Stock Returns Memories: a "Stardust" Memory?

Julien Fouquau*, Philippe Spieser†

October 12, 2011

Abstract

This article aims to investigate if stock market index returns present any type of memory. We study the dependence structure of four market indices between 1959 - 2010 and 1970 -2010. We used three different methodologies to obtain Hurst exponent, starting from the basic and old "R/S" approach, continuing with ARFIMA models and ending with the new and innovative wavelet analysis. Our findings are coherent according to the various methods, leading to the conclusion of absence or very short memory dynamics. Those results are in accordance of the weak efficiency financial theory, restraining successful forecasts and arbitrage opportunities.

Key Words: Hurst exponent, ARFIMA models, Wavelet models.

J.E.L Classification: C22, C58, G14, G17.

*Rouen Business School and CGEMP LEDa, E-mail: jfq@rouenbs.fr.
†ESCP Europe, E-mail: spieser@escpeurope.eu

Introduction

Financial markets are supposed to check a central hypothesis, namely the "efficiency market hypothesis" (EMH), defined in seminal papers by Fama (1965, 1972). Fama's important assumption states that in any given time, the market prices immediately reflect all public and available information. Therefore, no one could outperform the market by using the same information disposal to all investors and identifying possible trends or patterns of price changes can't be used to predict the future value of assets. In other words, as news are unpredictable, price changes are unexpected and the returns are realizations of a random process.

There is a closely linked assumption with the previous one, namely that stock returns are non autocorrelated. The popularity of this assumption stems from the fact that normally distributed returns are an implication of the "random walk" theory defined for the first time by Louis Bachelier (1900). In case of sufficient observations and if stock prices actually follow a random walk, if the returns are actually independent and identically distributed, then their distribution should be gaussian by applying the central limit theorem. According to the martingale property of daily stock prices P_t, the time series of the natural logarithm of returns $R_t = ln(P_t/P_{t-1})$, which is referred to as a geometric Brownian motion, shows no serial correlation.

The critiques addressed to those "stationary, independent and Gaussian" processes are twofold. On first hand, Empirical time series are obviously not Gaussian, the two first moments vary over time and the moments of order three and four do not fit to normal distribution. On the second hand, the conventional stationarity does not properly model a possible "long term dependence"[1] which could be revealed by a careful and deep study of serial autocorrelation. In the time space, the notion of long term dependence means

[1] We will use also "long term memory process".

that a time series present serial autocorrelation slowly decreasing when the orders increase. The presence of a long term memory implies lasting consequences of a external shock on that series. Long memory is one of the most "attractive" nonlinear representation of the return dynamics. The movements in the market might be influenced by the furthest past and price movements are then endogeneous phenomena. And if series of financial returns have a long memory, a reliable prediction of future returns based on the past is then possible and predictability of the returns, properly measured and utilized, allows to set up profitable trading strategies.

There are different techniques designed to reveal the dependence structure of time series. The most common one is the Hurst (1951) or R/S statistics, which comes from hydrology and has been thereafter adapted to financial problems by Mandelbrot (1965,1972). Considering the theoretical and practical limitations of this statistics (see further in section 2), another method has been proposed, namely the ARFIMA models. Those representations have the interesting and challenging particularity to provide a frame for short term as well as long term dynamics. More recently a promising technique has been developed, particularly in physics (signal, geophysics, computer science) but already tested those last years in some economic papers: the wavelet analysis. Like the Fourier transform, wavelet analysis decomposes signal into elementary periodic functions. The advantage of that technique is to describe the observed data in time space and frequency space but with flexible bandwidths. The main contribution of our paper lies precisely in the use of a specific wavelet applied to the determination of the depth of the memory.

Our paper presents the abovementioned approaches of long term memory estimation applied to four stock market indices of important countries: the Standard and Poors 500 and the Dow Jones for USA, the Dax 30 for Germany and the Nikkei for Japan. Since the litterature has proved the sensitivity of results to the lenght of the data sample and in order to comfort our results, we analyzed two different long periods respectively 1959 -2010 and 1970-2010 of daily returns. The purpose was to have significant long sample to check the presence of a long term memory effect. Our first main contribution is that, for the first time, the studied statistics are estimated by different techniques and different time periods to see if they give robust and coherent results. Another contribution is that we used new types of wavelets in finance, the Daubechies wavelets, applied on long time series of financial data. In ordrer to avoid an arbitrary choice, we used a large set of Daubechies moment and select the best. We so contributed to the debate of the market efficiency, at least in developed countries.

We follow the main step of previous studies investigating the long term persistence of markets indices that are featured by contradictory results. Those results are in accordance of the weak efficiency financial theory, restraining successful forecasts and arbitrage opportunities.

The paper is organized as follows: the first part is dedicated to present the three main techniques, the Hurst approach, the ARFIMA model and the wavelet analysis. In that section, there will be a short formal presentation followed by an overview of the existing related literature. The second section is the state of art methodology. The third section will be a presentation of our data and findings. We will conclude in the last section.

1 Capture the long memory

A lot of theoretical models are based on the efficient market hypothesis. So to validate empirically this assumption is of high importance in finance. This hypothesis is rejected when for example, one can put into evidence regular cycles or trends in the returns series allowing for forecasting market performances. This study can be led through different techniques, from the most simple and empirical one (Hurst methodology) to the most adapted one for signal capture processing (wavalet analysis). We will present those approaches which has been already used in specific fields of financial markets as exchange rates (Diebold et al., 1991, Baillie, 1996 ...) or GDP growth rates.

Three different approaches

The first attempt to capture long term memory (LT memory) in a purely heuristic way is due to the hydrologist H.E Hurst dealing with the data of water heights of the Nile river. He constructed the so called "Range Statistics" (R/S) denominated thereafter the Hurst coefficient or exponent "H". This approach was transposed in the field of finance by Mandelbrot (1965, 1972). The value of the exponent H, which lies between 0 and 1, allows to sort the time series. When $H = 0.5$, this latter can be modeled by the white noise, there is no long term dependence, the financial market might be efficient; when $0.5 < H < 1$ there is a long term memory, the market is inefficient - one can speak of a persistence effect, the time series will show clusters of comparable values. On the other hand, when $0 < H < 0.5$, there is a short term memory, one can speak of a antipersistence effect. In the eighties and nineties, the development of new econometric approaches showed the limitations of the Hurst method and more precisely that this technique cannot properly discriminate between short term and long term memories.

The second technique refers to ARFIMA processes which generalizes the well-known Box and Jenkins' ARIMA processes and allows to take into account the short and long term dynamics (Cochrane, 1988). Granger and Joyeux (1980) and Hosking (1981) had studied their properties in the time state space (autocovariance or autocorrelation functions) as well as in the frequency domain (spectral density function). In $ARFIMA(p, d, q)$ processes, the integration d of the time series (i.e. the number of times the series must be differenced) is fractional, contrarily to ARIMA model where the parameter is an integer. The main contribution of that approach is that fractional differenced parameter "d" captures the behavior of LT dependence. There is a remarkable relation between the Hurst exponent and the so called ARFIMA processes as proved by Hosking (1981) and Geweke Porter Hudak (GPH, 1983). The fractional integration parameter d and the hurst exponent H are bounded with the simple equation $H = d + \frac{1}{2}$. The time series can be sorted according to "d" in the same way as H did: if $0 < d < \frac{1}{2}$ the ARFIMA process is stationary with long term memory. If $d = 0$ the ARFIMA process is reduced to a standard ARMA process. If $-1/2 < d < 0$ the process is antipersistent, autocorrelations alternate. In the frequency domain, the spectral density tends to 0 when frequency tends to 0.

The third technique is the newest one, namely the wavelet approach. It belongs to the signal and data processing methods and consequently particularly adapted to the detection of periodic cycles. It is important to capture in a more accurate way the different frequencies fluctuations in time series analysis and to distinguish between high and low frequencies respectively associated to short and long term memories (Crowley, 2005). The Fourier transform is not absolutely a performing method to seize the shocks and changing patterns which affect the returns, especially in presence of non stationnarities. It decomposes the time series into infinite length trigonometric measures sines and cosines, discarding all time localization information. The wavelet method circumvents those limits notably because it allows analyzing data simultaneously in flexible time and in frequency domains. It decomposes a time series in terms of elementary functions which are referred to as "wavelets", small waves that grow and decay in a limited time period. They result from a mother wavelet and father which can be expressed as function of the time position(translation parameter) and the scale related to frequency (dilatation parameter). The father wavelet will help us to detect the low frequency dynamics and consequently the features of the long memory process.

Literature

As already mentioned, there is a vast literature concerning the detection of long term memory. The different contributions deal with many types of assets, exchange rates, stock indices... essentially in developed markets. Contrasted and sometimes contradictory results have been obtained by the different authors (for a survey see Sewell, 1997). A good example is given by a recent contribution of Serletis and Rosenberg (2009) who used daily data on four US stock market indices (SP500, Nasdaq, DJ and NYSE composite index), concluded that US stock markets returns display persistence, which was contested by heavily Kritoufek (2010) who described this conclusion as "spurious" due to an "incorrect use of detrending moving average method". We organize the review according to the conclusions regarding persistence or non persistence effects.

Among the relevant papers concluding with the existence of a long term memory in returns, one can note the following ones. Goetzmann (1993) studied the detrended London Stock Exchange and New York Stock Exchange prices with R/S analysis, on a three centuries prices data base. Nawrocki (1995) considered the CRSP monthly value-weighted index and the SP500 daily index with R/S and Hurst analysis and actually found a persistent finite memory. Huang and Yang (1999) concluded also positively with intraday NYSE and NASDAQ indices. With spectral analysis, Barkoulas, Baum and Travlos (2000), put also into evidence a significant positive LT persistence in the Greek market. Chen (2000) calculated a Hurst coefficient in seven Pacific and Asian countries and concluded also with the presence of a long term memory effect. According to Nath (2001) who used a R/S analysis, Indian stock exchange reveals a long-term memory and Panas (2001) concluded identically for the Athens stock exchange. Cajueiro and Tabak (2004) investigate the long-range dependence phenomena in three Asian markets, namely Hong Kong, Singapore and China and find evidence that these markets present long range-dependence. They went further by giving some explanations to the phenomena, above all liquidity and market capitalization. They suggest that more research is needed on testing for long range dependence for different markets. The same authors in 2005, investigated the H coefficient with six different methods. They pointed out that in the Brazilian Stock Market,the presence of long-range dependence in the asset returns "seems to be a stylized fact".

The following papers conclude with a negative or at least very weak evidence of a long term effect, using Hurst analysis : Mill (1993) in daily UK stock returns, Chow, Pan & Sakano (1996) for twenty-two international equity market indices, Lux (1996) studied the German market, Lardic & Mignon (1996) the canadian market and Hiemstra & Jones (1997) studied the return series of 1952 stocks. Lobato and Savin (1998) tested with semi-parametric procedure the presence of long memory in daily individual stock returns and studied also the square returns in the frame of volatility inquiry. For them, spurious results can be produced by non-stationarity and aggregation procedure. The test results show no evidence of long memory in the returns but by contrast, there is strong evidence in the squared returns. Grau-Carles (2005) found no evidence of long term memory in two major daily indices, the SP500 and DJ Industrials average. She used R/S analysis, modified R/S, GPH technique, tested the short-memory processes with AR, MA and ARCH models and long memory ones with ARFIMA model. Oh, Um and Kim (2006) studied stock market indices and foreign exchange market, for all daily and high frequency market data, no significant long-term memory could be detected, while a strong long-term memory property was nevertheless found in the volatility time series which was reviewed in the same paper.

Some papers at last proposed balanced conclusions. Barkoulas and Baum (1996) applied the spectral regression method and found no evidence in either aggregate or sectorial stock indices. They found nevertheless evidence of long memory in five companies out of thirty of the DJIA, intermediate memory in three and fractal structure in twenty-two. Willinger, Taqqu, Teverovsky (1999) dealing with stock price returns and using an ARFIMA technique, had also very cautious and balanced conclusions. Sadique and Silvapulle (2001) took the weekly stock returns of seven countries and found contradictory results: negative for three countries, namely Japan, USA, Australia, positive long term dependence for Korea, New Zealand, Malaysia, Singapore. Henry (2002) studied stock returns of nine countries and his results are negative for five countries (UK, USA, Hong Kong, Singapore, Australia), positive long term dependence for Germany, Japan, SouthKorea and Taiwan. Tolvi (2003) found long memory effect in the Finish market and also, in another study, Denmark and Ireland, but no effect for thirteen other markets. In another study, he used non-parametric methods, the data set has daily returns on six indices and forty companies. Depending on the method used, statistically significant long memory is detected in 24% to 67% of the series. Still in 2003, enlarging his study, he also carried out analysis on stock market returns of a monthly data set among sixteen OECD countries, statistically significant long memory is found for three countries. Limam (2003) analyzed stock market returns in fourteen markets and concluded for long-term memory only for the small ones. His study concerned stock index returns in fourteen markets with diverse levels of maturity - Japan, UK and USA, Brazil, India and Mexico, eight Arab countries using parametric and semi-parametric estimation procedures, the results show that the property of long-range dependence in stock index returns tend to be associated with relatively narrow stock markets.

The global conclusion we might draw from that literature is that most papers find no persistence effects in mature markets, whereas younger and consequently smaller stock exchanges do not check the efficiency hypothesis. Long memory has been widely found when using Hurst coefficient. So, empirical results seem support the hypothesis of more frequent presence of long memory in small markets but the limited length of the data set is a restriction to the accuracy of tests. Some evidence was also found that the presence of long memory in the returns of individual stocks is correlated with the sample moments. It is consequently desirable to set up the different techniques to put into evidence possible long term memory in order to comfort the results. Three estimation methods are now developed successively, the Hurst exponent, the ARFIMA models with GPH and Whittle methodology and the wavelet analysis.

2 Specification and Estimation methodology

2.1 The Hurst coefficient "H"

The Hurst coefficient has introduced for the first time a statistics allowing to detect the presence of a long term memory, classifying time series according to the nature of their dependence. Besides its historical interest and altogether present relevance in hydrology (Koirala et al., 2009), it gives a simple, rapid and rather easy calculation of the order of differentiating in continuous ARFIMA processes *cf. infra.* Hurst defined a rescaled range "R/S" statistics, which can be described as the span or expanse of the partial sums of the gap of a time series to its average divided by its standard error:

$$\left(\frac{R}{S}\right)_t = \frac{1}{\frac{1}{T}\sum_{t=1}^{T}\left(X_t - \bar{X}\right)^2}\left(\underset{1<k<T}{Max}\sum_{j=1}^{k}\left(X_j - \bar{X}_T\right) - \underset{1<k<T}{Min}\sum_{j=1}^{k}\left(X_j - \bar{X}_T\right)\right) \tag{1}$$

It has been proved (Mandelbrot, 1972) that this statistics is asymptotically proportional to T^H where T is the number of observations and the constant H, $0 < H < 1$ is precisely the Hurst exponent. So we can write:

$$\left(\frac{R}{S}\right)_t = c.T^H \tag{2}$$

where c is a constant. Then, the Hurst exponent is the estimated slope coefficient obtained by regressing with the OLS technique the logarithm of $\left(\frac{R}{S}\right)_t$ versus the logarithm of the time T:

$$ln\left(\frac{R}{S}\right)_t = ln(c) + H\ln(T) \tag{3}$$

This "H" has originally nevertheless major inconveniences. It is highly sensitive to the short term dependence (Lo 1991), its theoretical distribution is unknown (Hosking 1984). It is so impossible to have the standard error and the probability law associated with the estimator. Hosking (1984) stresses that H, when estimated with "R/S" analysis, is biased upwards if theoritical H is lower than 0.7 and downwards if superior to 0.7. Lo (1991) himself modified H based on the range of the partial sum by replacing the standard variance estimator with a heteroscedastic-autocorrelation consistent variance estimator. So he developed another test to short range dependence[2]. However, Teverovsky Taqqu and Willinger (1995) showed that Lo's improvement might nevertheless eliminate a long term memory which is actually present. For that reason we will stick to the "H" statistics and not to the Lo's development. We turn now to other methods like ARFIMA model.

2.2 The ARFIMA specification

A series X_t follows an $ARFIMA(p,d,q)$ process if:

$$(1-L)^d\Phi(L)X_t = \Theta(L)u_t \tag{4}$$

where Φ and Θ are lag operators of order p and q respectively, d is a fractional order of integration, u_t a white noise.

[2]The modified RS statistic is:

$$\frac{1}{\hat{\omega}(q)}\left(\underset{1<k<T}{Max}\sum_{j=1}^{k}\left(X_j - \bar{X}_T\right) - \underset{1<k<T}{Min}\sum_{j=1}^{k}\left(X_j - \bar{X}_T\right)\right)$$

with $\hat{\omega}(q)$: non parametric estimator of the weighted sum of autocovariances at lag q; if $q=0$ Lo's statistic is similar to Hurst's.

Once the stationnarity has been checked, in other words the integer order has been removed, it is compulsory to estimate "d". The parameter is then included between [-0.5,0.5], the process is then stationary and invertible. As abovementioned, when "d" lies between 0 and 0.5, the process has a long memory and can be defined either in time space or in frequency (or spectral) domain. In time space, such processes are featured by an autocorrelation function (ACF) which is hyperbolically decreasing along the lags whereas short term memory processes show an exponential decreasing slope. In other words, ACF decays slowly to 0 at a polynomial rate as lags increase. In the spectral domain, the density is not limited to a finite value at the frequency 0. The short memory part of the process lies in the ARMA representation.

We consider three procedures in order to estimate the ARFIMA model and more precisely the parameter "d". In the two first ones, it is common and accurate to use a two step methodology which consists in firstly estimating "d" parameter and secondly the coefficients of the ARMA process. The simplest way is to obtain "d" thanks to Hurst coefficient since they are linked with the relation[3] $H = d + \frac{1}{2}$. The second solution is to use the Geweke-Porter-Hudak (GPH) (1983) technique.

Geweke Porter Hudak (1983) (GPH)

GPH used the spectral behavior properties at the low frequencies of the time series to estimate "d". It is justified by Ohlshen (1967) who proved that the spectral density has better sample properties than the autocorrelation function when the spectral coordinates at frequencies $2\pi_j/T$ $(j = 1, \ldots, T/2)$ are asymptotically non correlated. Facing the behavior of the spectral den-

[3]The autocorrelation functions have the same power in the hyperbolic decline function:

$$\gamma(k) \approx \frac{\Gamma(1-2d)}{\Gamma(1-d)\Gamma(d)} k^{2d-1} \approx H(2H-1)k^{2H-2}. \text{ When } \; k \to \infty$$

we have asymptotically the important result: d= H - $\frac{1}{2}$

sity function around 0, GPH dealt with low frequencies and came to a linear univariate regression of the log periodogram related to the log frequency (for more details, see Mignon and Lardic (2002) or Robinson (2003)). Considering the spectral density function:

$$f(\lambda) = |1 - e^{i\lambda}|^{-2d} f_u(\lambda) \quad \text{with } f_u(\lambda) = \frac{\sigma^2 |\Theta(e^{i\lambda})|^{-2d}}{2\pi |\Phi(e^{i\lambda})|^{-2d}} \tag{5}$$

Where $f_u(\lambda)$ is the spectral density function of the ARMA (p, q), $u_t = \bigtriangledown^d X_t$. Immediately, we can obtain the periodogram of the series:

$$I(\lambda) = \frac{1}{2\pi T} |\sum_{j=1} e^{ij\lambda}|^2 \tag{6}$$

Taking the logarithm of (6) and replacing λ by its Fourier frequency $\lambda_j = \frac{2\pi_j}{T}$ with $j = 1, ..., O(T^{0.8})$ (Hurvitch et al., 1988), adding the natural logarithm of the periodogram $\ln(I(\lambda_j))$ on each side of the equation developed around frequency 0, we get:

$$\ln(I(\lambda_j)) = \ln(f_u(0)) - d \ln\left(|1 - e^{-i\lambda_j}|^2\right) + \frac{I(\lambda_j)}{f(\lambda_j)} + \ln\left(\frac{f_u(\lambda_j)}{f_u(0)}\right) \tag{7}$$

If λ is close to zero the last term is negligeable compared to the others terms (right and left) and so we can rewrite equation 7 under the form of a simple equation:

$$\ln(I(\lambda_j)) = \ln(f_u(0)) - d \ln\left(|1 - e^{-i\lambda}|^2\right) + \zeta_j \tag{8}$$

where $f_u(0)$ is an intercept and "d" is the regression coefficient estimated by OLS analysis. GPH have shown that, when $d \in [-\frac{1}{2}, \frac{1}{2}]$, the law of the estimator $\hat{d}$ of d tends to a normal distribution when T goes to infinity.

$$\hat{d} \rightsquigarrow N\left(d, \pi^2 \left[6 \sum_{j=1}^{m} Z_j - \hat{Z}\right]^2\right) \quad \text{with} \quad Z_j = \ln\left(|1 - e^{-i\lambda_j}|^2\right) \tag{9}$$

With the help of "d", the interesting parameter for us, we can obtain the parameters (p, q) of the ARMA part of the model.

Whittle 's approach

The last solution is to estimate the (p, d, q) triplet in one step. One popular approach is a method due to Fox and Taqqu (1986), based on gaussian maximum likelihood approximated function given by Whittle (1956). This type of technique requests adapted initial conditions which ensure the convergence of the algorithm. In the result part, we will use the GPH findings as starting values. Their estimators are convergent and asymptotically normal. We give her only a brief overview[4].

The Gaussian likelihood of the series X expressed in (4) equation maximizes on β and σ:

$$L(\beta, \sigma) = (2\pi\sigma^2)^{-T/2}(\tau_0, \cdots, \tau_{t-1})^{1/2} e^{\frac{1}{2\sigma^2}\sum_{j=1}^{T}\frac{(X_j-\hat{X}_j)^2}{\tau_{j-1}}} \tag{10}$$

where $\beta = (d, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_p)'$ with $-\frac{1}{2} < d < \frac{1}{2}$, $\hat{X}_j$ for $j = 1, \ldots, T$ forecast series with an increment of 1 and $\tau_{t-1} = \sigma^{-2}E(X_j - \hat{X}_j)^2$. Brockwell and Davis (1991) proved that this maximisation leads to:

$$\hat{\sigma}^2 = T^{-1}S(\hat{\beta}) \quad \text{where} \quad S(\hat{\beta}) = \sum_{j=1}^{T} \frac{(X_j - \hat{X}_j)^2}{\tau_{t-1}} \tag{11}$$

where $\hat{\beta}$ is the value of β that minimizes the following expression:

$$I_a(\beta) = \ln \frac{1}{T} \sum_j \frac{I_T(\omega_j)}{g(\omega_j, \beta)} \tag{12}$$

Fox et Taqqu (1986) proved that the estimator $\tilde{\beta}$ minimizing $I_a(\beta)$ is convergent if $d > 0$, it has the same asymptotic distribution as $\hat{\beta}$. The variance of the white noise is estimated:

$$\tilde{\sigma}^2 = \frac{1}{T} \sum_j \frac{I_T(\omega_j)}{g(\omega_j, \tilde{\beta})} \tag{13}$$

where I_T correspond to the periodogram, $\sigma^2 g(\omega_j, \tilde{\beta})/2\pi$ the spectral density of the model – the sum is on all Fourier frequencies except 0.

[4]Readers are referred to Fox and Taqqu (1986) or Mignon Lardic (2002)

2.3 The wavelet methodology

Wavelet analysis belongs to the data signal processing and is adapted to analysis of changes in time series. It is used for the detection of cycles or recurrent dynamics in large data samples. In other words, wavelet transforms can be understood as a more flexible form of a Fourier transform where the studied variable is transformed not into a frequency domain but into a time-scale wavelet domain. It allows the observation of a time series at a full range of different scales whilst retaining the time dimension of the original data. The (sinusoidal) functions of Fourier are replaced by wavelet basis functions.

We will restrain here in discrete wavelet transform, which assumes a signal consisting of points sampled at regularly time length - in our case, daily observations. Mathematically, the transform into wavelet consists in computing the convolution between a time series and two filters which have the following theoretical properties:

$$\int \phi(t)dt = 1 \quad \text{and} \quad \int \psi(t)dt = 0 \tag{14}$$

where the father wavelet $\phi(t)$ captures low frequency components which represent the trend (long time intervals) and the mother wavelet $\psi(t)$ is used to describe high frequency and captures small irregularities and noises (short time intervals). This particularity will allow us to estimate the Hurst coefficient and so to discriminate between long term memory component and noisy residuals. Both are necessary to extract the entire signal.

There are a large number of filters and consequently different wavelets, most of them having no analytical formula and so being based on simulations. We use here a specific type of dyadic wavelets, namely the Daubechies wavelet ("Daublets", due to Ingrid Daubechies, 1992). This choice is explained by the advantage to allow a large span of coefficients which allows for a more precise multiresolution and a better fit to the data. The Daubelets have also the interesting properties to be asymmetric (which is precisely empirically a stylized fact feature of financial returns) and also to be smooth. The mother wavelet is defined by the following equation:

$$\psi(x) = \sqrt{2}\sum_j h_j \phi(2x - j) \tag{15}$$

where $j = 1, \ldots, \frac{log(N)}{log(2)}$ is called an octave and N the data length. No information is lost if we sample continuous wavelet coefficients at a sparse set of points (the dyadic grid). $\phi(x)$ is the father wavelet or the dilation equation:

$$\phi(x) = \sqrt{2}\sum_j l_j \phi(2x - j) \tag{16}$$

For an octave j, the wavelet coefficients l_j and h_j are respectively the low-pass and high-pass filter coefficients defined as:

$$l_j = \frac{1}{\sqrt{2}} \int \phi(t)\phi(2t - j)dt \tag{17}$$

$$h_j = \frac{1}{\sqrt{2}} \int \psi(t)\phi(2t - j)dt \tag{18}$$

$$\tag{19}$$

The number n_j of coefficients available at octave "j" halves with each increase of "j". Non-zero wavelet coefficients at high scales characterize the noise inherent in the data. For illustration, if a stock price does not change during a week, the wavelets coefficients applied at the daily scale will be zero during that week. Based on those coefficients, it is possible to obatin the Hurst exponent.

More precisely, it is calculated from the wavelet spectral density plot (*i.e* scalogram) which is generated from the wavelet power spectrum. For octave j, the power spectrum is obtained here thanks the average of the squares of the wavelet coefficients minus the bias correction proposed by Abry and Veitch (1998). Then, the Hurst exponent[5] is deduced by the slope coefficient α of the linear regression on the wavelet power spectrum for all octaves.

[5] The Hurst exponent is : $H = \frac{\alpha+1}{2}$.

Abry and Veitch (1998) proposed to build the regression on a subsample of the octaves $[j_1, j_2]$ to eliminate the smallest high frequency (noises). The upper bound j_2 is defined by $\frac{log(N)}{log(2)}$ and the moment of the Daublets. The lower bound j_1 depends on the goodness-of-fit test (Abry and Veitch, 1998). The choice for j_1, *i.e* the cutoff between short-range dependence and the long-range dependence is similar to the bandwidth selection problem in semi-parametric statistics. The goodness-of-fit function is a generalized Pearson statistic:

$$Q = \sum_{j=j_1}^{j_2} \frac{(y_j - \hat{\alpha} w_j - \hat{a})^2}{\sigma_j^2} \tag{20}$$

where $\hat{\alpha}$ is the slope estimated coefficient defined above, $\hat{a}$ is the unbiaised estimator of the intercept of the linear regression, σ_j^2 is the variance of the series y_j (see Teyssière and Abry, 2005). The regression range is defined by the lowest octave j_1 and the highest octave j_2. This statistics follows $\chi^2(j_2 - j_1 - 1)$. The logics to obtain the optimal j_1 is to select the minimum octave k for which the p-value of the test differs from zero. The test checks if the following octave has a p-value ten times superior to the previous one; if it is not the case, the "optimal" j_1 is the current value.

3 Data and Results

We worked with five daily stock price indices (closely value) transformed into continuous returns, namely SP500, DJIA, DAX and Nikkei (source: Bloomberg). Two samples have been selected: one between 1959 and 2010 and the other 1970 and 2010. For the Nikkei, there were no daily data available before 1970. Choosing two samples avoids the criticism of the possible unstability of results due to the range sample. This data frequency is explained by the necessity of a large number of observations. We did not take into account intraday observations because we wanted to exclude very high frequency data in order not to pollute long memory dynamics.

We reported the descriptive statistics in Table 1. All returns show a large excess of kurtosis meaning that the distribution tails are larger than normal, a conclusion strongly reinforced by the Jarque Bera test which proves there are non-normality beyond any doubt. The results for skewness are more mitigated: DJ and SP show negative dissymmetry (around -1.20) whereas DAX, NIKKEI are closer to normal Gaussian distribution. Those findings are in accordance with many previous studies on stock market indices. This does not disqualify the market efficiency hypothesis but makes it less simple and tractable: one has to distinguish among the different markets .

The first important step is to test for the presence of unit roots in returns[6]. To this aim, we apply alternatively three standard tests: the Augmented Dickey Fuller (ADF), the Phillips Perron (PP) and the Kwiatkowski et al. (1992) (KPSS) test. Contrarily to the two first tests, the KPSS test has the advantage to be based on the null hypothesis of series stationnarity. In performing an ADF test, we will face two practical issues. We have to decide to include a constant, a constant and a linear time trend or neither in the test regression. Taking into account irrelevant regressors in the regression will reduce the power of the test: to reject the null hypothesis of a unit root. The optimal specification include an intercept, except for the Nikkei where there is a trend. The choice of alternative specification does not modify the conclusion.

Then, we will have to specify the number of lagged difference terms of the dependent variable to be added to the test regression. The usual advice is to include a number of lags sufficient to remove serial correlation in the residuals. In respect of it, we retain the number of lags that minimize information criteria. To specify the PP or KPSS test, we have to select the regression form to test as in the ADF test. We must equally choose the kernel

[6]With classical test, we found that the stock price indices are integrated of order (1).

and the bandwidth parameter needed for estimating the residual spectrum at frequency zero. The usual solution is to use Bartlett kernel and Newey West (1994) data-based automatic bandwidth parameter method. In Table 1, we have reported the results of the three unit root tests. The conclusion is that all returns are stationary at the 1% significance level.

The second step is to apply the different abovementioned methods in the search of a possible long memory process. We will successively present the structure of dependence obtained with the Hurst coefficient, the ARFIMA model and the wavelet approach.

In Table 2, we see that the Hurst coefficients obtained with R/S statistics are around 0.5 for all series under review. The implied consequence is that there seems to be a clear absence of memory in the financial markets. But we do not have a precise distribution law of the statistics and neither a confidence interval. The choice of the estimation period has almost no impact on the results. For example, the H value for DJ amounts to 0.552 against 0.542 when the period is reduced to 1970 instead of 1959. As for the dependence structure, because the exponent is near 0.5, the results are more in favor of the market efficiency.

The second measure is obtained with ARFIMA models which can be clearly separated into two categories, the first one in one step –the GPH procedure, the second one in two steps –the Whittle estimation. In order to make comparisons easier in Table 2, we directly reported the Hurst coefficient (instead of the "d" parameter) and also their confidence intervals at 95% level. In the GPH procedure, we get primarily the parameter "d" and thereafter the pair (p,q) is identified[7]. However, in the Whittle approach, the couple (p, q) is important since all parameter are calculated in one step

[7] We report it the couple, but by construction, it has no impact on the "d" parameter estimation.

simultaneously with “d”. The results are coherent with the previous conclusions: the 0.5 value is included in the interval confidence, which is in favor of the market efficiency theory. For illustration in the period 1959-2010, The "H" coefficient with GPH methodology is 0.493 with an confidence interval of [0.464,0.521] and 0.496 with [0.467, 0.564] for SP and DJ respectively.

The results of Whittle are very close to the previous one. This strengthens our conclusion, it proove robustness since they are not affected by the ARFIMA methodology and the sample time period. Moreover, they are not sensitive to a possible bad specification of the (p, q) couple. We indeed made different simulations with different ARMA parameters and again the results are remarkably stable. But in the two methodology, since the ARMA parameters being significant[8], taking account a short term dynamics is necessary.

Even if returns are uncorrelated or weakly correlated, it does not imply that absolute returns or square returns might not show a positive, significant, and slowly decaying autocorrelation function, for a time lag ranging from a few minutes to several weeks. In other words, uncorrelated returns can be observed with volatility clustering described with GARCH or FIGARCH family models and stochastic volatility models. This kind of heteroskedasticity does not imply any bias in the integration parameter.

The wavelet model estimations request three preliminary steps. The first one is to select the regularity parameter of the Daubechies mother function. A high regularity parameter implies a clearer signal but the very high frequency dynamics are eliminated, leading to a reduction of precision. Then, it is necessary to find an effective balance between the two constraints. We propose to select the regularity parameter according to the goodness of fit test proposed by Abry and Veitch (1998). As previously explained, this test shows the quality of the log linear regression between octaves and average

[8] They are not presented here but available for the interested readers.

square wavelet coefficient (a high p-value is according to a good linear fit). So the chosen regularity parameter is the one that gives the highest p-value. To comfort our results, we present in table 3 our conclusions regarding the depth of the memory when using other regularity values.

According to the goodness-of-fit procedure, the second step is to select the subsample of the octaves $[j_1, j_2]$ on which the Hurst exponent will be estimated. The figure 1 plots this selection for the regularity optimal parameter (Table 3). In x-axis, we have the different possible values of j_1 octave, and in y-axis the logarithm (base 10) of the goodness-of-fit (Q). For illustration, we clearly see that the goodness-of-fit value is out of the critical bounds for $j_1 = 2$ in the case of SP 500 and Dow Jones.

The third step is to implement the regressions in order to obtain the H exponent, in this objective, figure 2 displays the scalogram. The linear regression on the wavelet power spectrum for all octaves fits good and then we report the numerical results in the table 3. We have summed up those information: the Daubechies regularity parameter, the Hurst coefficient with its confidence interval (at 95%). The three last columns are to strengthen our conclusion in case of misspecification of the regularity parameter. Actually, we report the number of cases which respectively prove a short memory process, the absence of a memory process and the long memory process for all possible regularity parameters of the Daubechies wavelet (1 to 10).

In table 3, all p-values of goodness-of-fit test associated with the optimal range of octaves and the regularity parameter are largely superior to 10%. So, we cannot reject[9] the null hypothesis that the linear regression fits well with the power spectral density and consequently the Hurst exponent estimation is of relatively good quality. The p-values span from 0.33 to 0.97 and

[9]Not suitable regularity parameter or octave range could imply a reject of the null hypothesis.

the strongest refers to the SP 500. All values of Hurst exponents are close to 0.5, the highest is for the DAX (1970-2010) but the confidence interval [0.50, 0.61] includes 0.5 which infers that this process has no memory. If we continue with all confidence intervals, we can point out two cases: those which include 0.5 (SP500: 1959-2010, DJ: 1970-2010 and DAX: 1959-2010 and 1970-2010) and those inferior to 0.5 which exhibit short memory process (SP500: 1970-2010, DJ:1959-2010, Nikkei: 1970-2010). We focus now on the last three columns with regularity parameters (from 1 to 10). The strongest result is that we never find long memory dependence. Another conclusion is that when we change the regularity parameter, we alternate evidence of no memory and short memory processes; in this latter case however, the depth of the short memory is low (the confidence interval is very close to 0.5). These results are coherent with the ARFIMA representation where the coefficients for autoregressive and moving average dynamics are significant and the "d" parameters are close to 0 (H=d+1/2).

In conclusion, we noticed that the three different methods led to comparable results of no or very short memory dependence. From a financial point of view, this approach can be viewed as a weak efficiency tests: a process with structure dependence would invalidate the efficiency market hypothesis because it would allow to take advantage of the knowledge of past prices or returns to make abnormal profits. Our results support the weak efficiency theory according to Fama's definition for a sufficient long period of time but there is a possible forecast in a very short term. In other words, the idea of very short memory could imply a possible forecast for day to day arbitrager. Besides, the notion of memory is strongly bounded with the notion of liquidity. Indeed, when there is a long term memory, this is an evidence of illiquid market, so when no memory are found, it is in favor of a sufficient liquidity level in the market.

Conclusion

In This papper, we have investigated if stock market index returns present any type of memory. We have used three different methodologies to obtain Hurst exponent, starting from "R/S" approach, continuing with ARFIMA models and ending with wavelet analysis.

Our findings are coherent according to the various methods namely never long term memory but on the contrary absence or small antipersistence memory. Those results are in accordance of the weak efficiency financial theory, restraining successful forecasts and arbitrage opportunities.

Possible extension is to work with individual stocks instead of indices because one can suspect that some equities might be subject to semi-strong inefficiency. But the largest field of application could be to deal with volatility as Jiang and Tian (2010). To this respect we could use the same methodolgies applied to absolute values of return or also to square returns.

Bibliography

Abry, P. and Veitch, D., (1998), "Wavelet analysis of long-range-dependent traffic", IEEE Trans. Inform. Theory, 44, 2-15.

Baillie, R. T. (1996), "Long Memory Processes and Fractional Integration in Econometrics," Journal of Econometrics, 73, 5-59.

Bardet, J.M., Lang, G., Moulines, E. and Soulier, P., (2000), "Wavelet estimator of long range-dependant processes", Statistic Inference Stochastic Processes 3 85–99.

Barkoulas, J. T. and Baum, C. F. (1996), Long-term dependence in stock returns, Economics Letters 53(3), 253–259.

Barkoulas, J. T., Baum, C. F. and Travlos, N. (2000), "Long memory in the Greek stock market", Applied Financial Economics 10(2), 177–184.

Cajueiro, D. O. and Tabak, B. M. (2004), "Evidence of long range dependence in Asian equity markets: The role of liquidity and market restrictions", Physica A 342(3–4), 656–664.

Cajueiro, D. O. and Tabak, B. M. (2005), "Possible causes of long-range dependence in the Brazilian stock market", Physica A 345(3–4), 635–645.

Chen, S.-H. (2000), Lecture 7: Rescale range analysis and the Hurst exponent, Financial Economics (I), Department of Economics, National Chengchi University.

Chow, K. V., Pan, M.-S. and Sakano, R. (1996), "On the long-term or short-term dependence in stock prices: Evidence from international stock markets", Review of Quantitative Finance and Accounting 6(2), 181–194.

Cochrane J.H. (1988), "How big is the random walk in GNP", Journal Political Economy 86, 893-920.

Crowley, P., (2005), "An intuitive guide to wavelets for economists", Research Discussion Papers, Bank of Finland.

Diebold, F.X., Husted, S. and Rush M. (1991), "Real exchange rates under gold standard" Journal of Political Economics, 99 , 1252-1271.

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

Fama, E.F., (1965), "Random Walk in Stock Market prices", Financial Analysts Journal, 21(5), 55–59.

Fama, E.F., (1970), "Efficient Capital Markets : a review of theory and empirical work", Journal of Finance, 25, 383-417.

Fama E.F., "Efficient capital markets", Journal of Finance, 5.

Fox, R. and Taqu, M.S., (1986),"Large-sample properties for parameter estimates for strongly dependent stationary gaussian time series", Annals of statistics, 14, 517-532.

Geweke J. and Porter-Hudak, S., (1983), "The Estimation and Application of Long Memory Time Series Models", Journal of Time Series Analysis, 4, 221-238.

Goetzmann, W. N. (1993), "Patterns in three centuries of stock market prices", The Journal of Business 66(2), 249–270.

Granger, C. W. J. and Joyeux, R., (1980), "An introduction to long-memory time series models and fractional differencing". Journal of Time Series Analysis 1, 15–30.

Grau-Carles, P. (2005), "Tests of long memory: A bootstrap approach", Computational Economics 25(1–2), 103–113. Henry, O´ . T. (2002), "Long memory in stock returns: Some international evidence", Applied Financial Economics 12(10), 725–729.

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

Hosking, J.R.M., (1981), "Fractional Differencing", Biometrika, 68, 165-176.

Hosking, J. (1984), "Modeling persistence in hydrological time series using fractional differencing", Water Resources Research, 20, 1898-1908.

Huang, B.-N. and Yang, C. W. (1999), "An examination of long-term memory using the intraday stock returns", Technical Report 99–03, Clarion University of Pennsylvania, Clarion.

Jiang, G.J., and Tian, Y.S. (2010), "Forecasting Volatility Using Long Memory and Comovements: An Application to Option Valuation under SFAS 123R", Journal of Financial and Quantitative Analysis, 45(02), 503-533.

Lardic, S. and Mignon, V. (1996), "Les tests de mémoire longue appartiennent-ils au "camp du démon" ?", Revue économique 47(3), 531-540.
Lardic S., Mignon V. (2002), "Econometrie des séries temporelles macroéconomiques et financières", Economica. Limam, I. (2003), "Is long memory a property of thin stock markets? International evidence using arab countries", Review of Middle East Economics and Finance 1(3), 251–266.
Lo, A. W. (1991), "Long-term memory in stock market prices", Econometrica 59(5), 1279–1313.
Lobato, I. N. and Savin, N. E. (1998), "Real and spurious long-memory properties of stock-market data", Journal of Business & Economic Statistics 16(3), 261–268.
Lux, T. (1996), "Long-term stochastic dependence in financial prices: Evidence from the German stock market", Applied Economics Letters 3(11), 701–706.
Mills, T. C. (1993), "Is there long-term memory in UK stock returns?", Applied Financial Economics 3(4), 303–306. Nawrocki, D. (1995), "R/S analysis and long term dependence in stock market indices", Managerial Finance 21(7), 78–91.
Nath, G. C. (2001), "Long memory and Indian stock market – an empirical evidence", National Stock Exchange of India Ltd., Mumbai.
Oh, G., Um, C.-J. and Kim, S. (2006), Long-term memory and volatility clustering in daily and highfrequency price changes, arXiv:physics/0601174.
Olshen, R. (1967), "Asymptotic properties of the periodogram of a discrete stationary process", Journal of Applied Probababilty, 5, 508-528.
Panas, E. (2001), Estimating fractal dimension using stable distributions and exploring long memory through ARFIMA models in the Athens Stock Exchange, Applied Financial Economics 11(4), 395–402.
Sadique, S. and Silvapulle, P. (2001), "Long-term memory in stock market returns: International evidence", International Journal of Finance & Economics 6(1), 59–67.
Robinson, P.M. (2003) "Time Series with Long Memory", Oxford University

Press.
Serletis, A. and Rosenberg, A.A. (2009), "Mean reversion in the US stock market", Chaos, Solitons & Fractals 40(4), 2007–2015.
Taqqu M. S., Teverovsky V. and Willinger W., (1995), "Estimators for Long-range Dependence : An Empirical Study". Fractals, 3, 785-798.
Teyssiere, G. and ABRY, P., (2005), "Wavelet Analysis of Nonlinear Long-Range Dependent Processes. pplications to Financial Time Series", in Long-Memory in Economics, G. Teyssière and A. Kirman editors, 173-238, Springer Verlag.
Teverovsky V., Taqqu M.S., and Willinger, W., (1999), "A critical look at Lo's modified R/S statistics", Journal of Statistical Planning and Inference, 80, 211-228.
Tolvi, J. (2003), "Long memory in a small stock market", Economics Bulletin 7(3), 1–13.
Whittle (1956), "On the variation of yeld with plot size", Biometrika, 43, 337-343.

Tables and Figures

Table 1: Descriptive statistics and unit-root tests results

	DAX		DJ		SP 500		Nikkei
Period	1959	1970	1959	1970	1959	1970	1970
Mean	0.00023	0.00023	0.00023	0.00026	0.00023	0.00025	0.00015
Median	0.00039	0.00053	0.00031	0.00036	0.00031	0.00042	0.00049
Maximum	0.12	0.11	0.11	0.11	0.11	0.11	0.13
Minimum	-0.14	-0.14	-0.26	-0.26	-0.26	-0.23	-0.16
Std. Dev.	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Skewness	-0.13	-0.24	-1.28	-1.31	-1.28	-1.04	-0.37
Kurtosis	11.13	10.49	42.03	39.79	42.03	29.78	13.15
Jarque-Bera	35661***	24311***	825285***	590066***	825285***	312510***	43857***
ADF	-82.55	-74.16	-82.9	-75.1	-82.97	-74.91	-74.27
PP	-110.79	-101.8	-110.79	-100.53	-110.43	-100.29	-100.99
KPSS	0.10***	0.13***	0.16***	0.13***	0.10***	0.16***	0.06***

Notes: (*): significant at the 1% level; (**): significant at the 5% level and (***): significant at the 1% level. For all the unit root test (ADF, PP and KPSS), the optimal specification include an intercept, except for the Nikkei where there is a trend.

Table 2: ”R/S” Statistics and ARFIMA results

		R/S t	GPH	Whittle
	Order (p,q)	-	(0,2)	(0,2)
SP (1959-2010)	Hurst exponent	0.545	0.493	0.492
	Confidence Interval	-	[0.464 , 0.521]	[0.521 , 0.463]
	Order (p,q)	-	(2,0)	(2,0)
DJ (1959-2010)	Hurst exponent	0.542	0.496	0.496
	Confidence Interval	-	[0.467 , 0.524]	[0.524 , 0.467]
DAX (1959-2010)	Order (p,q)	-	(3,2)	(3,2)
	Hurst exponent	0.530	0.491	0.491
	Confidence Interval	-	[0.462 , 0.519]	[0.519 , 0.463]
SP (1970-2010)	Order (p,q)	-	(0,2)	(0,2)
	Hurst exponent	0.553	0.481	0.48
	Confidence Interval	-	[0.45 , 0.512]	[0.512 , 0.449]
DJ (1970-2010)	Order (p,q)	-	(0,2)	(0,2)
	Hurst exponent	0.552	0.493	0.492
	Confidence Interval	-	[0.464 , 0.521]	[0.521 , 0.464]
DAX (1970-2010)	Order (p,q)	-	(2,2)	(2,2)
	Hurst exponent	0.553	0.499	0.499
	Confidence Interval	-	[0.468 , 0.53]	[0.53 , 0.468]
Nikkei (1970-2010)	Order (p,q)	-	(2,0)	(2,0)
	Hurst exponent	0.545	0.518	0.516
	Confidence Interval	-	[0.486 , 0.549]	[0.549 , 0.485]

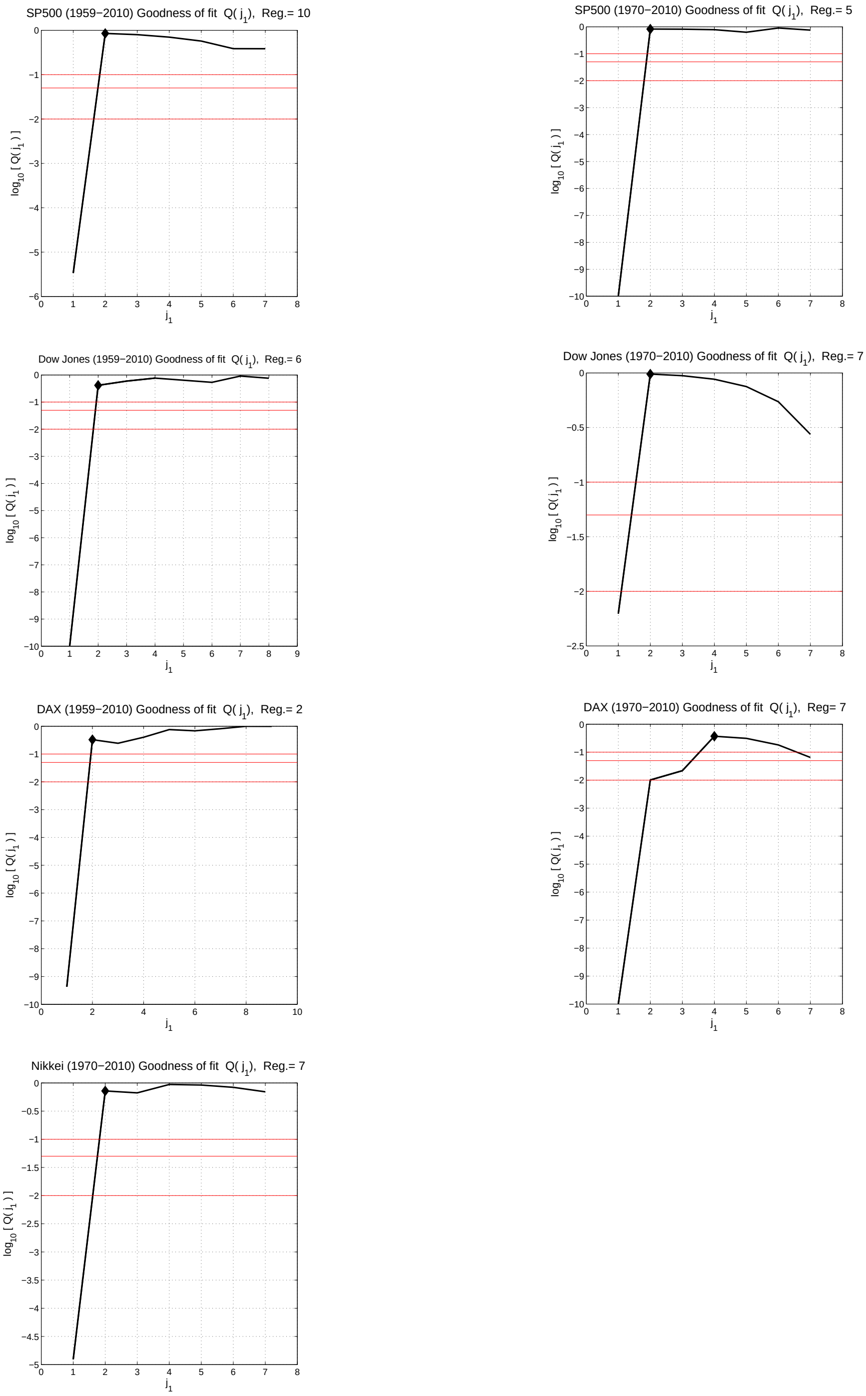
Notes: In the ARFIMA proccedure, we estimate the ”d” parameter of integration, but in this table we replace it directly by the Hurst exponent to make the comparison easier (H=d+1/2). The confidence interval is computed at 95%.

Table 3: Wavelet results

	Regularity	Goodness-of(fit	Octaves	Hurst Coef.	Conf. Interval	Memory class (reg.=1:10)		
						No	Short	Long
SP500 (1959-2010)	10	0.85	[2, 9]	0.475	[0.45; 0.50]	6	4	0
Dow Jones (1959-2010)	6	0.41	[2, 10]	0.467	[0.45; 0.49]	5	5	0
DAX (1959-2010)	2	0.33	[2, 11]	0.502	[0.48; 0.52]	7	3	0
SP500 (1970-2010)	5	0.82	[2,9]	0.462	[0.44; 0.48]	2	8	0
Dow Jones (1970-2010)	7	0.97	[2, 9]	0.475	[0.45; 0.50]	7	3	0
DAX (1970-2010)	7	0.37	[4, 9]	0.557	[0.50; 0.61]	8	2	0
Nikkei (1970-2010)	7	0.72	[2, 9]	0.463	[0.44; 0.49]	5	5	0

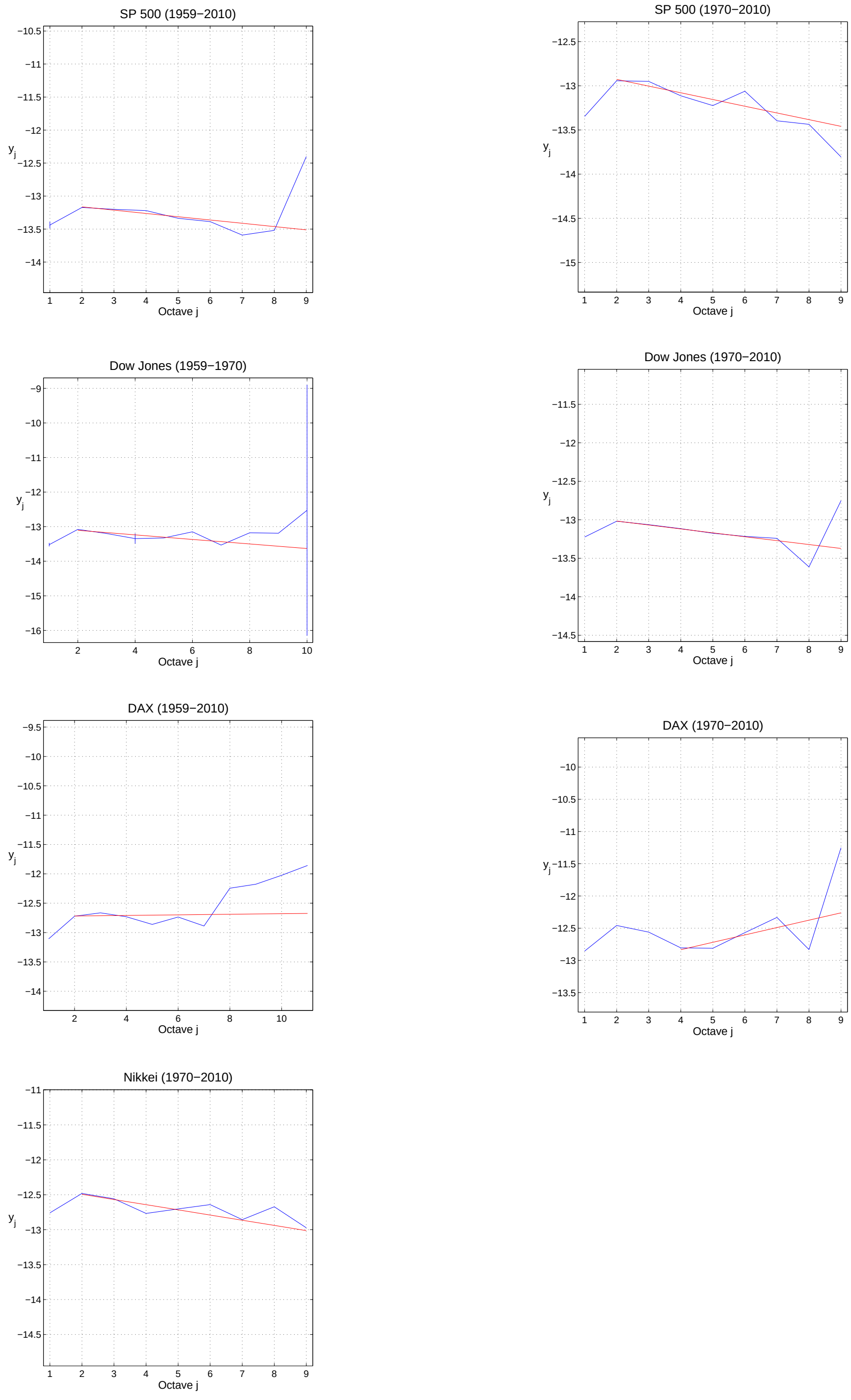
Notes: Regularity (reg.) parameter indicates the Daubechies used; the goodness of fit is the p-value of the test where the null hypothesis is the goodness of linear fit. Octave is the range of the octaves $[j_1, j_2]$ on which the Hurst exponent will be estimated. The confidence interval is done at 95 %. Class memory represent the number of processes where no memory or short memory or long memory is found for all other regularity values.

Figure 1: Goodness-of-fit tests



Notes: j_1 represents the octave of the lower bound. The dotted lines represent the critical bounds. Reg. is the regularity parameter.

Figure 2: Scalogram applied to Hurst exponent



Notes: This figure is done for the "optimal" regularity parameter. y_j is the logarithm (base 2) of the average of the squared wavelet coefficients minus the bias correction (Abry and Veitch, 1998) for all octaves. Hurst Exponent is deduced of the slope α of the linear regression (red line) $H = \frac{\alpha+1}{2}$.