Long term memory in the Karachi Stock Exchange 100 Index

Faisal H. Zai

5th November 2012

1 Introduction

Long term memory in financial data has been a common subject of empirical research. Mandelbrot (1971) was among the first to study the implications. It has been referred to as the 'Joseph effect' by Mandelbrot and Wallis (1968) in relation to the story of Joseph who foretold a seven years of plenty followed by seven years of drought in Egypt in the Old Testament and the Koran. Mathematically long term memory can be defined as the presence of autocorrelations at long lags which die slowly. It is a phenomenon which has been widely observed and documented in hydrology, meteorology and other natural sciences hence it is not surprising that financial time series also exhibit long term dependence. The presence of this dependence implies that market prices do not react immediately to information and thus past returns can be used to predict future returns. This in turn raises questions about the validity of the Efficient Market Hypothesis (EMH) which underlies most of modern day financial economics. Indeed martingale based derivative pricing theory which depends on efficient memoryless financial markets stands invalidated in the presence of long term memory. Similarly investment decisons tend to become more sensitive to investment horizons in the presence of long term dependence. Presence of long term memory in returns highlights the need for non-linear pricing models as opposed to the common linear ones. Hence it is imperative to understand the presence of long term dependence in financial markets before embarking on investment decisions particularly those dependent on martingale pricing methods.

2 Long term memory

A process is said to exhibit long term memory if its autocorrelation function decays at a hyperbolic rate thus making it un-integrable. The autocorrelation function for such a stationary process looks like:

$$\rho(k) \to C_\rho k^{-\alpha} \text{as } k \to \infty$$

where $\rho(k)$ is the autocorrelation at lag k, C_ρ is a positive constant and α is real number between 0 and 1. A smaller value for α means more long term memory in the process.

H, the Hurst exponent is defined as:

$$H = 1 - \alpha/2$$

H is linearly related to α and a higher value of H means higher long term memory in the process. The usual method is to calculate H rather than α. H has usually been calculated using the R/S statistic. This statistic was used by Hurst in 1951 to test for long term memory in the flooding of the Nile. It was later modified but basically it compares the minimum and maximum values of running sums of deviation from the sample mean. The sums are standardised by dividing them by the sample standard deviation. More details of the R/S statistic along with its relative merits and demerits can be found in Mandelbrot and Wallis (1969a), Mandelbrot and Taqqu (1979) and Lo(1991). The R/S statistic is calculated as follows:

$$Q_T = \frac{R_T}{s_T} = \frac{1}{s_T}[\max_{1 \leq k \leq T} \sum_{j=1}^{k} (y_i - \bar{y}) - \min_{1 \leq k \leq T} \sum_{j=1}^{k} (y_i - \bar{y})]$$

where

$$s_T = [\frac{1}{T} \sum_{j=1}^{T} (y_i - \bar{y})^2]^{\frac{1}{2}}]$$

and

$$\bar{y} = \frac{1}{T} \sum_{j=1}^{T} y_i$$

where y_i is the observed time series upto time T

Lo (1991) showed that asymptotically the R/S statistic follows:

$$\frac{R_T}{s_T} \sim cT^H,$$

where c is a constant. This can be written as

$$log\frac{R_T}{s_T} = logc + HlogT$$

Hence H can be estimated using simple linear regression. If R_T is a simple random walk (RW) then the plot is a straight line with slope 0.5. If the returns have long term memory the slope is >0.5, in case of anti-persistence or negative long term dependence the slope is <0.5. However, Weron (2002) mentions that for small T there can be significant deviation from 0.5 for RW and hence the R/S statistics for RW is approximated by

$$E(R_T/s_T) = \begin{cases} \frac{T-\frac{1}{2}}{T}\frac{\Gamma((T-1)/2)}{\sqrt{\pi}\Gamma(T/2)}\sum_{i=1}^{T-1}\sqrt{\frac{T-i}{i}}, & \text{T} \leq 340 \\ \\ \frac{T-\frac{1}{2}}{T}\frac{1}{\sqrt{T\pi/2}}\sum_{i=1}^{T-1}\sqrt{\frac{T-i}{i}}, & \text{T}>340 \end{cases}$$

where Γ is the Euler gamma function. H is thus calculated as 0.5 plus $R_T/s_T - E(R_T/s_T)$. Lo(1991) provided a different adjustment to the R/S statistic however this paper uses the Weron (2002) adjustment specified above.

A long term memory process can also be modelled using a fractional Auto Regressive Integrated Moving Average (FARIMA or ARFIMA) process i.e between stationary and unit root process. It can mean revert like a process with finite memory but unlike the autoregressive stationary process it decays at the hyperbolic rate which is much slower than the exponential decay. Hence it takes a longer time to return to equilibrium. A time series with unit root at level but stationary at first difference is called $I(1)$ process. A stationary process is called $I(0)$. A long memory process (ARFIMA) is $I(d)$, where d lies between 0 and 1. An ARFIMA process of order (p,d,q) with mean μ can be written as:

$$\Phi(L)(1-L)^d(y_t-\mu) = \Theta(L)\epsilon_t,$$

where y_t is the observed time series,

$$\epsilon_t \sim iid(0, \sigma^2),$$

L is the lag operator,

$$L^i y_t = y_{t-j},$$

$$\Phi(L) = 1\text{–}\phi_1 L - \ldots.. - \phi_p L_p,$$

$$\Theta(L) = 1 + \lambda_1 L + \ldots.. + \lambda_q L_q$$

and $(1-L)^d$ is the fractional differencing operator defined by

$$(1-L)^d = \sum_{k=0}^{\infty}\binom{d}{k}(-1)^k L^k$$

d can be any real number. For standard ARIMA models d is an integer. The

process y_t is stationary and invertible if the roots of $\Phi(L)$ and $\Theta(L)$ lie outside the unit circle and $|d| \leq 0.5$. For d:
(i) between 0 and 0.5, the ARFIMA process is persistent and exhibits long term memory. The autocorrelations are all positive and decay hyperbolically to zero as the lag length increases. It also implies that the market inefficiency
(ii) between -0.5 and 0 the process is anti-persistent and exhibits intermediate or negative long term dependence. The autocorrelations are all negative, and implies an ultra-efficient market
(iii) equal to 0 the process exhibits short memory and is the standard ARMA process

The parameter d can be determined by using maximum likelihood estimation. This paper uses the method described by Haslett and Raftery (1989).

There are many ways to test for long term dependence, the analysis here estimates both H and d using the methods described. For a time series with long term memory, $d = H - 0.5$ (Mukherjee et al 2011). If the analysis is consistent the estimates of H and d should be consistent with this relationship.

3 Previous studies on long term memory in financial markets

Lo (1991) carried out an analysis of long term memory on the US data and found no evidence of long term memory in the returns. Similar studies were carried out for other developed markets and there was no evidence for long term memory.

Emerging markets are a good source of diversification for asset managers. Traditionally these markets have been seen as inefficient and thus expected to display significant long term memory. Recent studies carried out for some of these reported evidence of long term memory for example Greece (Barkaoulas et al (2000)) and Finland (Tolvi (2003)). Surprisingly, the stock market returns in India did not exhibit any long term memory according to a recent paper by Mukherjee, Sen and Sarkar (2011).

This paper carries out an analysis of long term memory in the stock returns in Pakistan using the Karachi Stock Exchange 100 (KSE100) Index. Results are also produced for the US market using S&P500 data to contrast and compare with those of the KSE100. The paper analyses the raw returns as well as the absolute and squared-return data. The presence of long term memory in the raw return implies that quantitative trading strategies can be set up to profit from market inefficiency

Weekly returns	KSE100	S&P500
Mean	0.28%	0.06%
Stand. deviation	3.88%	2.68%
Skewness	-0.96	-0.71
Kurtosis	3.44	5.51

Table 1: Summary statistics

4 Statistical Analysis

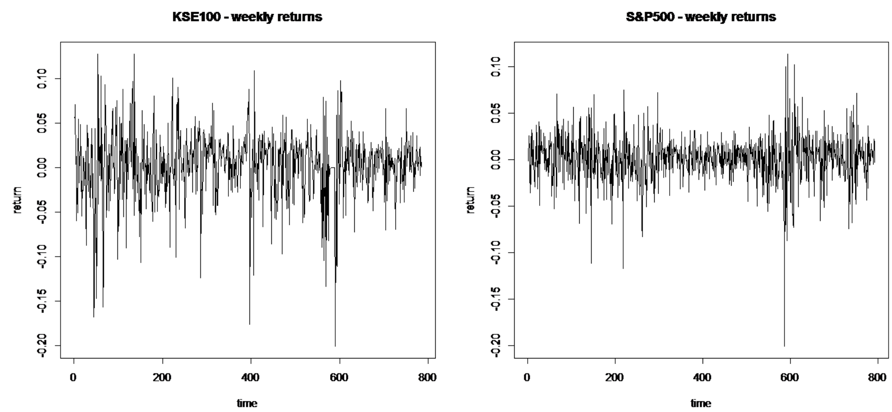
The analysis is carried out in R Project. This is a statistical software which is freely available on the internet and widely used by statisticians. R has a function to calculate the Hurst exponent using the R/S analysis. This function uses the methodology defined in Weron(2002) There is also a function to calculate the maximum likelihood estimators (MLE) of an ARFIMA model. The likelihood is approximated using the method described in Haslett and Raferty (1989).

4.1 Data and summary statistics

Weekly index prices from July 1997 till October 2012 for the KSE100 and S&P500 have been taken from Yahoo finance. The (local currency) return R_t for each dataseries is calculated as the difference is log index prices:

$$R_t = \ln(P_t) - \ln(P_{t-1})$$

Table 1 shows the weekly return statistics. As expected KSE100 has the higher average weekly return with higher volatility, measured as the standard deviation of return. Both the indices exhibit high skewness and kurtosis.



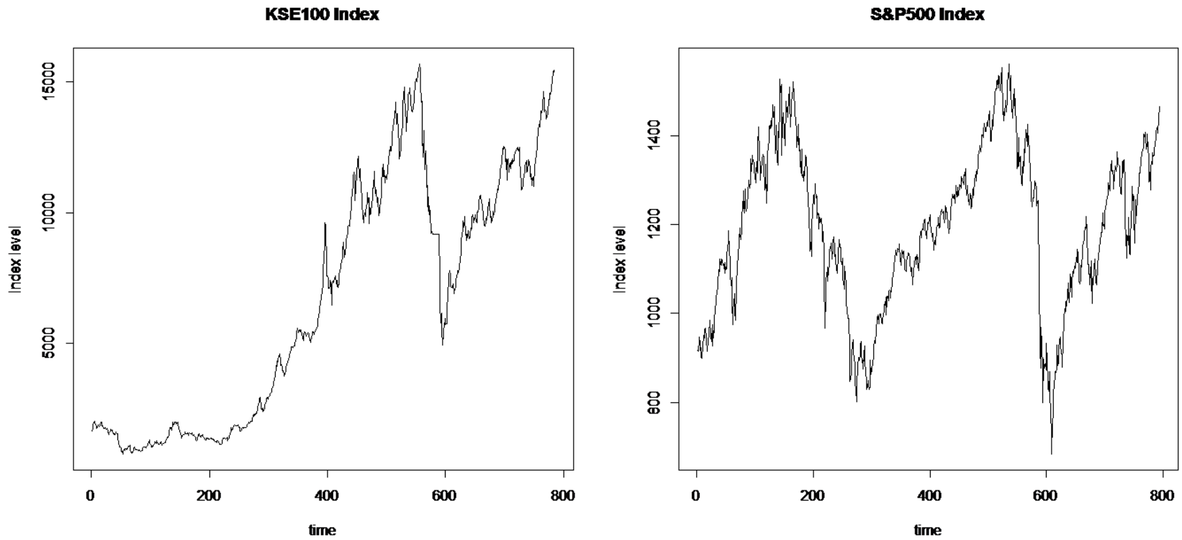


Figure 1: Weekly index prices from 2 July 1997 to 20 September 2012

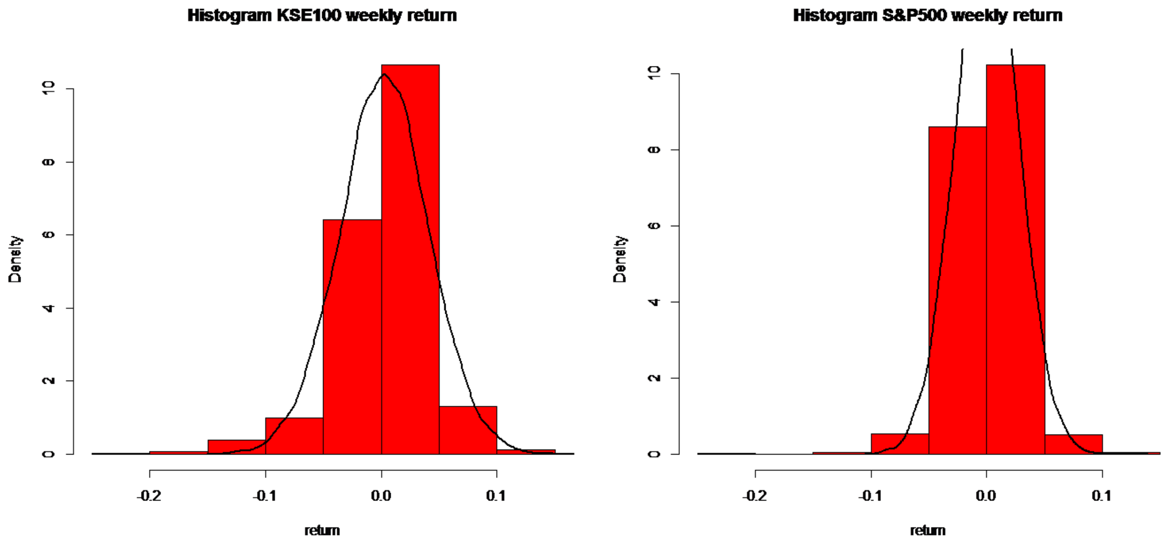


Figure 2: Comparing the index return distributions with a normal density (black line)

Figure 3: Weekly returns

Figure 1 shows the weekly levels of the KSE100 and S&P500 indices since 1997. The two appear to stationarity and there is no obvious trend in the indices. Stationary is tested statistically later.

Figure 2, illustrates the effect of negative skewness and high kurtosis. The red histograms show the weekly return distribution and the black line represents the density of a normal distribution with mean and standard deviation similar to the relevant index. The tails for the actual returns are heavier than we would expect under normally distributed returns. The KSE100 data displays higher skewness and heavier tails as borne out by the statistics in Table 1.

Figure 3 shows the weekly returns for the two markets. The return on the KSE100 appears to be more volatile as borne out by the higher standard deviation of returns in Table 1. There appears to be periods of high volatility followed by periods of relatively lower volatility i.e. volatility clustering in both the KSE100 and S&P500.

4.2 Analysis of long term memory in raw returns

The section analyses the KSE100 and S&P500 for long term memory in the raw returns. We first test for stationarity in the time series, then visually inspect the ACF for long term memory (i.e. slow decaying ACF) and finally apply the long term memory tests.

4.2.1 Augmented Dickey-Fuller (ADF) Test of stationarity

The(ADF) procedure is used to test the stationarity of the time series. Stationarity in the time series is important as we can apply the long term memory analysis to a stationary time series. The Augmented Dickey Fuller test is used to check stationarity in the time series and the results showed that both the KSE100 and S&P500 are stationary (at the 10% significance level). The output from R is in the appendix.

4.2.2 Visual inspection of the Auto Correlation Functions (ACF)

Graph 3 shows the ACFs of the two time series for weekly returns. The graphs show that there might be some long term memory in the KSE100 returns but probably not in the S&P500 return series. The blue lines represent the confidence intervals. Significant autocorrelations lie outside the interval.

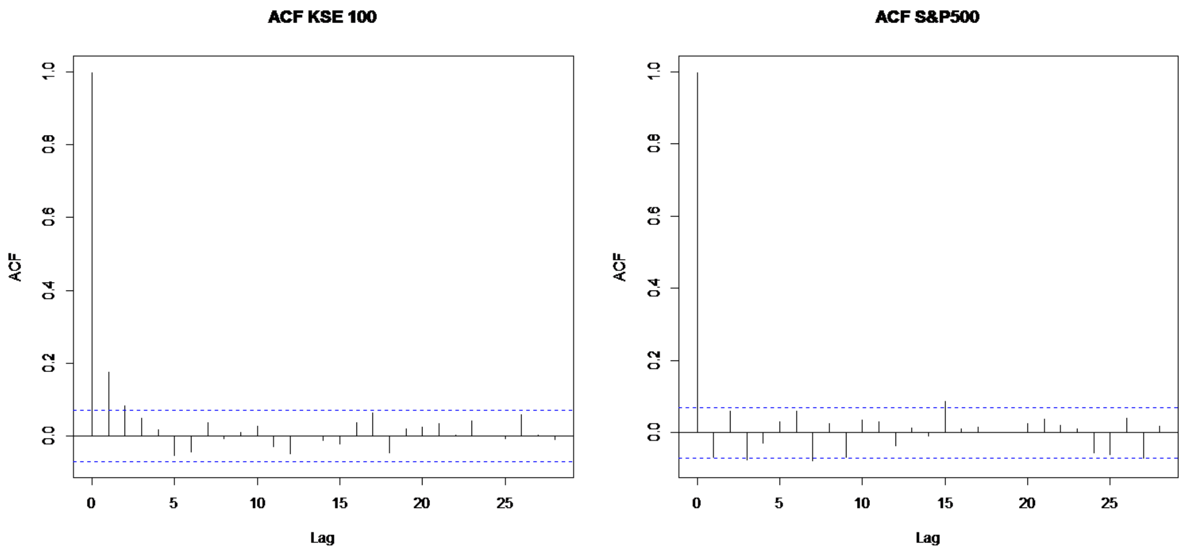


Figure 3: Autocorrelations

Market	Hurst Exponent
KSE100	0.618
S&P500	0.50

Table 2: Empirical Hurst exponents

4.3 Long term memory tests

This tests for the presence of long term memory in the time series.

4.3.1 Hurst exponent using the R/S statistic

This tests for the presence of long term memory in the time series. The empirical Hurst exponents (H) for the two time series are shown in Table 2

If $H = 0.5$ there is no evidence of long term memory in the process. The S&P500 has H equal to 0.5 suggesting the absence of long term memory. This was also shown by the visual ACF evidence. KSE100 has a higher H thus exibiting stronger longer term memory.

4.3.2 ARFIMA test

This section tests for the presence of the fractional difference component d in the ARFIMA model using the MLE method mentioned before. The analysis produces

Market	Lower limit	Higher limit
KSE100	0.1350	0.1352
S&P500	$9.787e^{-06}$	$9.187e^{-05}$

Table 3: Confidence interval for the fractional difference parameter d

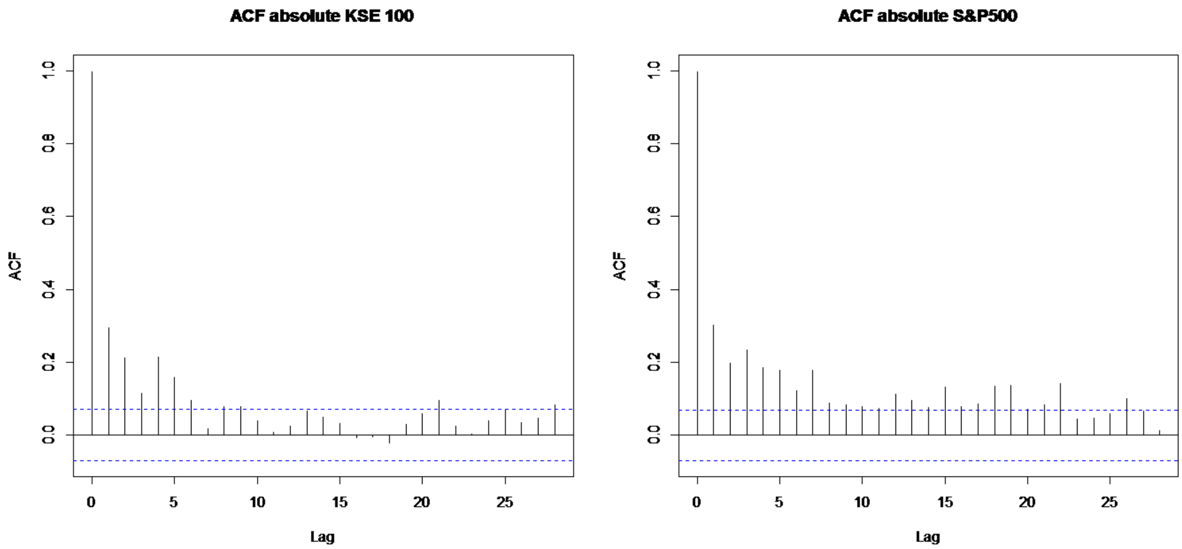


Figure 4: Autocorrelations of absolute returns

the 95

Table 3 shows that d is very close to 0 for S&P500 suggesting the absence of long term memory in the S&P500 returns as shown by the preceding analysis here and in Lo(1991). There is however significant long term memory in the KSE100 returns. It is also important to note that d and H satisfy the relationship mentioned previously i.e. $d \sim H - 0.5$.

4.4 Analysis of long term memory in absolute and squared returns

Absolute and squared return data is used as a measure of volatility in the data. This analysis looks at the presence of long term memory in return volatility.

Figures 4 and 5 show the ACFs for the absolute and squared returns for KSE100 and S&P500. The slow decaying ACFs suggest the presence of long term memory in volatility (as measured by absolute and squared return proxies)

Table 4 shows the empirical Hurst coefficients for the different markets. The numbers for the Indian market are taken from Mukherjee, Sen and Sarkar (2011).

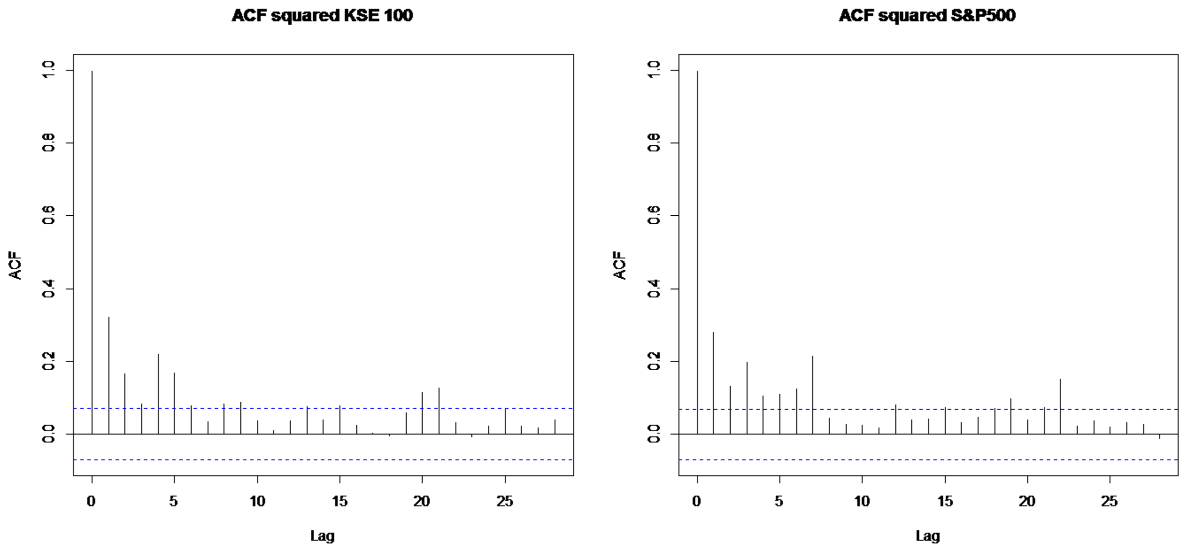


Figure 5: Autocorrelations of squared returns

Market	H(absolute returns)	H(squared returns)
KSE100	0.83	0.83
S&P500	0.57	0.56
India SENSEX	0.68	0.70

Table 4: Hurst coefficients for absolute and squared returns

The authors found that the Indian SENSEX market does not exhibit long term memory, however the absolute and squared returns do have long term memory. We would thus expect the SENSEX to lie between S&P500 and KSE100 in terms of market efficiency. The numbers in Table 4 seem to place the SENSEX in that range in terms of the long term memory in the volatility proxies.

5 Conclusion

The analysis shows the presence of a long term memory in the raw KSE100 returns. The KSE100 return data has a slow decaying ACF and the return is influenced by both recent and remote history. This is in contrast with the returns in the more efficient S&P500. This is an important distinction between emerging markets like the KSE100 and developed market such as the S&P500.

The long term dependence is even stronger in the absolute and squared returns implying that KSE100 volatility exhibits long term memory. Even the S&P500 return volatility exhibits long term dependence. This result is in line with the commonly observed volatility clustering in stock market data as shown in Figure 3.

The presence of long term memory suggests that better return forecasts can be made by building non-linear models for KSE100. Barkoulas, Baum and Travlos (2000) did this and compared forecasts of non-linear ARFIMA models with those of simple Random Walk and linear AR models for the Greek stock market and found that the ARFIMA models provided better forecasts. This establishes the usefulness of building non-linear models for emerging markets such as KSE100 which exhibit long term memory. Indeed long term dependence in the market means that better forecasting techniques can be used to generate higher returns and/or reduce volatility for funds invested in KSE100.

References

[1] A. W. Lo, Long-term memory in stock market prices, Econometrica, 1991

[2] I. Mukherjee, C. Sen and A. Sarkar, Long memory in stock returns: Insights from the Indian market, The Journal of App. Econ. and Fin, 2011

[3] B. Mandelbrot and J. Wallis, Noah , Joseph and operational hydrology, Water Resour. Res., 1968

[4] B. Mandelbrot and J. Wallis, Computer experiments with fractional gaussian noises Parts 1, 2, 3. Water Resour. Res, 1969

[5] B. Mandelbrot and M. Taqqu, Robust R/S analysis of long run serial correlation, Bull. Int. Stat. Inst., 1979

[6] H. Hurst, Long term storage capacity of reservoirs, Trans. Am. Soc. Civil Eng., 1951

[7] T. Barkoulas, C. F. Baum and N. Travlos, Long memory in the Greek market, Applied in Fin. Econ., 2000

[8] J. Tolvi, Long memory in a small stock market, Econ. Bul., 2003

[9] J. Haslett and A. E. Raftery, Space-time modelling with long-memory dependence: Assessing Ireland's wind power resource, Applied Statistics, 1989

[10] R. Weron, Estimating long range dependence: finite samplet properties and confidence intervals, Physica A, 2002

[11] J. P. Bauchaud and M. Potters, Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management (2nd Ed.), Cambridge University Press, 2004

6 Appendix- R Output

```
x -  KSE100 raw returns
x1 - S\&P500 raw returns
xx - KSE100 absolute returns
xx1- S\&P500 absolute returns
xxx - KSE100 squared returns
xxx1 - S\&P500 squared returns

ADF test of stationarity
> adf.test(x)

        Augmented Dickey-Fuller Test

data:  x
Dickey-Fuller = -8.2506, Lag order = 9, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(x) : p-value smaller than printed p-value
> adf.test(xx)

        Augmented Dickey-Fuller Test

data:  xx
Dickey-Fuller = -7.6019, Lag order = 9, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(xx) : p-value smaller than printed p-value
> adf.test(xxx)

        Augmented Dickey-Fuller Test

data:  xxx
Dickey-Fuller = -7.2661, Lag order = 9, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(xxx) : p-value smaller than printed p-value
> adf.test(x1)
```

```
        Augmented Dickey-Fuller Test

data:  x1
Dickey-Fuller = -9.1677, Lag order = 9, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(x1) : p-value smaller than printed p-value
> adf.test(xx1)

        Augmented Dickey-Fuller Test

data:  xx1
Dickey-Fuller = -6.446, Lag order = 9, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(xx1) : p-value smaller than printed p-value
> adf.test(xxx1)

        Augmented Dickey-Fuller Test

data:  xxx1
Dickey-Fuller = -7.1291, Lag order = 9, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(xxx1) : p-value smaller than printed p-value


Hurst exponent
> hurstexp(x)
Corrected R over S Hurst exponent:   0.6034453
Theoretical Hurst exponent:          0.5398235
Corrected empirical Hurst exponent:  0.5816325
Empirical Hurst exponent:            0.6182045

> hurstexp(x1)
Corrected R over S Hurst exponent:   0.623938
```

```
Theoretical Hurst exponent:          0.5250791
Corrected empirical Hurst exponent:  0.533264
Empirical Hurst exponent:            0.5004422
Warning message:
In matrix(x, n, m) :
  data length [792] is not a sub-multiple or multiple of the number of rows [61]

> hurstexp(xx)
Corrected R over S Hurst exponent:   0.7388683
Theoretical Hurst exponent:          0.5398235
Corrected empirical Hurst exponent:  0.7906822
Empirical Hurst exponent:            0.8314759

> hurstexp(xx1)
Corrected R over S Hurst exponent:   0.8234984
Theoretical Hurst exponent:          0.5250791
Corrected empirical Hurst exponent:  0.5934594
Empirical Hurst exponent:            0.5682212
Warning message:
In matrix(x, n, m) :
  data length [792] is not a sub-multiple or multiple of the number of rows [61]

> hurstexp(xxx)
Corrected R over S Hurst exponent:   0.7147583
Theoretical Hurst exponent:          0.5398235
Corrected empirical Hurst exponent:  0.7856645
Empirical Hurst exponent:            0.828376

> hurstexp(xxx1)
Corrected R over S Hurst exponent:   0.7177272
Theoretical Hurst exponent:          0.5250791
Corrected empirical Hurst exponent:  0.5890543
Empirical Hurst exponent:            0.5633297
Warning message:
In matrix(x, n, m) :
  data length [792] is not a sub-multiple or multiple of the number of rows [61]

Estimation of Fractional differencing parameter
> ts.test<-fracdiff(x,
+ ar = 2, ma = 0,
```

```
+ dtol = NULL, drange = c(0, 0.5),  M = 100, trace = 0)
> confint(ts.test)
      2.5 %   97.5 %
d 0.1351012 0.135161


> ts.test1<-fracdiff(x1,
+ ar = 0, ma = 0,
+ dtol = NULL, drange = c(0, 0.5),  M = 100, trace = 0)
> confint(ts.test1)
        2.5 %       97.5 %
d 9.78762e-06 8.187264e-05
```