PORTUGUESE STOCK MARKET: A LONG-MEMORY PROCESS?

Sameer Rege[1], Samuel Gil Martín[2]

[1]Universidade dos Açores, CEEAplA & Departmento de Economia e Gestão, Rua da Mãe de Deus, 58 Ponta Delgada,
9501 801 São Miguel, Açores, Portugal
[2]Universidad Autónoma de San Luis Potosí Avenida Pintores CP 78213 San luis Potosí, México
E-mails: [1]sameer@uac.pt; [2]samuel.gilmartin@gmail.com

Abstract. This paper gives a basic overview of the various attempts at modelling stochastic processes for stock markets with a specific application to the Portuguese stock market data. Long-memory dependence in the stock prices would completely alter the data generation process and econometric models not considering the long-range dependence would exhibit poor forecasting abilities. The Hurst exponent is used to identify the presence of long-memory or fractal behaviour of the data generation process for the daily returns to ascertain if the process follows a fractional brownian motion. Detrended fluctuation analysis (DFA) using linear and quadratic trends and the Geweke Porter-Hudak methods are applied to detect the presence of long-memory or persistence. We find that the daily returns exhibit a small amount of long memory and that the quadratic trend used in the DFA overestimates the value of the Hurst exponent. These findings are corroborated by the use of the Geweke Porter-Hudak method wherein the Hurst exponent is close to the DFA using the linear trend.

Keywords: geometric brownian motion, Hurst Exponent, Long-Memory, Detrended Fluctuation Analysis, Geweke Porter-Hudak method, stable distributions.

PORTUGALIJOS AKCIJŲ RINKA: AR INERTIŠKAS KAINŲ KITIMAS?

Sameer Rege[1], Samuel Gil Martín[2]

[1]Azorų universitetas, CEEAplA & Ekonomikos e Gestão katedra, Rua da Mãe de Deus, 58 Ponta Delgada,
9501 801 São Miguel, Açores, Portugalija
[2]Autonominis de San Luis Potosiuniversitetas, Avenida Pintores, CP 78213 San luis Potosí, Meksika
El. paštas: [1]sameer@uac.pt; [2]samuel.gilmartin@gmail.com

Santrauka. Straipsnyje aprašomi įvairūs vertybinių popierių rinkų stochastinio procesų modeliavimo bandymai. Autoriai atlieka skaičiavimus naudodami Portugalijos vertybinių popierių rinkos duomenis. Identifikuotas vertybinių popierių kainų kitimo inertiškumas. Šis inertiškumas iš esmės keičia duomenų prognozavimo procesą. Ekonominiai modeliai, neatsižvelgiantys į inertiškumą ilgalaikėje perspektyvoje, tampa netinkami prognozuoti. Hursto eksponentinis metodas taikomas inertiškumo reiškiniui arba jo daliniam pasireiškimui identifikuoti. Tikrinama, ar tos dienos duomenys patvirtina hipotezę apie vertybinių kainų Bruno kitimo inertiškumą. Gevek Porter-Hudako metodai taip pat taikomi inertiškumui identifikuoti. Autoriai nustato, remdamiesi išvardytais metodais ir skaičiuoti imdami kiekvienos dienos vertybinių kainų duomenis, kad keliama hipotezė pasitvirtino.

Reikšminiai žodžiai: geometrinis Bruno judėjimas, Hursto eksponentė, kainų kitimo inertiškumas, išlyginta svyravimų analizė, Geweke Porter-Hudak metodas, stabilus paskirstymas.

1. Introduction

The attempt to model the data generation process for financial data dates back to Bachelier (1900) wherein he attempted to model the French government bond and its futures. For further details on Bachelier's work refer Voit (2005). Bachelier (1900) and Black Scholes (1973) are consistent with the Efficient Market Hypothesis, first formulated by Samuelson (1965) and Fama (1970). The efficient market is an idealized, complex system, wherein the essential information about a traded asset is instantaneously incorporated in its price. As a foremost implication, the serial correlation of the rate of return is zero for any short-time scale, so that the return time series are random walks.

Most of the empirical works carried during the 1960s support the random-walk hypothesis. Over the last 40 years, however, financial markets have witnessed significant changes, and data showed important discrepancies between the Bachelier model and real markets. After the collapse of the Bretton Woods system, the value of currencies, together with other financial prices, commodity prices, including oil, and land prices were displaying fluctuations of an order of magnitude never experienced before. By this time, the electronic revolution started to adapt to financial markets and capital facilitating global capital movements all around the globe. The volume of financial transactions has since overwhelmed current account transactions by a factor of several hundreds to 1.

There are some striking features that do not fit with the Geometric Brownian motion Bouchaud (2002). Empirical evidence from stock markets around the world shows that the returns

$$r(t) = \ln\left(\frac{P_t}{P_{t-1}}\right) = \frac{P_t - P_{t-1}}{P_{t-1}} \qquad (1)$$

do not follow a Gaussian distribution but are fat tailed and skewed. Also the volatility of the returns $r(t)^2$ shows heteroskedasticity with periods of high volatility and low volatility. Also financial series often show long memory wherein hyperbolic decay rates of the autocorrelation function of the log price, at odds with the efficient market hypothesis, were first observed by Greene and Fielitz (1977) and Taylor (1986). The major implication is the possibility of earning speculative profits by means of remote information.

Second, contrary to the Gaussian model, the data generating process display fat tails. The presence of kurtosis suggests that rare events should not be assumed away when it comes to managing risk. These distributions can be accurately described by power-law distributions. The exponent corresponding to emerging markets can be less than two, in which case the variance diverges to infinity.

Moreover, periods of hectic activity and relatively quiescent ones coexist. Such a clustering in the volume of activity and volatility leads to a multifractal-like behavior of returns. The leverage effect, a negative correlation between (past) returns and (future) volatilities in turn leads to negative skewness in the distribution of returns.

In this paper we investigate the presence of some of these facts, with a special focus on long-term dependency. The intuition behind long memory is that the longest cycle in a sample will be proportional to the number of observations (Mandelbrot *et al.* 1997). There is no definitive conclusion about the existence of long memory in financial returns. Green and Fielitz (1977), Taylor (1986), Barkoulas *et al.* (2000), Taqqu *et al.* (1999), Ding *et al.* (2001), Sadique and Silvapulle (2001) and others claim that financial markets exhibit long memory. Other scholars do not reach a clearcut conclusion, such as Lo (1999). However, in the post Bretton-Woods era the dominant view is that long-term dependence exists in liquid markets up to a lag of a ten-minute order.

In face of this evidence it is important to investigate different stochastic processes and fit statistical distributions that mimicked the actual data as close as possible.

The first step toward this is to identify whether the process exhibits long or short memory or the assumption that the data does not exhibit memory holds. The presence of memory will then dictate the choice of models used to forecast the underlying process.

The mathematical definition of a stationary process with long-memory or long-range dependence or persistence is given by its autocorrelation function ρ_k such that $\lim_{k\to\infty} \rho_k = ck^{-\lambda}$ for some $0 < c$ and $0 < \lambda < 1$ (Cowpertwaite and Metcalfe 2009: 160). For a long-memory process the autocorrelation function decays slowly at a hyperbolic rate as opposed to an exponential rate for a Brownian motion. This implies that the autocorrelations are not summable or in other words $\sum_{k=-\infty}^{\infty} \rho_k = \infty$. The spectral density defined as: $f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \rho_k e^{ik\omega}$ (Fourier frequency ω) tends to infinity at zero frequency $f(\omega) \to C_f \omega^{\lambda-1}$ as $\omega \to 0$. We use the Hurst (1951, 1955) exponent (H) to identify the presence or absence of memory. Hurst used the rescaled-range statistic over a period k and found it proportional to k^H for some $H > 1/2$. The Hurst parameter is defined by $H = 1 - \lambda/2$ and hence ranges from $1/2 < H < 1$. If $\lambda = 1 \Rightarrow H = 1/2$ and the process is a Brownian motion with no long-range dependence.

The next section gives a succinct introduction to the various approaches used in literature and to justify the methodology adopted to identify the plausible data generation process. The empirical evidence follows with the methodology and results. The final section concludes.

2. Literature Survey

Two main approaches are used to fit models to financial time series like stock prices or options data.

1. Identifying the underlying distribution for the data generation process by calibrating the actual observations to Stable Distributions
2. Fitting econometric models like ARCH, GARCH, FARIMA, FIEGARCH based on the existence of memory in the evolution of prices. The existence of memory in the process is based on the value of the Hurst exponent.

Stable Distributions

The data generation process of the stock prices is assumed to be a random walk of size $x_i \, \forall i = 1,2,..., n$ with n i.i.d. changes at each instant of time δt. The position of the random walk in time $n\delta t$ equals the sum of the n i.i.d x_is. Thus $S_n = \sum_{i=1}^{n} x_i$. The simplest example is $x_i = s; \, \forall i = 1, 2, ..., n$. The question is what happens to the probability distribution of S_n as n increases? If the functional form of the density function is invariant under the summation then the distribution is classified as *stable*. Thus if x_i follows a normal distribution with mean μ and variance σ^2, then $S_n = \sum_{i=1}^{n} x_i$, follows a normal distribution with mean $n\mu$ and variance $n\sigma^2$.

Are there distributions that are stable with finite moments? Khintchine and Lévy (1936) derived the general formula for the entire class of stable distributions. Lévy stable distributions lack closed form density functions except for normal, Cauchy or Lorentzian and Lévy-Smirnov distributions. These distributions can be easily expressed in terms of the characteristic function, which is the Fourier transform of the distribution function (p(x)) given by $\phi_x(t) = E[e^{itx]} = \int p(x)e^{itx}dx$.

The general form of the characteristic function of stable distributions is given by

$$\log{}_x(t) = \begin{cases} -\sigma^\alpha |\, t\,|^\alpha \{i\beta \mathrm{sign(t)} \tan \frac{\pi\alpha}{2} + i\mu t\}, \alpha \neq 1, \\ -\sigma |\, t\,|\{i\beta \mathrm{sign(t)} \frac{2}{\pi}\log(t) + i\mu t\}, \alpha = 1, \end{cases} \quad (2)$$

where index of stability (tail index, tail exponent or characteristic exponent) α in (0,2) (for α > 2, p(x)<0), a skewness parameter β in [-1,1], a scale parameter σ > 0 and location parameter μ in R. For details refer Mantegna and Stanley (2004); Weron (2001):

Lévy-Smirnov: α = 1/2, β = *1*,

Cauchy or Lorentzian: α = 1, β = 0,

Normal or Gaussian: α = 2, β = 0,

when β = 0, the distribution is symmetric about μ. The pth moment of a Lévy stable distribution is finite if $p < \alpha$. Thus all Lévy stable distributions have infinite variance except the normal. This has implications for risk management as Value at Risk studies normally attempt to estimate the probability of loss beyond a certain number of standard deviations below the mean.

Self-Similarity

Since one is modelling the distribution of the returns, the analysis may be sensitive to the scaling of the time factor. The point to consider is whether the distribution of returns $r(\Delta t) = \dfrac{P(t + \Delta t) - P(t)}{P(t)}$ self-similar? In other words is the distribution of returns taken over different time intervals (Δt = \$ 1, 2, 5, 10 minutes, 1 hour, 1 day, 2 days etc.) different? Mantegna and Stanley (2004) show that non-Gaussian stable distributions are self-similar when appropriately scaled. The next question is to find the appropriate scaling factor that reflects self-similarity. The approach to finding the scaling factor is to find the probability of return to the origin; $p(S_n) = 0$ and show that the rescaled distribution $\tilde{S}_n = \dfrac{S_n}{n^{1/\alpha}}$ satisfies $\int_{-\infty}^{\infty} \tilde{p}(\tilde{S}_n)d\tilde{S}_n = 1$.

Truncated Lévy Flight (follows follows Mantegna and Stanley (2004)).

When each step takes time that is proportional to its length it is termed as a random walk. However when each step takes the same time regardless of the length, the random walk is termed as flight. When the steps are distributed according to a Lévy process it is termed as Lévy Flights. Except for the Gaussian distribution which is a stable Lévy distribution and hence scalable having a finite variance, no other Lévy distribution has finite variance though all are stable and scalable. Student's t distribution does not possess scaling properties but has finite variance. The only distribution that possesses a finite variance and scaling behaviour over a large range is the Truncated Lévy Flight defined by:

$$p(x) = \begin{cases} 0 & x > \mathrm{l}, \\ cp_L(x) & -1 \leq x \leq \mathrm{l}, \\ 0 & x < -\mathrm{l}, \end{cases} \quad (3)$$

where $pL(x)$ is a symmetric Lévy distribution and c is a normalising constant. Mantegna and Stanley (2004) show that TLF distribution converges to the gaussian for large values of *n* i.e.

$$S_n = \sum_{i=1}^{n} x_i,$$

$$pS_n = \begin{cases} p_L(S_N), & \text{when } n \ll n_x, \\ p_G(S_n), & \text{when } n \gg n_x. \end{cases}$$

Estimation of Tail Index α (follows Weron (2001)). When α< 2, the tails of the Lévy distribution are asymptotically equivalent to a Pareto law, i.e. if X ~ Sα(σ, β, μ), α < 2, σ = 1, μ = 0, then $x \to \infty$

$$p(X > x) = 1 - F(x) \to C_\alpha(1 + \beta)x^{-\alpha},$$

$$p(X < -x) = F(x) \to C_\alpha(1 - \beta)x^{-\alpha},$$

where $C_\alpha = \left[2\int_0^\infty x^{-\alpha}\sin(x)dx\right] = \dfrac{1}{\pi}\Gamma(\alpha)\sin\left(\dfrac{\pi\alpha}{2}\right)$.

Log-log linear regression

To estimate the tail index, a linear regression is fit to the dependent variable $\log(1 - F(x))$, where $F(x)$ is the cumulative density function of $x > 0$ v/s the independent variable $\log(x)$; $\forall x > 0$. This estimator is sensitive to the sample size and choice of number of observations.

Hill estimator

It is the non-parametric method to estimate the tail behaviour based on order statistics, where the upper tail is of the form $1 - F(x) = Cx^{-\alpha}$. The sample is ordered so that $X_{(1)} \geq X_{(2)} \geq ... \geq X_{(N)}$, the Hill estimator based on k largest order statistics is $\alpha_{Hill}(k) = \left[\dfrac{1}{k} \sum_{i=1}^{k} \log\left(\dfrac{X_n}{X_{k+1}} \right) \right]^{-1}$.

Weron (2001) finds that Hill estimator also over estimates the tail index parameter α and one needs to use high frequency data for asset returns and analyse only the most outlying values to correctly estimate α.

Brownian and Fractional Brownian motion

A Fractional Brownian Motion (FBM) (Vasconcelos 2004), is a Gaussian process $\{W_H(t), t > 0\}$ with zero mean and stationary increments whose variance and covariance are given by

$$E[W_H^2(t)] = t^{2H}$$

$$E[W_H(s)W_H(t) = 1/2(s^{2H} + t^{2H} - |t - s|^{2H}),$$

where $0 < H < 1$. It is a self similar process $W_H(at)^d = a^H W_H(t) \; \forall \; a > 0$. The parameter H is called the self-similarity exponent or the Hurst exponent. For $H = 1/2$, the FBM reduces to the usual Brownian motion where increments $\Delta W_t = W_H(t + \Delta t) - W_H(t)$ are i.i.d when $H \neq 1/2$, increments ΔW_t are known as fractional white noise displaying long-range correlation

$$E[\Delta W_{t+k} \Delta W_t] = 2H(2H - 1)k^{2(H-1)} \text{ for; } k \to \infty$$

Processes with lower H have a greater volatility than those processes with a higher H. Fractional Differencing: FARIMA (p,d,q) process A fractionally differenced ARIMA process $\{x_t\}$, FARIMA(p,d,q) has

the form $\phi(\mathbf{L})(1 - \mathbf{L})^d x_t = \psi(\mathbf{L})w_t$ for some $-\dfrac{1}{2} < d < \dfrac{1}{2}$. We fit $y_t = (1 - \mathbf{L})^d x_t = [\phi(\mathbf{L})]^{-1} \psi(L)w_t$ where

$$(1 - L)^d = 1 - dL + \frac{d(d-1)}{2!}L^2 - \frac{d(d-1)(d-2)}{3!}L^3 + ...$$

and L is the backward lag operator. The autocorrelation function ρ_k of a FARIMA (0,d,0) process tends to

$\dfrac{\Gamma(1-d)}{\Gamma(d)} |k|^{2d-1}$ for large n. The process is stationary provided $-\dfrac{1}{2} < d < \dfrac{1}{2}$ and provides a relationship between

the differencing parameter d and the long-memory parameter λ when $0 \leq d$. $2d - 1 = -\lambda$ implying, if $\lambda = 1$; $d = 0$. Thus the series has no long-range dependence if $d = 0$.

3. Methodology

We use two methods to estimate the long-range dependence in the daily returns of the PSI20.
1. Heuristic method called the Detrended Fluctuation analysis.
2. Semi non-parametric approach using the GPH test based on Geweke and Porter-Hudak (1983).

Detrended Fluctuation Analysis

The Hurst exponent was initially estimated by using the rescaled-range (RS) analysis of Hurst (1951). We adopt the Detrended Fluctuation Analysis (DFA) methodology as described in Peng *et al.* (1994), Moreira *et al.* (1994) and Vasconcelos (2004) for estimating the Hurst exponent. Costa and Vasconcelos (2003) find the DFA more reliable than the RS analysis for estimation of the Hurst exponent. Various studies have been carried out to estimate the Hurst exponent to determine the existence of fractional Brownian motion and multi fractality. Razdan (2001) finds that the Bombay stock exchange exhibits fractional Brownian motion with mono-fractality using the RS analysis. Da Silva *et al.* (2007) estimate the Hurst exponent for the Brazilian exchange rate market using the RS analysis find it close to 0.5 implying a Brownian motion.

1. Given a time series $r(t)$, $t = 1, 2 ..., T$ of say daily returns, obtain the cumulative time series X(t).

2. $X(t) = \sum_{k=1}^{T} r(k) - \bar{r}, t = 1, 2, ..., T$
   $\bar{r} = \dfrac{1}{T} \sum_{k=1}^{T} r(k)$

3. Break $X(t)$ into N non-overlapping intervals of equal length τ where $N = int\left[\dfrac{t}{T} \right]$, where N is an integer.

4. For each of the intervals τ, fit a linear regression $Y_\tau(t) = a_n + b_n t \forall t \in \tau$ where a_n and b_n are obtained from an OLS estimation procedure.

5. Compute the rescaled function F_τ for each τ.

6. $F_\tau = \dfrac{1}{S} \sqrt{ \dfrac{1}{N_\tau} \sum_{t=1}^{N_\tau} [X(t) - Y_\tau(t)]^2 }$

   $S = \sqrt{ \dfrac{1}{T} \sum_{t=1}^{T} \left[r(t) - \bar{r} \right]^2 }$

7. Repeat steps 3,4,5 for different values of τ and obtain F_τ for each τ.

8. The Hurst exponent H is obtained from the scaling behaviour of F_τ, $F_\tau = C_H \tau^H$ where C_H is a constant independent of the time lag τ.

9. Use OLS regression on the $log\ [F_\tau] = log\ [C_H] + Hlog\ [\tau]$ to obtain H.

To check for multi fractality modify step 6 to

$$F_\tau = \left\{ \frac{1}{N_\tau} \sum_{t=1}^{N_\tau} | X(t) - Y_\tau(t)|^{2q} \right\}^{\frac{1}{2q}}.$$

Spectral Density using the GPH test

A time series $Y = \{y_t\}_{t=1}^N$ is said to be integrated of order d, signified as $I(d)$ if it has a stationary, invertible autoregressive moving average (ARMA) representation after applying the difference operator $(1 - L)^d$ where L is the backward lag operator. The series is fractionally integrated when d is not an integer Geweke and Porter-Hudak (1983) suggested a semi-parametric estimator of d in the frequency domain. They consider the data generating process $(1 - L)^d y_t = z_t$ where $z_t \sim I(0)$.

Representing the process in frequency domain $f_y(\omega) = |1 - exp(-i\omega)|^{-2d} f_z(\omega)$ where $f_z(\omega)$ and $f_y(\omega)$ are spectral densities of z_t and y_t respectively. The spectral density of the fractionally integrated process y_t is given by

$$f_y(\omega_j) = \left[4 \sin^2\left(\frac{\omega_j}{2}\right) \right]^{-d} f_z(\omega_j) \qquad (4)$$

for $j = 1, 2, ..., n_f$ where $\omega_j = \frac{2\pi j}{N}$ and $f_z(\omega_j)$ is the spectral density corresponding to z_t. The fractional difference parameter d can be estimated using the regression:

$$\ln [f_y(\omega_j)] = \beta - d \ln\left[4 \sin^2\left(\frac{\omega_j}{2}\right) \right] + e_j, \qquad (5)$$

where $\beta = \ln\{f_z(0)\}$ and $e_j = \ln\left\{ \frac{f_z(\omega_j)}{f_z(0)} \right\}$.

Geweke and Porter-Hudak (1983) showed that using a periodogram estimate, the least squares estimate of d using the above regression is distributed in large samples if the number of observations $n_f(T) = T^\alpha$ with $0 < \alpha < 1$ as a normal distribution

$$\hat{d} \sim N\left(d, \frac{\pi^2}{6 \sum_{j=1}^{n_f} (U_J - \bar{U})^2} \right), \text{ where}$$

$$U_j = \ln\left[4 \sin^2\left(\frac{\omega_j}{2}\right) \right] \text{ and } \bar{U} = \frac{\sum_{j=1}^{n_f} U_j}{n_f}.$$

Under the null hypothesis of no long memory (d = 0), the

t-statistic $\quad t_{d=0} = \hat{d}\left(\frac{\pi^2}{6 \sum_{j=1}^{n_f} (U_J - \bar{U})^2} \right)^{-1/2} \quad$ has a limiting
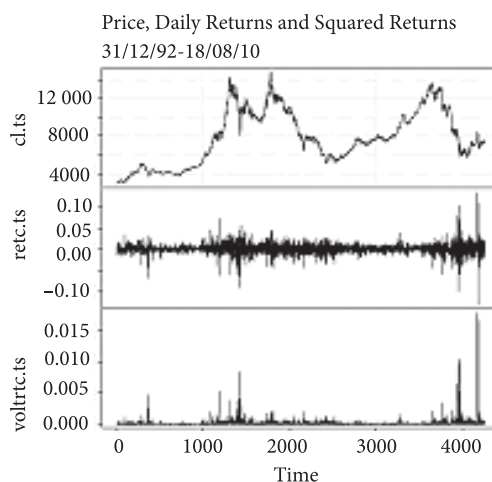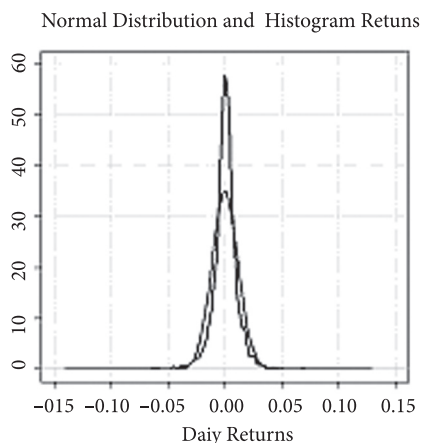
normal distribution.

4. Results

Data and Simulation Presentation

The section begins with the presentation of the data and distribution of the returns. This is followed by the analysis of the results.

The maximum return we find is 13.3486% while the minimum return is –12.8696%. Figure 1A shows the daily movement of the PSI20 (top) along with the daily returns (middle) and the volatility (bottom) of the returns. Figure 1B shows the distribution of the actual returns with a normal distribution superimposed on it. The actual distribution has a greater kurtosis and fatter tails which the normal distribution is unable to capture. The normal distribution has been fit using the maximum likelihood method with the parameters equated to the mean and the variance of the actual distribution.

A: price, returns, volatility

B: actual distribution and normal distribution

Fig. 1. PSI20: Level, Returns, Volatility and density of daily returns and estimated normal fit

Detrended Fluctuation Analysis

The modified DFA method is used to identify the possible multi-fractal behaviour of the returns data and to find the evolution of the Hurst exponent over time across different time horizons with a moving window of 1day.
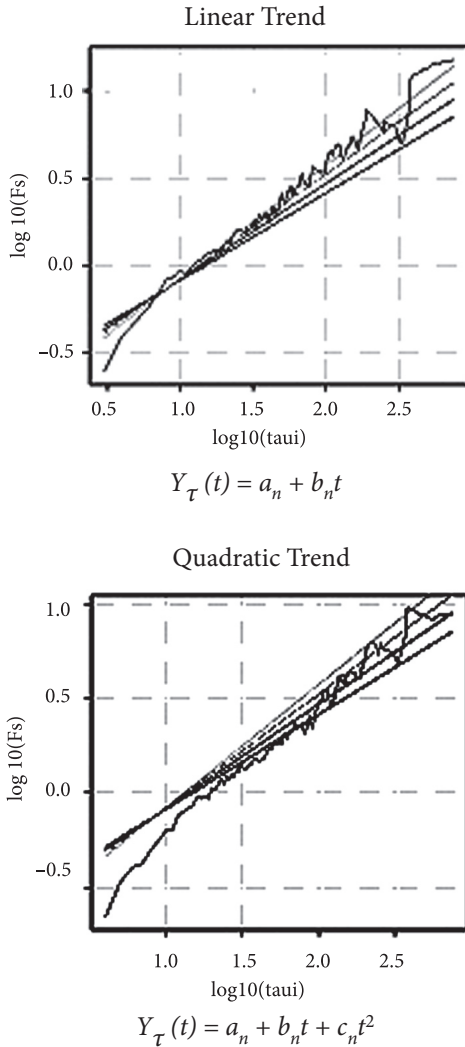
Linear Trend

$$Y_\tau(t) = a_n + b_n t$$

Quadratic Trend

$$Y_\tau(t) = a_n + b_n t + c_n t^2$$

Fig. 2. Hurst exponent with Linear and Quadratic Trends

Figure 2 shows the variation of $log(F_\tau)$ on the y-axis vs $log(\tau)$ on the x-axis for linear and quadratic trends when estimated using the whole data series. The straight lines are obtained for various values of the Hurst exponent 0.5, 0.55, 0.6 and 0.65 using

$$\log(F_\tau) = \log(C_H) + H \log(\tau), \qquad (6)$$

where $C_H = \left[\dfrac{2}{2H+1} + \dfrac{1}{H+2} - \dfrac{2}{H+1} \right]^{1/2}$. The quadratic

trend tends to overestimate the exponent over the entire data set. When using the entire data set we obtain a single

value of the exponent. Whether this is consistently true over the entire range was checked using a moving window of 1 day with a period of 3 years (750 days) for estimating each exponent.

Figure 3 gives a graphic representation of the linear and quadratic trend polynomials fit using the entire range of data consisting of 4251 points. The difference between X_t and Y_t is used for F_τ where X_t is the cumulative series of the daily returns minus the mean of the series of daily returns. Eventually the nature of the errors between the series X_t and its fit Y_t will determine the Hurst exponent. In our case we find that the quadratic trends tend to overestimate the Hurst exponent implying that the quadratic trend may not be the appropriate fit for the local trends.
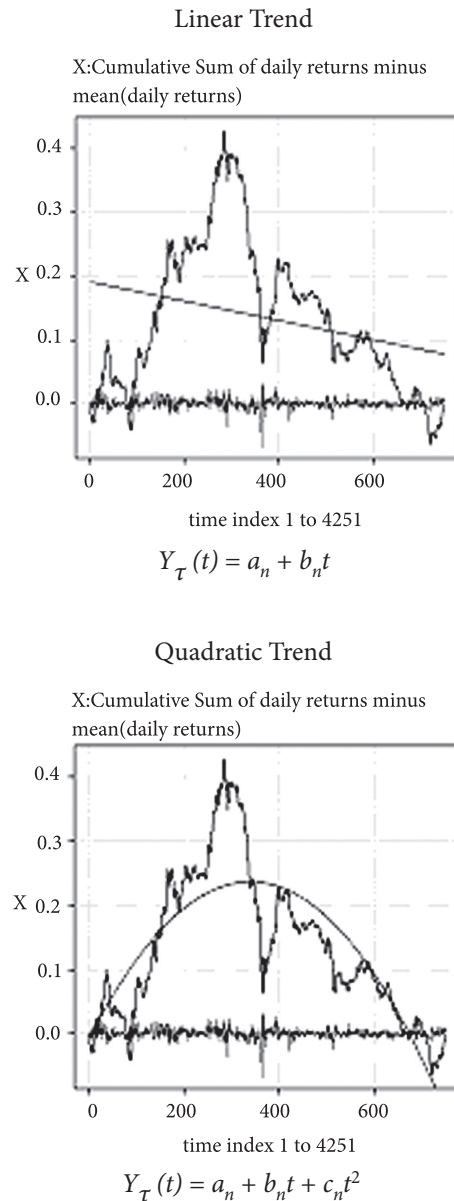
Linear Trend

X:Cumulative Sum of daily returns minus mean(daily returns)

time index 1 to 4251

$$Y_\tau(t) = a_n + b_n t$$

Quadratic Trend

X:Cumulative Sum of daily returns minus mean(daily returns)

time index 1 to 4251

$$Y_\tau(t) = a_n + b_n t + c_n t^2$$

Fig. 3. Cumulative Sum of daily returns minus mean dauly returns with Linear and Quadratic Trends

Figure 4 shows the movement of the Hurst exponent (y axis) over the same time periods (x axis) for 3500 periods with a moving window of 1 day. Initially we use 750 days to estimate the Hurst exponent and then advance one day till the end of the data period. We have used the linear and quadratic trends to estimate the exponent. The Hurst exponent (lower series) estimated using the linear trend exhibits a lower average value as opposed to the quadratic trend (higher series). Except for a small period where the values of the Hurst exponent exhibit opposite behaviour (fall for the quadratic trend and rise for the linear trend) they both exhibit similar behaviour. The Hurst exponent based on the quadratic trend lies completely above 0.5 implying an unequivocal long-range dependence in the daily returns while the exponent based on the linear trend shows some periods when the exponent falls below 0.5. This implies anti-persistence and a faster return to the original level.
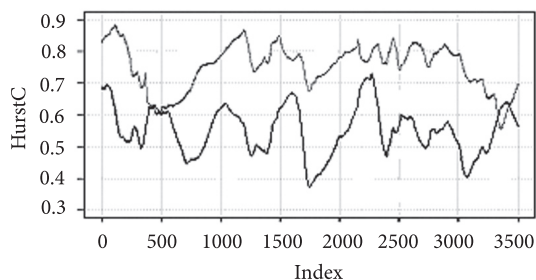

Fig. 4. Hurst Exponent for various durations (1 day moving window)

Table 1. Summary of Hurst Coefficients over time

Trend	mean	Std deviation
Linear	0.5515	0.0757
Quadratic	0.7520	0.0718

Spectral Density using the GPH test

We see from Figure 5 that the autocorrelations and the partial autocorrelations do not decay exponentially but show some persistence even at lags close to 35.
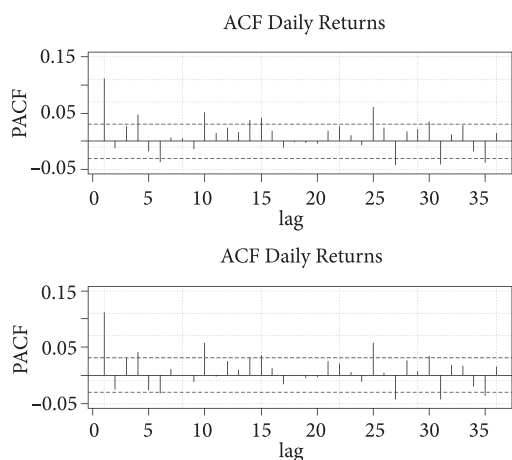

Fig. 5. ACF and PACF of Daily returns

Table 2. GPH Test for estimating fractional difference parameter d

	Estimate	Std Error	t-value	Pr(>|t|)
β	−9.9135	0.0537	−184.675	<2e−16
d	−0.0786	0.0137	−5.747	1.04e−8

We have the following relationships $\lambda = 1 - 2d$ and $H = 1 - \lambda / 2$. Thus from Table 2 we infer

$$H = 1 - \frac{(1-2d)}{2} = 1 - \frac{(1-2*0.0786)}{2} = 0.5786 \cdot$$

This shows a small persistence in the behaviour of daily returns. The Hurst exponent estimated using the Geweke and Porter-Hudak method is in close approximation with the linear trend used to estimate the Hurst exponent using the detrended fluctuation analysis.

Comparison with Other Studies

Podobnik *et al.* (2006) use the linear trend to estimate the Hurst exponent. Using their terminology, the DFA follows a scaling law $F_\tau \propto \tau^H$ they estimate the power-law auto-correlations H and find that all the indices for the ten transition economies in Europe exhibit power-law auto-correlations or in other words have a Hurst exponent different from 0.5 implying long-range dependence. Alvarez-Ramirez *et al.* (2008a) find the Hurst exponent varies substantially from persistence (above 0.5) to anti-persistence (below 0.5). They infer that the end of Bretton-Woods era in 1972 had a major impact on the efficiency of the markets wherein they use the Hurst exponent as a proxy for the market efficiency and conclude that markets became more efficient. Contrary results to Alvarez-Ramirez *et al.* (2008a) are obtained by Onali and Goddard (2009) where they find no evidence of long-range dependence in the returns of the Italian Mibtel index. Wang, Liu and Gu (2009) studies the improvement in efficiency of the Shenzen stock market using the multi-fractal DFA and find that the Hurst exponent falls consistently across time thus concluding that the market became more efficient over time. Cajueiro and Tabak (2007) investigate the long-range dependence of LIBOR interest rates on maturities of fixed income instruments for six countries. They find that the long-range dependence falls with increased maturity for four countries out of six and rises for the remaining two. Serlitis and Rosenberg (2007) estimated the Hurst exponent for the NYMEX futures and found the series to be anti-persistent and thus price corrections occur much faster. Since the futures prices are intricately linked to the spot and option prices, the anti persistence may be a result of the movements in the other markets where in the traders move much faster to rebalance their positions. On the contrary, with crude oil prices, Alvarez-Ramirez *et al.* (2008b) find the existence of persistence or a Hurst exponent in the range of 0.6–0.7 implying that the spot markets take time to adjust to information in the short run.

5. Conclusion

We have used two approaches to investigate the presence of long-range dependence in the daily returns of the PSI20. The detrended fluctuation analysis, a heuristic approach with a liner and quadratic trends over a large range of the returns series exhibit long-range dependence with a Hurst exponent greater than 0.5. This implies that the market is slow to respond to the shocks on the whole. It depends to be seen if the Hurst exponents are different during the rise and the fall as normally markets are quick to fall but slow to rise. The quadratic trend in the DFA method tends to obtain higher values of the Hurst exponent and may not be an appropriate fit.

We have used another semi-parametric approach to corroborate our estimates of the Hurst exponent using the Geweke and Porter-Hudak method. We find that the Hurst exponent estimated using this approach is closer to the linear trend used in the DFA supporting the claim that the quadratic trend may not be an appropriate fit to be used in the DFA.

We propose the use of Fractional GARCH models to estimate the differencing parameter d and their use for forecasting as opposed to the traditional GARCH models.

Although there is a vast amount of empirical findings dealing with the issue of long-term dependence, its underlying causes remain obscure. Anti-persistence can be more easily interpreted on the grounds, for instance, of a learning process leading to price overreactions that are immediately adjusted. Indeed, to the eye, short-memory processes appear indistinguishable from a white noise (Mandelbrot *et al.* 1997). Long memory, on the contrary, indicates the existence of importance pieces of information that are not immediately incorporated in the price.

This fact suggests that there can be sources of information easily captured by prices, while others do not. There are several reasons why this may occur. A fractal or a multi-fractal series suggests the action of interacting systems generating positive feedback. During 'normal' periods in which those systems operate rather independently, the 'low-scale' information that becomes operative is that giving rise to short memory. Suddenly, the high-range information dominates the markets and long-term dependence appears. For instance, in the benchmark of the current financial crisis, few scholars call into question that the huge amount of liquidity created by expansive monetary policies applied in the US and the Euro zone gave rise to a house bubble. This process, which resembled that occurred in Japan in the mid eighties, was for a long time compatible with a good performance of financial markets, growth, trade and other macroeconomic variables. According to Kindleberger and Aliber (2005) a situation like this becomes unsustainable whenever the ratio of the price of urban land to wage rises above a threshold level. In such a case, market adjustments push down land prices, giving rise to a scarcity of liquidity. Financial markets then collide with the real side of the economy making valuable units of information that were not operative prior to the process of revulsion

References

Alvarez-Ramirez, J.; Alvarez, J.; Rodrigues, E.; Fernandez-Anaya, G. 2008. Time-varying Hurst exponent for US stock markets, *Physica A* 387: 6159–6169. doi:10.1016/j.physa.2008.06.056

Alvarez-Ramirez, J.; Alvarez, J.; Rodrigues, E. 2008b. Short-term predictability of crude oil markets: A detrended fluctuation analysis approach, *Energy Economics* 30: 2645–2656. doi:10.1016/j.eneco.2008.05.006

Bachelier, L. 1964. Speculation, in P. H. Cootner, *The Random Character of Stock Market Prices*, MIT Press. ISBN 13:978-026-203-009-0 Théorie de la.

Barkoulas, J. T.; Christopher, F. B.; Nickolaos, T. 2000. Long Memory in Greek Stock Markets, *Applied Financial Economics* 10: 177–184. doi:10.1080/096031000331815

Black, F.; Scholes, M. 1973. The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* 81: 637–654. doi:10.1086/260062

Bouchaud, J. P. 2002. An Introduction to Statistical Finance, *Physical A* 313: 238–251. doi:10.1016/S0378-4371(02)01039-7

Cajueiro, D. O. Tabak, B. M. 2007. Long-range dependence and multifractality in the term structure of LIBOR interest rates, *Physica A* 603–614. doi:10.1016/j.physa.2006.04.110

Costa, R. L.; Vasconcelos, G. L. 2003. Long-range correlations and nonstationarity in the Brazilian stock market, *Physica A* 329: 231–248. doi:10.1016/S0378-4371(03)00607-1

Cowpertwaite, P. S. P.; Metcalfe, A. V. 2009. *Introductory Time Series with R.* Springer. ISBN 13: 978-0-387-88697-8.

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.

Ding, Z.; Granger, C. W. J.; Engle, R. F. 1993. A Long Memory Property of Stock Market Returns and a New Model, *Journal of Empirical Finance* 1(1): 83–106. doi:10.1016/0927-5398(93)90006-D

Fama, E. F. 1970. Efficient Capital Markets: A Review of Theory and Empirical Work, *Journal of Business* 38: 383–417.

Geweke, J.; Porter-Hudak, S. 1983. The Estimation and Application of Long Memory Time Series Models, *Journal of Time Series Analysis* 4: 221–237. doi:10.1111/j.1467-9892.1983.tb00371.x

Greene, M. T.; Fielitz, B. D. 1977. Long-Term Dependence in Common Stock Returns, *Journal of Financial Economics* 5: 339–349. doi:10.1016/0304-405X(77)90006-X

Hurst, H. 1951. Long-term storage capacity of reservoirs, *Transactions of the American Society of Civil Engineers* 116: 770–808.

Hurst, H. 1955. Methods of using long-term storage in reservoirs, in *Proceedings of the Institution of Civil Engineers* I: 519–577.

Khintchine, A. Y.; Lévy, P. 1936. Sur les loi stables, in *C. R. Acad. Sci. Paris* (202): 374–376.

Kindleberger, C. P.; Aliber, M. 2005. *Panics and Crashes: A History of Financial Crises*, 5th ed. Wiley NY. ISBN 13: 978-1-4039-3651-6.

Lo, A. W. 1991. Long-Term Memory in Stock Markets, *Econometrica* 59(5): 1279–1313. doi:10.2307/2938368

Mandelbrot, B. B.; Fisher, A. J.; Calvet, L. E. 1997. *A Multifractal Model of Asset Returns*. Cowles Foundation Discussion Paper No. 1164.

Mantegna, R. N.; Stanley, H. E. 2004. *An Introduction to Econophysics: Correlations and Complexities in Finance*. Cambridge University Press. ISBN 13: 978-052-162-008-68.

Moreira, J. G.; Kamphorst Leal da Silva, J.; Oliffson Kamphorst, S. 1994. On the fractal dimension of self-affine profiles, *Journal of Physics A* 27: 8079–8089. doi:10.1088/0305-4470/27/24/018

Onali, E.; Goddard, J. 2009. Unifractality and multifractality in the Italian stock market, *International Review of Financial Analysis* 18: 154–163. doi:10.1016/j.irfa.2009.05.001

Peng, C. K.; Buldyrev, S. V.; Havlin, S.; Simons, M.; Stanley, H. E.; Goldberger, A. L. 1994. Mosaic organization of DNA nucleotides, *Physical Review E* 49(2): 1685–1689. doi:10.1103/PhysRevE.49.1685

Podobnik, B.; Fu, D.; Jagric, T.; Grosse, I.; Stanley, H. E. 2006. Fractionally integrated process for transition economies, *Physica A* 362: 465–470. doi:10.1016/j.physa.2005.09.051

Razdan, A. 2002. Scaling in the Bombay stock exchange index, *Pramana Journal of Physic* 58(3): 537–544.

Samuelson, P. A. 1965. Proof that Properly Anticipated Prices Fluctuate Randomly, *Industrial Management Review* 6: 41–45.

Sadique, S.; Silvapulle, P. 2001. Long-Term Memory in Stock market Returns: International Evidence, *International Journal of Finance and Economics* 6(1): 59–67. doi:10.1002/ijfe.143

Serlitis, A.; Rosenberg, A. A. 2007. The Hurst exponent in energy futures prices, *Physica A* 380: 325–332. doi:10.1016/j.physa.2007.02.055

Taqqu, M. S.; Teverosky, V.; Willinger, W. 1999. Stock Market Prices and Long-Range Dependence, *Finance and Stochastics* 3(1): 1–13. doi:10.1007/s007800050049

Taylor, S. 1986. *Modelling Financial Time Series*. Wiley. ISBN 13: 978-981-277-084-4.

Vasconcelos, G. L. 2004. A Guided Walk Down Wall Street: An Introduction to Econophysics, *Brazilian Journal of Physics* 34(03B): 1039–1065. doi:10.1590/S0103-97332004000600002

Voit, J. 2005. *The Statistical Mechanics of Financial Markets*. 3 rd. Springer. ISBN 10-3-540-26285-7.

Wang, Y.; Liu, L.; Gu, R. 2009. Analysis of efficiency for Shenzen stock market based on multifractal detrended fluctuation analysis, *International Review of Financial Analysis* 18: 271–276. doi:10.1016/j.irfa.2009.09.005

Weron, R. 2001. Lévy-stable distributions revisited: tail index > 2 does not exclude the Lévy-stable regime, *International Journal of Modern Physics C* 12(2): 209–223. doi:10.1142/S0129183101001614

Appendix: R Code

```
# code to calculate the Hurst Exponent using Detrended
# Fluctation Analysis
# file name is psi20-close.txt
# it has 4 columns, dd mm yyyy cl for day, month, year
# and close price
file.choose()
pt20<-read.table("psi20-close.txt", header=T)
names(pt20)
attach(pt20)

dp    <- length(cl)
dr = dp-1
retc   <- seq(0,0,length.out=dr)
volretc <- seq(0,0,length.out=dr)
fdc    <- seq(0,0,length.out=dr)

for (j in 1:dr)
{
retc[j]   = (cl[j+1]-cl[j])/cl[j]
fdc[j]    = cl[j+1]-cl[j]
volretc[j] = retc[j]^2
}

dretc <- retc - mean(retc)
m1 <- mean(retc)
m2 <- sd(retc)

# drawing graphs from here
par(mfrow=c(3,1))
plot(cl,type="l",panel.first=grid())
plot(retc,type="l",panel.first=grid())
plot(volretc,type="l",col='green',panel.first=grid())

# obtain the cumulative sum of returns
# reduced by mean of returns
X <- seq(0,0,length.out=dr)
X <- cumsum(dretc)
S <- sd(dretc)

# Detrended fluctuation Analysis
DFA <- function(tau,X1)
{
dr = length(X1)
In = as.integer(dr/tau)

Yn <- matrix(nr=In,nc=tau)
for (i in 1:In)
for (j in 1:tau)
Yn[i,j] = X1[(i-1)*tau+j]

coeffx <- matrix(nr=In,nc=2)
tempy  <- seq(0,0,length.out=tau)
tempx  <- seq(0,0,length.out=tau)
for (i in 1:In)
{for (j in 1:tau)
{
tempy[j] = Yn[i,j]
tempx[j] = j
}
}
```

```
y.ld <- lm(formula = tempy ~ tempx)
coeffx[i,1] = coef(y.ld)[1]
coeffx[i,2] = coef(y.ld)[2]
}

lYt = In*tau
Yt <- seq(0,0,length.out=lYt)
Xt <- seq(0,0,length.out=lYt)

for(i in 1:In)
for (j in 1:tau)
{
Yt[(i-1)*tau+j] = coeffx[i,1]+coeffx[i,2]*j
Xt[(i-1)*tau+j] = X1[(i-1)*tau+j]
}

XtYtSq <- sum((Xt-Yt)^2)
Ft <- sqrt((1/lYt)*XtYtSq)
return(Ft)
}

EH <- function(Ft,taui)
{
hurst.ld <- lm(formula = log(Ft) ~ log(taui))
cCH = coef(hurst.ld)[1]
cHH = coef(hurst.ld)[2]

 return(cHH)
}

pX = 4
Ip = floor(length(X)/pX)
nX <- matrix(nr=Ip,nc=pX)
dnX <- matrix(nr=Ip,nc=pX)
cumdnX <- matrix(nr=Ip,nc=pX)

FtA <- matrix(nr=(Ip-2),nc=pX)
tauA <- matrix(nr=(Ip-2),nc=pX)
HurstC <- seq(0,0,length.out=pX)
nS <- seq(0,0,length.out=pX)

for (i in 1:Ip)
for (j in 1:pX)
{
nX[i,j] = retc[(j-1)*Ip+i]
}

nXmu <- seq(0,0,length.out=pX)
X1 <- seq(0,0,length.out=Ip)

for(j in 1:pX)
{
nXmu[j] <- mean(nX[,j])
nS[j] <- sd(nX[,j])
}
```

```
for (j in 1:pX)
dnX[,j] = nX[,j]-nXmu[j]

for (j in 1:pX)
cumdnX[,j] = cumsum(dnX[,j])

for(j in 1:pX)
{
X1[] <- cumdnX[,j]
ls = length(X1)-2
Ft <- seq(0,0,length.out=ls)
Fs <- seq(0,0,length.out=ls)
taui <- seq(0,0,length.out=ls)

for (i in 1:ls)
{
taui[i] = i+2
Ft[i] = DFA(taui[i],X1)
Fs[i] <- Ft[i]/nS[j]
}

FtA[,j] <- Fs
tauA[,j] <- taui
HurstC[j] <- EH(Fs,taui)

}

H <- seq(0,1,length.out=21)
Ch <- seq(0,0,length.out=21)

for (i in 1:21)
Ch[i] = sqrt(2/(2*H[i]+1)+1/(H[i]+2)-2/(H[i]+1))

tauH1 <- seq(0,0,length.out=(Ip-2))
tauH2 <- seq(0,0,length.out=(Ip-2))
tauH3 <- seq(0,0,length.out=(Ip-2))

tauH1 = taui^(0.5)
tauH2 = taui^(0.55)
tauH3 = taui^(0.6)

FH1 <- Ch[11]*tauH1
FH2 <- Ch[12]*tauH2
FH3 <- Ch[13]*tauH3

par(mfrow=c(2,2))
for (j in 1:4)
{
 plot(log10(taui),log10(FtA[,j]),type='l',col='dark red',panel.
first=grid())
lines(log10(taui),log10(FH1),type='l',col='blue')
lines(log10(taui),log10(FH2),type='l',col='dark green')
lines(log10(taui),log10(FH3),type='l',col='orange')
}
```

Sameer REGE. BE Mechanical Engineering, VJTI University of Mumbai and PhD Economics, IGIDR, Mumbai. Research Fellow – Universidade dos Azores. Interests: computable general equilibrium, financial derivatives, distributions in finance

Samuel Gil MARTÍN. PhD in Economics, European University Institute, Florence (Italy). Economics Professor at the Universidad Autónoma de San Luis Potosí. Interests: macroeconomics, international economics, financial markets and behavioral economics.