Are European equity markets efficient?

New evidence from fractal analysis

Enrico Onali*, John Goddard

Bangor Business School, Bangor University, Bangor, Gwynedd, LL57 2 DG, UK

Abstract: Fractal analysis is carried out on the stock market indices of six developed European countries. Evidence is found of long-range autocorrelation in the log return series of the Mibtel, the index of the Italian stock market, in contravention of the Random Walk Hypothesis. Long-range autocorrelation implies that predictable patterns in the log returns do not dissipate quickly, and may therefore produce potential arbitrage opportunities. No evidence contrary to the Random Walk Hypothesis is found for the other five stock markets.

Keywords: market efficiency, fractal analysis, random walk hypothesis.

JEL classification: G1, F3.

* Corresponding author: e.onali@bangor.ac.uk

1 Introduction

Traditional capital markets theory relies on the assumption that log prices are martingales, implying the expected value of log price is the log price in the previous period, and log returns are uncorrelated. Therefore, log prices follow random walks and log returns are unpredictable. This theory is called Random Walk Hypothesis (RWH) and represents one of the variants of the broader Efficient Market Hypothesis (EMH).

Mainstream financial economics also relies on Normality of the log returns. For instance, non-Normality is not consistent with the Mean-Variance approach of portfolio theory (Markovitz, 1952, 1959) and the related Capital Asset Pricing Model. The pricing of financial derivatives according to the models developed by Black and Scholes (1972, 1973) and the risk-management approach Value At Risk (RiskMetrics, 1996) also rely on the assumption of Normality.

In order to allow for non-Normality and autocorrelation in log returns, Peters (1994) introduces the Fractal Markets Hypothesis (FMH). The FMH does not reject a priori the assumption that returns are log-Normal and uncorrelated, but allows for a broader range of returns behaviour. As a result, the FMH does not necessarily constitute an alternative to the EMH, but rather a generalisation.

The FMH derives its name from the theory of fractals (Mandelbrot, 1982). A fractal is an object whose parts resemble the whole. Peters (1994) argues that markets have a fractal nature: when markets are stable, returns calculated over different time scales (daily, weekly, monthly, and so on) exhibit the same statistical properties. For instance, if daily returns are leptokurtic, so are monthly returns; if daily returns exhibit positive autocorrelation, so do monthly returns. This feature is called self-affinity.

The distributional properties and autocovariance structure of a self-affine time series can be represented by the Hurst exponent. In a financial time series, the Hurst exponent can be estimated to test the validity of the RWH. For Normal log returns, if the Hurst exponent is smaller (larger) than 0.5, negative (positive) autocorrelation exists for log returns calculated for any time scale. If the Hurst exponent is 0.5 the process is random, and the RWH is valid. A Hurst exponent larger than 0.5 suggests positive long-range autocorrelation in the log returns, and therefore the autocorrelation function decays slowly for log returns calculated for any time scale. A Hurst exponent smaller than 0.5 suggests negative autocorrelation for log returns calculated for any time scale. For independent returns, if the Hurst exponent is larger than 0.5 the distribution of returns calculated over any time scale is leptokurtic and the population variance is infinite. If the Hurst exponent is 0.5 the process is Normal. Often, the literature neglects the likely influence of non-Normality on the calculation of the Hurst exponent, or it simply assumes non-Normality increases the Hurst exponent and may thus lead to $H > 0.5$ even in the absence of long-range autocorrelation. The Rescaled Range Analysis (RRA) has been employed widely to calculate the Hurst exponent.

This paper presents tests for the validity of the RWH for six stock markets located in developed European countries. We examine stock market indices comprising a large number of stocks, and therefore predictability due to thin trading is very unlikely. Rejection of the RWH due to long-range autocorrelation would therefore suggest that patterns in the stock returns of individual stocks do not dissipate quickly, and are likely to generate arbitrage opportunities.

We contribute to the extant literature in several ways. First, we provide robust evidence of significant long-range autocorrelation for the Mibtel, the index of the Italian stock market. Significant long-range autocorrelation is contrary to the RWH, because it implies that correlation patterns in the log returns do not dissipate quickly and could therefore be used for arbitrage opportunities.

Second, unlike most of the extant literature, we employ Monte Carlo simulations to construct critical values for the null hypothesis of uncorrelated and Normal log returns. This enables us to assess whether the estimated Hurst exponents for the log return series are significantly different from the joint hypothesis of no long-range autocorrelation and Normality, and therefore do not comply with the RWH.

Third, for each of the six log return series we compare the estimated Hurst exponent with the Hurst exponent estimated for two surrogate series: a shuffled surrogate, with the same probability distribution of the original series, but no autocorrelation; and a Normalised surrogate, with the same autocorrelation properties of the original series, but whose distribution is Normal. A comparison with both series is necessary if one aims to determine whether statistical departure from the RWH is genuine, that is, it is due to long-range autocorrelation, rather than non-Normality.

Fourth, previous literature uses the RRA on the residuals of an autoregressive model of log returns, to avoid that short-range autocorrelation produce spurious detection of long-range autocorrelation, or persistence (Peters, 1994; Opong *et al.*, 1999). However, we argue that this procedure may impair the self-affine structure of the time series, which implies that the autocovariance function be the same at all time scales. We use both procedures (that is, we run the RRA both after pre-filtering is carried out, and on the log return series that have not been pre-filtered), and compare the results. The RRA applied to the pre-filtered log returns fails to reject the RWH for any of the indices. However, when the RRA is run on the log return series, evidence contrary to the RWH is found for the Mibtel, for which the Hurst exponent is significantly larger than 0.5. The Hurst exponent for a shuffled surrogate of the Mibtel log returns is found not to differ significantly from 0.5, while the Hurst exponent for a Normalised surrogate of the Mibtel log returns is found to differ from 0.5 at the 10% level of significance. Therefore, contravention of the RWH for the Mibtel is due to long-range autocorrelation. An examination of the fractal properties of the six indices over small and large time scales only has confirmed this finding.

The results of the RRA for the other five indices do not provide evidence contrary to the RWH. However, there is evidence of short-term autocorrelation and non-Normality for small scales. We also find that pre-filtering the log return series prior to the RRA may *increase* the chances of a spurious detection of long-range autocorrelation, contrary to widely held belief.

The rest of the paper is as follows. Section 2 reviews the properties of self-affinity, long-range autocorrelation, and the generalised Central Limit Theorem. Section 3 describes the methodology and data. Section 4 reports the results. Section 5 concludes.

2 Self-affinity, long-range autocorrelation, and the generalised Central Limit Theorem

Self-similarity is the distinguishing feature of fractals: each part comprising a fractal resembles the whole. In a financial time series, the weaker concept of self-affinity is employed: self-affine time series have the same properties irrespective of the time scale (for instance, daily, weekly, or monthly returns). Section 2.1 describes the concept of self-affinity and the implications of self-affinity on the autocorrelation structure of the log returns measured over different time scales. Section 2.3 describes the properties of the Stable Paretian distribution and the generalised Central Limit Theorem.

2.1 Self-affinity and the autocorrelation structure of a time series

Self-affinity can be described mathematically as follows (Calvet and Fisher, 2002, p. 383):

$$\{X(nt_1),\ldots,X(nt_k)\} \overset{d}{=} \{n^H X(t_1),\ldots,n^H X(t_k)\} \quad (1)$$

Where $H > 0$ and $n, k, t_1 \ldots t_k \geq 0$, and $\overset{d}{=}$ denotes equality in distribution.

Self-affinity in a time series with Normal increments implies that the variance, γ_0, scales proportionately with the time scale over which increments are measured, n, according to a factor of proportionality governed by the Hurst exponent, H:

$$\gamma_0^{(n)} = n^{2H} \gamma_0^{(1)} \quad (2)$$

The brackets encircling n and 1 are to denote that they are not exponents, but denote the time scale over which the increments are calculated. Self-affinity also implies that the autocorrelation function for lag $k, \rho_k^{(n)} = \gamma_k^{(n)} / \gamma_o^{(n)}$, does not depend on n. That is, $\rho_k^{(n)} = \rho_k^{(1)}$ for $n \geq 1$ and $k \geq 1$. The first-order autocorrelation is $\rho_1^{(n)} = 2^{(2H-1)} - 1$ for $n \geq 1$, and there are equivalent expressions for $\rho_k^{(n)}$ for $k > 1$ (Onali and Goddard, 2009).

Fractional Brownian Motion (FBM) satisfies the property of self-affinity[1]. The FBM is a generalisation of the Brownian Motion, used to define a random walk process, and was introduced by Mandelbrot and van Ness (1968). This model has the same features as the Brownian Motion, but its increments can be dependent[2]. Correlation (at all time scales) in a self-affine series is represented by the parameter $0 < H < 1$. For $0 < H < 0.5$, the series is negatively correlated at all time scales, or antipersistent. For $0.5 < H < 1$, the series is positively correlated at all time scales, or persistent. Temporal self-affinity ensures that '[…] the distribution of returns over different sampling intervals are identical except for a single, non-random contraction' (Mandelbrot, Fisher and Calvet, 1997, p. 8). This has important consequences for researchers, as a time-inconsistent model renders the results reliable only for the selected time scale. For example, if weekly data are analysed, the results might not be valid for monthly data.

Techniques based on the (temporal) self-affinity property have been employed to assess the degree to which returns are long-range dependent. If a log return series is found to be either persistent ($H > 0.5$) or antipersistent ($H < 0.5$), the Efficient Market Hypothesis

[1] Long-range dependence, and thus a slow decay for the autocorrelation function, is found in Autoregressive Fractionally Integrated Moving Average processes (ARFIMA, Granger and Joyeux, 1980; Hosking, 1981). ARFIMA(0,d,0) processes, where d = H-0.5, are asymptotically self-affine (Fisher, Calvet and Mandelbrot, 1997).

[2] The increments of the FBM constitute the Fractional Normal Noise (FGN), a stationary continuous process.

(EMH), in the form of the Random Walk Hypothesis (RWH), is violated. An extensive literature examines whether the RWH correctly represents the behaviour of stock market returns using fractal analysis. If the RWH is rejected, the FMH may represent a better explanation for the behaviour of stock returns.

Empirical studies on the long-range autocorrelation properties of returns precede the FMH (Peters, 1994). Initially, evidence of long-range autocorrelation is found for the US stock market (Greene and Fielitz, 1977; Peters 1991). Subsequent refinements of the methodology used to measure long-range autocorrelation have produced results consistent with the Random Walk Hypothesis (Lo, 1991). Recently, Serletis and Rosenberg (2009) do not find evidence of persistence for four US stock market indices. International equity markets have also been examined (Cheung and Lai, 1995; Opong *et al.*, 1999; Howe *et al.*, 1999; McKenzie, 2001; Costa and Vasconcelos, 2003; Kim and Yoon, 2004; Zhuang *et al.*, 2004; Norouzzadeh and Jafari, 2005; Onali and Goddard, 2009), as well as commodities markets (Cheung and Lai, 1993; Alvarez-Ramirez *et al.*, 2002; Serletis and Rosenberg, 2007), and exchange rates (Mulligan, 2000; Kim and Yoon, 2004; Da Silva *et al.*, 2007). Recently, the connection between the Hurst exponent and market crashes has been investigated (Grech and Mazur, 2004; Grech and Pamula, 2008). Cajueiro and Tabak (2004, 2005) endeavour to rank the degree of efficiency of emerging markets on the basis of the Hurst exponent for either stock returns or volatility of returns. Their findings differ according to whether stock returns or volatility of returns is examined.

Many of the studies above neglect that, if the assumption of Normality of the increments is not satisfied, $H > 0.5$ does not necessarily imply long-range autocorrelation, because the estimation of H can be affected by the presence of non-Normality. Certain studies compare the Hurst exponent of a log return series with the Hurst exponent of a shuffled surrogate of the series (for instance, Peters, 1991). If the Hurst exponent of the shuffled surrogate is found to be lower than the Hurst exponent for the log return series, it is argued that there is long-range positive autocorrelation. However, in our view, this interpretation assumes that non-Normality can only result in an increase of the Hurst exponent. In the absence of a properly defined theory as to the impact of non-Normality on the Hurst exponent of a long-range autocorrelated series calculated using RRA, one cannot rely on this comparison only. A comparison of the estimated Hurst exponent with the Hurst exponent of a Normally distributed series with the same autocorrelation structure as the original series may be useful to improve robustness, although it may still not be decisive. In the next section, we discuss the properties of the Stable Paretian distribution, which is able to account for self-affinity of independent variables that are non-Normally distributed.

2.2 The Stable Paretian distribution and the generalised Central Limit Theorem

For independent processes, the Hurst exponent is the reciprocal of the characteristic exponent for Stable Paretian distributions[3], $\alpha = 1/H$. Unlike the Normal distribution, the Stable Paretian distribution allows for high 'peakedness', fat tails, and nonzero skewness. The log of the characteristic function of a Stable Paretian distribution, $\psi(t)$, can be expressed as follows (Elton et al., 1975, p. 232):

[3] Stable Paretian distribution is also referred to as Mandelbrot-Lévy, L-stable, Lévy-stable, and Pareto-Lévy (Mulligan, 2004).

$$\begin{cases} \ln[\psi(t)] = i\delta t - \upsilon |t|^{\alpha} w(t,d) \\ w(t,d) = \left[1 - i\zeta(t/|t|) \begin{pmatrix} \tan\dfrac{\pi\alpha}{2} & if\ \alpha \neq 1 \\ -2/\pi \ln|t| & if\ \alpha = 1 \end{pmatrix} \right] \end{cases} \quad (3)$$

Where $i = \sqrt{-1}$. There are four parameters that define the shape of a Stable Paretian distribution:

$-\infty < \delta < \infty$	is an estimate of the central tendency of the distribution. For $\alpha > 1$, δ is the mean of the distribution.
$\upsilon > 0$	is a measure of the degree of dispersion about the central tendency parameter δ, and replaces the standard deviation when the distribution is not Normal (for $\alpha < 2$). For $\delta = 0$ and $\upsilon = 1$ the distribution is said to be in its reduced form (Peters, 1994).
$-1 < \zeta < 1$	is a measure of skewness; if $\zeta = 0$ the distribution is symmetrical around δ, if $\zeta > 0$ the distribution is skewed to the right and if $\zeta < 0$ the distribution is skewed to the left. As $\alpha \to 2$, the distribution tends to become symmetric and ζ becomes irrelevant for the shape of the distribution (Jin and Frechette, 2002).
$\alpha \leq 2$[4]	is a measure of the kurtosis of the distribution, i.e. the degree to which the data are clustered around the mean, and of the 'fatness' of the tails. $\alpha = 2$ for a Normal distribution and $\alpha < 2$ for a leptokurtic distribution. If $\alpha < 2$ the variance is not defined, and if $\alpha < 1$ the mean and the variance are not defined.

Infinite variance is a feature that might appear counter-intuitive as it implies non-stationarity in returns (i.e. absolute value of returns can be infinitely large). However, according to Mandelbrot, Fisher and Calvet (1997), there is no a priori justification for rejecting the hypothesis of infinite variance in returns. Early studies (Mandelbrot, 1963, 1967; Fama, 1965; Roll, 1970) find that for log return series $1 < \alpha < 2$ (or $1 > H > 0.5$). More recently, evidence of existing second moments is found (Jansen and de Vries, 1991; Loretan and Phillips, 1994; Mantegna and Stanley, 1995; Pagan, 1996; Hiemstra and Jones, 1997; Annaert *et al.*, 2001).

Independent variables that are distributed according to (3) satisfy the property of 'stability under addition', which is a generalisation of the Central Limit Theorem (CLT). The sum of independent and Stable Paretian distributed variables with α and ζ is a Stable Paretian distributed variable with the same α and ζ (Fama, 1965). In other words, the Stable Paretian distribution is able to account for self-affinity of a time series with independent non-Normal increments.

[4] According to Hols and de Vries (1991), non-integrated ARCH models and Student's t processes can produce $\alpha > 2$. However, Jamdee and Los (2005) argue that if $\alpha > 2$ the probability distribution is not properly defined, because one or both tails of the distribution are outside the interval [0,1]. According to Annaert *et al.* (2001), $\alpha > 2$ suggests that the Stable Paretian distribution simply cannot account for the properties of the time series examined.

3 Methodology and data

In this paper, we use the Rescaled Range Analysis (RRA) to examine the fractal properties of six European indices. Section 3.1, describes the steps of the RRA for calculation of the Hurst exponent. Section 3.2 reports some descriptive statistics for the log returns of the six indices.

3.1 Methodology

The RRA is based on an examination of the average rescaled range of the cumulative deviation of a time series from its mean value within each of a number of subperiods. The rescaled range statistic, denoted $(R/S)_n$, is specific to the time scale, n, equivalent to the number of daily returns observations included in each subperiod. To obtain an indication of the scaling behaviour of $(R/S)_n$ as n varies, $(R/S)_n$ is constructed so as to vary proportionately with n^H. The RRA is performed using all observations of the log return series, for each n. Most previous studies discard the last few observations of the series when the total number of observations is not a multiple of n. Finally, in order to assess to what extent $H \neq 0.5$ indicates violation of the RWH, Monte Carlo simulations of random Normal innovations are performed to obtain critical values for the significance of the departure from the null hypothesis, $H = 0.5$.

Previous literature argues that pre-filtering using an AR(p) model is needed to avoid detecting spurious long-range autocorrelation (Peters, 1994; Opong *et al.*, 1999). However, we argue that eliminating short-term autocorrelation might eliminate genuine long-range autocorrelation. For the properties of self-affine processes, the autocovariance structure for a certain lag k should be the same regardless of n. Therefore, eliminating autocorrelation for lag k and $n = 1$ could result in the elimination of autocorrelation for values of n other than 1. For this reason, and for comparability with previous literature, we run the RRA on both filtered and unfiltered log returns.

Let N denote the total number of observations in the series z_t, where z_t is either the residuals of an AR(p) model on the log return series, $r_t = \alpha + \beta r_{t-p} + z_t$[5], or the log return series itself, r_t. Starting from the first observation, subdivide N into M contiguous subperiods labelled $m = 1,...,M$, each containing n observations such that $N - n < Mn \leq N$. For the observations within subperiod m, the mean and standard deviation of z_t are:

$$\mu_m = n^{-1} \sum_{t=(m-1)n+1}^{mn} z_t \tag{4}$$

$$S_m = \sqrt{n^{-1} \sum_{t=(m-1)n+1}^{mn} (z_t - \mu_m)^2} \tag{5}$$

for $m = 1,...,M$.

The cumulative deviations of z_t from μ_m within subperiod m are:

[5] The number of lags of the AR model, p, is equal to the lags for which the Partial Autocorrelation Function (PACF) is statistically significant at the 5%. Only lags up to the tenth lag are considered.

$$x_t = \sum_{s=(m-1)n+1}^{t} (z_s - \mu_m) \quad (6)$$

Where $t = (m-1)n+1,...,mn-1$, and $x_{nm} = 0$.

The range for subperiod m is defined as the difference between the maximum and minimum values of x_t for the observations within subperiod m:

$$R_m = \max_{t\in m}(x_t) - \min_{t\in m}(x_t) \quad (7)$$

Commonly, N is not a multiple of n. If $Mn < N$, then $L = N - nM$ observations at the end of the observation period are unused in the above procedure. To avoid discarding these L observations, the procedure is repeated starting from the L+1th observation (rather than from the first observation). A second set of M calculated values of R_m and S_m is obtained, where $m=M+1,...,2M$. If $Mn = N$, R_m and S_m for $m = 1,...,M$ are identical to R_m and S_m for $m = M+1,...,2M$.

The $(R/S)_n$ statistic is the mean of the rescaled range values for $m = 1,...,2M$

$$(R/S)_n = (2M)^{-1} \sum_{m=1}^{2M} (R_m / S_m) \quad (8)$$

Finally, the scaling behaviour of $(R/S)_n$ can be investigated by examining the power-law relationship $(R/S)_n \sim n^H$, where H is the Hurst Exponent. Having obtained values of $(R/S)_n$ for several time scales n, H can be estimated by running the ordinary least squares (OLS) regression

$$\ln[(R/S)_n] = \ln(c) + H \ln(n) + \eta_n \quad (9)$$

where η_n is a disturbance term.

The theoretical property of self-affinity applies to the distribution of returns calculated over all time scales. In practice, however, the scaling behaviour summarized by $r_t^{(n)} \overset{d}{=} r_t^{(1)} \times n^H$ may vary with the time scale. In the case of the RWH (zero temporal correlation), $(R/S)_n \sim \sqrt{n}$. The V_n statistic is defined:

$$V_n = (R/S)_n / \sqrt{n} \quad (10)$$

A plot of V_n against n provides a convenient visual indication of the variation in the scaling behaviour of $(R/S)_n$ with $\sqrt{n}$. Asymptotically, V_n is constant over all n under the RWH, that is under the assumption of Normality and no autocorrelation. If there is persistence in the returns measured over a specific range of values for the time scale n, asymptotically V_n is increasing in n over this range. Similarly if there is antipersistence in returns measured over a specific range of values for n, asymptotically V_n is decreasing over this range.

As noted earlier, the Hurst exponent is able to identify long-range autocorrelation in a Normal series as well as non-Normality in an independent series. However, when there is both dependence and non-Normality the interpretation of the results is ambiguous. For this reason, I run the RRA on shuffled and Normalised surrogates of the log return series. The shuffle eliminates temporal correlation while preserving the original probability distribution. The Normalising procedure preserves temporal correlation but ensures that the log return series is distributed Normally.

A shuffle surrogate is created using a method described by Norouzzadeh and Rahmani (2006, p331). Two integers t_1 and t_2 ($1 \le t_1, t_2 \le N$, where N is the number of observations) are drawn randomly from a uniform distribution, and the positions of r_{t1} and r_{t2} in $\{r_t\}$ are exchanged. The same procedure is repeated $20 \times N$ times, ensuring that the shuffled series is devoid of temporal correlation.

A Normalising transformation is performed as follows. The original series $\{r_t\}$, for $1 \le t \le N$, is sorted in ascending order. A variable $s = 1,\ldots,N$ is generated where $1 \le s \le N$. The original order is then replaced by sorting $\{r_t\}$, $\{t\}$ and $\{s\}$ according to t. These steps ensure that, while the ordering of $\{r_t\}$ and $\{s\}$ is based on the values of t, $\{s\}$ can be used as an indicator for ranking $\{r_t\}$: the smallest value for $\{r_t\}$ is for $s = 1$, the largest value for $\{r_t\}$ is for $s = N$, and for $s_1 < s_2$, $r_{s1} < r_{s2}$. In the next step, two variables are generated: $p \sim N(0,1)$ with N observations; and $q = \bar{r} + \sigma_r \times p$, where $\bar{r} = \frac{1}{N}\sum_{t=1}^{N} r_t$ and $\sigma_r = \sqrt{\sum_{t=1}^{N} \frac{(r_t - \bar{r})^2}{N-1}}$. Finally, $\{r_t\}$ is replaced by corresponding values of q_s. For instance, the tenth observation for r_t sorted according to t ($r_{t=10}$) is replaced by the tenth observation of q sorted according to s ($q_{s=10}$). The resulting series will have the same autocorrelation structure as $\{r_t\}$, but the probability distribution will be Normal, because the variable q is Normal, with the same mean and standard deviation as $\{r_t\}$.

Several previous studies have noted that the application of the RRA in finite samples produces an upward-biased estimator of H (Feller, 1951; Anis and Lloyd, 1976; Peters, 1994; Qian and Rasheed, 2004; and Norouzzadeh and Jafari, 2005). Furthermore, in finite samples V_n is increasing in n under the RWH. Following Onali and Goddard (2009), Monte Carlo simulations are used to examine whether the estimated Hurst exponents for the log return series of the indices differ significantly from the value that is expected under the RWH[6], and similarly whether the realised V_n statistics for each n differ significantly from those expected under the RWH. This procedure involves generating 5,000 simulated returns series containing random Normal innovations. The RRA is repeated for each of the simulated series, and the sampling distributions are obtained for the estimated H, and for the realized V_n for each n.

3.2 Data and descriptive statistics

Daily log returns have been calculated, based on closing daily prices provided by Thomson Analytics, for the following stock market indices: Mibtel (Milan), CDAX (Frankfurt), FTSE 350 (London), Amsterdam S.E. all-share (henceforth, ASE), Madrid S.E. all-share (MSE), and Swiss Market Index (SWX, Zurich).

[6] For convenience, I consider the RWH in the form that requires log returns to be Normally distributed. As long as log returns are stochastic (instead of deterministic), the RWH is valid. Therefore, rejection of the Normality assumption does not necessarily imply rejection of the RWH. However, non-Normal log returns weaken asset pricing models and other important approaches of mainstream capital markets theory.

Figure 1 exhibits the daily log returns, r_t, of the six indices over time. Volatility clustering (heteroskedasticity) is apparent, as large returns (of either sign) tend to be concentrated within certain periods (high-volatility periods). This condition is believed to cause leptokurtosis in log returns (Chen *et al*, 2001). The assumption $r_t \sim N(\mu,\sigma^2)$ is substituted by $r_t \sim D(\mu,h_t)$, where D is some non-Normal probability distribution, h_t is the conditional variance, $r_t = u_t h_t^{1/2}$, and u_t are IID.

[insert figure 1 here]

Table 1 reports descriptive statistics for the log return series on the indices for four time scales: daily, weekly, monthly, and quarterly[7]. The total observations for the four series are 2,608, 522, 120, and 40, respectively. The descriptive statistics shown in Table 1 relate to the first four central moments of the distribution of the log returns: mean, standard deviation, skewness, and kurtosis.

[insert table 1 here]

The skewness is negative in most cases. As the time scale increases, the probability density functions of the log return series do not appear to become more symmetric. The departure from Normality for quarterly data is contrary to the alleged phenomenon of 'aggregational Normality' (Cont, 2001). The presence of negative skewness for long time scales suggests that negative daily returns tend to cluster, making large cumulative losses more likely than large cumulative gains. Negative skewness implies large losses are more likely than large gains. *Ceteris paribus*, negative skewness should encourage investors to require higher expected returns than if the skewness is zero.

The probability density functions of the log return series seem to become less leptokurtic as the time scale increases (although for the FTSE 350 the quarterly log returns are more leptokurtic than the monthly log returns). Leptokurtosis implies extreme returns of either sign are more likely than in a Normal distribution. *Ceteris paribus*, investors should require higher expected returns when the distribution is leptokurtic than in the case of a mesokurtic distribution (where the kurtosis is three).

Given a random Normal variable u, with mean μ and standard deviation σ, the variable $z = \frac{u-\mu}{\sigma}$ is a Normal variable with mean 0 and standard deviation 1. A plot of z against u, where the values of both variables are sorted in ascending order, is called Normal probability graph. Because z is a linear transformation of u, the graph is a straight line. However, for variables that are not Normal, the Normal probability graph is a curve that may assume a variety of shapes (Fama, 1965).

Figure 2 shows the Normal probability graphs for the log returns of the six indices at the daily and monthly frequencies. The comparison between the quantiles of the probability density functions of the log returns and the quantiles of a Normal distribution shows for which quantiles there is evidence of departure from Normality, and the degree of such departure. As said before, if the distribution of log returns were Normal, the plot should be a straight line. Leptokurtosis should cause the extreme parts of the plots to diverge from a

---

[7] Given n the number of trading days comprising each time scale, for the daily returns, $n = 1$. For the other time scales, there is no exact correspondence between n and the actual time scale used in the analysis, as the number of trading days may vary according to the week, month, or quarter.

straight line representing the Normal case: for lower quantiles the lower half of the plot should bend downwards, since large negative returns are more likely than in a Normal distribution; for upper quantiles the upper part of the plot should bend upwards, since large positive returns are more likely than in a Normal distribution.

[insert figure 2 here]

The plots show evidence of leptokurtosis in the daily log returns. The plots assume a typical inverted-***S*** shape. Leptokurtosis decreases as the time scale becomes longer. This is reflected in the plots for the monthly returns lying closer to the straight line representing the case for which a Normal variable is plotted against the standard Normal distribution. However, it can be noticed that for the lower quantiles outliers still exist. Outliers for the lower quantiles cause the plots to diverge from the straight line in the extreme part of the lower half of the plots. The lower quantiles correspond to large negative returns. Consistent with Table 1, there is asymmetry even for long time scales. Negative skewness for long time scales may be due to clustering of negative daily returns.

Provided the population variance is finite, the Central Limit Theorem (CLT) ensures that the sum of independent random variables converges to a Normal distribution as the number of variables increases, regardless of the distribution of each individual random variable. If the variance of the population is infinite, the standard CLT does not hold. In this case, the sum of IID variables converges to a Stable Paretian distribution with a characteristic exponent equal to that of the individual random variables. In order to investigate whether my data comply with the CLT once temporal correlation has been eliminated, the procedures described above are repeated on shuffled surrogates of the log return series, in which the ordering of the observations is randomised.

Table 2 reports descriptive statistics for the shuffled surrogate of the log return series for four time scales: daily, weekly, monthly, and quarterly. There is little evidence of leptokurtosis for long time scales. In comparison with the results reported in Table 1, the degree of skewness is also less for long time scales. The reduction in skewness for long time scales in the shuffled series supports the hypothesis that negative skewness for long time scales may be due to clustering of negative daily returns. These results support the validity of the CLT.

[insert table 2 here]

Similar to Figure 2, Figure 3 shows the Normal probability graphs for the shuffled log returns of the six indices at the daily and monthly frequencies. The plots for the daily frequency are the same as those exhibited in Figure 2, as the shuffle destroys temporal correlation leaving the probability density function unaltered. However, the shuffle may modify the probability density function of the monthly log returns. Before the shuffle, the clustering of positive (or negative) daily returns in a month may result in larger positive (or negative) monthly returns. On the other hand, periods of negative correlation in the daily returns tend to produce smaller monthly returns of either sign. The shuffle eliminates the effect of clustering, and therefore should also reduce skewness and leptokurtosis for large time scales.

[insert figure 3 here]

The plots for monthly returns in Figure 3 show a closer fit to the straight line than those exhibited in Figure 2. Consistent with Table 2, eliminating the temporal correlation through the shuffle procedure results in a significant reduction of leptokurtosis for log returns calculated for long time scales. The degree of asymmetry is also reduced.

4 Results

The Rescaled Range Analysis (RRA) is carried out over 40 values of the time scale parameter n, defined by increasing $\ln(n)$ in steps of 0.1 from a minimum of $\ln(n) = 1.6$ ($n = 5$) to a maximum of $\ln(n) = 5.7$ ($n = 299$). For each value of $\ln(n)$, n is obtained by rounding $e^{[\ln(n)]}$ to the nearest integer. Two of the 42 values of $\ln(n)$ in the range (1.6, 5.7) are discarded, because the integer values of $e^{[\ln(n)]}$ and $e^{[\ln(n-0.1)]}$ are identical. The results of the RRA for each index are compared with critical values obtained via Monte Carlo simulations. Section 3.1 reports the results for the pre-filtered log returns. Section 3.2 reports the results for the unfiltered log returns.

4.1 Results for pre-filtered returns

Tables 3, 4, and 5 report the estimation results for the RRA when the original log returns are pre-filtered using an AR(p) model before calculating the R/S statistic. The first column reports the name of the index. The second column reports the estimated Hurst exponent for the log return series of each index, H_i. The third column reports the estimated Hurst exponent for the shuffled log return series of each index, H_s. The fourth column reports the estimated Hurst exponent for the Normalised log return series of each index, H_N. The sixth column reports the average Hurst exponent estimated with respect to the 5,000 Monte Carlo simulations, μ_H, and the Hurst exponents associated with the quantiles: 0.005, 0.025, 0.050, 0.950, 0.975, and 0.999.

Table 3 reports the estimation results for the time scales $5 \leq n \leq 299$. For the log return series of all six indices, the Hurst exponent estimated over the time scales $5 \leq n \leq 299$ is higher than the average H obtained from the 5,000 Monte Carlo simulations ($\mu_H = 0.572$). However, none of such values is higher than the critical value associated with the 10% level of significance. Accordingly, a two-tail test fails to reject the null hypothesis of $H = 0.5$ (log returns are temporally uncorrelated at all time scales) in favour of the alternative $H \neq 0.5$ (long-range autocorrelation at all time scales) for any of the 6 indices. The estimated Hurst exponents for the shuffled log return series of all six indices are considerably lower than the estimated Hurst exponents for the original log return series. The estimated Hurst exponents for the Normalised log return series of all six indices are lower (but in some case the difference is very slight) than the estimated Hurst exponents for the original log return series. For all six indices, $H_s < H_N < H_i$. Therefore, temporal correlation seems to affect the estimated Hurst exponent for the log return series of the indices to a greater extent than non-Normality does. However, none of such values is lower than the critical value associated with the 10% level of significance.

[insert table 3 here]

To investigate whether there is any statistical evidence of departure from the RWH (due to either persistence or antipersistence) when returns are measured over small or large time scales only, the estimation of equation (9), $\ln[(R/S)_n] = \ln(c) + H \ln(n) + \eta_n$, and the Monte Carlo simulations are repeated, by fitting a spline function to allow for a change in the estimated Hurst exponent at the midpoint of the set of 40 values for ln(n) used in the estimation.

Table 4 reports the results for the spline regressions for the small time scales $5 \leq n \leq 40$. The format of the table is similar to that of Table 3. H_i is smaller (for the CDAX and SWX), larger (for the FTSE350 and the ASE), or equal (for the Mibtel and MSE) to the average Hurst exponent estimated via Monte Carlo simulations (μ_H = 0.606). However, as before, H_i is neither smaller nor larger than the critical values obtained via Monte Carlo simulations for any index. Unlike what observed for $5 \leq n \leq 299$, $H_s < H_i$ does not apply to all six indices, but only in four cases, and $H_s < H_N$ only in two cases. On the contrary, $H_N < H_i$ still applies to all six indices. Therefore, non-Normality influences the estimated Hurst exponent for $5 \leq n \leq 40$ to a greater extent than for $5 \leq n \leq 299$. This is consistent with a more leptokurtic distribution for small n than for large n, as shown in section 3.2.

[insert table 4 here]

Table 5 reports the results for the spline regressions for the large time scales $40 \leq n \leq 299$. The format of the table is similar to that of Table 3 and Table 4. H_i is neither smaller nor larger than the relevant critical values obtained via Monte Carlo simulations. However, H_i is larger than the average Hurst exponent estimated via Monte Carlo simulations (μ_H = 0.540) for all six indices. Therefore some pattern, due to either long-range autocorrelation or non-Normality, might be present. For all six indices $H_s < H_i$. Similar to what obtained for $5 \leq n \leq 299$, $H_N > H_i$ for all six indices. Temporal correlation appears to affect the time series properties of the log return series to a greater extent than non-Normality does. For the Mibtel and MSE $H_N > H_i$, indicating that, once non-Normality is taken into account, temporal correlation appears even stronger. Thus, it appears that non-Normality and long-range temporal correlation in a series might combine together in various ways, which may hinder the interpretation of the Hurst exponent.

[insert table 5 here]

4.2 Results for unfiltered returns

Tables 6, 7, and 8 report the estimation results for the RRA when the original log returns are not pre-filtered using an AR(p) model before calculating the R/S statistic.

Table 6 reports the estimation results for the time scales $5 \leq n \leq 299$. There are several differences with respect to the results reported in Table 3. H_i for the FTSE 350 is lower than μ_H and than H_i when pre-filtering of the log returns is carried out prior to the RRA. A higher H_i when pre-filtering is carried out is due to significantly negative autocorrelation for the lags 3, 5, 6, and 10 in the latter. Therefore, pre-filtering might *increase* the probability of a spurious detection of long memory, contrary to widely held beliefs (for instance, Lo, 1991). H_i for the Mibtel is significantly larger than 0.5 (at the 5% level of significance). This result is in contravention of the RWH, and implies that using pre-filtered log returns may prevent the detection of long-range autocorrelation. Moreover, H_s is not significantly different from 0.5, while H_N is very close the H_i, suggesting that long-range autocorrelation is correctly detected. Further discrepancies between the results reported in Table 6 and those reported in Table 3 are as follows:

- When pre-filtering is carried out on the log returns, $H_s < H_N < H_i$ for all six indices – when pre-filtering is not carried out, $H_s < H_N < H_i$ only for the Mibtel, CDAX, and SWX

- For the FTSE 350 $H_s > H_i > H_N$, suggesting that non-Normality has a large impact on the estimation of the Hurst esponent

- For the ASE $H_N < H_s < H_i$.

[insert table 6 here]

Table 7 reports the results for the spline regressions for the small time scales $5 \leq n \leq 40$. For all indices except the Mibtel, H_i for the original log returns is lower than H_i for the pre-filtered log returns, suggesting that pre-filtering may *increase* the chances of a spurious detection of long-range autocorrelation. For the Mibtel, H_i and H_N for the original log returns are higher than H_i and H_N for the pre-filtered log returns.

[insert table 7 here]

Table 8 reports the results for the spline regressions for the large time scales $40 \leq n \leq 299$. Consistent with what found for $5 \leq n \leq 299$, some departure from the RWH is found for the Mibtel, for which H_i and H_N are different from 0.5 at the 10% level of significance. Rejection of the null hypothesis for large time scales, even if for a relatively low level of significance, suggests long-range autocorrelation is not spurious. For the FTSE 350, similar to what found for $5 \leq n \leq 299$, H_i estimated without pre-filtering the log returns is lower than when pre-filtering is carried out. Therefore, as before, pre-filtering might cause detection of long memory when it does not exist, rather than avoiding detection of long memory when returns are not long-range autocorrelated.

[insert table 8 here]

The behaviour of the V-statistic, $V_n = (R/S)_n/n^{0.5}$, indicates how patterns in temporal correlation vary over different time scales.

Figure 4 shows the plot of the realized values of V_i against ln(n) for the log returns of the Mibtel, together with the plots of the shuffled and Normalised log return series, V_s and V_N respectively, and the plots of the mean values of the V-statistic for each n obtained from the 5,000 replications of the Monte Carlo simulation (V_m), and the 0.975 and 0.025 quantiles of the sampling distributions of the V-statistic for each n obtained from the Monte Carlo simulation ($V_{.975}$ and $V_{.025}$, respectively). The realised values of V_i are similar to those of V_m over low values of n, but the two plots tend to diverge for large n. Evidence contrary to the RWH is found for values for which the plot of V_i breaks through the upper critical value associated with the 5% level of significance: $49 \leq n \leq 67$ ($3.9 \leq \ln(n) \leq 4.2$), $99 \leq n \leq 110$ ($4.6 \leq \ln(n) \leq 4.7$), $n = 134$ ($\ln(n) = 4.9$), $n = 164$ ($\ln(n) = 5.1$), and $245 \leq n \leq 299$ ($5.5 \leq \ln(n) \leq 5.7$). Similar comments can be made with respect to the plot of V_N for the scales: $n = 67$ ($\ln(n) = 4.2$), $n = 99$ ($\ln(n) = 4.6$), $n = 164$ ($\ln(n) = 5.1$), and $245 \leq n \leq 299$ ($5.5 \leq \ln(n) \leq 5.7$). The plot of V_s tends to bend downward for large values of n. Therefore, the V-statistic confirms that long-range autocorrelation may exist for the Mibtel.

[insert figure 4 here]

Figure 5 shows the plots of V_i, V_s, and V_N for the CDAX, along with V_m, $V_{.975}$ and $V_{.025}$. Evidence contrary to the RWH is found for $n = 270$ ($\ln(n) = 5.6$), for which $V_i > V_{.975}$. The plot of V_s breaks through $V_{.025}$ for approximately $49 \leq n \leq 60$, or $3.9 \leq \ln(n) \leq 4.1$.

[insert figure 5 here]

Figure 6 shows the plots of V_i, V_s, and V_N for the FTSE 350, along with V_m, $V_{.975}$ and $V_{.025}$. The plots show that V_i and V_s, never fall outside the range of values defined in ($V_{.975}$, $V_{.025}$). Therefore, the FTSE 350 behaves consistently with the RWH for all time scales ($5 \leq n \leq 299$).

[insert figure 6 here]

Figure 7 shows the plots of V_i, V_s, and V_N for the ASE, along with V_m, $V_{.975}$ and $V_{.025}$. V_i lies within the boundaries defined by $V_{.975}$ and $V_{.025}$ for all time scales.

[insert figure 7 here]

Figure 8 shows the plots of V_i, V_s, and V_N for the MSE, along with V_m, $V_{.975}$ and $V_{.025}$. The plot of V_i suggests departure from the RWH for $n = 15$ ($\ln(n) = 2.7$), for which $V_i > V_{.975}$. For larger values of n, the results are consistent with the RWH. Therefore, some departure from the RWH appears due to short-range autocorrelation.

[insert figure 8 here]

Finally, Figure 9 shows the plots of V_i, V_s, and V_N for the SWX, along with V_m, $V_{.975}$ and $V_{.025}$. The plots of V_i and V_N tend not to diverge substantially from V_m, although $V_N > V_{.975}$ for $n = 5, 10$ ($\ln(n) = 1.6, 2.3$). While for very large scales V_i and V_N tend to be larger than V_m, the plots do not break through the boundary represented by the plot of $V_{.975}$. Departure from the RWH is therefore not detected according to conventional standard criteria.

[insert figure 9 here]

5 Conclusions

This paper presents an empirical analysis of the unifractal properties of the log returns of six European stock market indices. Preliminary tests have shown non-Normality in the probability distribution function of the log returns calculated for various time scales. However, once temporal correlation in the daily log returns is eliminated through a shuffle procedure, the distribution of the log returns calculated for large time scales tends towards Normality.

We have employed the Rescaled Range Analysis (RRA) to investigate the long-range properties of the log returns of the indices. The RRA enables the calculation of the Hurst exponent. If the estimated Hurst exponents for any of the six stock market indices were found to be significantly larger than 0.5, there would be long-range autocorrelation, or non-Normality, or both. In order to correctly identify long-range autocorrelation and/or non-Normality, two surrogate series for each index have been created: a Normal series with the same autocorrelation structure as the original log return series; and a series with no autocorrelation but the same probability distribution as the original log return series. For the properties of the Stable Paretian distribution, if $H > 0.5$ and log returns are independent the population variance is infinite. If $H > 0.5$ and log returns are Normally distributed, there is long-range positive autocorrelation, and returns are predictable.

We run the RRA on the log return series of the six indices and find evidence contrary to the RWH is for the Mibtel, for which the Hurst exponent is significantly larger than 0.5. The Hurst exponent for a shuffled surrogate of the Mibtel log returns is found not to differ significantly from 0.5, while the Hurst exponent for a Normalised surrogate of the Mibtel log returns is found to differ from 0.5 at the 10% level of significance. Therefore, contravention of the RWH for the Mibtel is due to long-range autocorrelation. Long-range autocorrelation in the log returns implies possible arbitrage opportunities that exploit predictability in the log returns. The results for the other five indices do not provide evidence contrary to the RWH.

References

Alvarez-Ramirez J., Cisneros M., Ibarra-Valdez C., Soriano A., 2002. Multifractal Hurst analysis of crude oil prices. Physica A, 313, 651-670.

Anis A. A., Lloyd E. H., 1976. The expected value of the adjusted rescaled Hurst range of independent normal summands. Biometrika, 63, 111-116.

Annaert J., De Ceuster M. J. K., Hodgson A., 2001. Moment condition failure Australian evidence. University of Antwerp, Faculty of Applied Economics, Working Paper No 10.

Black F., Scholes M., 1972. The valuation of option contracts and a test of market efficiency. Journal of Finance, 27, 399–418.

Black F., Scholes M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy, 81, 418–637.

Cajueiro D. O., Tabak B. M., 2004. Ranking efficiency for emerging markets. Chaos, Solitons and Fractals 22, 349-352.

Cajueiro D. O., Tabak B. M., 2005. Ranking efficiency for emerging equity markets II. Chaos, Solitons and Fractals 23, 671-675.

Calvet L., Fisher A., 2002. Multifractality in asset returns: theory and evidence. The Review of Economics and Statistics 84 (3), 381-406.

Cheung Y. W., Lai K. S., 1993. Do gold market returns have long memory? The Financial Review 28 (2), 181-202.

Cheung Y. W., Lai K. S., 1995. A search for long memory in international stock market returns. Journal of International Money and Finance 14 (4), 597-615.

Cont R., 2001. Empirical properties of asset returns: stylized facts and statistical issues. Quantitative Finance, 1, 223-236.

Costa R. L., Vasconcelos G. L., 2003. Long-range correlations and nonstationarity in the Brazilian stock market. Physica A 329, 231-248.

Da Silva S., Matsushita R., Gleria I., Figueiredo A., 2007. Hurst exponents, power laws, and efficiency in the Brazilian foreign exchange market. Economics Bulletin, 7 (1), 1-11.

Elton E. J., Gruber M. J., Kleindorfer P. R., 1975. A closer look at the implications of the Stable Paretian Hypotheses. The Review of Economics and Statistics, 57 (2), 231-235.

Fama E. F., 1965. The behaviour of stock-market prices. Journal of Business, 38 (1), 34-105.

Feller W., 1951. The asymptotic distributions of the range of sums of independent random variables. Annals of mathematical statistics, 22, 427-432.

Fisher A., Calvet L., Mandelbrot B., 1997. Multifractality of Deutschemark/US Dollar exchange rates. Yale University, Cowles Foundation Discussion Paper #1166.

Granger C. W. J., Joyeux R., 1980. An introduction to long memory time series models and fractional differencing. Journal of Time Series Analysis, 1 (1), 15-29.

Grech D., Mazur Z., 2004. Can one make any crash prediction in finance using the local Hurst exponent idea? Physica A, 336, 133-145.

Grech D., Pamula, G., 2008. The local Hurst exponent of the financial time series in the vicinity of crashes on the Polish stock exchange market. Physica A, 387, 4299-4308.

Greene M. T., Fielitz B. D., 1977. Long-term dependence in common stock returns. Journal of Financial Economics 4, 339-349.

Hiemstra C., Jones J. D., 1997. Another look at long memory in common stock returns. Journal of Empirical Finance, 4, 373-401.

Hols M. C. A. B., de Vries C. G., 1991. The limiting distribution of extremal exchange rate returns. Journal of Applied Econometrics, 6 (3), 287-302.

Hosking J. R. M., 1981. Fractional differencing. Biometrika, 68 (1), 165-176.

Howe J. S., Martin D. W., Wood B. G. Jr, 1999. Much ado about nothing – Long-term memory in Pacific Rim equity markets. International Review of Financial Analysis, 8 (2), 139-151.

Jansen D. W., de Vries C. G., 1991. On the frequency of large stock returns: putting booms and busts into perspective. The Review of Economics and Statistics, 73, 18-24.

Jin H. J., Frechette D. L., 2002. Fractal geometry in agricultural cash price dynamics. Annual Meeting of the American Agricultural Economics Association in Long Beach.

Kim K., Yoon S., 2004. Multifractal features of financial markets. Physica A, 344, 272-278.

Lo A. W., 1991. Long-term memory in stock market prices. Econometrica, 59 (5), 1279-1313.

Loretan M., Phillips P. C. B., 1994. Testing the covariance Stationarity of heavy-tailed time series: an overview of the theory with applications to several financial data series. Journal of Empirical Finance, 1, 211-248.

Mandelbrot B., 1963. The variation of certain speculative prices. The Journal of Business, 36 (4), 394-413.

Mandelbrot B., 1967. The variation of some other speculative prices. The Journal of Business, 40 (4), 393-413.

Mandelbrot B., 1982. The fractal geometry of nature. New York, W. H. Freeman.

Mandelbrot B., Fisher A., Calvet L., 1997. A Multifractal Model of Asset Returns. Yale University, Cowles Foundation Discussion Paper #1164.

Mandelbrot B., van Ness J. W., 1968. Fractional Brownian Motion, Fractional Noises and application. SIAM Review, 10, 422-437.

Mantegna R. S., Stanley H. E., 1995. Scaling behaviour in the dynamics of an economic index. Nature, 376, 46-49.

Markovitz, H., 1952. Portfolio selection. Journal of Finance, 7 (1), 77-99.

Markovitz, H., 1959. Portfolio selection: efficiency diversification of investments. Cowles Foundation, Monograph n.16, New York, John Wiley & Sons, Inc.

McKenzie, M. D., 2001. Non-periodic Australian stock market cycles: evidence from rescaled range analysis. The Economic Record 77 (239), 393-406.

Mulligan R. F., 2000. A fractal analysis of foreign exchange markets. International Advances in Economic Research, 6 (1), 33-49.

Mulligan R. F., 2004. Fractal analysis of highly volatile markets: an application to technology equities. The Quarterly Review of Economics and Finance, 44, 155-179.

Norouzzadeh P., Jafari G. R., 2005. Application of multifractal measures to Teheran price index. Physica A, 356, 609-627.

Norouzzadeh P., Rahmani B., 2006. A multifractal detrended fluctuation description of Iralian rial-US dollar exchange rate. Phisica A, 367, 328-336.

Onali E., Goddard J., 2009. Unifractality and multifractality in the Italian stock market. International Review of Financial Analysis, 18 (4), 154-163.

Opong K. K., Mulholland G., Fox A.F., Farahmand K., 1999. The behaviour of some UK equity indices: an application of Hurst and BDS tests. Journal of Empirical Finance, 6, 267-282.

Pagan A. R., 1996. The econometrics of financial markets. Journal of Empirical Finance, 3, 15-102.

Peters E. E., 1991. Chaos and order in the capital markets. New York, John Wiley & Sons.

Peters E.E., 1994. Fractal Markets analysis: applying chaos theory to investments and economics, John Wiley & Sons, New York.

Qian B., Rasheed K., 2004. Hurst exponent and financial market predictability. Financial Engineering and Applications.

RiskMetrics, 1996. Technical Document. Fourth ed. J.P. Morgan, London.

Roll R., 1970. The behavior of interest rates: the application of the efficient market model to US treasury bills. Basic Books, New York.

Serletis A., Rosenberg A. A., 2007. The Hurst exponent in energy futures prices. Physica A, 380, 325-332.

Serletis A., Rosenberg A. A., 2009. Mean reversion in the US stock market. Chaos, solitons and fractals, 40, 2007-2015.

Zhuang X.-t., Huang X.-y., Sha Y.-l., 2004. Research on the fractal structure in the Chinese stock market. Physica A, 333, 293-305.

Figure 1: Fluctuations of r_t over the period 31/08/1995-30/08/2005.

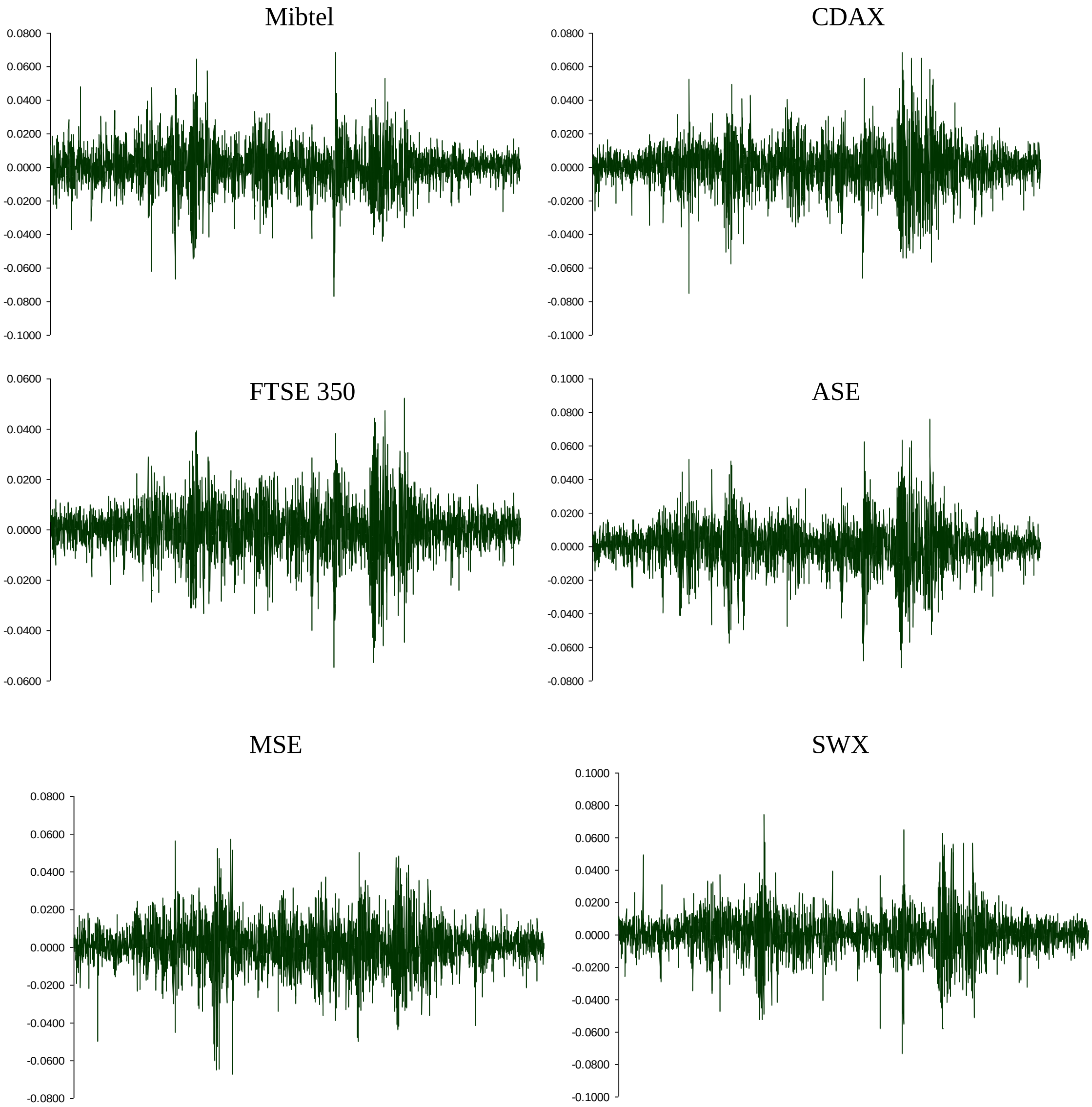


Figure 2: Normal probability graphs for the daily and monthly log returns.

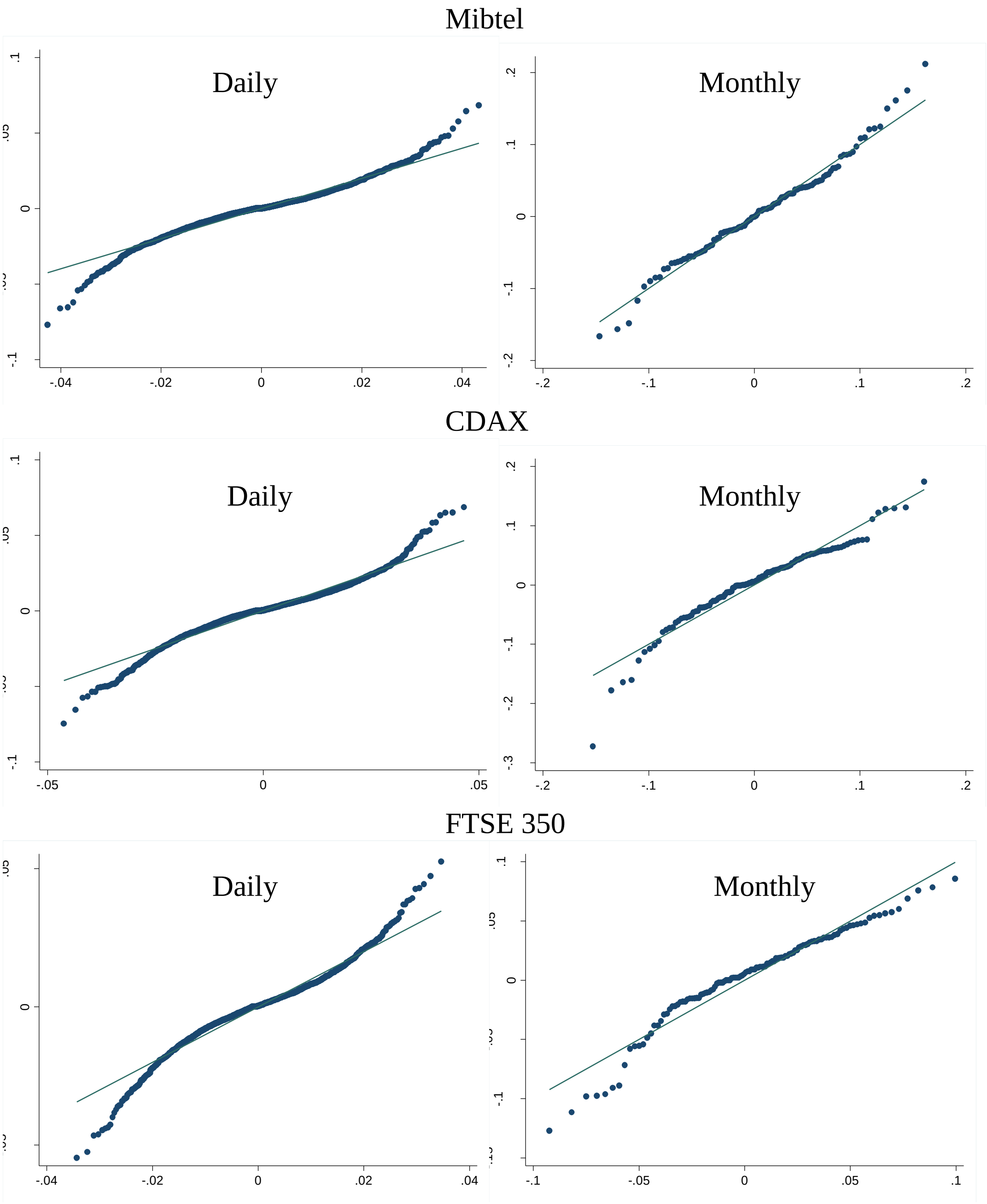


Figure 2 continued

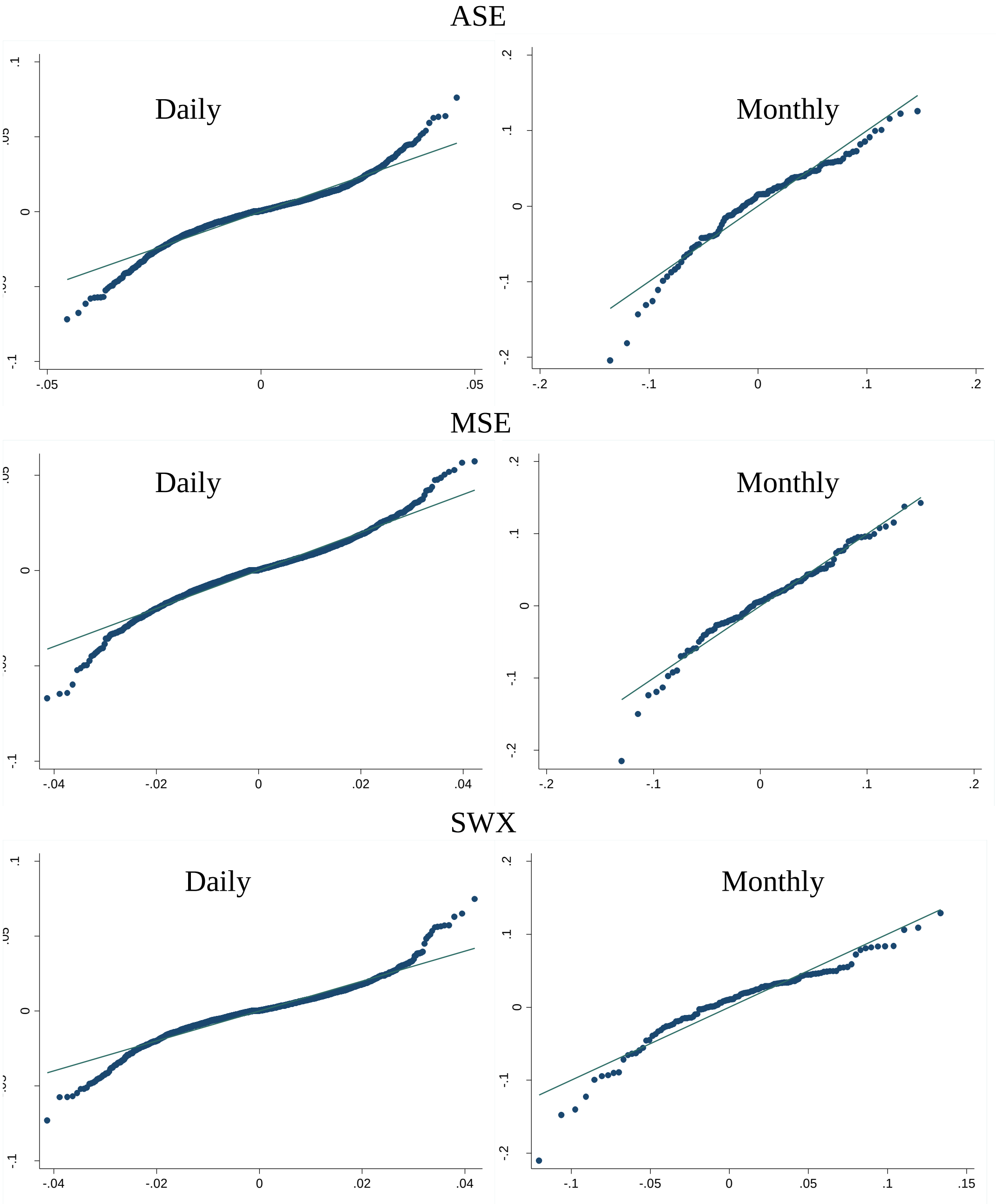


Figure 3: Normal probability plots for the daily and monthly shuffled log returns.

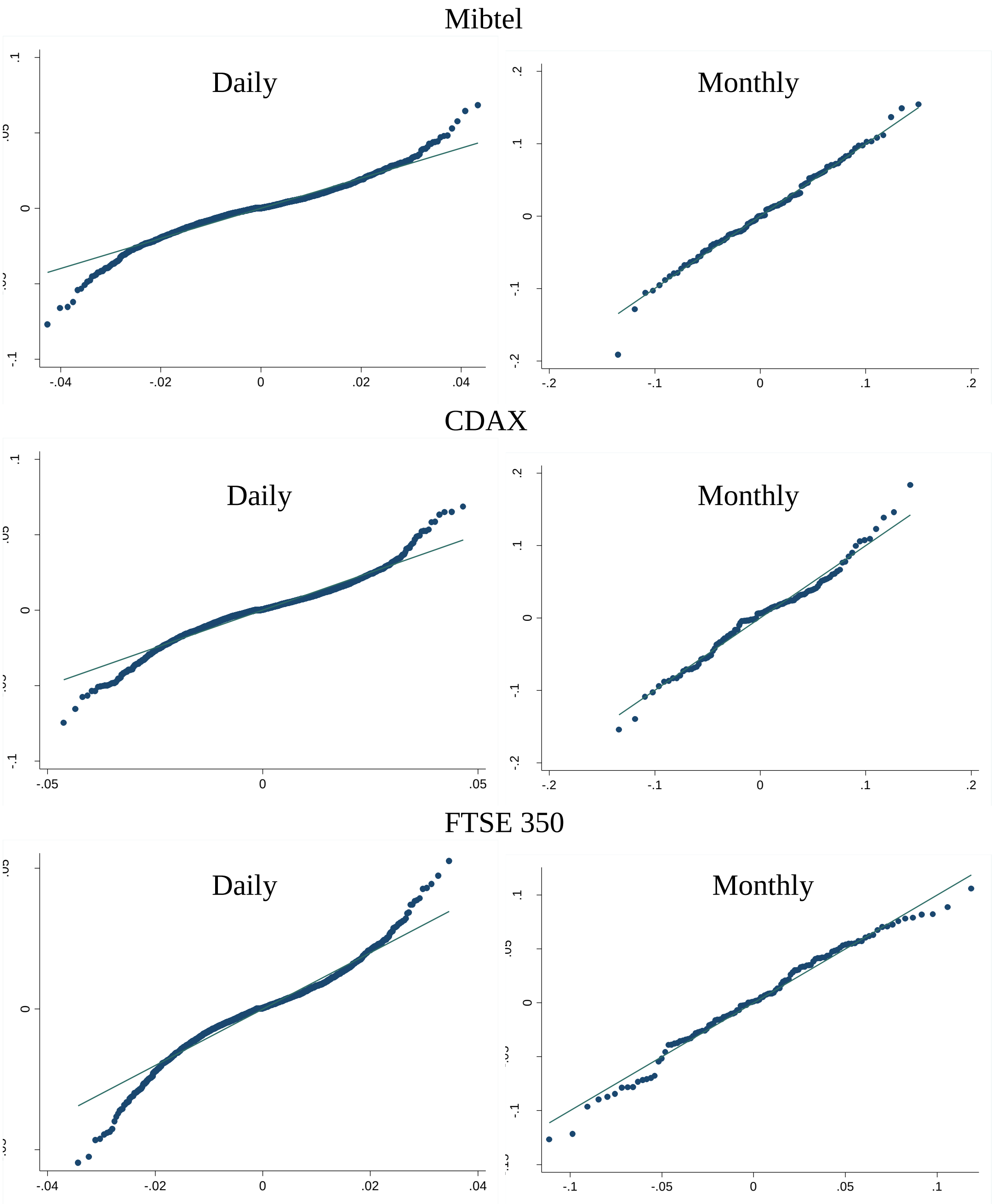
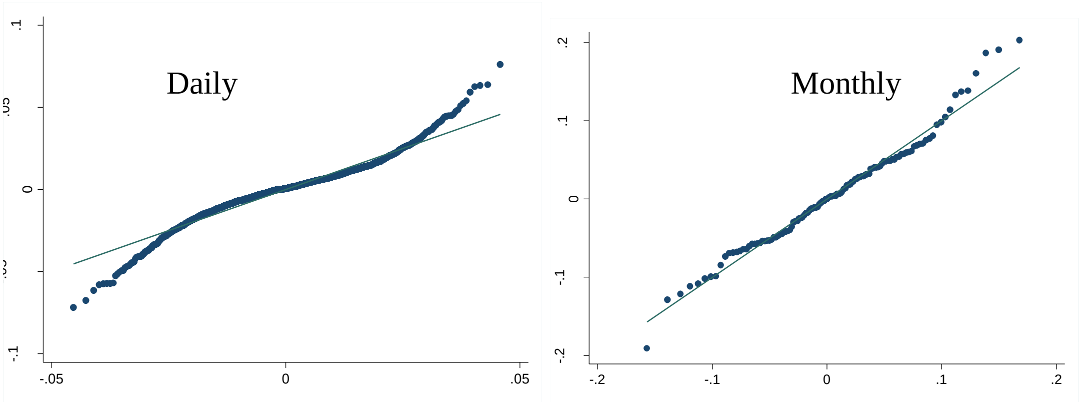
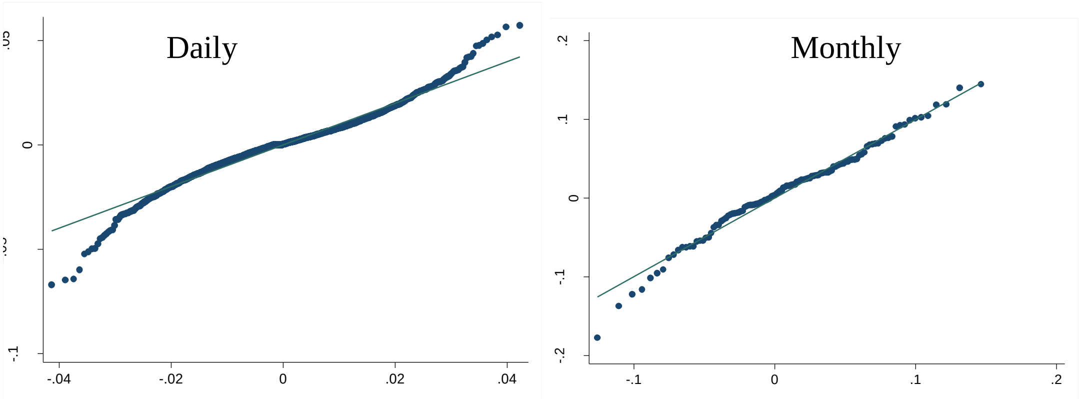


Figure 3 continued

ASE



MSE



SWX

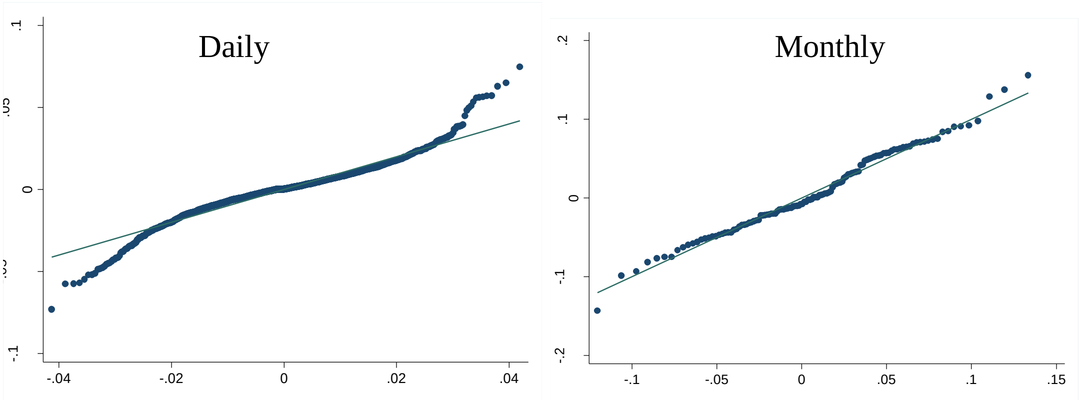
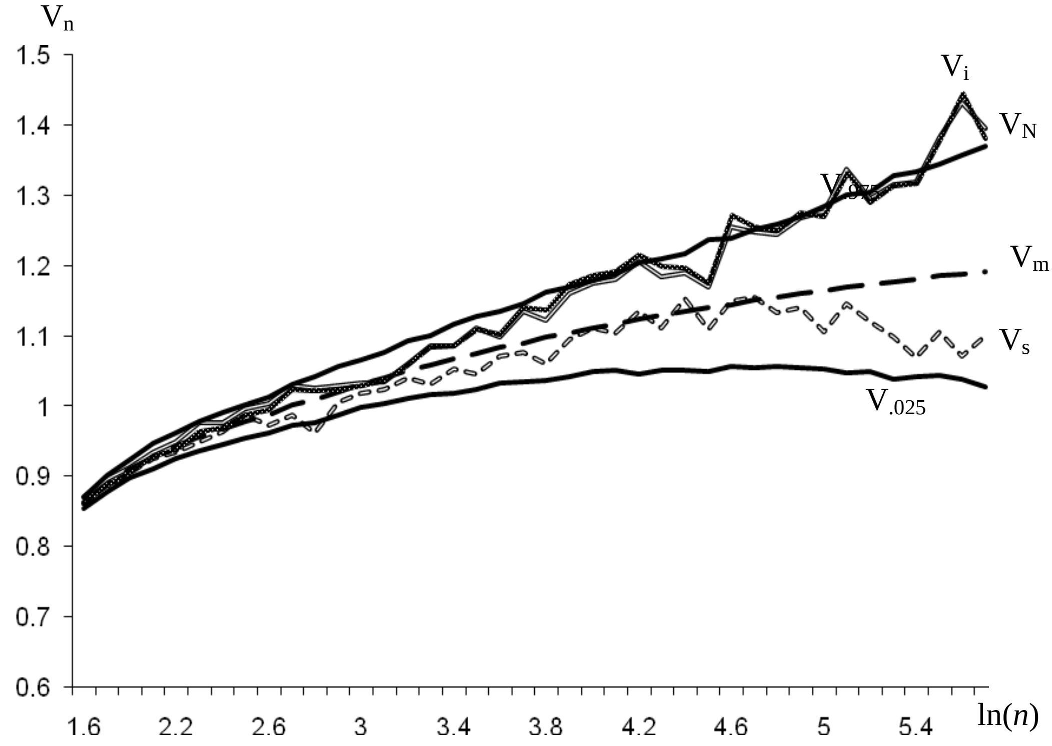


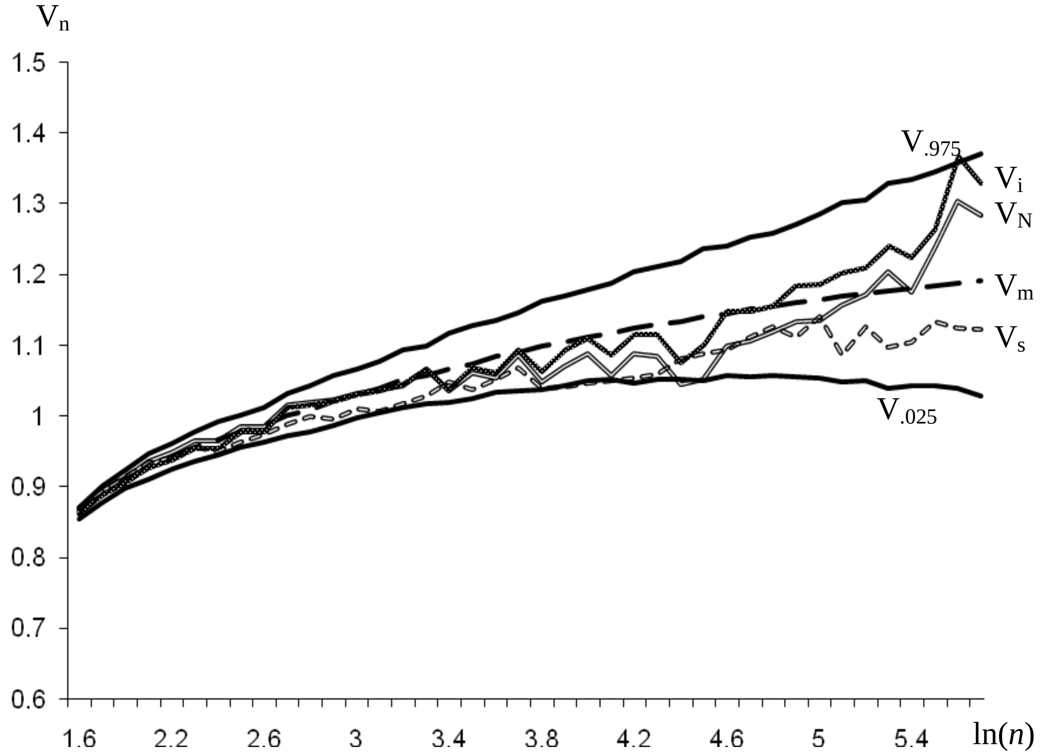
Figure 4: Plots of the realised values of V_i against V_s, V_N, V_m, $V_{.975}$ and $V_{.025}$ – Mibtel.



Notes:

V_i	refers to the V-statistic for the log return series of Mibtel for scale {n}.
V_s	refers to the V-statistic of the shuffled log return series of Mibtel for scale {n}.
V_N	refers to the V-statistic of the Normalised log return series of Mibtel for scale {n}.
V_m	refers to the average V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.
$V_{.975}$	refers to the upper critical value (5% level) of the V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.
$V_{.025}$	refers to the lower critical value (5% level) of the V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.

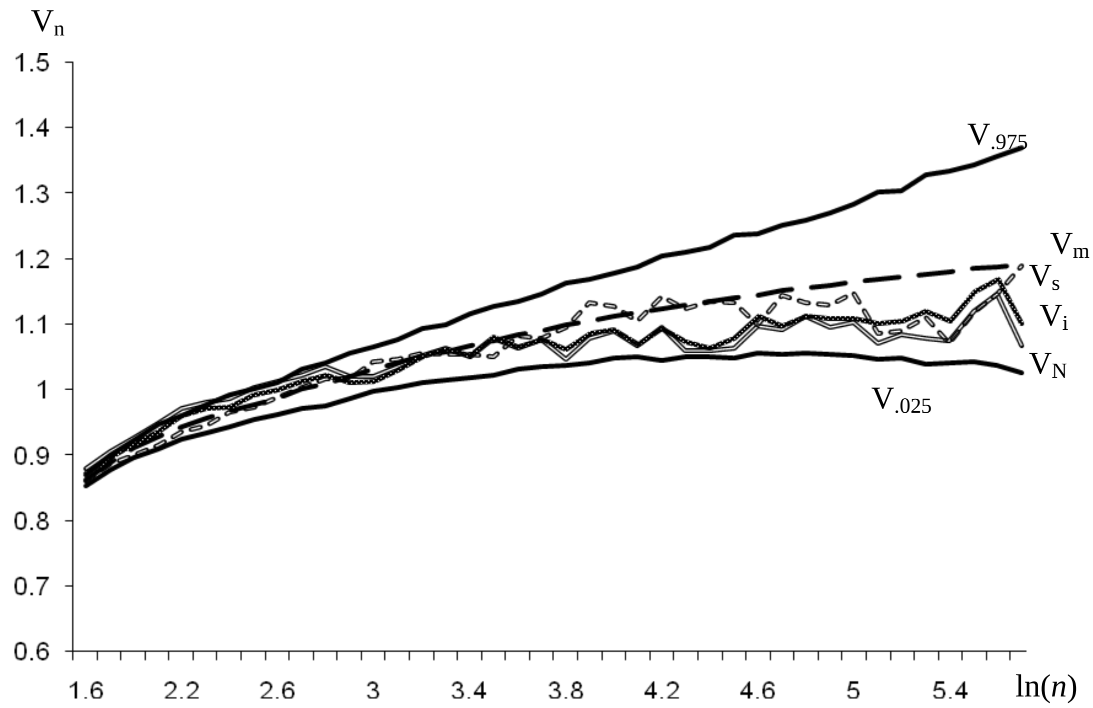
Figure 5: Plots of the realised values of V_i against V_s, V_N, V_m, $V_{.975}$ and $V_{.025}$ – CDAX.



Notes:

V_i	refers to the V-statistic for the log return series of CDAX for scale {n}.
V_s	refers to the V-statistic of the shuffled log return series of CDAX for scale {n}.
V_N	refers to the V-statistic of the Normalised log return series of CDAX for scale {n}.
V_m	refers to the average V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.
$V_{.975}$	refers to the upper critical value (5% level) of the V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.
$V_{.025}$	refers to the lower critical value (5% level) of the V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.

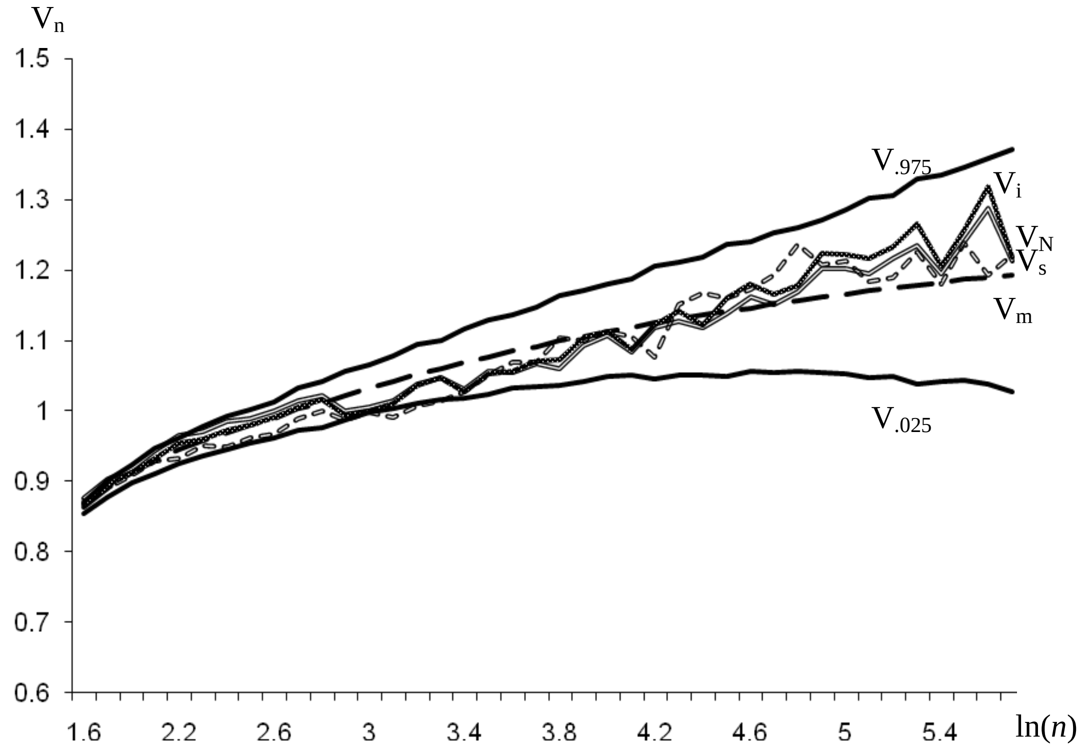
Figure 6: Plots of the realised values of V_i against V_s, V_N, V_m, $V_{.975}$ and $V_{.025}$ – FTSE 350.



Notes:

V_i	refers to the V-statistic for the log return series of FTSE 350 for scale {n}.
V_s	refers to the V-statistic of the shuffled log return series of FTSE 350 for scale {n}.
V_N	refers to the V-statistic of the Normalised log return series of FTSE 350 for scale {n}.
V_m	refers to the average V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.
$V_{.975}$	refers to the upper critical value (5% level) of the V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.
$V_{.025}$	refers to the lower critical value (5% level) of the V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.

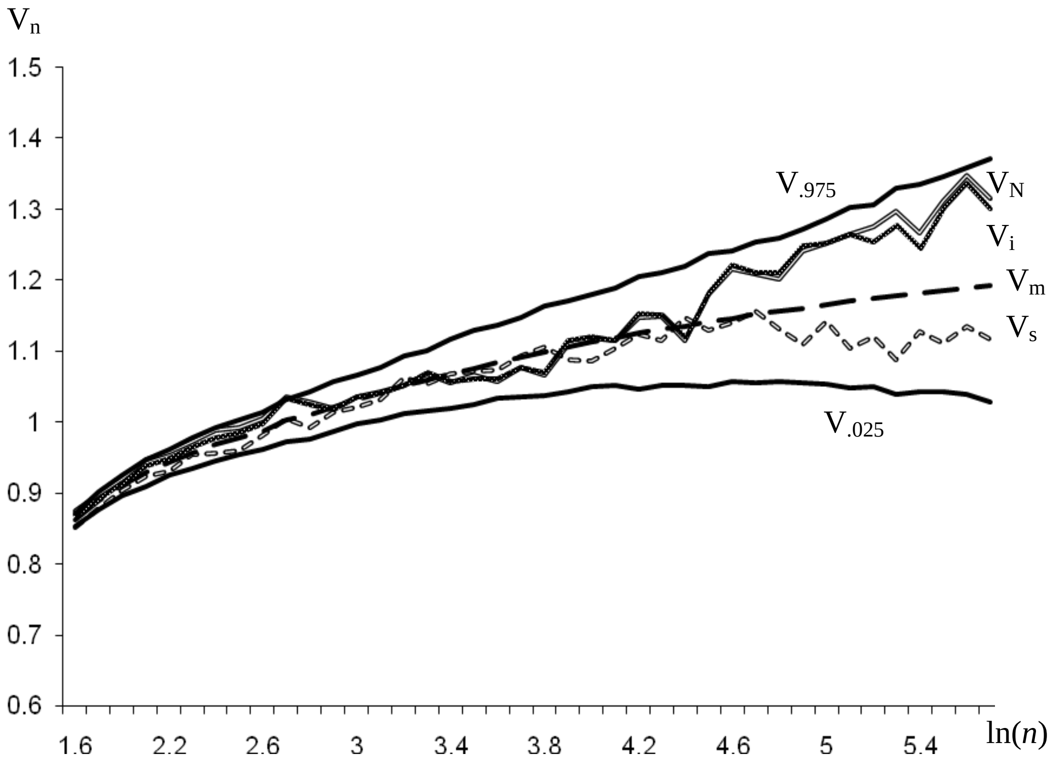
Figure 7: Plots of the realised values of V_i against V_s, V_N, V_m, $V_{.975}$ and $V_{.025}$ – ASE.



Notes:

V_i	refers to the V-statistic for the log return series of ASE for scale {n}.
V_s	refers to the V-statistic of the shuffled log return series of ASE for scale {n}.
V_N	refers to the V-statistic of the Normalised log return series of ASE for scale {n}.
V_m	refers to the average V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.
$V_{.975}$	refers to the upper critical value (5% level) of the V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.
$V_{.025}$	refers to the lower critical value (5% level) of the V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.

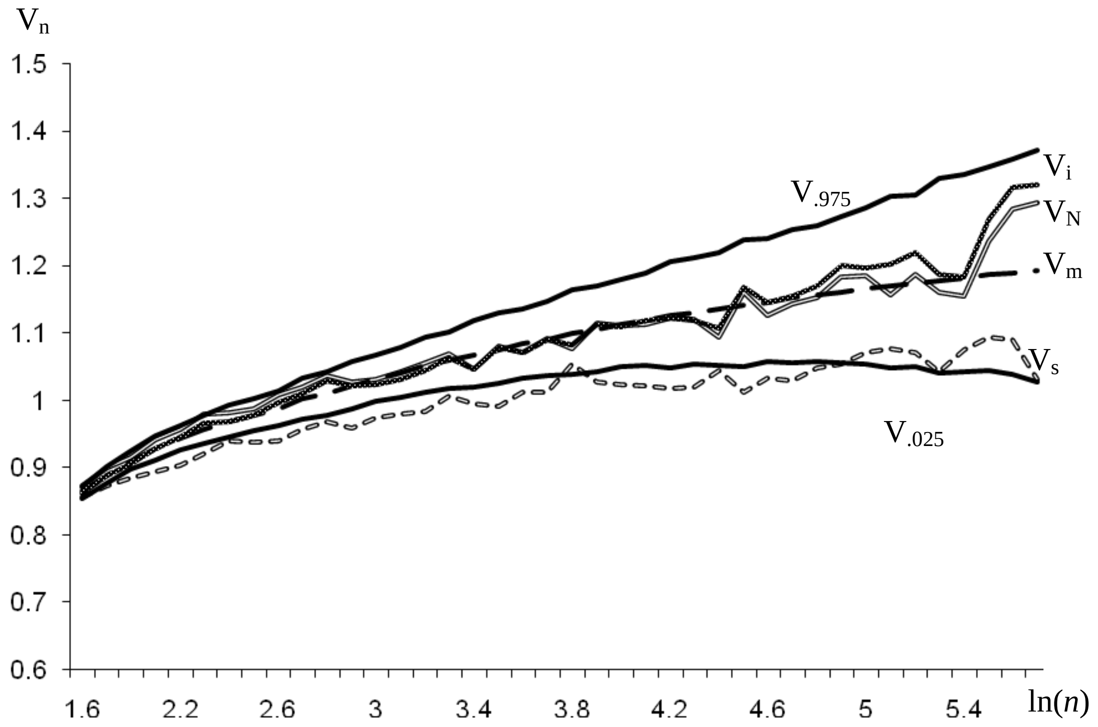
Figure 8: Plots of the realised values of V_n against V_m, $V_{.975}$ and $V_{.025}$ – MSE.



Notes:

V_i	refers to the V-statistic for the log return series of MSE for scale {n}.
V_s	refers to the V-statistic of the shuffled log return series of MSE for scale {n}.
V_N	refers to the V-statistic of the Normalised log return series of MSE for scale {n}.
V_m	refers to the average V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.
$V_{.975}$	refers to the upper critical value (5% level) of the V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.
$V_{.025}$	refers to the lower critical value (5% level) of the V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.

Figure 9: Plots of the realised values of V_n against V_m, $V_{.975}$ and $V_{.025}$ – SWX.



Notes:

V_i	refers to the V-statistic for the log return series of SWX for scale {n}.
V_s	refers to the V-statistic of the shuffled log return series of SWX for scale {n}.
V_N	refers to the V-statistic of the Normalised log return series of SWX for scale {n}.
V_m	refers to the average V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.
$V_{.975}$	refers to the upper critical value (5% level) of the V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.
$V_{.025}$	refers to the lower critical value (5% level) of the V-statistic calculated for each scale {n} and obtained from the 5,000 Monte Carlo simulations.

Table 1: Descriptive statistics for the log return series calculated for the time scales: daily, weekly, monthly, and quarterly.

	Mean				*Standard deviation*			
	D	W	M	Q	D	W	M	Q
Mibtel	0.0004	0.0018	0.0077	0.0230	0.0128	0.0285	0.0644	0.1037
CDAX	0.0002	0.0009	0.0041	0.0123	0.0138	0.0296	0.0654	0.1062
FTSE350	0.0002	0.0008	0.0036	0.0109	0.0103	0.0212	0.0400	0.0661
ASE	0.0003	0.0013	0.0055	0.0165	0.0135	0.0284	0.0589	0.0982
MSE	0.0005	0.0024	0.0103	0.0310	0.0124	0.0268	0.0585	0.0957
SWX	0.0003	0.0015	0.0066	0.0199	0.0124	0.0263	0.0530	0.0898
	Skewness				*Kurtosis*			
	D	W	M	Q	D	W	M	Q
Mibtel	-0.2069	-0.3173	0.2048	0.4449	6.1328	5.0851	3.8485	3.2967
CDAX	-0.2404	-0.4544	-0.8957	-0.4392	5.9983	5.9294	5.5214	3.3556
FTSE350	-0.2426	-0.3489	-0.9070	-0.8632	5.9985	5.0516	4.1338	4.5561
ASE	-0.2144	-0.6814	-0.8963	-1.0305	6.4857	5.6720	4.3286	4.2485
MSE	-0.2481	-0.5190	-0.6175	-0.4311	5.7933	5.5297	4.4316	3.1450
SWX	-0.1546	-0.3278	-1.0693	-0.8010	7.2722	6.9835	5.3010	4.0282

Notes:
D = daily, W = weekly, M = monthly, and Q = quarterly. The number of observations for each sample frequency is: 2,608 (D), 522 (W), 120 (M) and 40 (Q).

Table 2: Descriptive statistics for the shuffled log return series calculated for the time scales: daily, weekly, monthly, and quarterly.

	Mean				*Standard deviation*			
	D	W	M	Q	D	W	M	Q
Mibtel	0.0004	0.0018	0.0077	0.0230	0.0128	0.0284	0.0594	0.1053
CDAX	0.0002	0.0009	0.0041	0.0123	0.0138	0.0306	0.0576	0.1050
FTSE350	0.0002	0.0008	0.0036	0.0109	0.0103	0.0233	0.0479	0.0719
ASE	0.0003	0.0013	0.0055	0.0165	0.0135	0.0296	0.0678	0.1189
MSE	0.0005	0.0024	0.0103	0.0310	0.0124	0.0277	0.0568	0.1011
SWX	0.0003	0.0015	0.0066	0.0199	0.0124	0.0247	0.0529	0.1008
	Skewness				*Kurtosis*			
	D	W	M	Q	D	W	M	Q
Mibtel	-0.2069	0.0058	-0.1245	0.2896	6.1328	3.5781	3.3671	2.8103
CDAX	-0.2404	-0.2658	0.0869	-0.1896	5.9983	3.3644	3.6264	2.8605
FTSE350	-0.2426	-0.1545	-0.3739	-0.3915	5.9985	3.5712	2.7612	3.1666
ASE	-0.2144	0.0029	0.3306	0.3772	6.4857	3.1832	3.6534	2.6888
MSE	-0.2481	-0.1944	-0.3888	-0.2709	5.7933	3.6475	3.6577	2.2852
SWX	-0.1546	-0.2709	0.1806	-0.2209	7.2722	3.5318	2.9351	2.6644

Notes:
D = daily, W = weekly, M = monthly, and Q = quarterly. The number of observations for each sample frequency is: 2,608 (D), 522 (W), 120 (M) and 40 (Q).

Table 3: Rescaled Range Analysis results (with pre-filtering).

Index	H_i	H_s	H_N	***Monte Carlo simulations*** *Mean and critical values*	*Estimated H*
Mibtel	0.593	0.560	0.592	μ_H	0.572
CDAX	0.592	0.559	0.580	*quantiles*	
FTSE350	0.585	0.564	0.576	0.005	0.528
ASE	0.596	0.575	0.586	0.025	0.539
MSE	0.590	0.561	0.589	0.050	0.545
SWX	0.598	0.550	0.588	0.950	0.601
				0.975	0.612
				0.999	0.618

Notes:
The second column reports the estimated Hurst exponent for the log return series of each index, H_i. The third column reports the estimated Hurst exponent for the shuffled log return series of each index, H_s. The fourth column reports the estimated Hurst exponent for the Normalised log return series of each index, H_N. The sixth column reports the average Hurst exponent estimated with respect to the 5,000 Monte Carlo simulations, μ_H, and the Hurst exponents associated with various quantiles of the Monte Carlo simulations. The series have been pre-filtered using an AR(p) model, where p includes the lags (up to the tenth lag) for which the partial autocorrelation function of the log returns is significantly (α = 5% level at least) different from zero. The lags used for each index are: Mibtel – 4th lag; CDAX – 6th and 8th lag; FTSE350 – 3rd, 5th, 6th, 8th, 10th lag; ASE – 3rd, 5th, 8th, 9th lag; MSE – 3rd, 8th, 10th lag; SWX – 5th lag.

Table 4: Rescaled Range Analysis (with pre-filtering) with spline regressions for 5 ≤ n ≤ 40.

Index	H_i	H_s	H_N	***Monte Carlo simulations***	
				Mean and critical values	*Estimated H*
Mibtel	0.606	0.611	0.598	μ_H	0.606
CDAX	0.591	0.580	0.577	*quantiles*	
FTSE350	0.617	0.612	0.612	0.005	0.562
ASE	0.614	0.593	0.600	0.025	0.573
MSE	0.606	0.613	0.597	0.050	0.579
SWX	0.603	0.577	0.598	0.950	0.633
				0.975	0.645
				0.999	0.650

Notes:
The second column reports the estimated Hurst exponent for the log return series of each index, H_i. The third column reports the estimated Hurst exponent for the shuffled log return series of each index, H_s. The fourth column reports the estimated Hurst exponent for the Normalised log return series of each index, H_N. The sixth column reports the average Hurst exponent estimated with respect to the 5,000 Monte Carlo simulations, μ_H, and the Hurst exponents associated with various quantiles of the Monte Carlo simulations. The series have been pre-filtered using an AR(p) model, where p includes the lags (up to the tenth lag) for which the partial autocorrelation function of the log returns is significantly (α = 5% level at least) different from zero. The lags used for each index are: Mibtel – 4th lag; CDAX – 6th and 8th lag; FTSE350 – 3rd, 5th, 6th, 8th, 10th lag; ASE – 3rd, 5th, 8th, 9th lag; MSE – 3rd, 8th, 10th lag; SWX – 5th lag.

Table 5: Rescaled Range Analysis (with pre-filtering) with spline regressions for 40 ≤ n ≤ 299.

Index	H_i	H_s	H_N	***Monte Carlo simulations***	
				Mean and critical values	*Estimated H*
Mibtel	0.580	0.510	0.586	μ_H	0.540
CDAX	0.593	0.539	0.582	*quantiles*	
FTSE350	0.554	0.518	0.542	0.005	0.451
ASE	0.578	0.558	0.572	0.025	0.475
MSE	0.573	0.512	0.581	0.050	0.484
SWX	0.593	0.524	0.579	0.950	0.597
				0.975	0.618
				0.999	0.630

Notes:
The second column reports the estimated Hurst exponent for the log return series of each index, H_i. The third column reports the estimated Hurst exponent for the shuffled log return series of each index, H_s. The fourth column reports the estimated Hurst exponent for the Normalised log return series of each index, H_N. The sixth column reports the average Hurst exponent estimated with respect to the 5,000 Monte Carlo simulations, μ_H, and the Hurst exponents associated with various quantiles of the Monte Carlo simulations. The series have been pre-filtered using an AR(p) model, where p includes the lags (up to the tenth lag) for which the partial autocorrelation function of the log returns is significantly (α = 5% level at least) different from zero. The lags used for each index are: Mibtel – 4th lag; CDAX – 6th and 8th lag; FTSE350 – 3rd, 5th, 6th, 8th, 10th lag; ASE – 3rd, 5th, 8th, 9th lag; MSE – 3rd, 8th, 10th lag; SWX – 5th lag.

Table 6: Rescaled Range Analysis (without pre-filtering).

Index	H_i	H_s	H_N	*Monte Carlo simulations*	
				Mean and critical values	*Estimated H*
Mibtel	0.613**	0.559	0.610	μ_{H}	0.572
CDAX	0.588	0.557	0.573	*quantiles*	
FTSE350	0.553	0.561	0.543	0.005	0.528
ASE	0.588	0.587	0.580	0.025	0.539
MSE	0.595	0.562	0.595	0.050	0.545
SWX	0.585	0.551	0.575	0.950	0.601
				0.975	0.612
				0.999	0.618

Notes:
The second column reports the estimated Hurst exponent for the log return series of each index, H_i. The third column reports the estimated Hurst exponent for the shuffled log return series of each index, H_s. The fourth column reports the estimated Hurst exponent for the Normalised log return series of each index, H_N. The sixth column reports the average Hurst exponent estimated with respect to the 5,000 Monte Carlo simulations, μ_{H}, and the Hurst exponents associated with various quantiles of the Monte Carlo simulations.
** Denotes departure from the RWH at the 5% level of significance.

Table 7: Rescaled Range Analysis (without pre-filtering) with spline regressions for $5 \leq n \leq 40$.

Index	H_i	H_s	H_N	***Monte Carlo simulations*** *Mean and critical values*	*Estimated H*
Mibtel	0.623	0.610	0.613	μ_H	0.606
CDAX	0.586	0.579	0.572	*quantiles*	
FTSE350	0.586	0.613	0.580	0.005	0.562
ASE	0.588	0.600	0.580	0.025	0.573
MSE	0.593	0.616	0.585	0.050	0.579
SWX	0.594	0.577	0.589	0.950	0.633
				0.975	0.645
				0.999	0.650

Notes:
The second column reports the estimated Hurst exponent for the log return series of each index, H_i. The third column reports the estimated Hurst exponent for the shuffled log return series of each index, H_s. The fourth column reports the estimated Hurst exponent for the Normalised log return series of each index, H_N. The sixth column reports the average Hurst exponent estimated with respect to the 5,000 Monte Carlo simulations, μ_H, and the Hurst exponents associated with various quantiles of the Monte Carlo simulations.

Table 8: Rescaled Range Analysis (without pre-filtering) with spline regressions for 40 ≤ n ≤ 299.

Index	H_i	H_s	H_N	***Monte Carlo simulations*** *Mean and critical values*	*Estimated H*
Mibtel	0.602*	0.509	0.607	μ_H	0.540
CDAX	0.590	0.535	0.573	*quantiles*	
FTSE350	0.520	0.510	0.506	0.005	0.451
ASE	0.587	0.573	0.580	0.025	0.475
MSE	0.597	0.510	0.604	0.050	0.484
SWX	0.577	0.525	0.561	0.950	0.597
				0.975	0.618
				0.999	0.630

Notes:
The second column reports the estimated Hurst exponent for the log return series of each index, H_i. The third column reports the estimated Hurst exponent for the shuffled log return series of each index, H_s. The fourth column reports the estimated Hurst exponent for the Normalised log return series of each index, H_N.The sixth column reports the average Hurst exponent estimated with respect to the 5,000 Monte Carlo simulations, μ_H, and the Hurst exponents associated with various quantiles of the Monte Carlo simulations.
* Denotes departure from the RWH at the 10% level of significance.