RESEARCH PAPER | Management | Volume : 4 | Issue : 8 | August 2014 | ISSN - 2249-555X

Analyzing Long Range Dependence in Stock Markets of India

KEYWORDS: Long-range dependence, R/S, Hurst parameter, Fractional ARIMA

Prashant Joshi

Professor and Head, UKA Tarsadia University

ABSTRACT *Using daily stock return data of Nifty of National Stock Exchange and Sensex of Bombay Stock Exchange of India, we examine whether or not the stock market prices exhibit long range dependence. We employ modified R/S method of Lo(1991) and a variety of different time domain based as well as frequency domain based graphical and statistical methods described in Taqqua, Teverovskyand Willinger(1995, 1997) and Taqqu and Teverovksy(1997) for checking for long-run memory to the stock price series. We find empirical evidence of some- though rather weak (i.e. H values of around 0.57)long- range dependence in the stock price returns.*

1. Introduction

Long-range dependence is prevalent in nature (Mandlebrot, 1982). It has remained a topic of interest and extensive research in finance and economic time series (Lo, 1991, Cutland, Kopp and Willinger, 1995). Long-range dependence indicates the presence of nonperiodic cyclical patterns. Standard methods of statistical inference are not applicable to time series exhibiting such dependence(Lo, 1991).The existence of long range dependence in financial time series is inconsistent with the efficient market hypothesis which has served the basis of modern finance theory and its applications.

Hurst (1951) first proposed rescaled range or R/S statistics (range over standard deviation) to examine long range dependence which was further refined by Mandelbrot (1972, 1975). But both the methodssuffer a serious drawback that it does not differentiate short run and long run dependence in stock return. The R/S analysis turns out to be sensitive to short range dependence structure in data. It, therefore, has become less attractive to analyze dependence structure in data. The dependence structure of stock price returns has been at the centre of intense scrutiny for the last 30 years. Fama(1965, 1970) examined the dependence structure of asset return of thirty stocks of Dow Jones Industrial Average during 1957 to 1962 using simple significant tests for checking asset return with first few autocorrelation coefficient. He concluded that the successive returns follow random walk. Lo and MacKinlay (1988) investigated the random walk hypothesis after a careful analysis of market returns from a 25-year period during 1962-1987. They found substantial short-range dependence in the data and strongly rejected the hypothesis that asset returns are independent.

Lo (1991) came out with modified R/S statistics by replacing standard deviation of sample by square root of variance of partial sum. It also includes autocovariance of thepartial sum. Lo claimed that the modified R/S statics is robust with respect to short run dependence and it has reasonable power against certain long run dependence. Lo (1991) applied the modified R/S statistics to data sets of daily stock return indices from Centre for Research in Security Prices (CRSP). He found no evidence of long range dependence in the data. Tverovsky, Taqqu and Willinger(1997) identified numerous problems with Lo's method and its use. They found that Lo method is biased towards accepting null hypothesis of no long-range dependence irrespective of its presence or absence in data. Willinger, Taqqu and Teverovksy (1999) revisited the question of whether or not CRSP daily stock return data during 1962 to 1987 exhibit long range dependence using empirical investigations reported in Tverovsky, Taqqu and Willinger(1997).Their study advised against the sole use of Lo method to test for long range dependence in data structure and advocated the use of a diversified portfolio of time domain-based and frequency domain-based graphical and statistical methods like graphical R/S method, the modified R/S statistics with corresponding V_q-vs-q-plots and FARIMA methodto investigate long range dependence in data structure.

Our study tries to examine long-range dependence in stock market price data of National Stock Exchange and Bombay Stock Exchange of India using diverse portfolio of graphical and statistical methods discussed in Tverovsky, Taqqu and Willinger (1995, 1997) and Taqqu and Teverovsky (1997). The remaining part of the paper is organized as follows. In Section 2, we discuss methodology used in the study. Section 3 illustrates the results of our in-depth analysis of data set of two exchanges. In Section 4, we conclude.

2. Methodology

We offer discussions on Classical R/S analysis, Modified R/S statistics and FARIMA methods used in our analysis to test for long range dependence in our data set in the section.

2.1Classical R/S Analysis

The classical R/S analysis developed by Mandelbrot and Wallis (1969) tries to analyze historical information fully. It subdivided a given sample of N observations into K blocks, each of size N/K. For each lag n,n≤N, estimates $R(k_m,n)/S(k_m,n)$ of R/S statistics is defined by,

$$\frac{R}{S}(n) = \frac{1}{S(n)}\left[\max_{0\le t\le n}\left(Y(t) - \frac{t}{n}Y(n)\right) - \min_{0\le t\le n}\left(Y(t) - \frac{t}{n}Y(n)\right)\right], n \ge 1. \text{ Equation 1}$$

The estimates are computed by starting at the points, k_m= (m-1) N/K+1, m=1, 2, 3…., K, and such that k_m+n≤N. Thus, for a given m, all data points before k_m=mN/K+1 are ig-

Volume : 4 | Issue : 8 | August 2014 | ISSN - 2249-555X

nored. There are K different estimates of R(n)/S(n) for values of n smaller than N/K; for values of n approaching N, there are fewer values as few as 1 when n≥N-N/K. The R(k_m,n)/S(k_m,n) values corresponding to neighboring values of km and n are strongly interdependent; for a given n> N/K, the various estimates of R(n)/S(n) involve overlapping observations and estimates when evaluated at different lags. Hurst (1951) found that the relation E[$\frac{R}{S}(n)$] approaches to $c_1 n^H$ as $n \to \infty$ with typical value of Hurst Parameter (H) in the interval (0.5,1) suggesting long range dependence and a finite positive constant that does not depend on n. If the observations X_i come from a short-range dependent model, then $\frac{R}{S}(n) \sim c_2 n^{0.5}$ as $n \to \infty$ where c_2 is independent of n, finite and positive. The discrepancy between these two relations is generally referred to as the Hurst Effect or the Hurst parameter. The classical R/S analysis tries to find out value of the Hurst parameter H from time series.

The graphical approach consists calculating the estimates of R(k_m,n)/S(k_m,n) for logarithmically spaced values of n, and plotting log(R(k_m,n)/S(k_m,n)) vs. log(n) for all starting points k_m. This results are known as pox plot of R/S.

2.2Modified R/S Statistics

Lo(1991) modified the classical R/S by considering only the length of the series instead of multiple lags. He used weighted sum of autocovariances instead of sample standard deviation in the denominator of R/S. It is defined by

$$S_q(N)=\left(\frac{1}{N}\sum_{j=1}^{N}(X_j-\bar{X}_N)^2+\frac{2}{N}\sum_{j=1}^{q}\omega_j(q)\left[\sum_{i=j+1}^{N}(X_i-\bar{X}_N)(X_{i-j}-\bar{X}_N)\right]\right)^{1/2} \quad \text{Equation 2}$$

It tests the following null and alternative hypothesis:

H_o = { no long-range dependence, i.e., H=0.5}

Against the composite alternative

H_a = { there is long-range dependence, i.e. ½ < H < 1}

Here, $\bar{X}_N$ denotes the sample mean of the time series, and the weights $\omega_j(q)$ are given by

$$\omega_j(q) = 1 - \frac{j}{q+1}, q<N$$

Lo presented modified R/S statistics, $V_q(N)$ as follows

$$V_q(N)=N^{-1/2}R(N)/S_q(N) \quad \text{Equation 3}$$

with R given in equation 1. Lo used $S_q(N)$ in the denominator to compensate for the extra short-range dependencies that may be present in the data. If the time series contains both short and long range dependencies, the value of $V_q(N)$ lie outside the confidence interval of [0.809, 1.862]. Lo method is used to detect long-range dependence in a given time series. The choice of q in Lo's method is crucial since it influences the actual size of the test as well as its power.Tverovsky, Taqqu and Willinger(1997) show that for large q-values,

$$V_q(N) \sim q^{1/2-H} \quad \text{Equation 4}$$

It suggests, as q increases, the test statistic V_q decreases for H>1/2 and it will be within the confidence interval of [0.809, 1.862] for large enough q. The modified R/S Statistics with corresponding Vq-vs-q are plotted to examine dependence structure in the data. The very stable plot without any fluctuation suggests no long range dependence.

Tverovsky, Taqqu and Willinger(1997) performed detailed Monte Carlo simulation study. Using simulated purely long-range dependence time series also known as Fractional Gaussian Noise(FGN) with Hurst parameter (H), 0.5 < H < 1, purely short-range dependent sequences namely ARMA (p,q) process as well as combination of short-range and long-range dependent processes i.e fractional autoregressive-moving average stationary time series FARIMA (p,d,q) with d = H-1/2, 0<d<0.5, they found that the practical application of Lo's modified R/S-statistics requires care and has a problem of strongly accepting the null of no long-range dependence and found a strong dependence between the outcome of test and the choice of lag q. They recommended to use a wide range of different q-values and associated V_q values to get reliable results. Following the advice of the Tverovsky, Taqqu and Willinger, we use FARIMA or ARFIMA to test for the long range dependence in the time series in addition to other graphical and statistical methods described above.

2.3ARFIMA

Granger and Joyeux (1980) and Hosking (1981) introduced ARFIMAto test long memory property in the asset returns. The purpose of this model to consider fractionally integrated process I(d) in the conditional mean. The ARFIMA (p,d,q) model can be expressed as follows:

$$\psi(L)(1-L)^{\xi}(r_t-\mu)=\theta(L)\varepsilon_t \quad \text{Equation 5}$$

$$\Delta r_t = \alpha + \delta r_{t-1} + \sum_{i=1}^{p}\beta_i \Delta r_{t-i} + \varepsilon_t$$

$$\Delta r_t = r_t - r_{t-1}; r_t = \ln(R_t) \quad \text{Equation 6}$$

Where ε_t is independent and identically distributed with variance σ^2 and L denotes the lag operator and replacing with difference operator (1-L) of an ARIMA process with the fractional difference operator $(1-L)^d$, where d denotes the degree of fractional integration. The differencing parameter d need not be an integer, but the integer value of d leads to a traditional ARMA models. If $0<d<0.5$, all autocorrelations are positive implying long memory while they are negative if $-0.5<d<0$. The negative values indicate that the process exhibits negative dependence between distant observationsuggesting anti-persistence. The process is said to be stationary when $d=0$. For $d=1$, the process follows a unit root process. It is fractional autoregressive-moving average stationary time series FARIMA (p,d,q) with d = H-1/2, 0<d<0.5.If the value of H parameter is greater than 0.5, the series is said to have long range dependence.

The above model- ARFIMA is applied to stationary time series. Time series can be tested for stationarity using unit root test. A unit root test is a statistical test for the proposition that in an autoregressive statistical model of a time series, the autoregressive parameter is one. It is a test for detecting the presence of stationarity in the series. The early and pioneering work on testing for a unit root in time series was done by Dickey and Fuller (1979, 1981). If the variables in the regression model are not stationary, then it can be shown that the standard assumptions for asymptotic analysis will not be valid. In other words, the usual "t-ratios"will not follow a t-distribution, so we cannot validly

Volume : 4 | Issue : 8 | August 2014 | ISSN - 2249-555X

undertake hypothesis tests about the regression parameters.

Stationary time series is one whose mean, variance and covariance are unchanged by time shift. Nonstationary time series have time varying mean or variance or both. If a time series is nonstationary, we can study its behaviour only for a time period under consideration. It is not possible to generalize it to other time periods. It is, therefore, not useful for forecasting purpose.

The presence of unit root in a time series is tested with the help of Augmented Dickey-Fuller Test. It tests for a unit root in the univariate representation of time series. For a return series Rt, the ADF test consists of a regression of the first difference of the series against the series lagged k times as follows:

$$\Delta r_t = \alpha + \delta r_{t-1} + \sum_{i=1}^{p} \beta_i \Delta r_{t-i} + \varepsilon_t \quad \text{Equation 7}$$

$$\Delta r_t = r_t - r_{t-1}; r_t = \ln(R_t)$$

The null hypothesis is H0: and H1: $\delta < 1$. The acceptance of null hypothesis implies nonstationary.We can transform the nonstationary time series to stationary time series either by differencing or by detrending. The transformation depends upon whether the series are difference stationary or trend stationary.

2.4Period of study

We collected data on daily closing price Sensex of BSE and S&P CNX Nifty of NSE stock price indices from January 1, 2000 to March 15, 2014. The period is recent one and major changes took place during the period like internet trading, derivative trading, index based future and options, currency derivatives banning of carry forward facility etc. These changes might have influenced dependence structure of data and therefore it will be instructive to study it in this period.

3. Empirical Results

Before investigating the long range dependence in return, we check the series for a presence of unit root. Stationarity condition of the Sensex and Nifty daily return series were tested by Augmented Dickey-Fuller Test (ADF). The results of this test reported in the Table 1.

Table 1 ADF tests Results-Sensex and Nifty

Null Hypothesis	Test statistics				Mackinnon Asymptotic
Presence of Unit root	Level		First Difference (Return Series)		Critical value @ 1% level
	Sensex	Nifty	Sensex	Nifty	
Intercept	-0.37	-0.42	-20.64	-42.48	-3.13
Intercept and Trend	-2.74	-2.88	-20.65	-42.47	-3.96
None	-1.19	-1.08	-20.59	-42.45	-2.56

ADF statistics in level series shows presence of unit root in both the stock markets as their Mackinnon's value do not exceed the critical value at 1% level. It suggests that both the returns series are nonstationary. It is, therefore, necessary to transform the series to make it stationary by taking its first difference.

ADF statistics reported in the Table 1 show that the null hypothesis of a unit root in case of both Sensex and Nifty is rejected. The absolute computed values for both the indices are higher than the MacKinnon critical value at 1% level. Thus, the results of both the indices show that the first difference series are stationary.

A plot of Nifty and Sensex time series are shown in Fig 1 and 2 respectively. The preliminary analysis of the graph suggests that the series is stationary and the returns fluctuate around mean value.

Figure 1Traces of the Nifty daily data

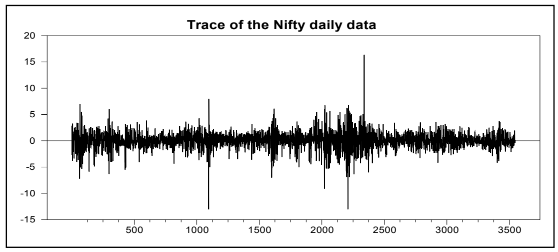


Figure 2Traces of the Sensex daily data

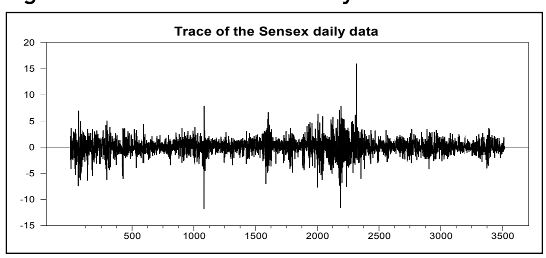


Figure 3R/S Analysis of Nifty Return, H = 0.5654

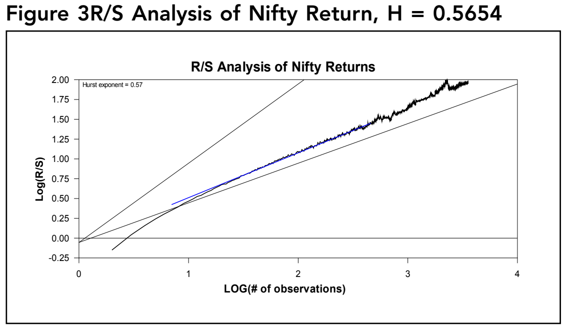


Figure 4R/S Analysis of Sensex Return, H=0.5653

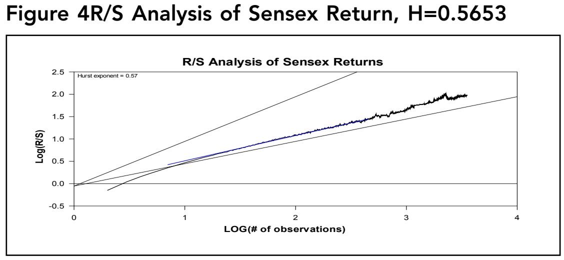


Figure 3 and 4 depict the pox plot of R/S corresponding to Nifty and Sensex series and results in an estimate of Hurst parameter of about 0.5654 and 0.5653 respectively.

Volume : 4 | Issue : 8 | August 2014 | ISSN - 2249-555X

Figure 5V(q) for the Nifty Series on a log-log scale

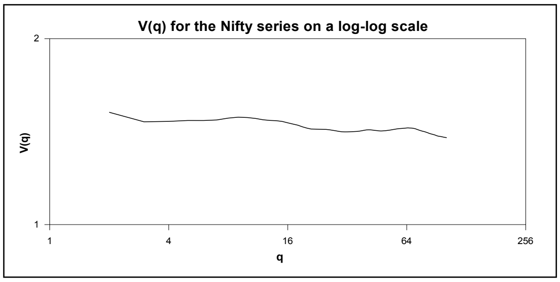
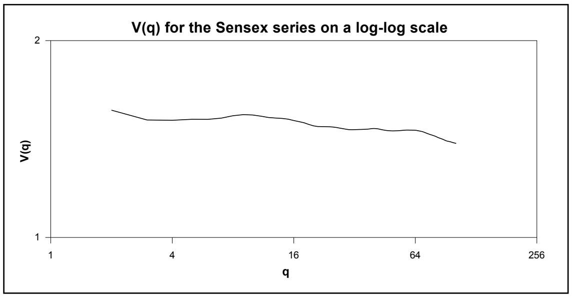


Figure 6V(q) for the Sensex Series on a log-log scale



Returning to modified R/S statistics, Figure 3 shows log-log version of V_q Vs. q for a range of q values. We observe some fluctuations in V_q suggesting weak long range dependence.

We estimated FARIMA (1, d, 1) model for Nifty and Sensex series. The resulting H estimates are 0.486 for both Nifty and Sensex, with AR and MA coefficients [(-0.024, -0.33)] for Nifty and [(-0.157, -0.24)] for Sensex and they are significantly different from zero. The values of H for both the stock indices indicate weak long range dependence. When we aggregate series over blocks of 10 and 20, respectively, and assuming now an underlying FGN model, we obtain H-estimate is about 0.5533 for blocks of size 10 and 0.5520 for blocks size of 20 for Nifty. H estimate for Sensex is about 0.5665 for blocks of size 10 and 0.5473 for blocks of size 20. All values are statistically significant and different from zero.

In order to remove any "extra" short range dependence and isolate hidden pure long range dependence effects that may be present in a given data set, we partition the time series into non overlapping blocks, m, of size 10, 20, 40 and shuffle the observations within each block. The shuffling destroys any particular lag structure of the auto-correlation function below m but to keep the high lags essentially intact.Analyzing the shuffled data using graphical method and modified R/S method, we obtain results which are very similar to the unshuffled data. We also performed a shuffling experiment where we shuffled blocks of different size but left the observations within each block intact. Such a block has an effect of maintaining the short range dependence by eliminating any long range memory in the data. We obtain statistically significant H-estimates of 0. 5494 and 0.5454 for Nifty and Sensex respectively indicating presence of long range dependence though weak.

4. Conclusion

To summarize the results of our analysis of NSE and BSE data, we have obtained some though weak evidence of long range dependence in the data structure. Moreover, when applying a variety of different methods for checking for long-run memory to both the series, the results consistently indicate the presence of some though rather weak long-range dependence since H-values come to around 0.57. This implies it is possible to predict future prices and gains could be obtained trading in these markets, contrary to what efficient market theory points out. We have not attempted to measure long range dependence structure in volatility. Future research can be undertaken to analyze long memory property in volatility using Fractionally Integrated ARCH and GARCH classes of models for risk analysis.

REFERENCE • AbryP. andVeitchD. (1998). Wavelet analysis of long range dependent traffic. IEEE Transactions on Information Theory, 44 (1), 2–15. | • AnnisA. and LloydE. H. (1976).The expected value of the adjusted rescaled Hurst range of independent normal summands.Biometrika 63, 111–116. | • B. Mandelbrot(1967). Forecasts of future prices, unbiased markets and martingale models.Journal of Business, 39, 242–255. | • BeranJ.(1994).Statistics for Long-Memory Processes. New York: Chapman & Hall. | • CampbellJ. Y., LoA. W. andMacKinlayA. C.(1997).The Econometrics of Financial Markets. Princeton University Press, NJ. | • CutlandN. J., KoppP. E., WillingerW.(1995).Stock price returns and the Joseph effect: a fractional version of the Black-Scholes model. In: E. Bolthausen, M. Dozzi, F. Russo (eds.) Seminar on Stochastic Analysis, Random Fields and Applications. Boston: Birkha¨user, pp 327–351 | • Dickey D. and Fuller W. A. (1979). Distribution of Estimates for Autoregressive Time Series with a Unit Root.Journal of American Statistical Association, 74, 427 – 31. | • Dickey D.and Fuller W. A. (1981). LikelyhoodRatro Statistics for for Autoregressive Time Series with a Unit Root. Econometrica, 49, 1057 – 72 | • FamaE.(1965). The behavior of stock market prices.Journal of Business, 38, 34–105. | • FamaE.(1970).Efficient capital markets: A review of theory and empirical work.Journal of Finance, 25, 383–417. | • Granger and Joyeux(1980). An introduction to long memory property of stock market returns and a new model.Journal of Time Series Analysis, 1, 15-39. | • Hosking (1981). Fractional Differencing.Biometrika, 68, 165-176. | • HurstH. E. (1951).Long-term storage capacity of reservoirs.Transactionof American Society of Civil Engineer, 116, 770–808. | • LoA. W. (1991). Long-term memory in stock market prices.Econometrica, 59, 1279–1313. | • LoA. W. andMacKinlayA. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test.The Review of Financial Studies, 1, 41–66. | • MandelbrotB. B. (1971).A fast fractional Gaussian noise generator.Water Resources Research, 7, 543–553 | • MandelbrotB. B. (1972).Statistical Methodology for non-periodic cycles:From the covariance to R/S Analysis. Annals of Economic and Social Measurement, 1, 259-290. | • MandelbrotB. B. (1975). Limit theorems on the self-normalized range for weakly and strongly dependent processes.Z. Wahrscheinlichkeitstheorieverw, Gebiete, 31, 271-285. | • MandelbrotB. B. (1982).The Fractal Geometry of Nature, W.H. Freeman, San Francisco. | • MandelbrotB. B. andTaqquM. S. (1979).Robust R/S analysis of long-run serial correlation. In Proceedings of the 42nd Session of the International Statistical Institute, Manila, Bulletin of the International Statistical Institute. Vol.48, Book 2, pp. 69-104 | • MandelbrotB. B. and WallisJ. R. (1968).Noah, Joseph and operational hydrology. Water Resources Research, 4, 909–918. | • MandelbrotB. B. and WallisJ. R. (1969)."Robustness of the rescaled range R/Sin the measurement of noncyclic long-run statistical dependence.Water Resources Research, 5, 967–988. | • MandelbrotB. B. and WallisJ. R.(1969).Some long-run properties of geophysical records.Water Resources Research, 5, 321–340. | • TaqquM. S. and TeverovskyV. (1997).Estimating long-range dependence in finite and infinite variance series.Published In AdlerR., FeldmanR. andTaqquM. S. (eds.),A Practical Guide to Heavy Tails: Statistical Techniques for Analyzing Heavy-Tailed Distributions. Boston: Birkha¨user. | • TaqquM. S., TeverovskyV. and WillingerW.(1995). Estimators for long-range dependence: an empirical study.Fractals, 3(4) 785–798. | • TeverovskyV.,Taqqu M. S. and WillingerW. (1997).On Lo's modified R/S statistic. Preprint. | • Willinger W., Taqqu M. S., and Teverovsky V. (1999). Stock market prices and long-range dependence.Finance and Stochastics, 3, 1-13. |