooperation agreements with several countries and regional groups across Asia and Europe. Even if, the Indian economy is closely associated with the rest of world through international trade and foreign direct investment, the extent of global integration of the Indian stock market is not clear. The recent crash in the BSE Sensex triggered by US financial crisis is simply an indicator of the impact of international contagion. It has created substantial interest in studying the interdependence and dynamic linkages among national stock markets. The main objective of this paper is to probes the existence of Long memory for the Indian Stock Market, with the developed markets in Germany, Japan, Switzerland, United Kingdom, and United States. It considers daily data of closing stock prices from these countries for the time period 1997 to 2011. It uses techniques which are considered robust for testing the long memory component of markets: the modified GPH test, along with a battery of other tests such as KPSS, GPH and the Rescaled Range. The study revealed that the data is log non-normal. More importantly, the results were split: the powerful GPH test indicated long memory, while the earlier generation of tests point towards its absence. The results indicated the existence of fractional integration of Indian Stock Market with those of the Developed world. This is suggestive of increasing integration of the Indian Markets with the developed markets. This also implies an absence of diversification benefits for these selected markets for the long run.

The plan for the rest of the paper is as follows: Section 2 presents a brief literature review on stock market cointegration and long memory, Section 3 discusses the data used and the
methodology employed for analysis and Section 4 discusses the results while Section 5 concludes.

2. Literature Review

As highlighted in the introduction, the interrelationship between various stock markets has important microeconomic and macroeconomic implications. The various stock markets in developing countries are having greater amounts of portfolio investment from mature markets as they are remunerative. Along with globalization, the other contributory factors were the loosening of financial regulations, introduction of innovative financial products, advancement in trading systems and improvements in communications networks (Wong et al., 2005).

Following the seminal work of Grubel (1968), there has been substantial interest in studying the relationships among the financial markets. Early work by Ripley (1973), Lessard (1976), and Hilliard (1979) found low correlations between national stock markets.

The October 1987 crash focused attention on the connections between various national equity markets. Eun and Shim (1989) found evidence of co-movements between the US stock markets and the other national stock markets by using vector autoregression models. Lee and Kim (1994) analyzed the October 1987 crash and found that national stock markets had become interrelated after the crash and that the movements among national stock markets were stronger when the US stock market were more volatile. Jeon and Furstenberg (1990) found that the degree of international co-movement in stock price indices saw a significant increase after the 1987 crash.

However, there are some studies indicating an absence of integration among the stock markets. Koop (1994) found no co-movement in stock prices by using Bayesian methods.

There have only been a few studies to look into the co-movement of the Indian stock market along with the world markets. Sharma and Kennedy (1977) in their study found no evidence of a cyclical component or periodicity of the Indian stock market along with that of the US and UK markets. Rao and Naik (1990) found a poor relationship of the Indian market with the international markets while using Cross-Spectral analysis. However, the Indian stock market has become much more open to the world since liberalization and there could be a change in the relationship. Wong et al. (2005) examined the nature of the co-movement between Indian stock market and main indices to find that the Indian stock market is statistically significantly cointegrated with the stock markets in the US, UK and Japan by using OLS estimation. They examined the long-run dynamics of these four markets by using fractionally integrating technique and found that the Indian stock index forms fractionally cointegrated relationship in the long run. They emphasize that since only the Johansen Multivariate model can generate the stationary error term, it shows the superiority of the Johansen method over others. Overall, they conclude that the Indian financial liberalization opened up the Indian stock markets to the outside world and its markets are being influenced by other markets. They believe that this reflects the heavy controls on the Indian economy through the seventies and the liberalization measures being initiated only in the late eighties. However, substantial changes have taken place since the time and Indian economy continues to be on a high growth path. Thus, there is a need to revisit the issue for evidence of Long memory.
Jagric et al. (2005) found mixed evidence of long memory presence in the stock indices of six Central and Eastern European (CEE) countries for the period between 1991-2004: strong long-range dependence was identified in the returns of the Czech, Hungarian, Russian and Slovenian stock markets, whereas weak or no long-range dependence in the returns of the Slovakian and Polish stock markets. Assaf (2006) investigated the long memory properties of stock market returns and volatility for countries of MENA region (that includes Egypt, Jordan, Morocco, and Turkey) and used fractional integration (non-parametric) procedures proposed by Lo (1991) and Giraitis et al. (2003). He finds evidence of long memory in the stock returns for Egypt and Morocco. For Jordan and Turkey, he finds evidence of anti-persistence. He also finds evidence of long memory in the volatility series for the entire sample countries examined. Jagric et al. (2006) investigated long memory in stock returns of ten CEE countries. Their results suggested strong long memory presence in eight stock markets, among them in the Croatian and Hungarian stock markets. Gurgul and Wójtowicz (2006) investigated the long-memory properties of trading volume and the volatility of stock returns (given by absolute returns and alternatively by square returns) for the daily stock data of German companies in the DAX segment of the German Stock Exchange. Authors, performed analysis for daily data covering period from January 1994 to November 2005 and in three sub-periods: January 1994 to December 1997, January 1998 to December 2001, and January 2002 to November 2005. The study found that for the equities listed in the DAX index the log-volume and returns volatility exhibit long memory. Moreover these two series have the same long-memory parameters for most of the equities. This common long memory of both series is especially strongly pronounced in the latest data. On the other hand, there is no evidence that log-volume and volatility share the same long memory component. Caporale and Gil-Alana (2007) examined the long-run dynamics
and the cyclical structure of various series (such as inflation, real risk-free rate, real stock returns, equity premium, and price/dividend ratio) related to the US stock market using fractional integration for the period from 1871 to 1993. They found that the estimated order of integration varies considerably, but nonstationarity is found only for the price/dividend ratio, when focusing exclusively on the long-run or zero frequency. When the authors taken into account the cyclical component, they found that series are stationary and exhibit long memory with respect to both components in almost all cases with the exception to the price/dividend ratio. Supportive evidence of long memory is provided by Ozdemir (2007) and Chan and Feng (2008) for the DJI, the S&P500, the FTSE, DAX and NIKKEI (for different time periods), Bilel and Nadhem (2009) for G7 countries stock indices for high frequency data between 2003-2004 and Mariani et al. (2010) for international stock indices. Kasman et al. (2009a) investigated the fractal structure of the CEE stock markets. The authors found the existence of long memory in five of eight studied markets. Furthermore, Kasman et al. (2009b), investigating four CEE stock markets, found significant long memory in the return series of the Slovak Republic, weak evidence of long memory for Hungary and the Czech Republic, and no evidence for Poland.

Lim (2009) tested the efficient market hypothesis for Middle East and African countries using non-linear models. He finds that the returns for emerging stock markets for both Middle East and Africa contain predictable components, and therefore are inefficient. Alagidede and Panagiotidis (2009) investigated the random walk behavior of stock market returns for Egypt, Kenya, Morocco, Nigeria, South Africa, Tunisia and Zimbabwe. They reject the null hypothesis that the stock market returns for the sample countries are random walk processes. However, using smooth transition and conditional volatility models they find evidence of volatility clustering, leptokurtosis and leverage effect in the African data. Karanasos and Kartsaklas (2009)
investigated the issue of temporal ordering of the range-based volatility and turnover volume in
the Korean market for the period 1995–2005 and also for the period before financial crisis and
after financial crisis. The authors examined the dynamics of the two variables and their
respective uncertainties using a bivariate dual long-memory model. The authors found that the
apparent long-memory in the variables was quite resistant to the presence of breaks and when
they take into account structural breaks the order of integration of the conditional variance series
decreases considerably. Further, the authors showed that the impact of foreign volume on
volatility was negative in the pre-crisis period but turns to positive after the crisis. The authors
found that before the crisis there was no causal effect for domestic volume on volatility whereas
in the post-crisis period total and domestic volumes affect volatility positively.

Tan et al. (2010) examined the efficient market hypothesis on the Malaysian stock market
under both bullish and bearish periods covering from 1985:01 to 2009:12. They first identified
the bullish and bearish periods by using the Bry and Boschan (1971) algorithm, following by
Geweke and Porter-Hudak’s (1983) test for the diagnosis of market efficiency. Their study
provided the evidence of persistent long memory under the earlier periods of the study,
suggesting the possibility to predict stock prices especially before the 1997 financial crisis. Tan
et al. (2011) examined the predictability in stock return in developed and emerging markets by
testing long memory in stock returns using wavelet approach. Wavelet-based maximum
likelihood estimator of the fractional integration estimator is superior to the conventional Hurst
exponent and Geweke and Porter-Hudak estimator in terms of asymptotic properties and mean
squared error. The authors used 4-year moving windows to estimate the fractional integration
parameter. They suggested that stock return may not be predictable in developed countries of the
Asia-Pacific region. However, predictability of stock return in some developing countries in this
region such as Indonesia, Malaysia and Philippines may not be ruled out. Stock return in the Thailand stock market appears to be not predictable after the political crisis in 2008. Festić et al. (2012) analyzed and compared the fractal structure of the Croatian and Hungarian stock market returns. The presence of long memory components in asset returns provides evidence against the weak-form of stock market efficiency. The study provided the evidence that the wavelet ordinary least squares (WOLS) estimator may lead to different conclusions regarding long memory presence in the stock returns from the KPSS unit root tests and/or Lo’s R/S test. Furthermore, it proved that the fractal structure of individual stock returns may be masked in aggregated stock market returns (i.e. in returns of stock index). The main finding of the study was that the Croatian stock index, Crobex, and individual stocks in this index exhibited long memory. Long memory is identified for some stocks in the Hungarian stock market as well, but not for the stock market index BUX. Based on the results of the long memory tests, the study concluded that while the Hungarian stock market is weak form efficient, the Croatian stock market is not.

The literature review above clearly shows that there is conflicting evidence on the issue of international stock market linkages, depending on the methodology, data, and time period studied. Unfortunately, little empirical evidence exists concerning linkages between Indian stock market and other markets of the world and that uses long memory tests. In this regard India has much less exposure in the stock market integration literature until recently. Given India’s fast-growing economic influence, research on the Indian stock market still seems to be inadequate and needs further investigation using approach of long memory tests. The present study extends the existing stock market integration literature in the following ways. First, to provide further evidence, we examine the stock price linkages i.e., interdependence between the stock market of India and that of the Germany, Japan, Switzerland, United Kingdom and the U.S. using daily
stock closing price indices data covering the period July 1, 1997 to April 29, 2011. Second, this research is based on the long memory approach and a battery of tests is utilized to examines interdependence of Indian stock market to the developed world and provide robust evidence. And lastly, from the implication and policy point of view, the results from this research provide implications regarding international diversification and market efficiency that are important for investors and fund managers who are interested in investing in these markets and from the macroeconomic and monetary policies perspective as stock market co-movements might lead to market contagion because investors incorporate into their trading decisions information about price changes in other markets.

3. Data and Methodology

The paper has used daily closing stock prices data for the period July 1, 1997 to April 29, 2011. We have used the following proxies for the markets: Sensex for India, DAX for Germany, Nikkei for Japan, SSMI for Switzerland, FTSE for United Kingdom and the NYSE Composite Index for the U.S. All data was taken from the Yahoo Finance (www.finance.yahoo.com). The total number of observations was 3371 after removing the data for the days in which trading have not occurred at least in one stock market. Our choice of country’s stock markets is guided by the consideration that India has significant trade and financial relations with these countries or they are big players in the international stock market and thereby able to transmit shocks to other stock markets if they experience it and also as Tiwari and Islam (2012) provided a comprehensive review in their study which show that most of the studies those have conducted test of cointegration to analyses the stock market integration have mostly tested with those group of countries we analyzed in the present work. The period of the data is dependent upon the
availability of the data from the Yahoo Finance. Data is converted into natural logarithms before any computation is performed. The features of the data are described in Table 1 and trend plot of the variables is presented in Figure 1.

Table 1: Summary Statistics of Daily Closing Prices of Stock Market Indices (values in natural log form)

<table>
<thead>
<tr>
<th></th>
<th>India</th>
<th>Germany</th>
<th>Japan</th>
<th>Switzerland</th>
<th>United Kingdom</th>
<th>USA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>8.832401</td>
<td>8.557555</td>
<td>9.464249</td>
<td>8.792090</td>
<td>8.576939</td>
<td>8.817620</td>
</tr>
<tr>
<td>Median</td>
<td>8.642667</td>
<td>8.583707</td>
<td>9.471687</td>
<td>8.794628</td>
<td>8.599750</td>
<td>8.808061</td>
</tr>
<tr>
<td>Minimum</td>
<td>7.863313</td>
<td>7.697557</td>
<td>8.861489</td>
<td>8.209417</td>
<td>8.097731</td>
<td>8.349085</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.656400</td>
<td>0.259106</td>
<td>0.259367</td>
<td>0.179181</td>
<td>0.157085</td>
<td>0.188554</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.293700</td>
<td>-0.500600</td>
<td>-0.066301</td>
<td>-0.200984</td>
<td>-0.537008</td>
<td>0.170084</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>1.534729</td>
<td>2.729016</td>
<td>1.801186</td>
<td>2.574537</td>
<td>2.311651</td>
<td>2.514957</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>350.0305 (0.000000)</td>
<td>151.1096 (0.000000)</td>
<td>204.3302 (0.000000)</td>
<td>48.12058 (0.000000)</td>
<td>228.5728 (0.000000)</td>
<td>49.29824 (0.000000)</td>
</tr>
</tbody>
</table>

It is evident from Table 1 that the sample means of level data set is positive for all the countries. The measure of skewness indicates that in the log level form stock indices of Germany, Japan, Switzerland and U.K. are negatively skewed whereas for India and USA is positively skewed. Data series in all log level series have demonstrated less kurtosis which indicates that distributions of stock prices of countries analysed are Platykurtic relative to a normal distribution. The Jarque-Bera normality test rejects normality of all series at any level of statistical significance indicating that stock prices are not log-normally distributed.
Figure 1: Time Series Plot of Daily Closing Prices of Stock Market Indices (July 1, 1997 to April 29, 2011) (values in natural log form)

From the graphical exposition (depicted in Figure 1), which depicts the movements of different stock market indices during the period July 1, 1997 to April 29, 2011, it is evident that there is quite close association between trends in Indian stock market and the stock markets of the developed World which has experienced the various phases of ups and downs during the study period.

Long memory describes the correlation of a series after long lags, which shows temporal dependence over distant observations. Long memory would arise from non-linear dependence in the first movement of the distribution and denotes existence of a predictable component in the time series distribution. The existence of long memory in stock prices gives evidence against the weak form of market efficiency. Long Memory process or long range dependence holds that past events that occurred a long time ago can have an influence on events today. It holds that
correlations between observations far apart in time decay to zero at a rate that is much slower than the one expected from independent data or data that is modeled according to the classical ARMA or Markov frameworks. In other words, the dependence between events far apart diminishes slowly with an increasing distance. Beran (1994, p.6) remarks “a stationary process with slowly decaying correlations which is un-summable is, therefore, called a stationary process with long memory or long-range dependence or strong dependence (in contrast to processes with summable correlations which are also called processes with short memory or short-range dependence or weak dependence)”.

The long memory process can be seen existing half-way between the \( I(0) \) and \( I(1) \), implying different long run predictions than conventional approaches. We discuss the techniques of the KPSS test, Fractional Cointegration Analysis, Mean reversion, Martingale Variance Ratio Test, Long memory, Mean Reversion and Rescaled Range Analysis below.

### 3.1 KPSS Test

Lee and Schmidt (1996)\(^1\) propose the test of Kwiatkowski et al. (1992), hereafter KPSS, as a test for the null of stationarity against the alternative hypothesis of fractional integration. The KPSS test involves regressing \( re_t \) against a constant \( \mu \), and a trend \( \tau \). The test is based upon

\[
\eta_T = T^{-2} \sum S_t^2 / \sigma_T(q) \tag{1}
\]

\(^1\) The Kwiatkowski, Phillips, Schmidt, and Shin (KPSS, 1992) test for stationarity of a time series differs from those “unit root” tests in common use (such as ADF, PP, and DF-GLS) by having a null hypothesis of stationarity. The test may be conducted under the null of either trend stationarity or level stationarity. Inference from this test is complementary to that derived from those based on the Dickey-Fuller distribution (such as ADF, PP, and DF-GLS). The KPSS test is often used in conjunction with those tests to investigate the possibility that a series is fractionally integrated (that is, neither \( I(1) \) nor \( I(0) \)); see Lee and Schmidt (1996). As such, it is complementary to GPH, and modified rescaled range statistic.
where $S_t^2$ represents the square of the partial sum of the residuals and $\sigma^2_t(q)$ represents an estimate of the long run variance of the residuals. Lee and Schmidt (1996) present Monte Carlo evidence that the KPSS test has power properties similar to the modified rescaled range statistic of Lo (1991).

### 3.2 Fractional Integration Analysis

The process was developed independently by Granger (1980), Granger and Joyeux (1980) and Hosking (1981). It is also termed as the ARFIMA($p, d, q$) model. Here, $d$ denotes the fractional differencing parameter, and is written as

$$
\Phi(L)(1-L)^d (y_t - \mu) = \Theta(L) \varepsilon_t, \quad \varepsilon_t \sim i.i.d.(0, \sigma^2_{\varepsilon})
$$

(2)

where $L$ is the backward-shift operator. The parameter $d$ assumes any real value. For $d \neq 0$ the data series is $I(d)$. The stochastic process $y_t$ is both stationary and invertible if all roots of $\Phi(L)$ and $\Theta(L)$ lie outside the unit circle and $|d| < 0.5$. The process is nonstationary for $d \geq 0.5$ as it possessed infinite variance. For $d \in (0,0.5)$, the ARFIMA process exhibits long memory or long-range positive dependence. The process exhibits intermediate memory or long-range negative dependence for $d \in (-0.5,0)$ and short memory for $d = 0$.

While the usual notion of integration has the strict $I(0)$ and $I(1)$ distinction, fractional integration allows the variables to be fractionally integrated. A flexible and parsimonious way to model both the short-term and the long-term behavior of a time series $Z_t = (Z_1, \ldots, Z_T)$ and its first

---

2 In this study we computed a modified form of the Geweke and Porter-Hudak (GPH, 1983) estimate of the long memory (fractional integration) parameter, $d$, of a time series, proposed by Phillips (1999a, 1999b). This is because if a series exhibits long memory, then distinguishing unit-root behavior from fractional integration may be problematic, given that the GPH estimator is inconsistent against $d>1$ alternatives. This weakness of the GPH estimator is solved by Phillips’ (1999a, 1999b) Modified Log Periodogram Regression estimator, in which the dependent variable is modified to reflect the distribution of $d$ under the null hypothesis that $d=1$. The estimator gives rise to a test statistic for $d=1$ which is a standard normal variate under the null. Phillips suggests (1999a, 1999b) that deterministic trends should be removed from the series before application of the estimator.
difference, \( X_t \), denoted \( X_t = (1 - L)Z_t \), where \( L \) is the lag (or backward-shift) operator by way of an autoregressive fractionally-integrative moving-average (ARFIMA) process of order \( (p, d, q) \) becomes:

\[
\Phi(L)(1 - L)^d Z_t = \Phi(L)(1 - L)^d X_t = \Theta(L) \varepsilon_t, \varepsilon_t \sim iid(0, \sigma^2),
\]

where \( d \) is a fractional difference operator and possibly a non-integer; and \( \Phi(L) \) and \( \Theta(L) \) are the autoregressive and moving average lag polynomials of orders \( p \) and \( q \), respectively, [i.e., \( \Phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p, \Theta(L) = 1 + \phi_1 L + \ldots + \phi_q L^q, d = 1 + d' \), and \( (1 - L)^d \) is the fractional differencing operator defined by \( (1 - L)^d = \sum_{k=0}^{\infty} \Gamma(k - d) L^k / \Gamma(k + 1) \Gamma(-d) \) with \( \Gamma(\cdot) \) denoting the gamma or generalized factorial function] with roots lying outside the unit circle. An integer value of \( d = 0 \) yields a standard ARMA process (i.e., the process exhibits short memory for \( d = 0 \), whereas \( d = 1 \) gives rise to the unit-root nonstationary process. For \( d \in (0, 0.5) \) and \( d \neq 0 \), the ARFIMA process is said to exhibit long memory and the autocorrelations of the ARFIMA process decay hyperbolically to zero as \( k \to \infty \). For \( d \in (-0.5, 0) \), the ARFIMA process is said to exhibit intermediate memory (see Granger and Joyeux, 1980). Although for \( d \in (0.5, 1) \) the process implies mean reversion, however, it is not a covariance stationary process. Finally \( d > 1 \) implies that the process is not mean reverting. The fractional integration test due to Geweke and Porter-Hudak (1983) is used to detect the nature of stationarity. Geweke and Porter-Hudak (1983) demonstrate that the fractional differencing parameter \( d \) can be estimated consistently from the least squares regression at frequencies near zero:

\[
\ln(I(\omega_j)) = \alpha + \beta \ln(4 \sin^2(\omega_j / 2)) + v_j, \quad j = 1, \ldots, J
\]
where $\alpha$ is the constant term, $\omega_j = 2\pi j / T (j = 1, \ldots, T - 1)$, $J = T^n \ll T$, $T$ is the number of observations, and $I(\omega_j)$ is the periodogram of time series $X$, at frequency $\omega_j$. With a proper choice of $J$, the negative of the OLS estimate of $\beta$ coefficient gives a consistent and asymptotically normal estimate of the order of integration $d$.

Moreover, this is true regardless the orders and the estimates of the parameters of the ARMA process. While $\eta = 0.5$ is suggested in the empirical analysis, we set $\eta$ also equal to 0.475 and 0.525 for checking the sensitivity of the results to the selection of $\eta$. GPH test is conducted on the log transformed series of the data. As $d$ of the level series equals $1 + d'$, a value of $d'$ equal to zero corresponds to a unit root in $Z$. Thus, the GPH test is used as a unit root test as we apply it to the first differences of the relevant series. The unit root null hypothesis $d = 1 (d' = 0)d$ can be tested against the one-sided, long-memory, fractionally-integrated alternative $d < 1 (d' < 0)$. The rejection of the unit-root null hypothesis would suggest the existence of fractional integration.

An alternative semi-parametric approach is that of Robinson (1992), who considers the discretely averaged Periodogram

$$F(\theta) = \int f(\lambda) d\lambda$$

Robinson estimates the Hurst coefficient, or equivalently the degree of fractional integration, $d$, as

---

3 Cheung and Lai (1993) conducted a Monte Carlo experiment and found better performance for $\eta = 0.55$, 0.575, and 0.6. In another investigation, Cheung (1993) used $\eta = 0.5$ (which is commonly used to test for fractional integration), and also reports results for $\eta = 0.45$ and 0.55 to check the sensitivity of the estimates. Overall, it may be inferred that, irrespective of sample size, a value of $\eta$, between 0.5 and 0.6 appears to be the ideal choice.
\[ H_q = \{1 - 2\log(q)\}^{-1} \log \left\{ F(q^{w_m}) / F(w_m) \right\} \]  

where \( q \in (0,1) \), \( \hat{F}(w_m) \) is the average of periodogram with \( I(w_j) \), with \( w_j = 2\pi j/T \) for \( j = 1, \ldots, m < T/2 \), and \( n \) is the sample size. Specifically, \( \hat{F}(w_m) = 2\pi n^{-1} \sum_{j=1}^{\{wT/2\pi\}} I(w_j) \) where \([ \cdot ]\) means the integer part. Robinson (1994) shows that the measure \( d \) is consistent as \( T \) gets large. We use \( q = 0.5 \) in this study. As with the rescaled range and GPH approaches, there is a substantial amount of evidence documenting the poor performance of the Robinson’s semi-parametric estimator in terms of bias (see Baillie, 1996).

3.3 Mean Reversion, Martingale and Variance Ratio Test

To examine further, whether the cointegrating series of stock markets follows a martingale, we employ a joint variance ratio test based on Lo and MacKinlay (1988). The Lo and MacKinlay (1988, 1989) overlapping variance ratio test, examines the predictability of time series data by comparing variances of differences of the data (returns) calculated over different intervals. If we assume the data follow a random walk, the variance of a \( q \)-period difference should be \( q \) times the variance of the one-period difference. Evaluating the empirical evidence for or against this restriction is the basis of the variance ratio test. Diebold (1989) show that the joint variance ratio test used in this study has good power against fractionally integrated alternatives.

The variance ratio test makes use of the fact that the variance of the increments in a martingale is linear in the sampling interval. That is, if the time series \( Z_t \) follows a martingale

\footnote{Lo and MacKinlay (1988) formulate two test statistics for the random walk properties that are applicable under different sets of null hypothesis assumptions about \( \varepsilon \). First, Lo and MacKinlay make the strong assumption that the \( \varepsilon \) are i.i.d. Gaussian with variance \( \sigma^2 \) (though the normality assumption is not strictly necessary). Lo and}
process, the variance of its $q$th differences would be $q$ times the variance of its first differences. Using a daily interval as the base in our study, the variance ratio at lag $q$, $VR(q)$, is defined as

$$VR(q) = \frac{\sigma_q^2}{q \sigma_1^2},$$

(7)

where $\sigma_q^2$ is an unbiased estimator of the variance of the $q$th difference of $Z_t$ and $\sigma_1^2$ is an unbiased estimator of the variance of the first difference of $Z_t$. Thus, under the martingale hypothesis, $VR(q)$ equals one for any $q$ chosen. The variance ratio estimate is closely related to the serial correlation of the series as the variance-ratio estimate for a given $q$ is approximately a linear combination of the first $q-1$ autocorrelation coefficients (Cochrane, 1988). Accordingly, rejecting the null hypothesis of martingale with a variance ratio of less (greater) than one implies negative (positive) serial correlations, or mean-reversion (mean-aversion). Thus, the $VR$ statistic provides an intuitively appealing way to search for mean reversion, an interesting deviation from the martingale hypothesis. The hypothesis is tested under the asymptotic distributions of both homoskedasticity-and heteroskedasticity-robust variance ratio estimators developed by Lo and MacKinlay (1988). Nevertheless, Lo and MacKinlay’s (1988) asymptotic distribution might not be an accurate approximation when $q$ is large and sample size is small (Richardson and Stock, 1990). We thereby provide the $p$ value of the $VR$ statistics in testing the null by using the bootstrap method.

MacKinlay term this the homoskedastic random walk hypothesis, though others refer to this as the i.i.d. null. Alternately, Lo and MacKinlay outline a heteroskedastic random walk hypothesis where they weaken the i.i.d. assumption and allow for fairly general forms of conditional heteroskedasticity and dependence. This hypothesis is sometimes termed the martingale null, since it offers a set of sufficient (but not necessary), conditions for $\mathbb{E}_t$ to be a martingale difference sequence (m.d.s.).
Furthermore, in earlier works (i.e., Liu and He, 1991), the martingale hypothesis is considered rejected when at least some of the VR statistics provide evidence against it. Richardson (1993) notes that the failure of including a joint test that combines all of the information from several VR statistics would tend to yield stronger results. To provide a joint test that takes account of the correlations between VR statistics at various horizons, we consider the two approaches one by Chow and Denning (1993) and second one by Richardson and Smith (1991). This is because the variance ratio restriction holds for every difference \( q > 1 \), it is common to evaluate the statistic at several selected values of \( q \). To control the size of the joint test, Chow and Denning (1993) propose a (conservative) test statistic that examines the maximum absolute value of a set of multiple variance ratio statistics. The \( p \)-value for the Chow-Denning statistic using \( m \) variance ratio statistics is bounded from above by the probability for the Studentized Maximum Modulus (SMM) distribution with parameter \( m \) and \( T \) degrees-of-freedom. Following Chow and Denning, we approximate this bound using the asymptotic \( (T = \infty) \) SMM distribution. A second approach is available for variance ratio tests of the \( i.i.d. \) null. Under this set of assumptions, we may form the joint covariance matrix of the variance ratio test statistics as in Richardson and Smith (1991), and compute the standard Wald statistic for the joint hypothesis that all \( m \) variance ratio statistics equal 1. Under the null, the Wald statistic is asymptotic Chi-square with \( m \) degrees-of-freedom.

However, there is possibility that the empirical distributions of VR statistics might have a large degree of positive skewness, suggesting that inference based on the \( \chi^2 \) distribution will be misleading. Accordingly, we calculate the Wald statistic for each bootstrapped VR estimator.
vector and use the bootstrapped distribution of Wald statistics for hypothesis testing by using

3.4 Long memory, mean reversion and rescaled range analysis

To further examine the characteristics of long-memory and mean-reversion in the cointegrating
series of stock markets, we employ the rescaled range (R/S) test. The R/S statistic is formed by
measuring the range between the maximum and minimum distances that the cumulative sum of a
stochastic random variable has deviated from its mean and then dividing this by its standard
deviation. An unusually small (large) R/S statistic signifies mean-reversion (mean-aversion).
Mandelbrot (1972) demonstrates that the R/S statistic can uncover not only periodic dependence
but also non periodic cycles. He further shows that the R/S statistic is a more general test of
long-memory in time series than examining either autocorrelations (i.e., the variance ratio test) or
spectral densities.

Lo (1991) points out that the original version of the R/S analysis (which may be termed
as classical R/S test) has limitations in that it cannot distinguish between short and long-term
dependence, nor is it robust to heteroskedasticity. Lo (1991) developed a test for short memory
versus long memory based on a simple modification of the rescaled range or RS statistic
introduced by Hurst (1951). The modified RS statistic is

5 Kim (2006) offers a wild bootstrap approach to improving the small sample properties of variance ratio tests. The
approach involves computing the individual (Lo and MacKinlay) and joint (Chow and Denning, Wald) variance
ratio test statistics on samples of observations formed by weighting the original data by mean 0 and variance 1
random variables, and using the results to form bootstrap distributions of the test statistics. The bootstrap p-values
are computed directly from the fraction of replications falling outside the bounds defined by the estimated statistic.
E-views offer three distributions for constructing wild bootstrap weights: the two-point, the Rademacher, and the
normal. Kim’s simulations indicate that the test results are generally insensitive to the choice of wild bootstrap
distribution.
\[ Q_n = \frac{1}{\delta} \left\{ \max_{1 \leq k \leq T} \sum_{i=1}^{k} (X_i - \bar{X}) - \min_{1 \leq k \leq T} \sum_{i=1}^{k} (X_i - \bar{X}) \right\}, \quad (8) \]

Under the null hypothesis, Lo (1991) showed that

\[ RS = \frac{Q_n}{\sqrt{n}} \Rightarrow U_{RS} = \max_{1 \leq k \leq n} W^0(t) - \min_{1 \leq k \leq n} W^0(t). \]

The distribution function of the random variable \( U_{RS} \) was derived by Feller (1951) and has the formula

\[ F_{U_{RS}}(x) = 1 + 2 \sum_{k=1}^{\infty} (1 - 4k^2x^2)e^{-2k^2x^2}. \]

This distribution function is used to calculate the critical values of the modified RS test. Critical values at various significance levels are available in Table 2 in Lo (1991). Further, for estimation in our analysis we considered two lag truncation procedures: a lag truncation equals to zero and the Andrews’ formula.\(^6\)

---

\(^6\) In practice, however, a bandwidth value \( q \) has to be selected in the construction of the tests. Consequently, the finite sample performance of these tests depends on the choice of the bandwidth. The most popular bandwidth choice is probably the data-dependent automatic bandwidth \( q = \mu_k \hat{\delta}(f, k)n^{1/(2p+1)} \) where \( \mu_k \) is a constant associated with the kernel function \( k \), \( \hat{\delta}(f, k) \) is a function of the unknown spectral density and is estimated using a plug-in method, and \( p \) is the characteristic exponent of \( k \). This bandwidth choice has been studied by Andrews (1991) in the estimation of a covariance matrix for stationary time series and is now widely used in econometrics applications. It has the advantage that it partially adapts to the serial correlation in the underlying time series through the data-dependent component \( \hat{\delta}(f, k) \). Lima and Xiao (2010) find that the problem of choosing the optimal \( q \) is not solved yet, but providing a bandwidth procedure that is robust against both the null and alternative models may help practitioners in their investigation on the presence of long memory in financial time series.
To achieve our objective i.e., to test for fractional cointegration, if it exists, we first run simple OLS regression using Indian stock market index as dependent variable and other stock market of the developed world as independent variables, sequentially, (both stock indices are measured in natural log form) and then saved residuals as error-correction term. In the second step we applied a modified GPH test, along with a battery of other tests such, GPH and the Rescaled Range to the first differences of the error correction terms test to test for the evidence of cointegration. Evidence of fractional cointegration will be found if null of unit root is rejected for the first differenced OLS residual series. We presented in Table 2 the summary statistics for the OLS residuals estimated from simple OLS regression analysis for the period under consideration in this study.7

Table 2: Summary Statistics and KPSS Unit Root Results for the OLS Regression Residuals

<table>
<thead>
<tr>
<th>Countries</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Jarque-Bera (p-values)</th>
<th>KPSS</th>
<th>ηₐ</th>
<th>ηₓ</th>
</tr>
</thead>
<tbody>
<tr>
<td>India-Germany</td>
<td>-4.59E-14</td>
<td>0.546</td>
<td>-0.3238</td>
<td>1.7967</td>
<td>262.277 (0.00)</td>
<td>36.7(6)</td>
<td>2.41(6)</td>
<td></td>
</tr>
<tr>
<td>India-Japan</td>
<td>7.39E-15</td>
<td>0.650</td>
<td>0.26161</td>
<td>1.57615</td>
<td>323.211 (0.00)</td>
<td>33.4(6)</td>
<td>4.96(6)</td>
<td></td>
</tr>
<tr>
<td>India-Switzerland</td>
<td>2.39E-14</td>
<td>0.621</td>
<td>0.22040</td>
<td>1.800454</td>
<td>229.399 (0.00)</td>
<td>39.6(6)</td>
<td>3.43(6)</td>
<td></td>
</tr>
<tr>
<td>India-United kingdom</td>
<td>1.07E-14</td>
<td>0.635</td>
<td>0.18151</td>
<td>1.544387</td>
<td>316.115 (0.00)</td>
<td>40.7(6)</td>
<td>4.63(6)</td>
<td></td>
</tr>
<tr>
<td>India-USA</td>
<td>-2.78E-14</td>
<td>0.463</td>
<td>0.80571</td>
<td>3.02676</td>
<td>364.833 (0.00)</td>
<td>26.1((6)</td>
<td>4.24(6)</td>
<td></td>
</tr>
</tbody>
</table>

Note: ***, ** and * indicate significance at the 0.01, 0.05 and 0.1 level, respectively. In case of KPSS test the lag length selection is based on the truncation for the Bartlett Kernel, as suggested by the Newey-West test (1987). Critical values for KPSS test at 1%, 5% and 10% level of significance 0.739, 0.463 and 0.347 respectively.

7 Diebold and Rudebusch (1991) demonstrate that the standard unit root tests have low power against fractional alternatives.
It appears from Table 2 that OLS residual series are extremely non-normal. India’s stock market cointegrating series with Germany is negatively skewed, possibly due to the large downturn associate with this stock market and with the rest of the developed countries it is positively skewed, possibly due to the large upturn associate with international market. The data also display a relatively less degree of excess kurtosis except with USA, indicating that OLS residual series are Platykurtic relative to a normal distribution. Such skewness and kurtosis are common features in such financial dataset. For all country case OLS residual series fail to satisfy the null hypothesis of normality of the Jarque-Bera test at the 1% level of the significance. However, non-stationarity does not preclude the possibility of long-memory in the India’s stock market cointegrating series with major developed countries.

Table 2 also presents the results of the KPSS (Kwiatkowski et al., 1992). Lee and Schmidt (1996) argue that KPSS test is consistent against stationary long memory alternatives and may therefore be used to distinguish between short and long memory processes. The KPSS test is similar to Lo’s modified rescaled range statistic in both power and construction (see Lee and Schmidt, 1996 and Baillie, 1996 for further details). However, through KPSS test we find that India’s stock market cointegrating series with major developed countries have long memory as in none of the case the stock market cointegrating series with major developed countries satisfy the null hypothesis of stationarity about a (possibly non-zero) mean of the $\eta_\mu$ and $\eta_\tau$ test at the 1% level of significance. These results suggest that India’s stock market cointegrating series with major developed countries contain a nonstationary component. In short, the evidence from the KPSS tests is in favor of the presence of long memory in the all series of India’s stock market cointegrating series with major developed countries.
Baillie (1996) comments that the GPH estimator is potentially robust to non-normality as our data set also exhibits that data (both raw data as well as residual series) is non-normal. Table 3 displays the results of the GPH and MGPH estimates of $d$, the degree of fractional integration, along with F-statistics for the null hypotheses of $d = 1$. A concern in the application of the GPH and MGPH estimator is the choice of $J$ (as in most cases, the results vary across the different values of $\eta$), the number of spectral ordinates from the periodogram of $X_t$, to include in the estimation of $d$. Here results are presented for the first differenced OLS residual series for $J = T^\eta$; where $\eta = 0.5, 0.55, 0.6$, where $T$ represents the sample size (i.e., number of observations).

Table 3: Results of the Geweke-Porter-Hudak (GPH) and Modified GPH (MGPH) Test for Fractional Integration and Robinson Estimates of Rescaled Range

<table>
<thead>
<tr>
<th>Countries</th>
<th>MGPH estimates</th>
<th>Rescaled range (R/S) analysis of long memory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\eta = 0.5$</td>
<td>$d$</td>
</tr>
<tr>
<td>India-Germany</td>
<td></td>
<td></td>
</tr>
<tr>
<td>India-Japan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>India-Switzerland</td>
<td></td>
<td></td>
</tr>
<tr>
<td>India-United</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kingdom</td>
<td></td>
<td></td>
</tr>
<tr>
<td>India-USA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[\text{Agiakloglou et al. (1993) suggest that the estimate of the order of fractional integration from the GPH method could be biased for a model with large ARMA parameters. However, if the estimates } d \text{ remain stable for different } \eta \text{'s, there is no hint of a bias due to ARMA parameters (Hassler and Wolters, 1995).}\]
### Table 1: GPH Estimates

<table>
<thead>
<tr>
<th>Country Pairs</th>
<th>GPH Estimates</th>
<th>MGPH Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>India-Germany</td>
<td>0.44</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>(0.45)</td>
<td>(0.64)</td>
</tr>
<tr>
<td></td>
<td>0.20</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>0.22</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.48</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>(0.59)</td>
<td>(0.58)</td>
</tr>
<tr>
<td>India-Japan</td>
<td>0.97</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>(0.30)</td>
<td>(0.10)</td>
</tr>
<tr>
<td></td>
<td>1.63</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>1.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.52</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>1.40</td>
<td>(0.16)</td>
</tr>
<tr>
<td>India-Switzerland</td>
<td>0.46</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>(0.76)</td>
<td>(0.10)</td>
</tr>
<tr>
<td></td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.73</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.12</td>
<td></td>
</tr>
<tr>
<td>India-United Kingdom</td>
<td>-0.32</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>-0.27</td>
<td>(0.04)</td>
</tr>
<tr>
<td></td>
<td>-2.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.72</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-2.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.93</td>
<td></td>
</tr>
<tr>
<td>India-USA</td>
<td>0.29</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>0.31</td>
<td>(0.84)</td>
</tr>
<tr>
<td></td>
<td>-0.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.65</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.63</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.51)</td>
<td>(0.52)</td>
</tr>
</tbody>
</table>

Note: The GPH and MGPH test statistics are F-statistics from the spectral regression. The null hypotheses of \( d = 1 \) and \( d = 0 \) are tested against the alternatives of \( d \neq 1 \) and \( d \neq 0 \), respectively. Values in the parenthesis are the p-values those are reported below the F-statistics.

Results reported in Table 3 of the GPH and MGPH estimates suggest that the results are not very much sensitive to the choice of \( \eta \) at 5% level of significance (with only exception for U.K. for choice of \( \eta = 0.55,0.6 \)). Interestingly, we have contrary results for GPH and MGPH estimates. The null hypotheses of \( d = 0 \) and \( d = 1 \) are rejected in all cases when we applied MGPH whereas these null hypotheses are not rejected in all cases when we applied GPH estimates. Worthy to mention that the null hypothesis of \( d = 0 \) is rejected for GPH estimates and the null hypothesis of \( d = 0 \) is not rejected for MGPH estimates only for U.K. for the choice of \( \eta = 0.55,0.6 \). As we know that MGPH is more powerful test statistics, we relied upon the results of MGPH only to draw conclusions. The rejection of the null hypothesis of \( d = 0 \) in all cases imply the presence of unit root in the first differenced OLS residual series i.e., nonstationary of India’s stock market cointegrating series with major developed countries.
However, the rejection of the null hypothesis of \( d = 1 \) irrespective of the value of \( \eta \), implying that the error-correction in all cases are clearly fractionally integrated (i.e., long memory), suggesting that the Indian stock market is fractionally cointegrated with the developed countries. Further, for the cases where the value of \( \eta \) implies that though the error-correction term is mean reverting (i.e., cointegrated), however, it is not a covariance stationary process.

Table 3 also displays the estimates of \( d \) calculated using the Gaussian semiparametric estimator of Robinson (1992). For \( \eta \) (i.e., for \( 0 < \eta < 0.5 \)) the series exhibits long memory in the cointegration series. The Robinson’s (1992) estimates of Rescaled range (R/S) analysis of long memory of \( d \) are outside the long memory range and close to 1 for all countries (and for Lag=0 and Lag=1), supporting the findings obtained from GPH test.

5. Conclusions

The study used the techniques of fractional co-integration involving the modified GPH test, along with a battery of other tests such as KPSS, GPH and the Rescaled Range for India, Germany, Japan, Switzerland, United Kingdom, and United States for the time period 1997 to 2011. Though the study gave contradictory result, with the earlier tests pointing to an absence of long memory, the powerful GPH test indicated long memory. The estimation confirms our hypothesis of the existence of fractional cointegration among the five pairs of stock markets, namely India-Germany, India-Japan, India-Switzerland, India-U.K. and India-U.S.A.
The results indicate that they conclude that the Indian financial liberalization opened up the Indian stock markets to the outside world and its markets are being influenced by other markets.

The confirmation of fractional cointegration by this paper should guide investors in their investment decisions regarding the Indian stock market. The results could be specifically used for both fundamental and technical analysis of stock markets. Since stock markets are a barometer for the economy, the results can also help us examine the overall economic situation and the existence of problems in the financial system.

It should be kept in mind that the use of daily closing stock indices would mean that the limits for daily price could be affected by the closing price of the previous day. It is plausible that the data would have hit a circuit breaker and thus be contaminated. Even so, daily price data is crucial to the understanding of the behavior of the series. It should be pointed out that relations among stock markets are also affected by macroeconomic variables such as trade and levels of foreign exchange. Future research should incorporate relevant macroeconomic series which might provide more insight into the working of equity market and/or some new tests of fractional cointegration can be used those incorporates structural breaks in the data set.

Acknowledgement

We are very thankful to an anonymous referee for his/her constructive comments and suggestions on the earlier version of the manuscript. We strongly feel that this paper would have remained flawed unless those suggestions were incorporated. We have incorporated all of his/her comments to the best of our abilities. Error remains, if any, is the authors sole responsibility. The usual disclaimer applies.
References


