The long memory of volatility

Author:

Despoina Giatra

Supervisor:

Theodore Simos

UNIVERSITY OF
IOANNINA
DEPARTMENT OF ECONOMICS

Department of Economics

University of Ioannina

Greece

January,2020

The long memory of volatility

Author:

Despoina Giatra

Supervisor:

Theodore Simos

UNIVERSITY OF IOANNINA

DEPARTMENT OF ECONOMICS

Department of Economics

University of Ioannina

Greece

January,2020

Preface

I would like to thank my supervisor Professor Theodore Simos and Pofessor Spyridon Symeonides, for the excellent guidance and support during this process. Also, I would like to thank my family and my friends for their support.

Abstract

This dissertation provides a survey and review of the major models that help us analyse volatility, stochastic volatility, long memory stochastic volatility and fractional differencing.We also make a review on AR, MA, ARMA, ARIMA and ARFIMA models and other tools such as spectrum density that are needed to analyse the above. In section 4 we will show how the fractional differencing is connected with parameter d and long-term memory. Section 5 presents the empirical analysis, which shows whether long memory appears in the U.S.A. market and in particular in the S& P500, Dow Jones, Nasdaq and Russel 2000 indices. Furthermore, we see the volatility which was created by the Dot-com bubble and the financial crisis that have occurred in America and affected these four indices.

*Keywords:*Volatility; Stochastic volatility; Long memory stochastic volatility; Fractional differencing; ARFIMA; Spectrum density

Contents

List of Figures

List of Tables

1 Introduction

In this dissertation we analyse the long memory of stochastic volatility.Volatility is a rate at which the price of a security increases or decreases for a given set of returns. It is measured by calculating the standard deviation of the annualized returns over a given period of time. It shows the range to which the price of a security may increase or decrease. Volatility measures the risk of a security. It is used in option pricing formula to gauge the fluctuations in the returns of the underlying assets. Volatility indicates the pricing behaviour of the security and helps estimate the fluctuations that may happen in a short period of time.

We will focus on long memory in stochastic volatility. The financial variables that were used are serially uncorrelated and can be modelled by means of a GARCH class such as the one Engle (1982) proposed, and hypothesized that the conditional variance of the observations is an exact function of the squares of the previous ones. So, in the first section we analyse the definitions of volatility and stochastic volatility and at the second the long memory stochastic volatility models .

In the third section we will see the AR (auto-regressive) and MA (moving average) models and how these could lead us to an ARMA (auto-regressive moving average) model and then to an ARIMA and in the end how an ARIMA (auto-regressive-integrated moving average) model can be come an ARFIMA (autoregressive fractional-integrated moving average) using the fractional differencing. We will analyse the spectrum density and how this is connected with the ARFIMA models and the Hurst exponent that we analysed in the previous chapter. The forth section is about fractional differencing. We will show the connection with the parameter d and the long-term memory as represented by Hosking (1981) and how the d parameter connected with ARFIMA models and FIGARCH models. In the fifth section we present the empirical analysis using figures and tables in the U.S.A. market for the S&P500, Dow Jones, Nasdaq and Russel 2000 indices, whether they present long memory and how the crisis affected them. And in the sixth section, we will sum up with a review of everything presenting above.

2 Stochastic Volatility

In this section we will see the long memory in stochastic volatility. We will determine what volatility is and what a stochastic volatility model is.

2.1 Volatility and Stochastic volatility model

There are many definitions according to which we can define volatility.

The first proposed that volatility is a statistical measure of the dispersion of returns for a given security. It is often measured as either standard deviation or variance between returns from the same security or market index.

The second said that volatility is a rate which the price of security increases for a given set of returns. It is measured by calculating the standard deviation of the annualized returns over a given period of time. It is shows the range to which the price of a security may increase or decrease.

If we want to be more descriptive, volatility express the pricing behaviour of the security and helps estimate the fluctuations that may happen in a short period of time. Stochastic volatility models are those in which the variance of a stochastic process is itself randomly distributed.They are used in the field of mathematical finance to evaluate derivative securities.

To model serial correlation in volatility we often use GARCH models. Another way that Harvey mention is to consider the logarithm of σ^2 created by a linear stochastic process such as AR(1).

$$\sigma_t^2 = \sigma^2 exp(h_t) \tag{1}$$

where

$$h_{t+1} = \phi h_t + \eta_t,\ \eta \sim NID(0, \sigma^2),\ 0 \leq \phi \leq 1 \tag{2}$$

So, σ^2 is a scale factor, ϕ is a parameter and η is a disturbance term. This stochastic volatility model has two advantages, as Harvey(2007) noted. Firstly, the natural discrete time analogue of the continuous time model used in works on option pricing. Secondly, the statistical properties of a SV (stochastic volatility) model are easy to determine. The only disadvantage is that the maximum likelihood can be carried out by a computer intensive technique.

3 Long memory stochastic volatility model (LMSV)

In this section we will see what a long memory stochastic volatility model is and the dynamic properties of a LMSV (long memory stochastic volatility).

F. Jay Breidt , Nuno Crato and Pedro de Lima (1998) supported that the LMSV model constructed by ARFIMA process in a standard stochastic volatility scheme and the parameters can be estimated by a frequency domain likelihood estimator. In a LMSV model the U_t is a stationary long memory process, Y_t is covariance and it is stationary and because Y_t acquired from the properties of lognormal distribution is a white noise sequence. One characteristic of this model is the excess kurtosis of y_t. They defined the stochastic volatility model as:

$$y_t = \sigma_t \xi_t, \sigma_t = (u_t/2) \tag{3}$$

where u_t is independent of ξ_t, ξ_t is i.i.d. and $u_t$is an ARMA model. Also $E[y_t] = 0,$

$$Var(y_t) = exp\{\gamma(0)/2\}\sigma^2 \tag{4}$$

and $Cov(y_t, y_{t+h}) = 0, \ for \ \ h \neq 0 \ y_t^2$ is also covariance and stationary and also follows the properties of lognormal distribution.

$$E[y_t^2] = exp\{\gamma(0)/2\}\sigma^2 \tag{5}$$

$$Var(y_{t^2}) = \sigma^4[\{1 + Var(\xi_t^2)\{2\gamma(0)\} - exp\{\gamma(0)\}] \tag{6}$$

$$Cov(y_t^2, y_{t+h}^2) = \sigma^4\{\gamma(0) - \gamma(h)\} - exp\{\gamma(0)\}], forh \neq 0 \tag{7}$$

When the series transformed to the stationary process with E_t i.i.d., mean zero and variance σ^2

$$x_t = \log y_t^2 = \log \sigma^2 + E[\log \xi_t^2] + u_t + (\log \xi_t^2 - E[\log \xi_t^2]) = \mu + u_t + \varepsilon_t \tag{8}$$

A long-memory model for u_t is a fractionally integrated Gaussian noise which is defined by :

$$\gamma_x(h) = Cov(x_t, x_{t+h}) = \gamma(h) + \sigma_\epsilon^2 I_{h=o} \tag{9}$$

The spectral density , the ACVF (auto-covariance function) and the ACF (auto-correlation function) of u_t are given by the equations :

$$f(\lambda) = \frac{\sigma_\eta^2}{2\pi}|1 - e^{-i\lambda}|^{-2d}, \ -\pi \leq \lambda \leq \pi \tag{10}$$

$$\gamma(0) = \sigma_\eta^2 \Gamma(1-2d)/I^{-2}(1-d) \tag{11}$$

$$\rho(h) = \frac{\Gamma(h+d)\Gamma(1-d)}{\Gamma(h-d+1)\Gamma(d)}, h = 1, 2, \ldots \tag{12}$$

And as an ARFIMA (p,d,q) the u_t can be modelled like:

$$(1-B)^d \phi(B) u_t = \theta(B)\eta_t,\ \{\eta - t\} i.i.d, N(0.\sigma_\eta^2) \tag{13}$$

There are two dynamic properties of a long-memory stochastic volatility model. According to Harvey(2007) an auto-regressive stochastic volatility model y_t is a martingale difference and the stationarity of entail stationarity to y_t and follows the properties of lognormal distribution.

4 ARFIMA Process

In this section we will see some very important processes, the White Noise , the Moving Average (MA) in first and q^{th} order, the Auto-regressive (AR) in first and p^{th} order. Then we will see the mixed Autoregressive Moving Average (ARMA) processes and we will make a report in the auto-regressive-generating function. Afterwards,we will analyze the population spectrum and the power spectral density, we will see what they are and describe their properties.What follows is a brief description of Autoregressive-integrated moving average (ARIMA) processes and then an extensive description of autoregressive fractionally integrated moving average (ARFIMA) processes.

4.1 White Noise

chronological series is white noise if it has virtually no distinct shape or pattern. We can symbolized with ε_t . It is white noise if it has stable mean and usually equal to zero, stable variance and its values are uncorrelated.

$$E(\varepsilon_t) = 0 \,\forall\, t \;\; \gamma_0 = [E(\varepsilon_t)]^2 = [\sigma_t]^2 \,\forall\, t \;\; \gamma_\kappa = E(\varepsilon_t \varepsilon_{t-\kappa}) = 0 \,\forall\, t \, and \; \kappa \neq 0 \tag{14}$$

The white noise process is stationary. A basic characteristic is that the auto-covariance and the auto-correlation coefficients are equals to zero.

4.2 Moving Average (MA)

The moving average process of order q has the following form:

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \varepsilon + \ldots + \theta_q \varepsilon_{t-q} \tag{15}$$

With μ and θ as constants and E_t as white noise. Y_t : is the weighted average of the random errors of q previous periods and it's called moving average model of q orders and we can symbolize it as MA(q). The mean of the above equation could be given by:

$$E(Y_t) = \mu + E(\varepsilon_t)\zeta + \theta_1 E(\varepsilon_{t-1}) + \theta_2 E(\varepsilon_{t-2}) + \ldots + \theta_q E(\varepsilon_{t-q}) \tag{16}$$

And the variance of MA(q) is :

$$\gamma_0 = E(Y_t - \mu)^2 = E(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q})^2 \tag{17}$$

If we had a first order moving average process or MA(q), the Y_t is written:

$$Y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1} \tag{18}$$

The mean of Y_t:

$$E(Y_t) = E(\mu + \varepsilon_t + \theta\varepsilon_{t-1}) = \mu + E(\varepsilon_t) + (\varepsilon_{t-1}) = \mu \tag{19}$$

The variance:

$$\begin{aligned} E(Y_t - \mu)^2 &= E(\varepsilon_t + \theta\varepsilon_{t-1})^2 \\ &= E(\varepsilon_t^2 + 2\theta\varepsilon_t\varepsilon_{t-1} + \theta^2\varepsilon_{t-1}^2) \\ &= \sigma^2 + 0 + \theta^2\sigma^2 \\ &= (1+\theta^2)\sigma^2 \end{aligned} \tag{20}$$

Because the white noise has stable variance and its values don't autocorrelation and the variance of Y_t is finite and independent of time t, the MA(1) presents stability of mean and variance.

The first auto-covariance :

$$\begin{aligned} E(Y_t - \mu)(Y_{t-1} - \mu) &= E(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-1} + \theta\varepsilon_{t-2}) \\ &= E(\varepsilon_t\varepsilon_{t-1} + \theta\varepsilon_{t-1}^2 + \theta\varepsilon_t\varepsilon_{t-2} + \theta^2\varepsilon_{t-1}\varepsilon_{t-2}) \\ &= 0 + \theta\sigma^2 + 0 + 0 \end{aligned} \tag{21}$$

And the higher are all equal to zero. The auto-correlation function for a MA(1) has the form:

$$\rho_1 = \theta\sigma^2/[(1+\theta^2)\sigma^2] \tag{22}$$

Higher auto-correlations are equal to zero.

4.3 Autoregressive Process (AR)

The general form of an Auto-regressive (AR) model p^{th} order is:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t \tag{23}$$

Where ε_t is the white noise, ϕ and c are constants and $Y_{t-1}, \ldots$ are independent of orders.

The roots $1 - \phi_1 z - \phi_2 z^2 - \ldots - \phi_p z^p$ are equal to zero so a covariance-stationary has the follow form:

$$Y_t = \mu + \psi(L)\varepsilon_t \tag{24}$$

Where

$$\psi(L) = (1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p)^{-1} \tag{25}$$

And $\sum_{j=0}^{\infty} |\psi_j|$

The mean is:

$$Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2} - \mu) + \ldots + \phi_p(Y_{t-p} - \mu) + \varepsilon_t \tag{26}$$

The auto-covariances are:

$$\gamma_j = \begin{cases} \phi_1\gamma_{j-1} + \phi_2\gamma_{j-2} + \ldots + \phi p\gamma_{j-p} & for\ j = 1, 2, \ldots \\ \phi_1\gamma_1 + \phi_2\gamma_2 + \ldots + \phi_p\gamma_p + \sigma^2 & for\ j = 0 \end{cases} \tag{27}$$

Using the Yule-Walker equations the auto-correlations are:

$$\rho_j = \phi_1\rho_{j-1} + \phi_2\rho_{j-2} + \ldots + \phi_p\rho_{j-p}\ for\ j = 1, 2, \ldots \tag{28}$$

The first order auto-regressive process (AR(1)) given by the equation:

$$Y_t = c + \phi Y_{t-1} + \varepsilon_t \tag{29}$$

When the $|\phi| < 1$, exist a covariance- stationary process and the Y_t could be written:

$$\begin{aligned} Y_t &= (c + \varepsilon_t) + \phi(c + \varepsilon_{t-1}) + \phi^2(c + \varepsilon_{t-2}) + \ldots \\ &= [c/(1 - \phi)] + \varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \ldots \end{aligned} \tag{30}$$

Shows that a stationary AR could be written as equation of current error. This capacity of an AR(1) model to convert in a MA(∞) called invertibility and applied for all stationary AR models.

The mean for an AR(1) is written as :

$$\mu = c/(1-\phi) \tag{31}$$

The variance is:

$$\begin{aligned}\gamma_0 &= E(Y_t-\mu)^2 \\ &= E(\varepsilon_t+\phi\varepsilon_{t-1}+\phi^2\varepsilon_{t-2}+\ldots)^2 \\ &= (1+\phi^2+\phi^4+\ldots)\sigma^2 \\ &= \sigma^2/(1-\phi^2)\end{aligned} \tag{32}$$

The auto-covariance is:

$$\begin{aligned}\gamma_j &= E(Yt-\mu)(Y_{t-j}-\mu) \\ &= E[\varepsilon_t+\phi\varepsilon t-1+\phi^2\varepsilon_{t-2}+\ldots+\phi^j\varepsilon_{t-j}+\phi^{j+1}\varepsilon_{t-j-1}+\phi^{j+2}\varepsilon_{t-j-2}][\varepsilon_{t-j}+\phi\varepsilon_{t-j-1}+\phi^2\varepsilon_{t-j-2}+\ldots] \\ &= [\phi^j+\phi^{j+2}+\phi^{j+4}+\ldots]\sigma^2 \\ &= \phi^j[1+\phi^2+\phi^4+\ldots]\sigma^2 \\ &= [\phi^j/(1-\phi^2)]\sigma^2\end{aligned} \tag{33}$$

The auto-correlation function is:

$$\rho_j = \gamma_j/\gamma_0 = \phi^j \tag{34}$$

So, the ρ_j for stationary $|\phi| < 1$ AR(1) start for the unit and decreases geometrically and tending to zero as j grows.

4.4 Autoregressive Moving Average Processes (ARMA)

Auto-regressive Moving Average Processes (ARMA) interpret a large number of real data. They are based on the idea of a strong correlation between different series in economics, natural sciences and other scientific fields. The general form is:

$$Y_t = c+\phi_1 Y_{t-1}+\phi_2 Y_{t-2}+\ldots+\phi_p Y_{t-p}+\varepsilon_t+\theta_1\varepsilon_{t-1}+\theta_2\varepsilon_{t-2}+\ldots+\theta_q\varepsilon_{t-q} \tag{35}$$

C is a constant, $\phi_1\ldots,\phi_p$ coefficients of AR, $\theta_1,\ldots,\theta_q$ coefficients of MA and ε_t is the white noise. It is a regressive model with Y_t as dependent variable and interpretive the lags of p-order of Y_t and a number of previous errors (lags q-order of Et), and that's why it is called ARMA(p,q).

4.5 The Autocovariance-Generating Function

This is a function that is necessary in order to analyse the population spectrum. It is a sum of the jth auto-covariance multiplied by a number z raised in j^{th} power as we can see below:

$$g_Y(z) = \sum_{j=-\infty}^{\infty} \gamma_j z^j \tag{36}$$

The number z given from the below equation :

$$z = \cos(\omega) - i\sin(\omega) = e^{-i\omega} \tag{37}$$

Where $i = \sqrt{-1}$ and the ω is the radian angle.

If $z = e^{-i\omega}$ and divide the auto-covariance-generating function by 2π, the result gives the population spectrum that we will discuss in the next subsection.

4.6 Population Spectrum and the Power spectral density

As we have seen, the form of population spectrum of Y_t is given if the auto-covariance-generating function, when $z = e^{-i\omega}$,divided with 2π.

$$S_Y(\omega) = \frac{1}{2\pi} g_Y(e^{-i\omega}) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_j e^{-i\omega j} \tag{38}$$

Using the De Moivre's theorem we can write the e^{-i} as:

$$e^{-i\omega j} = \cos(\omega j) - i\sin(\omega j) \tag{39}$$

So we can rewrite the first equation like:

$$S_Y(\omega) = \frac{1}{2\pi} \sum_{-\infty}^{\infty} \gamma_j [\cos(\omega j) - i\sin(\omega j)] \tag{40}$$

As we now the covariance- stationary process $\gamma_j = \gamma_{-j}$, so the $S_Y()$ is equal :

$$S_Y(\omega) = \frac{1}{2\pi}\gamma_0[\cos(0) - i\sin(0)] + \frac{1}{2\pi}\sum_{j=1}^{\infty}\gamma_j[\cos(\omega j) + \cos(-\omega j) - i\sin(\omega j) - i\sin(-\omega j)] \tag{41}$$

From the trigonometry is known that:

$\cos(0) = 1$

$\sin(0) = 0$

$\sin(-\theta) = -\sin(\theta)$

$\cos-(\theta) = \cos(\theta)$

Using all of the above the $S_Y(\omega)$ is written :

$$S_Y(\omega) = \frac{1}{2\pi}\gamma_0 + 2\sum_{j=1}^{\infty}\gamma_j \cos(\omega j) \tag{42}$$

The calculation of the population spectrum for various processes such as:

$$S_Y(\omega) = (2\pi)^{-1}\sigma^2\psi(e^{-i\omega})\psi(e^{i\omega}) \tag{43}$$

For the MA(1) model if the white noise process $\psi(z) = 1 + \theta z$, the population spectrum is:

$$S_Y(\omega) = (2\pi)^{-1}\sigma^2[1 + \theta^2 + 2\theta\cos(\omega) \tag{44}$$

For an AR(1), the white noise is $\psi(z) = 1/(1 - \phi z)$ with $|\phi| < 1$ the population spectrum is:

$$\begin{aligned} S_Y(\omega) &= \frac{1}{2\pi}\frac{\sigma^2}{(1-\phi e^{-i\omega})(1-\phi e^{i\omega})} \\ &= \frac{1}{2\pi}\frac{\sigma^2}{(1-\phi e^{-i\omega}-\phi e^{i\omega}+\phi^2)} \\ &= \frac{1}{2\pi}\frac{\sigma^2}{[1+\phi^2-2\phi\cos(\omega)]} \end{aligned}$$

(45)

For an ARMA(p,q) The population spectrum is defined as :

$$S_Y(\omega) = \frac{\sigma^2\prod_{j=1}^{q}[1+\eta_j^2 - 2\eta\cos(\omega)]}{2\pi\prod_{j=1}^{p}[1+\lambda_j^2 - 2\lambda_j\cos(\omega)]} \tag{46}$$

With the population spectrum we can calculate the auto-covariances. James D.Hamilton(1994) has a proposition in which $\{\gamma_j\}_{j=-\infty}^{\infty}$ be an absolutely summable sequence of auto-covariances, and define $S_Y(\omega)$ Then

$$\int_{-\pi}^{\pi} S_Y(\omega) e^{i\omega k} \, d\omega = \gamma_k \tag{47}$$

or

$$\int_{-\pi}^{\pi} S_Y(\omega) \cos(\omega k) \, d\omega = \gamma_k \tag{48}$$

The power spectral density is used, for example, for stationary processes. The power spectral density describes how the power of time series is divided over frequency. It is given by the function:

$$S_{xx}(\omega) = \lim_{T\to\infty} E[|\hat{x}(\omega)|^2] \tag{49}$$

Where E is the expected value and $x(\omega)$ is the signal or a time series. The auto-correlation function is :

$$R_{xx}(\tau) = \langle X(t)X(t+\tau)\rangle = E[X(t)X(t+\tau)] \tag{50}$$

Where X(t) is the complex-valued. And the spectral density can be written:

$$S_{xx}(\omega) = \int_{-\infty}^{\infty} R_{xx}(\tau) e^{-i\omega\tau} \, d\omega = \hat{R}_{xx}(\omega) \tag{51}$$

4.7 Autoregressive-integrated moving average (ARIMA) processes

The ARIMA model has three parameters :

$$\Phi(B)(1-B)^d y_t = \mu + \Theta(B)\varepsilon_t \tag{52}$$

The first, is the parameter of p which is for the auto-regressive terms of lags: $\Phi(B) = 1 + \rho_1 B + \rho_2 B^2 - \ldots - \rho_p B^p$ The second, is the parameter of q , for the moving-average terms of lag $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_q B^q$ The third is the parameter d which express the order of differencing.

4.8 Hurst Exponent

The Hurst exponent according to Bo Qian and Khaled Rasheed (2005) is a statistical measure used to classify time series of financial data in different periods. The estimation of Hurst exponent requires the reset of the rescaled range in the time span of a time series.If the parameter H is equal to 0.5 then we have a random series and if the H is more than 0.5 then we have a reinforcing series. The rescaled range is a method that is used to define the Hurst exponent.

$$E[\frac{R(n)}{S(n)}]n \longrightarrow \infty = C_{n^H} \tag{53}$$

E(.) is the value of observations ,R(n) is the width of the first n values,S(n) standard deviation and C is the constant.

4.9 Autoregressive Fractionally Integrated Moving Average (ARFIMA) processes.

The ARFIMA model estimates the fractional differencing parameter by Maximum Likelihood. An ARFIMA model is the generalization of an ARIMA and ARMA model as described by Kai Liu, Yang Quan Chen and Xi Zhang (2017). It can capture short-range and long-range dependence but it gives better results in LRD (long-range dependence). The LRD is based on Hurst's analysis as we see above and in ARFIMA and FIGARCH models which were created to analyze this process. It is also could be defined by the autocorrelation function analysis or ACF :

When $x(t); t \in (-\infty, \infty)$

$$\rho_k = \frac{Cov(x(t), x(t-k))}{Var(x(t))} \tag{54}$$

An AR and a MA models can be written as:

$$(1 - \sum_{i=1}^{p} \phi_i B^i)(1-B)^d(x(t) - \mu) = (1 + \sum_{i=1}^{q} \theta_i B^i)\varepsilon_t \tag{55}$$

$(1-B)^d$ is the difference operator ∇^d

$$(1-B)^d = \sum_{k=o}^{\infty} \binom{d}{k} (-B)^k = \sum_{k=o}^{\infty} \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d+1-k)} (-B)^k \tag{56}$$

The ARFIMA (p,d,q) could be defined as:

$$\Phi(B)(1-B)^d x(t) = \Theta(B)\varepsilon_t \, , d \in (-0,5,0,5)$$

where p denotes the auto-regressive order, q denotes the moving average order and d defined as the differencing parameter Jin Xin and Yao Jin (2007) proposed three steps to create an ARFIMA model. Firstly, the long-term memory in a time series must be analyzed and the fractional differencing parameter d must be defined. Secondly, the fractional difference of the series is found and an ARMA process with a mean that equals to zero is used. Lastly, parameters p and q are determined.

5 Fractional Differencing

In this section we will see what is the fractional differencing when we use it and the connection with ARFIMA and FIGARCH models.

Fractional derivative offers the knowledge for description of memory and hereditary properties of various material and processes including natural and physical phenomena. The theory of fractional derivatives used in fractal theory, theory of control of dynamic systems, diffusion processes and many others. Richard Pierse argued that the fractionally integrated models could be stationary or non-stationary. The fractional integrated process can lead from an ARMA to an ARFIMA model and conditional volatility models to fractionally integrated GARCH models and fractionally integrated stochastic volatility models. The fractionally integrated models have the following form:

$\triangle_t^d = (1-B)^d y_t = u_t$

The meaning of fractional values d is been given by the Binomial Theorem. It is that the $(1-B)^d$ could be explained as an infinite series for $d > -1$.

$\triangle^d = (1-B)^d = \sum_{k=0}^{\infty} d(-d)^k$

Where $\binom{d}{k} = \frac{d(d-1)(d-2)...(d-k+1)}{k!}$

$k! = 123\cdots(k-1)k$ with $0! = 1$

So, $(1-B)^d = 1 - dB + \frac{d(d-1)}{2!}B^2 - \frac{d(d-1)(d-2)}{3!}B^3 + \ldots$

According to Hosking (1981) the fractionally differenced processes demonstrate long -term persistence and anti-persistence and also shows the connection between parameter d and long term memory.

To continue, we analyse the derivation of fractionally differenced which is given by the Brownian motion, a continuous time stochastic process B(t) with independent Gaussian steps and spectral density equal to ω^{-2} The derivative of this stochastic process is the continuous-time white noise process and with constant spectral density . This has three basic properties:

1. The parameter H (0<H<1) in fractional Brownian motion is ($\frac{1}{2}$-H)th derivate
2. The spectral density is ω^{-2H-1}
3. The covariance function is $|k|^{2H-2}$

Hosking (1981) proves that an ARIMA (0,d,0) process with d, 0<d<$\frac{1}{2}$, is a long-memory stationary process. When d is a real value and be in an interval $[-\frac{1}{2}, \frac{1}{2}]$ is stationary and invertible, but if d is equal to $\pm\frac{1}{2}$ then can't be both. To modelled the ARIMA (0,d,0),combine the fractional differencing with the family of Box-Jenkins models and that shows the effect on the d parameter which is that the distant observations decays when the lag increases and the parameters ϕ and θ decay exponentially.

To estimate the d parameter we can estimate it as part of maximum likelihood estimation of the parameters of the ARFIMA model.

$$\log L(y; d; \phi; \theta; \Sigma) = -\frac{T}{2}\log 2\pi - \frac{1}{2}\log|\Sigma| - \frac{1}{2}y'\Sigma^{-1}y \tag{57}$$

In the long memory models we use the FIGARCH model to analyse better the variance of the financial data series as Richard Pierse proposed.

The GARCH (1,1) model

$$\sigma^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{58}$$

The GARCH(1,1) can be expressed as an ARMA process in

$$u_t^2 = \alpha_0 + (\alpha_1 + \beta_1)u_{t-1}^2 + \varepsilon_t - \beta_1 \varepsilon_{t-1}$$

(59)

where

$$\varepsilon_t = u_t^2 - \sigma_t^2 \tag{60}$$

The FIGARCH (1,d,1) model is defined as follows

$$\Delta^d u_t^2 = \alpha_0 + (\alpha_1 + \beta_1)\Delta^d u_{t-1}^2 + \varepsilon_t - \beta_1 \varepsilon_{t-1} \tag{61}$$

or can be written as:

$$\sigma_t^2 = \alpha_0 + (1 - \Delta^d)u_t^2 - (\beta_1 - \alpha_1 + \beta_1)\Delta^d)u_{t-1}^2 + \beta_1 \sigma_t^2 \tag{62}$$

where $\alpha_0 > 0$, $\alpha_1 + d \geq 0$ and $1 - 2(\alpha_1 + \beta_1) \geq 0$.

6 Empirical Results

In this section I estimate the ARFIMA(p,d,q) model for the indicators SP500, Dow Jones, Nasdaq and Russel 2000. I use daily data from 31 December 1992 to 31 December 2018 (6547 observations). The data are derived from Yahoo Finance.

First of all, I describe the history of the four indicators which are in the market of the United States of America. In the United States of America there are two known crises that affected these indicators. The first, is the Dot-com bubble or else known as the internet bubble and it was a rapid rise of U.S technology stock equity valuations fuelled by investments in internet-based companies in the 1990s. The value of equity markets grew very fast. The second, is the financial crisis in 2007-2008. It began in 2007 with a crisis in subprime mortage market in the U.S.A. and it caused as a result the international banking crisis with the breakdown of Lehman Brothers on September 15 2008.

The first indicator is the S&P500 which is the abbreviation of Standard Poor's 500 index, which is the biggest index for the biggest (valuable)American corporations for the markets NYSE and NASDAQ. The index includes 500 corporations and captures about 80% of the available market capitalization. It includes Apple, Microsoft and Exxon as the biggest corporations. The second index is Dow Jones, which is the most well known and important index.Its full name is Dow Jones Industrial Average (DJIA). The owner is the Dow Jones corporation and created by Charles Dow. Important contributions had been made by the statitian Edward Jones. Both indices, track the stock prices of the thirty companies which are selected by the editors of the Wall Street Journal and include some of the world's most prominent companies. For example, Apple, Nike, Microsoft, Disney, Coca-Cola and American Express. The Dow Jones index, differentiates the weight of each stock in the index, according to its nominal value, which means that companies that have not split in their stock, have bigger impact on the index. This index, first calculated in May 1896 and has been referenced in the financial markets and a point of reference the political economy around the world. The third index is Nasdaq or NASDAQ Composite and is one of the most well-known stock market indices in the U.S.A. stock markets. Because of its composite it is heavily weighted towards information technology companies. Furthermore, Nasdaq was affected by the Dot-com bubble. In March 10 2000 it peaked at 5,132.52 but after April 17 of the same year it fell and continued to fall for thirty more months over 78% from its peak in March. Moreover, it was affected by the financial crisis of 2007-2008 where it dropped in September 2008 almost 200 points, since the tech bubble burst, losing 9.14%. Finally, the fourth index is Russell 2000. It measures the performance of 2000 smallest- cap American companies. It is a market-cap weighted index. It was created in 1984 by Frank Russsell Company. Investors compare small- cap mutual fund performance with the Russell 2000 index because it reflects the return opportunity presented by the entire sub-section of that market rather than the opportunities offered by narrower indices, which may contain biases or more stock-specific risk that distort a fund manager's performance.

In table 1 we can see the descriptive statistics of returns. The number of observations are 6547. The standard deviation shows how deviations deviate from the mean. In the results the standard deviations are low so, they show low volatile. Furthermore, we observe that they have negative skewness and from the kurtosis, that they are platykurtic. Using the Jarque Bera test we can see that we do not have normal distribution. Finally, the variables are statistically significant, at the significance level of 1%.

Figure 1: Stock Prices of SP500, Dow Jones, Nasdaq and Russel 2000

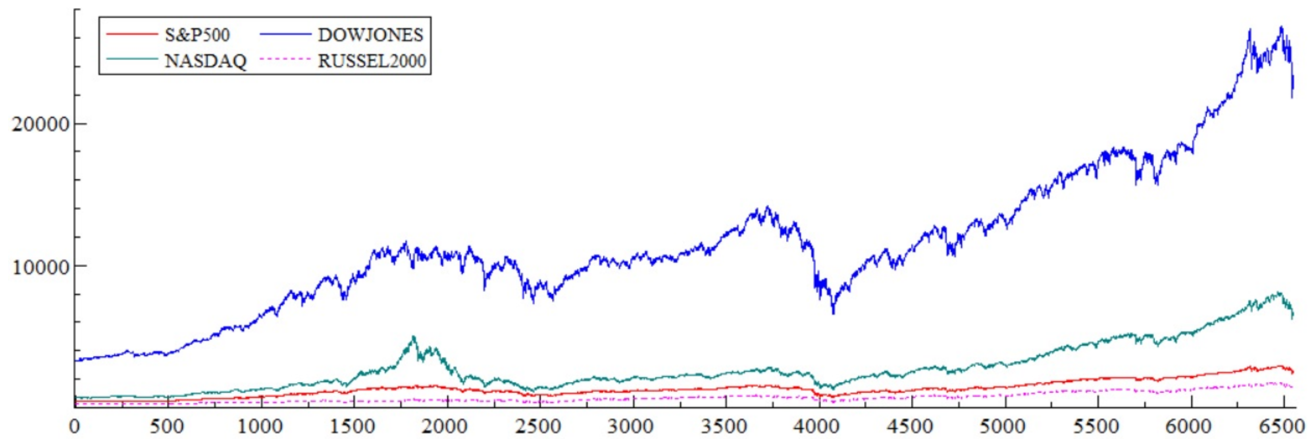


Table 1: Descriptive Statistics of Returns

Variable	OBS	min	mean	max	std.dev	Skewness	p-value	Kurtosis	p-value	Jarque-Bera	p-value
S&P500	6547	0.095	0.000	0.110	0.011	-0.268	0.000	8.845	0.000	21419	0.000
DOWJONES	6547	0.082	0.000	0.105	0.011	-0.197	0.000	8.351	0.000	19067	0.000
NASDAQ	6547	0.102	0.000	0.133	0.015	-0.095	0.002	6.127	0.000	10249	0.000
RUSSEL2000	6547	0.126	0.000	0.089	0.013	-0.361	0.000	6.048	0.000	10121	0.000

In table 2 we can see the descriptive statistics of squared returns. In this table the number of observations are the same, 6547. The standard deviation is very small. That means that the variables have low volatile. The skewness in this table is positive and the kurtosis is leptokurtic. From the Jarque Bera test we see that we do not have normal distribution. And finally, the variables are statistically significant, at the significance level of 1%.

Table 2: Descriptive Statistics of squared Returns

Variable	OBS	min	mean	max	std.dev	Skewness	p-value	Kurtosis	p-value	Jarque-Bera	p-value
S&P500	6547	0	1E-04	0.012006	0.00042248	13.108	0.0000	259.24	0.0000	1.85E+07	0.0000
DOWJONES	6547	0	1E-04	0.011043	0.00037344	13.385	0.0000	280.99	0.0000	2.17E+07	0.0000
NASDAQ	6547	0	2E-04	0.017569	0.0006313	10.031	0.0000	166.22	0.0000	7.65E+06	0.0000
RUSSEL2000	6547	0	2E-04	0.015911	0.00051593	11.03	0.0000	208.8	0.0000	1.20E+07	0.0000

Table 3: ARFIMA (p,d,q) Estimation of SP500, Dow Jones, Nasdaq and Russel 2000

S&P500	Coefficient	Std.Error	t-prob
d parameter	0.3211	0.0130	0.0000
AR-1	-0.4757	0.0486	0.0000
MA-1	0.1945	0.0613	0.0010

DowJones	Coefficient	Std.Error	t-prob
d parameter	0.3483	0.0112	0.0000
AR-1	-0.3300	0.0145	0.0000

Nasdaq	Coefficient	Std.Error	t-prob
d parameter	0.3178	0.0110	0.0000
AR-1	-0.2559	0.0150	0.0000

Russel 2000	Coefficient	Std.Error	t-prob
d parameter	0.3543	0.0109	0.0000
AR-1	0.2784	0.0148	0.0000

Figure 2: Actual and fitted Values, ACF and PACF and Spectral Density of SP500

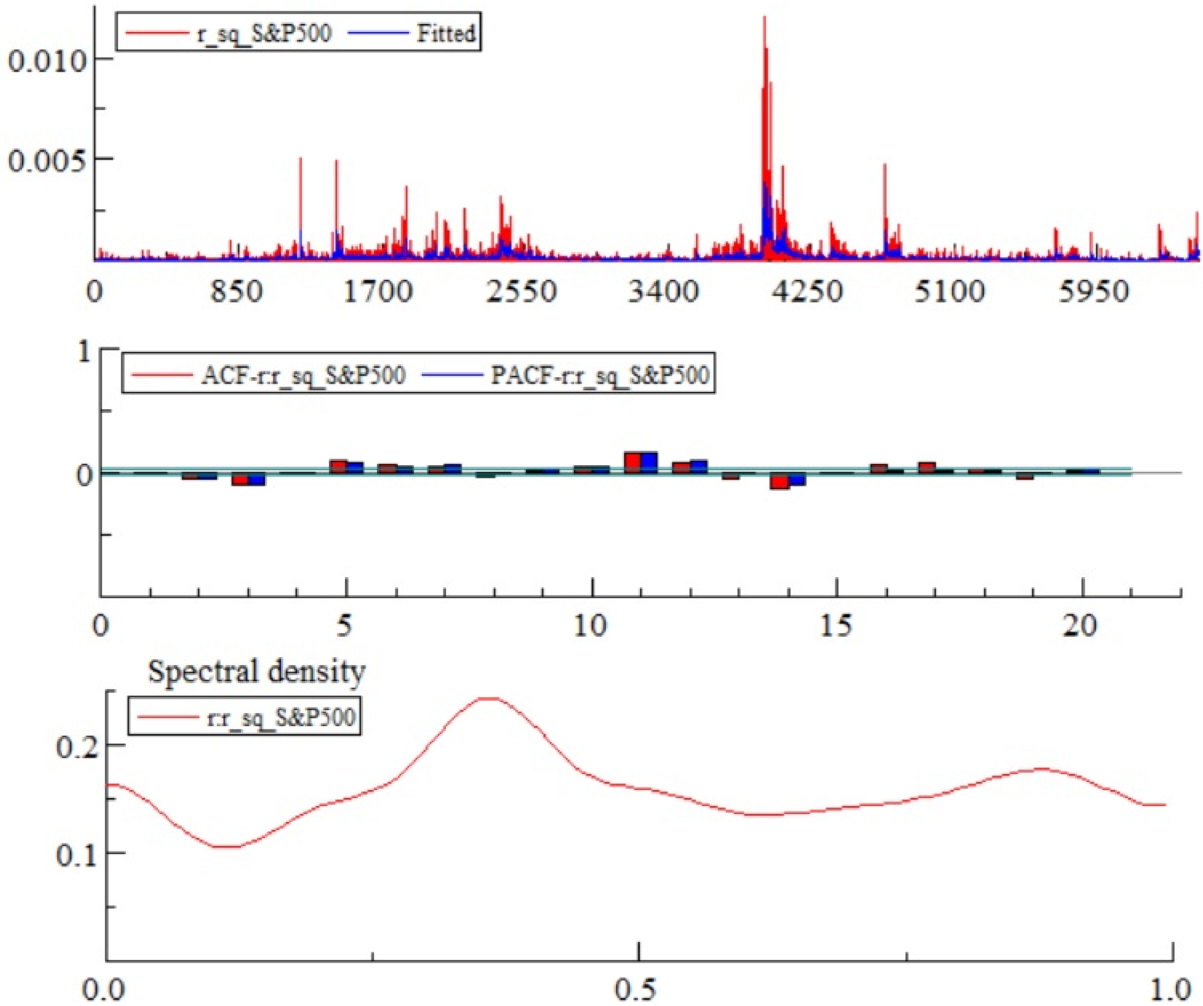


Figure 3: Actual and fitted Values, ACF and PACF and Spectral Density of DowJones

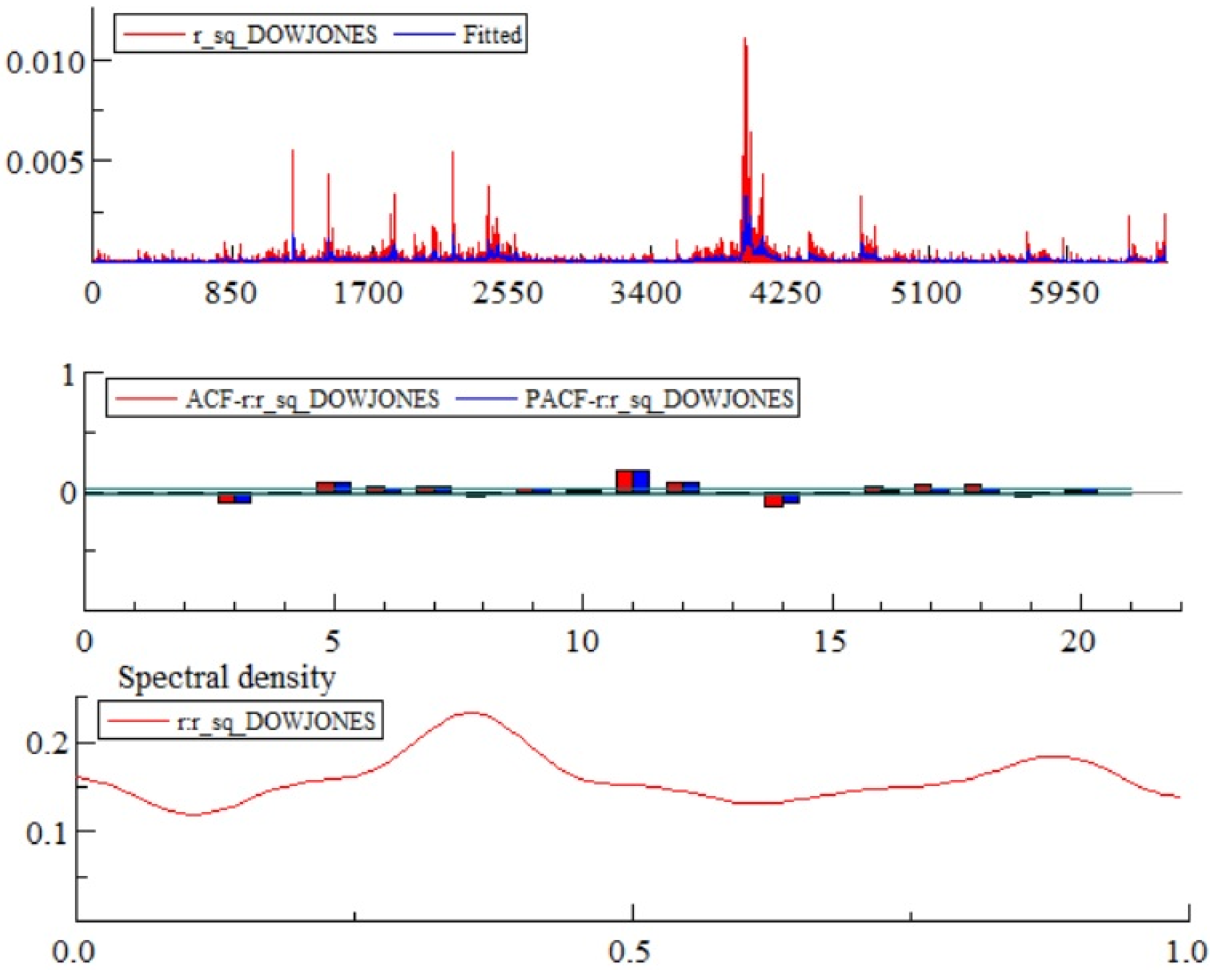


Figure 4: Actual and fitted Values, ACF and PACF and Spectral Density of Nasdaq

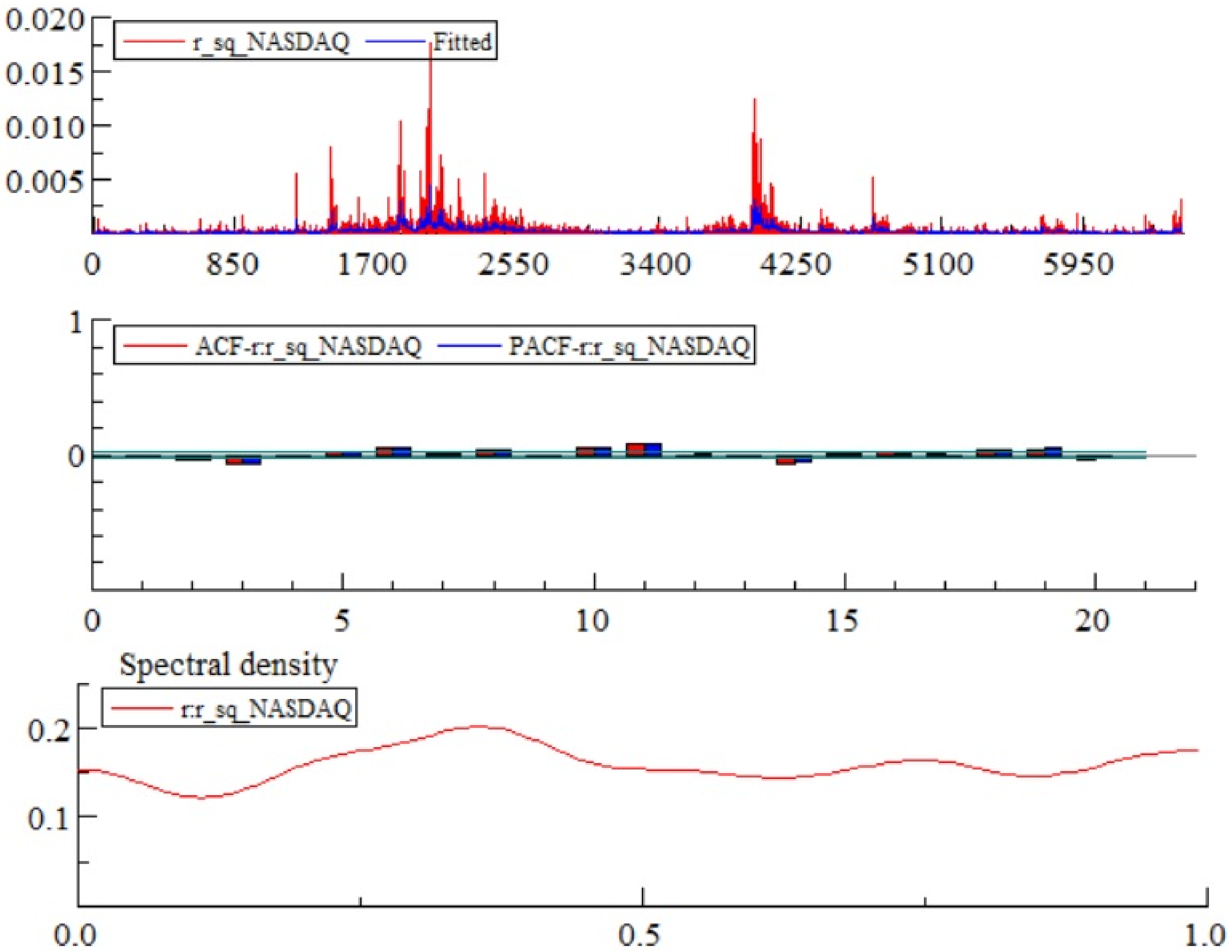
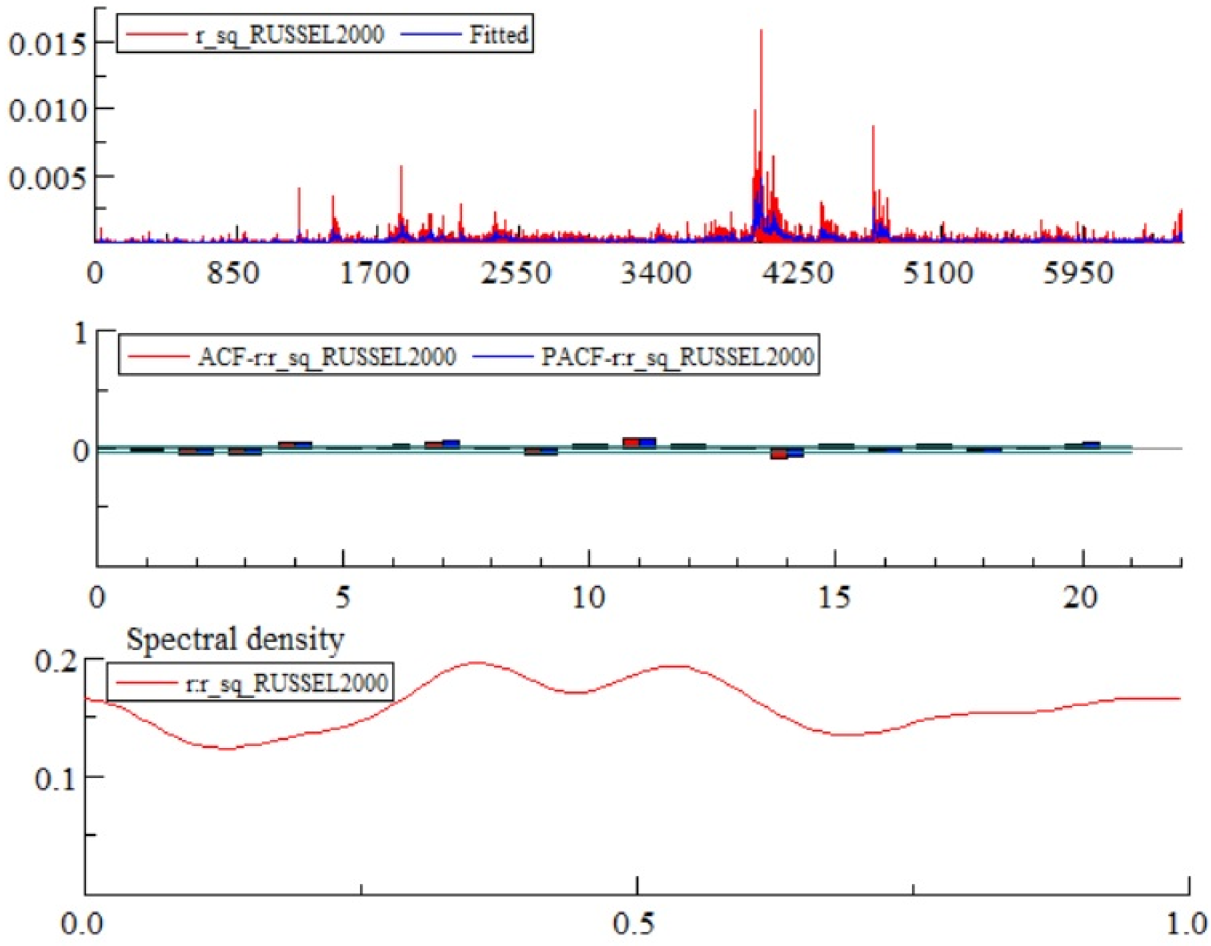


Figure 5: Actual and fitted Values, ACF and PACF and Spectral Density of Russel 2000



In table 3 we can see the results of the estimation of the ARFIMA(p,d,q) for each index.

1)For the S&P500 we use an ARFIMA(1,d,1)

2)For the DowJones we use an ARFIMA(1,d,0)

3)For the Nasdaq we use an ARFIMA(1,d,0)

4)For the Russel2000 we use an ARFIMA(1,d,0)

For persistent, d, is between (0 , 0,5) in the space that is specified by long memory. The parameter d and the AR(1) and MA(1) in each index are statistically significant, at the significance level of 1%. The specialization was made by the information criterion, AIC. We use low lags because all the empirical analysis in the bibliography use models like these. In figures 2,3,4 and 5 we see the actual and fitted values which show the volatility of the values of S&P500, Dow Jones, Nasdaq and Russel 2000. They show the volatility that happened in periods 1999-2000 and 2007-2008 which corresponds to the two big crisis the Dot-com bubble and the Financial crisis in the U.S.A. that we analyse above. Finally in the same figures we can see the spectral density. As we can see, the frequencies that were generated correspond to the volatility which was created by the values in the above diagrams. In figures 1,2 and 3 the high frequencies are at 0,4 but in the last figure they are at 0,4 and 0,6 .

7 Conclusion

In this dissertation we analysed what is volatility and stochastic volatility and more specifically the long memory stochastic volatility models. We make a review on AR, MA, ARMA, ARIMA and ARFIMA models as well as other tools needed to analyse the above. Fractional differencing was also analysed to determine the parameter d. We then examined whether long memory appears in the US market and in particular in the indices S&P500, Dow Jones, Nasdaq and Russel 2000.The empirical analysis was performed where the four indices, S&P500, Dow Jones, Nasdaq and Russel 2000, exhibited long memory as the parameter d was within the necessary limits. Also in the US market, where these four indices belong, we see the volatility which was created by the two major crises, the Dot-com bubble and the financial crisis that have occurred in America and affected them.

8 References

Agiakloglou, C., Newbold, P., and Wohar, M., 1993, Bias in an estimator of the fractional difference parameter, *Journal of Time Series Analysis*, 14(3), 235-246

Andersson, M. K., and Gredenhoff, M. P. ,1997, Bootstrap testing for fractional integration, *Stockholm School of Economics*, the Economic Research Inst

Baillie, R. T., 1996, Long memory processes and fractional integration in econometrics, *Journal of Econometrics*, 73(1), 5-59

Barkoulas, J. T., Baum, C. F., 1997, Fractional Differencing Modeling and Forecasting of Eurocurrency Deposit Rates, *Journal of Financial Research*, 20(3), 355-372

Breidt, F. J., Crato, N., and de Lima, P., 1998, The detection and estimation of long memory in stochastic volatility, *Journal of Econometrics*, 83(1-2), 325-348

Brockwell, P. J., and Davis, R. A., 1991, *Time Series: Theory and Methods: Theory and Methods*, Springer-Verlag

Ellis, C., 1999, Estimation of the ARFIMA (p, d, q) fractional differencing parameter (d) using the classical rescaled adjusted range technique ,*International Review of Financial Analysis*, 8(1), 53-65

Engle, R. F. 1982, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50(4), 987

Hamilton, J. D., 1994, *Time series analysis* (*Vol. 2, pp. 690-696*), Princeton, NJ: Princeton university press

Harvey,A.C., 2007, *Long memory in stochastic volatility* ,Forecasting volatility in the financial markets,351-363 Hosking, J. R. M., 1981 , Fractional differencing , *Biometrika*, 68(1), 165-176

Korkmaz, T., Cevik, E. I., Ozatac, N., 2009, Testing for long memory in ISE using ARFIMA-FIGARCH model and structural break test

Liu, K., Chen, Y., Zhang, X., 2017, An Evaluation of ARFIMA (Autoregressive Fractional Integral Moving Average) Programs, *Axioms*, 6(4), 16

Mandelbrot, B. B., Van Ness, J. W., 1968, Fractional Brownian motions, fractional noises and applications, *SIAM review*, 10(4), 422-437

Newbold, P., Agiakloglou, C.,1993, Bias in the sample autocorrelations of fractional noise, *Biometrika*, 80(3), 698-702 Pierse,R.G.,Long memory and fractional integration, *Financial Econometrics*

Qian,B., and Rasheed, K., 2005, Hurst Exponent and Financial Market Predictability

Rachev, S. T., Mittnik, S., Fabozzi, F. J., Focardi, S. M., 2007, *Financial econometrics: from basics to advanced modeling techniques* (*Vol. 150*), John Wiley Sons

Raimundo, M. S., Okamoto Jr, J., 2018, Application of Hurst Exponent (H) and the R/S Analysis in the Classification of FOREX Securities, *Int. J. Model. Optim*, 8, 116-124
Robinson, P. M.,2001, The memory of stochastic volatility models, *Journal of econometrics*, 101(2), 195-218
Robinson, P. M. (Ed.), 2003, *Time series with long memory*, Advanced Texts in Econometrics
Xiu, J., Jin, Y., 2007, Empirical study of ARFIMA model based on fractional differencing, *Physica A: Statistical Mechanics and Its Applications*, 377(1)

Links:
https://www.investopedia.com/terms/v/volatility.asp
https://www.investopedia.com/terms/d/dotcom-bubble.asp
https://en.wikipedia.org/wiki/S%26P_ 500_ Index
https://en.wikipedia.org/wiki/Russell_ 2000_ Index
https://en.wikipedia.org/wiki/Dow_ Jones_ Industrial_ Average
https://en.wikipedia.org/wiki/Nasdaq
https://en.wikipedia.org/wiki/Dot-com_ bubble
https://en.wikipedia.org/wiki/Financial_ crisis_ of_ 2007%E2%80%9308
https://en.wikipedia.org/wiki/Spectral_ density# Power_ spectral_ density
https://en.wikipedia.org/wiki/Hurst_ exponent
https://en.wikipedia.org/wiki/Stochastic_ volatility