PORTUGUESE STOCK MARKET: A LONG-MEMORY PROCESS?

Sameer Rege† and Samuel Gil Martín
CEEAplA & Departmento de Economia e Gestão
Universidade dos Açores
E-mail: samrege@gmail.com, 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 returns to ascertain if the process follows a fractional brownian motion. The multifractal or long-memory process wherein the hurst exponent differs from $\frac{1}{2}$ implies using Auto Regressive Fractionally Integrated Moving Average ARFIMA process as opposed to ARCH, GARCH

Keywords: Hurst Exponent, Long-Memory

JEL Classification Codes: G10

*The authors would like to thank Prof Gualter Couto for his help in procuring the Portuguese stock market data and the Fundação para a Ciência e a Tecnologia for funding the research
†corresponding author
1 Introduction

The attempt to model the data generation process for financial data dates back to [1] wherein he attempted to model the French government bond and its futures.\(^1\) Empirical evidence from stock markets around the world shows that the returns \(^2\) do not follow a gaussian distribution but are fat tailed and skewed. Also the volatility of the returns \(^3\) show a behaviour that is heteroskedastic with periods of high volatility and low volatility. 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. We use the Hurst exponent to identify the presence or absence of memory.

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, ARFIMA 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.

2.1 Stable Distributions

The data generation process of the stock prices is assumed to be a random walk of size \(x_i, i = 1, 2,., n\) with \(n\) i.i.d. changes at each instant of time \((\delta t)\). The position of the random walk in time \(n\delta t\) equals the sum of the \(n\) i.i.d \(x_i\)s.

\[
S_n = \sum_{i=1}^{n} x_i
\]  

(1)

The simplest example is where \(x_i = s \ \forall i\). 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.

---

\(^1\) for further details on Bachelier’s work refer [5]

\(^2\) \(r(t) = \log \frac{P_t}{P_{t-1}} = \frac{P_t - P_{t-1}}{P_{t-1}}\)

\(^3\) \(r^2(t)\)
Thus if $x_i$ follows a normal distribution with mean $\mu$ and variance $\sigma^2$, 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? [2] 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

$$\phi_x(t) = \mathbb{E}[e^{itx}] = \int p(x)e^{itx} dx$$

(2)

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

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

(3)

where index of stability (tail index, tail exponent or characteristic exponent) $\alpha \in (0, 2]$ (for $\alpha > 2$, $p(x) < 0$), a skewness parameter $\beta \in [-1, 1]$, a scale parameter $\sigma > 0$ and location parameter $\mu \in \mathbb{R}$.

- Lévy-Smirnov: $\alpha = \frac{1}{2}, \beta = 1$
- Cauchy or Lorentzian: $\alpha = 1, \beta = 0$
- Normal or Gaussian: $\alpha = 2, \beta = 0$

when $\beta = 0$, the distribution is symmetric about $\mu$. The $p^{th}$ 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 a certain number of standard deviations below the mean.

2.1.1 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 $r(\Delta t) = \frac{P(t+\Delta t) - P(t)}{P(t)}$ self-similar? In other words is the distribution of returns taken over different time intervals ($\Delta t = 1, 2, 5, 10$ minutes, 1 hour, 1 day, 2 days etc.) different?

[3] 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 $\tilde{S}_n = \frac{S_n - \mu}{\sigma}$ satisfies $\int_{-\infty}^{\infty} \tilde{p}(\tilde{S}_n) d\tilde{S}_n = 1$

\[\text{for details refer [3][6]}\]
2.1.2 Truncated Lévy Flight

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

\[ p(x) = \begin{cases} 
0 & x > l \\
cp_L(x) & -l \leq x \leq l \\
0 & x < -l 
\end{cases} \]

where \( p_L(x) \) is a symmetric Lévy distribution and \( c \) is a normalising constant. [3] show that TLF distribution converges to the gaussian for large values of \( n \) i.e. \( S_n = \sum_{i=1}^{n} x_i \)

\[ p(S_n) = \begin{cases} 
p_L(S_n) & \text{when } n \ll n_x \\
p_G(S_n) & \text{when } n \gg n_x 
\end{cases} \]

2.1.3 Estimation of Tail Index \( \alpha \)

When \( \alpha < 2 \), the tails of the Lévy distribution are asymptotically equivalent to a Pareto law, i.e. if \( X \sim S_\alpha(\sigma, \beta, \mu) \), \( \alpha < 2, \sigma = 1, \mu = 0 \), 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]^{-1} = \frac{1}{\pi} \Gamma(\alpha) \sin \frac{\pi \alpha}{2} \]

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

Hill estimator 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 \( X(1) \geq X(2) \ldots \geq X(N) \), the Hill estimator based on \( k \) largest order statistics is

\[ \alpha_{Hill}(k) = \left[ \frac{1}{k} \sum_{n=1}^{k} k \log \frac{X(n)}{X(k+1)} \right]^{-1} \]
[6] finds that Hill estimator also over estimates the tail index parameter $\alpha$ and one needs to use high frequency data for asset returns and analyse only the most outlying values to correctly estimate $\alpha$.

### 2.2 Brownian and Fractional Brownian Motion

A Fractional Brownian Motion (FBM)\(^7\) is a gaussian process $\{W_H(t), t > 0\}$ with zero mean and stationary increments whose variance and covariance are given by

\[
\begin{align*}
\mathbb{E}[W_H^2(t)] &= t^{2H} \\
\mathbb{E}[W_H(s)W_H(t)] &= \frac{1}{2}(s^{2H} + t^{2H} - |t - s|^{2H})
\end{align*}
\]  

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

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

\(\text{Figure 1: Fraction Brownian Motion for } H = 0.3, 0.5, 0.7, 0.9\)

\(\text{\(^7\)refer [4] for details}\)
Figure 1 shows the variation in the data generation process for different Hurst exponents 0.3, 0.5, 0.7 and 0.9. The series with an exponent of 0.3 is the most volatile and as the Hurst exponent approaches 1, the volatility of the series decreases as seen by the scales on the y-axis.

3 Methodology

We adopt the Detrended Fluctuation Analysis (DFA) methodology as described in [4] for estimating the Hurst exponent for the Fractional Brownian Motion.

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

\[
X(t) = \sum_{k=1}^{t} [r(k) - \bar{r}], \quad t = 1, \ldots, T
\]  

(10)

\[
\tau = \frac{1}{T} \sum_{k=1}^{T} r(k)
\]  

(11)

2. Break \( X(t) \) into \( N \) non-overlapping intervals of equal length \( \tau \) (\( N = \text{int} \left[ \frac{T}{\tau} \right] \)), where \( N \) is an integer.

3. For each of the intervals \( \tau \), fit a linear regression

\[
Y_\tau(t) = a_n + b_n t \quad \text{for} \quad t \in \tau
\]  

(12)

where \( a_n \) and \( b_n \) are obtained from an OLS estimation procedure

4. Compute the rescaled function \( F_\tau \) for each \( \tau \)

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

(13)

\[
S = \sqrt{\frac{1}{T} \sum_{t=1}^{T} [r(t) - \bar{r}]^2}
\]  

(14)

5. Repeat steps 2,3,4 for different values of \( \tau \) and obtain \( F_\tau \) for each \( \tau \)

6. The Hurst exponent \( H \) is obtained from the scaling behaviour of \( F_\tau \)

\[
F_\tau = C_H \tau^H
\]  

(15)

where \( C_H \) is a constant independent of the time lag \( \tau \)

7. Use OLS regression on the

\[
\log[F_\tau] = \log[C_H] + H \log[\tau]
\]  

(16)

to obtain \( H \)

To check for multifractality modify step 4 to

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

(17)
4 Results

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

4.1 Data and Simulation Presentation

The modified DFA method is used to identify the possible multifractal 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 1 month.

Figure 2 shows the opening values of the Portuguese stock index from January 1993 to October 2006. The returns and the volatility are shown with retopen on the y-axis. The histogram shows the density of the returns and the logarithm of the density of the returns is compared to the simulated gaussian which resembles a parabola $\log(10(pdfreto))$ v/s index.

Figure 2: Portuguese Stock Index: Level, Returns, Volatility & Density

Figure 3 shows the cumulative distribution function [cdf] of the gaussian and returns data and the double logarithmic plot of $1-cdf$ v/s right tail of the gaussian and returns distribution. The pictures indicate the returns to be concentrated in a very narrow region.
Figure 3: CDF, double-log:tail, double-log PDF: Gaussian & PSI20

Figure 4 is a double logarithmic plot of \([qH^q \text{ v/s } q]\) for the simulated brownian motion and the returns data. The returns data shows a slight non-linearity exhibiting the presence of multi-fractility, while the brownian shows a linear curve indicating the presence of mono-fractility.

<table>
<thead>
<tr>
<th>log (q)</th>
<th>log (qH^q_{BM})</th>
<th>log (qH^q_{PSI:20})</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000</td>
<td>-0.148074</td>
<td>-0.579526</td>
</tr>
<tr>
<td>0.301</td>
<td>0.147457</td>
<td>-0.281713</td>
</tr>
<tr>
<td>0.477</td>
<td>0.319707</td>
<td>-0.109100</td>
</tr>
<tr>
<td>0.602</td>
<td>0.441443</td>
<td>0.010420</td>
</tr>
<tr>
<td>0.699</td>
<td>0.535549</td>
<td>0.099520</td>
</tr>
<tr>
<td>0.778</td>
<td>0.612220</td>
<td>0.169554</td>
</tr>
<tr>
<td>0.845</td>
<td>0.676888</td>
<td>0.228197</td>
</tr>
<tr>
<td>0.903</td>
<td>0.732785</td>
<td>0.279982</td>
</tr>
<tr>
<td>0.954</td>
<td>0.781989</td>
<td>0.326862</td>
</tr>
<tr>
<td>1.000</td>
<td>0.825908</td>
<td>0.369658</td>
</tr>
<tr>
<td>1.041</td>
<td>0.865545</td>
<td>0.408901</td>
</tr>
<tr>
<td>1.079</td>
<td>0.901638</td>
<td>0.445050</td>
</tr>
<tr>
<td>1.114</td>
<td>0.934747</td>
<td>0.478509</td>
</tr>
<tr>
<td>1.146</td>
<td>0.965312</td>
<td>0.509627</td>
</tr>
<tr>
<td>1.176</td>
<td>0.993681</td>
<td>0.538697</td>
</tr>
</tbody>
</table>

Figure 4: \(\log(qH^q) \text{ v/s } \log(q)\) PSI20 & Brownian Motion

Figure 5 shows the variation in the Hurst exponent over time estimated using a 2-year and 3-year data period and a moving period of 1 month. The Hurst exponent shows small variations across time with values below and above 0.5 indicating the presence of multi-fractality. As far as dependence is concerned it shows both short-memory \((0 < H < \frac{1}{2})\) and long-memory \((\frac{1}{2} < H < 1)\).
4.2 Analysis of Simulations

The Hurst exponent shows a variation that largely depends on the size of the time period used for estimation. The lack of high frequency data prevents the analysis of the impact of time scaling on the Hurst exponent. The variation in the Hurst exponent is on account of the large volatility observed in the PSI:20 which affects the local trends. As the time period for estimating the Hurst exponent increases the trends tend to get affected by the volatility and hence impact the value of the Hurst exponent. The larger the volatility in the data series the lower is the value of the Hurst exponent.
5 Conclusion

We find that the PSI:20 exhibits a behaviour that has both long-memory and short-memory depending on the scale of the time period used and believe that an ARFIMA process would be better suited for estimation of the data generation process as opposed to ARCH, GARCH which do not account for memory.

6 References


