STOCHASTIC MODELING OF STOCK EXCHANGE MARKETS: THREE ESSAYS

YAVUZ YILDIRIM

JANUARY, 2021

STOCHASTIC MODELING OF STOCK EXCHANGE MARKETS:

THREE ESSAYS

YAVUZ YILDIRIM

DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILISOPHY

IN

DEPARTMENT OF FINANCIAL ECONOMICS

YEDİTEPE UNIVERSITY

JANUARY, 2021

PLAGIARISM

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Date: 25/02/2021

Name/Surname: YAVUZ YILDIRIM

Signature :

ÖZET

Bu tez çalışmasında, BIST100 ve büyük borsa endekslerinin günlük kapanış fiyatlarının zaman serilerini analiz etmek için stokastik modelleme uygulanmıştır. Yatırımcılar ve portföyleri için yeni uygulanabilir araçlar ve sonuçlar hedeflenmiştir. Çalışma üç bölümden oluşmaktadır. İlk makalede ağır kuyruklu dağılımların genelleştirilmiş otoregresif koşullu heteroskedastisite (GARCH) modeli ve sürekli zamanlı, COGARCH, BIST100 için Meixner dağılımlı modeli ile uygulanabilirliği tanıltıdı. İkinci bölümde, makale de beş büyük borsadaki uzun bellek özelliğini farklı modeller göz önünde bulundurarak inceliyoruz ve otoregresif kesirli olarak entegre hareketli ortalama (ARFIMA) modelinin modellemede fraksiyonel olarak entegre genelleştirilmiş otoregresif koşullu heterossedastikten (FIGARCH) daha iyi bir aday olduğunu gösteriyoruz. Son makale, DAX ve NIKKEI hisse senedi endeksleri arasında son zamanlarda biraz popülerlik kazanan dalgacık dönüşümüne odaklanıyor ve ikisi arasında bir miktar korelasyon ve tutarlılık olduğunu gösteriyor. Hurst değerleri aylık olarak tahmin modellemesi yapıldı ve zaman serilerinde multifraktal süreç belirtileri gözlemlendi.

Anahtar Kelimeler:stokastik modelleme, COGARCH, Wavelet

ABSTRACT

In this study, stochastic modelling is applied to analyze time series of daily closing prices of BIST100 and major stock exchange indices to gain new insight and help develop new applicable tools for investors and market participants alike for their portfolios and investments. The study is comprised of three parts. In the first essay we present the applicability of heavy-tailed distributions with the generalized autoregressive conditional heteroskedasticity (GARCH) model and continuous-time, COGARCH, model with Meixner distribution for BIST100. In the second part, we examine the long memory property of five major stock exchanges by considering different models and show that the autoregressive fractionally integrated moving average (ARFIMA) model is a better candidate than the fractionally integrated generalized autoregressive conditional heteroscedastic (FIGARCH) in modelling volatility of the stock indices studied. The final essay focuses on wavelet transformation, which has gained some popularity recently, between DAX and NIKKEI stock indices, and show that there are some level of correlation and coherency between the two. The Hurst exponent also estimated and there exist signs of multifractal process in the time series.

Keywords:stochastic modelling, COGARCH, Wavelet

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my academic advisor Prof. Dr. Gazanfer Ünal for his guidance through each stage of the process of my Ph.D research, and I am grateful to him for taking the time to share his extensive knowledge. I am also very grateful to the members of my thesis committee: Prof. Dr. Burak Saltoğlu and Doç. Dr. Sema Dube, for their insightful comments and encouragements.

I would also like to express my gratitude to my family, especially my wife for her unwavering support throughout writing this thesis.

Table of Contents

List of Tables

List of Figures

1. Introduction

The aim of this research study is to analyze financial time series of major stock exchange indices and BIST100 to give new insights to researchers and provide new applicable tools for investors and risk managers to employ for risk management. To achieve this goal, it is essential to adopt mathematical frameworks and models that are able to capture the stylized facts or the most significant empirical features of the stock markets that are of interest.

Perhaps one of the most widely used measure of risk, if not the most popular, is volatility of returns. Numerous studies have shown that there exists considerable degree of autocorrelation in squared returns of many financial time series and clustering of volatility in time space. In Chapter 2 we will study the BIST100 daily returns, where we employ the generalized ARCH processes to model the volatility of the index. In our methodological approach, we first assumed a normal distribution for the innovations, but then sought a more appropriate model to capture the frequency of extreme events, and found that Meixner distribution is a more appropriate description for BIST100. To the best of author's knowledge, this is the first attempt to model BIST100 volatility with Meixner driven GARCH process.

In the second study, which is discussed in Chapter 3, we build on our work in Chapter 2 and investigate the existence of persistence or long-range dependence in BIST100 returns, and extend this analysis to four major stock indexes. What sets this study apart from some of the earlier works is that we apply a range of persistency tests rather than focus on a single one, and adopt fractionally integrated GARCH processes to capture the persistence in conditional variance, as well as the ARFIMA process to model persistency in the first-order moment (mean). We also quantify the predictive powers of our models, based on the historical data set we have.

In our third and final study, our contribution to the literature is on the cross-correlation analysis of the two major stock indices located in Europe and Asia using Wavelet coherence and Hurst analysis with the aim to develop a framework for Hurst forecast. In all three studies we relied on widely available statistical and computational software applications for our analysis and computations, for the most part, the *R* statistical programming language.

1.1. Background

The financial incentive for money managers and individual investors to maximize return, while at the same time, minimize risk when putting capital into publicly traded assets such as company stocks and indexes (via ETFs or funds), is arguably the main reason to analyze financial time series with new mathematical models and tools. Hence the investors are always looking forward to be able to predict or even obtain a small hint on the next index value. The size of capital that is traded in the stock exchanges are enormous which drives the investors and traders to get hold of a key information towards predicting future market movements and, or directions. To be able to predict market movements, there is a need for mathematical models to capture and process the markets' underlying properties. Having a good candidate model for a stock exchange market becomes very meaningful as this would help investors to measure investment risks and make more informed decisions on their investments.

Minimizing the financial risk has always been one of the main focuses of investment theories and practices. In 1952, Markowitz [142] pioneered the modern portfolio theory (MPT) in his mean-variance model, since then researches on financial risk and other related issues depended on mathematical methods, amongst the most prominent ones are Efficient Market/Random Walk Theory (EMH), Capital Asset Pricing Model (CAPM), Capital Structure Theory (CST) and Arbitrage Pricing Theory (APT). Other than EMH and CST, the rest are associated with volatility of asset prices.

P. A. Samuelson and E. F. Fama developed the EMH over a half-century ago and the theory held up to be accurate by many practitioners of economics for a long time. The theory states that new information becomes accessible to everyone participating in the markets at almost zero cost and that the participants react to information rationally,competing against one another in predicting the future value of financial instruments and assets to make profit. In other words, the securities market is very efficient in incorporating historical and current information into existing security prices. And it is practically impossible for individuals to perform better than the overall market when trading based on news (or information) alone, and that to obtain higher returns one must increase their risk of capital such as by investing in riskier securities. Consequently, given that security prices already take the current information into account, the next change in price can only be driven by the arrival of future news, which, by its very definition, is not predictable. Following this argument, it is fair to say that the EMH theory is akin to the random walk theory, which states that security price moves are random in nature and therefore must be unpredictable. This is demonstrated by Kendall [102], Roberts [164], Osborne [152], [153] and Samuelson [168] who modified Bachelier's model and showed that the future stock price is unpredictable. Bachelier's work is known as Arithmetic Brownian Motion that defined as the price of securities follow a Brownian motion. Whereas Samuelson's model states that the rate of returns follow Brownian motion which is known as Geometric Brownian Motion. Samuelson argued that security prices follow a log-normal distribution rather than normal distribution.

Weiner [202], Kolmogorov [111] and Ito [95] developed their studies through Bachelier's works. Bachelier's work is remarkable in the sense that he appears to be the first to provide mathematical rigour to value security warrants (options), and the model's association to the theory of diffusion.

Samuelson's GBM approach in modeling stock prices became the foundation that the works of Robert Merton, Fisher Black and Myron Scholes based their option pricing formulation, which has since become the first widely used model to value options. Based on these fundamental frameworks, significant amount of academic work has been carried out in development of more complex models to better understand and model the dynamics of asset prices. And with growth of global financial markets and increasing availability of financial data, development and application of time series analysis tools gained popularity to test the accuracy of these theories.

1.2. Outline

Chapter 2; Discrete-Time And Continuous-Time Modelling of Istanbul Stock Exchange Market (BIST100) volatility characteristic will be looked into with considering Meixner heavy tailed distribution. After constructing the model,a simulation study is developed.

Chapter 3; Long Memory Modelling of Major Stock Exchange Markets. A detailed analysis is carried out to pinpoint the long dependants using different tools. Then ARFIMA, FIGARCH ad FIEGRACRH models are compared.

Chapter 4; Wavelet Coherence and Hurst Analysis of DAX and NIKKEI Stock Markets daily log returns and volatility are focused. Given the selected period and values of Hurst, VARIMA model is applied to forecast the next 30 days for the stock indices.

2. Discrete-Time And Continuous-Time Modeling

2.1. Introduction

Financial time series analysis is the study of financial data such as asset prices or returns, ordered sequentially in time, to identify trends or structure the data might exhibit and to make inferences and predictions. Earlier theories of distribution of asset returns assumed price changes from a point in time to the next are independently and identically distributed (iid) random variations with log of returns normally distributed [152]. However, studying wheat price returns from 1883-1934, Kendall [102] questioned the normality assumption and identified a fairly leptokurtic profile, which was also confirmed later by Mandelbrot [132] in an independent study. Furthermore, a number of econometric studies identified the limitations of using constant variance (volatility) in forecast models [144] and [104]. To address this, Engle [64] introduced a new class of stochastic processes: serially uncorrelated, time-varying conditional (on the past) variances - leaving the unconditional variance constant, called autoregressive conditional heteroskedastic (ARCH) processes. Later, a more general class of processes, as an extension to ARCH, were introduced by Bollerslev [28], known as GARCH. Many subsequent financial markets data studies have since confirmed the existence of volatility clustering in time series (periods of high and low volatility) and the assumption of iid - normal returns (constant volatility) lost its popularity.

Stock prices were firstly modelled as continuous-time Ito processes by Merton [145] in 1973 and later developed by Black-Scholes. The Black-Scholes and Merton model

(BSM) has been at the center of modern quantitative finance since its inception, due to its elegance (its precise) and simplicity (its limitations clearly stated) in pricing derivative financial instruments, mainly options. The model governs the evolution of the value of a European-style option contract over its lifetime. Central idea underlying the model is based on delta-hedging the option continuously by either buying or selling the asset underlying the contract to achieve delta-neutral position and eliminate price risk. Therefore, the price sensitivity to the underlying asset is eliminated and the option value will depend only on time (to expiry) and other observed constants. In the BSM world, the evolution of the asset value in the risk-neutral measure is given by

$$dS_t = \mu S_t d_t + \sigma S_t dW_t \tag{2.1}$$

Where:

S_t is price of the underlying asset (stock) at time t

μ is the drift term signifies the annualized change in expected value

σ is the volatility of the stock price (σ^2 is its variance rate)

W_t is the Weiner process or Brownian motion

The most significant assumption underlying the model is volatility being constant over time. The option value is calculated considering the above assumptions and also the property of no-arbitrage gain and frictionless markets.

But later developments in the financial econometrics undermined the applicability of this model, since the seminal work of Engle [64] and Bollerslev [28] volatility has been considered as a stochastic process itself and needed to be modelled as well. Heston [90] introduced a two-factor stochastic model that price and volatility processes evolve according to stochastic differential equations. and driven by two correlated Brownian motions.

$$\begin{aligned} dS_t &= \mu S_t d_t + \sqrt{V_t} S_t dW_t^1 \\ dV_t &= \kappa(\theta - V_t) d_t + \xi \sqrt{V_t} dW_t^2 \end{aligned} \tag{2.2}$$

Where:

V is the variance term

Another approach in modelling the volatility is the continuous-time diffusion limit of the discrete time GARCH processes by Nelson [150]. In his model volatility process evolve independently of the driving process in price equation. Therefore, feedback mechanism between the price and volatility in the GARCH process is lost due to two sources of randomness. Also Brownian process do not capture the jumps and tail-heaviness of the volatility.

$$\begin{aligned} dG_t &= \sigma_t dB_t^1 \\ \sigma_t^2 &= (\beta - \eta\sigma_t^2)d_t + \varphi\sigma_t^2 dB_t^2 \end{aligned} \tag{2.3}$$

B^1 and B^2 refer to Brownian motions that are independent and the constant terms β, η and φ.

To address these issues; in 2004, Klüppelberg, Lindner and Maller [109] introduced a GARCH process continous in time (COGARCH) as an extension to the discrete time GARCH model. Like the GARCH model, it is an SDE driven (based on a single random variable) by a Lévy process.

The COGARCH process $(G_t)_{t\geq 0}$ is defined in terms of its stochastic differential dG, such that

$$\begin{aligned} dG_t &= \sigma_t dL_t \\ d\sigma_t^2 &= (\beta - \eta\sigma_t^2)d_t + \varphi\sigma_t^2 d[L,L]_t^{(d)} \end{aligned} \tag{2.4}$$

The need to replace constant variance (volatility) framework, which is not consistent with the real world, with stochastic volatility, gave birth to family of SDE driven ARCH models such as GARCH, E-GARCH, T-GARCH etc. Having said that, the BSM model remains to be prevalent with option traders and used as a standard quoting convention partly because of its simplicity (BSM has a closed form solution).

The market crashes in 1929, 1987 and in 2000 could not be described by Gaussian Distribution because of the existence of fat tails in the stock price distribution. Mandelbrot [132] argued that power law distribution is a more accurate description of the tails of distribution of asset price changes. By the time Mandelbrot published his study, it became a common knowledge among practical economists that the normal distribution did a poor job accounting for the peakness of the empirical financial data. All these studies indicated that the distribution of asset price variations are more complex than the Gaussian framework.

Mandelbrot's model is based on a class of probability distributions defined by Paul Lévy: the principle property of this family of distributions are that the random variable made up of a linear combination of iid random variables with common distribution, said to have a stable distribution. A particular case of a Lévy distribution is the Gaussian distribution.

The contribution of this section of the thesis to the field is to identify a stochastic model that explains the behavioural properties of the Istanbul Stock Exchange market (BIST). As discussed above, there are many candidate models deemed appropriate to analyze volatility as a continuous-time stochastic process [17], [37], [109],[38], [39]. For this study, we apply COGARCH(1,1) Meixner Lévy process to model the volatility of the datasets. The main focus is on the volatility (price risk) in the stock exchange index.

Here, we consider descriptive statistics to identify the general features of the BIST data. As COGARCH is an analogue version of GARCH model, the volatility of discrete and continuous time will be compared. Various estimation methods will be used to work out the parameters of COGARCH models. The dataset will be modelled with

COGARCH Meixner Lévy process. Simulation based coefficient estimation and Pseudo Likelihood methods are applied.

2.2. Methodology

2.2.1. Discrete-Time Modeling

Engle's [64] ARCH model capture many empirical features of financial time series including heavy tails, time varying volatility and clustering of large and small variations (the serial correlation in the volatility). In developing an ARCH model, definitions for conditional mean and conditional variance are required.

As shown by Engle [64], let y_t be a stationary time series such as stock price returns, with a known zero mean, y_t may be expressed as

$$y_t = \varepsilon_t \sigma_t \tag{2.5}$$

where ε is taken to be the error term (white noise) with $\varepsilon_t \sim WN(0,1)$ and σ_t denotes the conditional variance, $Var(y_t|y_{t-1}, y_{t-2}...) = \sigma_t^2$, which by definition is a function of history and given as

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} (\alpha_1 y_{t-i}^2) \tag{2.6}$$

where q is a positive integer and denotes the lag value in the time series. The above equation is defined as ARCH(q) process given that the squared returns, y_t^2 follows an autoregressive, AR(q) process.

ARCH was developed to address the limitations of constant variance models as well as models where variance is changed by exogenous variable when analyzing econometric and financial time series, including evaluating the risk of an asset or portfolio in terms of its volatility. Although there is elegance in its simplicity, often times ARCH model requires many parameters (high lag, q) to sufficiently model asset returns. However, in general, fixed lag must be imposed to avoid computational issues with negative variance estimates. Bollerslev [28] introduced the Generalized ARCH process (GARCH) to address such issues.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q}(\alpha_i y_{t-i}^2) + \sum_{j=1}^{p}(\beta_j \sigma_{t-j}^2) \tag{2.7}$$

where $p \geq 0$, $q > 0$. To keep the conditional variance generated by the $GARCH(p,q)$ model positive, Bollerslev [13] imposed the following conditions

$\alpha_0 > 0$

$\alpha_i \geq 0$

$\beta_j \geq 0$

$\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \alpha_j < 1$

The above conditions can be weakened and should not be forced in estimation. Empirical examples illustrating the importance of relaxing these constraints are shown by Nelson and Cao [151]. Also note that, from the above equation, $GARCH(p,q)$ process reduces to $ARCH(q)$ when $p = 0$.

As part of any time series model construction, it is necessary to test the appropriateness of the model specified in describing the observations. As for the $ARCH(q)$ process, it is first necessary to check for the so called ARCH effects or more specifically conditional heteroskedasticity - otherwise OLS would suffice in modeling the time series. One of the two tests used for this is the LM (Lagrange Multiplier) test as outlined in Engle [64] (the other being the Ljung-Box statistics). The LM test is similar to the ordinary F-statistic in that we check whether the proposed model gives a superior fit in comparison to the one without the independent variables. This is expressed, in terms of the alternative and null hypotheses : if we define $e_t = y_t - \mu_t$ as the residuals of the mean equation where μ_t stands for the mean of y_t, then

$$H_0 : \alpha_1 = ... = \alpha_q = 0 \tag{2.8}$$

$$H_a : e_t^2 = \alpha_0 + \alpha_1 e_{t-1} + ... + \alpha_q e_{t-q} + \varepsilon_t \tag{2.9}$$

where ε denotes the error term. Given the test statistic

$$LM = TR^2 \; X^2(p) \tag{2.10}$$

where T refers to the size of the sample and R^2 is the usual coefficient of determination. H_0 is then rejected if $LM > \chi_q^2$ i.e. ARCH effects exist in the time series.

2.2.2. Continuous-Time Modelling

The discrete-time ARCH models discussed above are based on stochastic difference equations and is purposed for time series with equally spaced data points, such as daily closing price of stocks. However, they are not well suited for time series with arbitrary or unevenly spaced financial time series, as well as option pricing applications. One way to model volatility continuous in time is to use Lévy driven continuous-time ARMA (CARMA) such as Poisson or Gamma processes [17], [37], [38]. Another tool for continuous volatility modeling is to use the values of the realized process as continuous time GARCH (COGARCH) [109], [40].

By assuming infinitesimally small time increments, Nelson [150] suggested continuous time form of discrete time processes. This is expresed in equation 2.3.

However, unlike the discrete form of the GARCH process for which single source of randomness sufficient to represent the stylized facts of asset returns. Nelson's model of continuous-time GARCH diffusion approximation is driven by two independent Brownian motions and therefore introducing two sources of randomness.

The COGARCH model of Brockwell, Chadraa and Lindner [40] is a single Lévy driven SDE process, L and expressed in equation 2.4. More rigorous mathematical description can be found in [40].

2.2.3. Lévy Processes

Lévy processes has a wide use in various scientific disciplines where the area of interest involves study of stochastic processes in continuous-time such as turbulent flow in physics, machine learning in computer science & statistics, and continuous time series in finance and economics [18].

In mathematical finance, (jump-diffusion) Lévy processes have somewhat become omnipresent due to their ability to capture the empirical features of the financial markets more accurately than the models based on the classical Geometric Brownian motion. Empirically, asset prices exhibit sharp variations or jumps, which need to be taken into account by investors and risk managers alike to capture the underlying risks more accurately. For example, sudden variations in the price of BIST100 can be observed in Figure 1(a). Moreover, historical financial time series of asset returns tend to exhibit sharper peaks, heavier tails (leptokurtic) and skewness than predicted by the normal distribution, see Figure (5) for a characteristic plot. Also, in pricing of options, unlike what the BSM model assumes, we observe skew or smile like volatility surfaces across expiry and strike (moneyness), as depicted in Figure (3). Lévy processes with well defined mathematical properties allow us to accurately capture the aforementioned attributes of financial time series.

More recently, the use of infinitely divisible distributions in modelling financial time series has been proposed. Madan and Senata [127] introduced the Variance Gamma distribution as a model for stock returns. Other distributions to model stock returns include Normal Inverse Gaussian distribution by Barndorff-Nielsen [16] and the Meixner distribution by Schoutens and Teugels [173].

Definition 2.1. A stochastic process $L = \{L_t : t \geq 0\}$ defined on a filtered probability space $(\Omega, F, (F_t)_{0\leq t<\infty}, P)$, is called a Lévy process if

- the initial condition $L_0 = 0$ almost surely.
- L paths are cádl'ag (right continuous with left limits)
- $L_t - L_s$ is independent of $\{L_u : u \leq s\}$ for any $0 \leq s \leq t$
- L has stationary increments: $L_t - L_s$ is equal in distribution to L_{t-s} for any $0 \leq s \leq t$
- L_t is continuous in probability, $\lim_{\delta\to 0} P(\mid L_{t+\delta} - L_t \mid \geq \varepsilon) = 0$

The Wiener and the Poisson processes share the same properties stated above, and are examples of Lévy process, also known as additive process listed above. Also note that a linear combination of independent Lévy processes, such as Brownian motion with drift and a compound Poisson process is again a Lévy process, such a process is often referred to as jump-diffusion process.

The characteristic function of $L(t)$, $\phi t(\theta) : E[e^{(i\theta L(t))}]$ has the Lévy-Khintchine representation

$$\phi_t(\theta) \quad = \quad e^{t\xi(\theta)}, \theta \in \Re \tag{2.11}$$

where

$$\xi(\theta) \quad = \quad (i\theta m - 0.5\theta^2\sigma^2 + \int_{\Re}(e^{i\theta x} - 1 - \frac{i\theta x}{1+x^2})v(dx) \tag{2.12}$$

For $m \in \Re, \sigma > 0$ and the measure v is on the Borel subsets of $\Re_0 = \Re/\{0\}$, known as the Lévy measure, which governs the jump process, of L satisfies the constraint

$$\int_{\Re_0}(\frac{u^2}{1+u^2})v(du) < \infty \tag{2.13}$$

$v = 0$ is Brownian motion

$m = \sigma^2 = 0, \int_{\Re} \frac{|u|}{1+u^2} v(du) < \infty \rightarrow$ compound Poisson with drift

$v(du) = \alpha u^{-1} e^{-\beta u} du \rightarrow$ gamma process with i

$$\xi(\theta) = \int (e^{i\theta x - 1} - 1) v(dx) = (1 - \frac{i\theta}{\beta})^{-\alpha t} \tag{2.14}$$

$v(du) = 0.5\alpha |u|^{-1} e^{-\beta|u|} du \rightarrow \alpha$ symmetrized gamma process $(L_1 - L_2)$

$\xi(\theta) = e^{-c|\theta|^{\alpha}}, 0 < \alpha \leq 2 \rightarrow$ symmetric stable process

The feature of Lévy processes over Gaussian driven processes is that Lévy processes can model skewness, excess kurtosis and more importantly the jumps that occur in the volatility.

2.2.4. Meixner Density Function

Definition 2.2. A random variable X is Meixner distributed with parameters $\theta = (\alpha, \beta, \delta)$, if it has the pdf [177]

$$f(x;\theta) = \frac{(2\cos(\beta/2))^{2\delta}}{2\alpha\pi\Gamma(2\delta)} e^{\frac{\beta x}{\alpha}} \left| \Gamma(\delta + \frac{ix}{\alpha}) \right|^2 \tag{2.15}$$

where

$$\left| \Gamma(\delta + \frac{x}{\alpha}) \right| = (\int_0^{\infty} \cos(\frac{x}{\alpha} \log y) y^{\delta-1} e^{-y})^2 + (\int_0^{\infty} \sin(\frac{x}{\alpha} \log y) y^{\delta-1} e^{-y})^2 \tag{2.16}$$

where $x \in \Re$, $\alpha \in (0, \infty)$, $\beta \in (-\pi, \pi)$, $\delta \in (0, \infty)$ Meixner distribution is denoted by as $X \overset{d}{\sim} Meixner(\alpha, \beta, \delta)$.

Lemma 2.3. *Let $X \overset{d}{\sim} Meixner(\alpha, \beta, \delta)$. The moments of the distribution is given by*

$$E[X] = \alpha\delta \tan\frac{\beta}{2} \tag{2.17}$$

$$Var[X] = \frac{\alpha^2 \delta}{2} \sec^2 \frac{\beta}{2} \tag{2.18}$$

$$Skew[X] = \sqrt{\frac{2}{\delta}} \sin \frac{\beta}{2} \tag{2.19}$$

$$Kurtosis[X] = 3 + \frac{2 - \cos(\beta/2)}{\delta} \tag{2.20}$$

Then the characteristic function $\phi(u)$, and the cumulant generating function, $\kappa(u) = \log E[e^{uX}] = \log \phi(-iu)$ are respectively given by

$$\Phi(u;\theta) = (\frac{\cos(\beta/2)}{\cosh((\alpha u - i\beta)/2)})^{2\delta} \tag{2.21}$$

and

$$\kappa(u;\theta) = 2\delta[\log(\cos(\beta/2)) - \log(\cos((\alpha u + \beta)/2))] \tag{2.22}$$

For proof see [81].

where $i = \sqrt{-1}$ and $u \in \Re$

The kurtosis of the density function is described by all the parameters of the distribution. Also note that, from equation 2.20, it is clear that Meixner distribution has heavier tails in comparison to the standard normal dist. with kurtosis 3. However, the skewness (symmetry) of the density function is described by the parameter β. The distribution is skewed to the left for $\beta < 0$, $\beta > 0$ indicates right skewed distribution and symmetric for $\beta = 0$.

2.2.5. Meixner Process

The described approach to the definition of Lévy processes in finance, and in particular of Meixner process, can be classified as a “classical – statistical” one, which finds its motivation in the attempt of providing an improvement to the lack of fit of Brownian motion based models. Its theoretical introduction in this sense dates back to 1998 with a paper by Schoutens and Teugels [173] dealing with orthogonal polynomials and martingales. Its applications and appropriateness for financial markets were demonstrated in [83], [175] The Meixner distribution is infinitely divisible - a required property to construct a Lévy process. Let $\{M_t : t \geq 0\}$ be a stochastic process with initial condition $M_0 = 0$ and has stationary and independent increments, then the distribution of M_t, as specified in definition (2.2), can be given by $Meixner(\alpha, \beta, \delta)$. This is referred to as the Meixner process.

Recall that, typically, Lévy process consists of a Brownian motion with drift and an independent jump process such as a compound Poisson process. As given in [176] there is no Brownian motion in $\{M_t : t \geq 0\}$ and the jump part is described by the following Lévy measure

$$v(dx) = \delta \frac{e^{\beta x/\alpha}}{xsinh(\pi x/\alpha)} dx \tag{2.23}$$

2.2.6. Pseudo Maximum Likelihood Estimation Method

To be able to estimate the parameters of COGARCH(1,1) model driven by a Meixner Lévy process, Pseudo Maximum Likelihood (PML) is used which is set out in Maller et.al. [131]. Taking the stationary value $\beta/(\eta - \varphi)$ as starting value for $\sigma^2(0)$, maximize L_M to get PML of the parameters (β, η, φ).

Using Eq.(2.21) we have

$$L_M = -\frac{\mu_z}{\rho_z}\sqrt{\sigma_t} + 2\delta \left[\log(\cos\frac{\beta}{2}) - \log(\cos(\frac{\alpha\sqrt{h_t}}{\rho_z} + \beta)/2) \right] \tag{2.24}$$

see Appendix for details. where $\theta = (\alpha, \beta, \delta)$, $u_t = \frac{Z_t - \mu_z}{\rho_z}$

2.3. Results and Analysis

2.3.1. Results: Descriptive Statistics

In this study, Istanbul Stock Market's (BIST100) volatility characteristics are analysed in detail using end of day prices of the index. We use continuously compounded daily

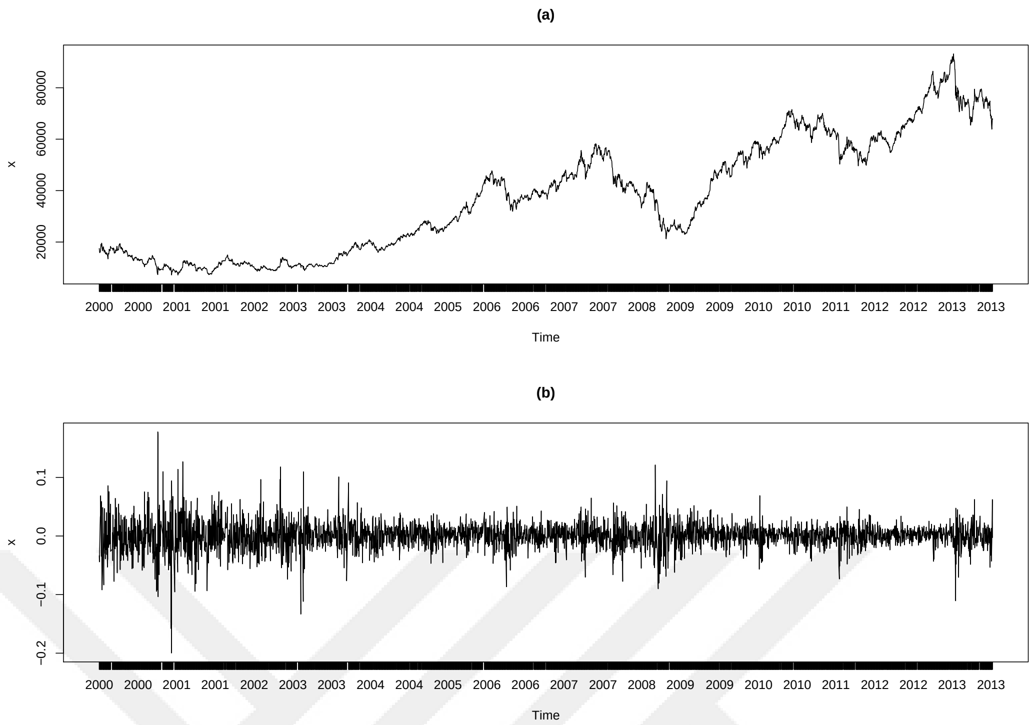

Figure 1. Timeseries and Log-return plot of BIST100 stock market indices. Top figure is the timeseries plot and the bottom figure is the plot of the log-return.

returns, expressed as the difference in the natural log of BIST100 closing prices, $X_t = ln(P_t) - ln(P_{t-1}) = ln(P_t/P_{t-1})$. The time period that analysed is from 04/01/2000 to 31/12/2013.

The BIST100 time series plot is shown in Figure 1(a) and the corresponding log-returns is shown in Figure 1(b). From the log-returns, by observation, the lack of any significant patterns suggests quasi-random behavior. The ACF plot of log-returns in Figure 2 suggests that there is not a strong correlation amongst the lags, supporting the stationarity of the log-returns.

The main characteristics of the log return of BIST100 are summarized in Table 1. The mean is positive which shows that there is a overall positive (nominal) gain investing in Istanbul Stock Exchange Market, i.e. profit for the investors. Kurtosis is a shape measure, and describes the relative concentration in the tail of the distribution to its center: higher kurtosis indicate heavier tails. Excess kurtosis being 6.669 tells us that

Table 1 Descriptive Statistics

Data Period	2000-2013
Sample Size	3504
Mean	0.003863
S.Dev	0.02327
Excess Kurtosis	6.669
Skewness	-0.06766
ADF Test	0
KPSS Test	0.0865
ARCH Effect Test	0
Jarque-Bera Test	0
Zivot-Andrews Test	0.0004373

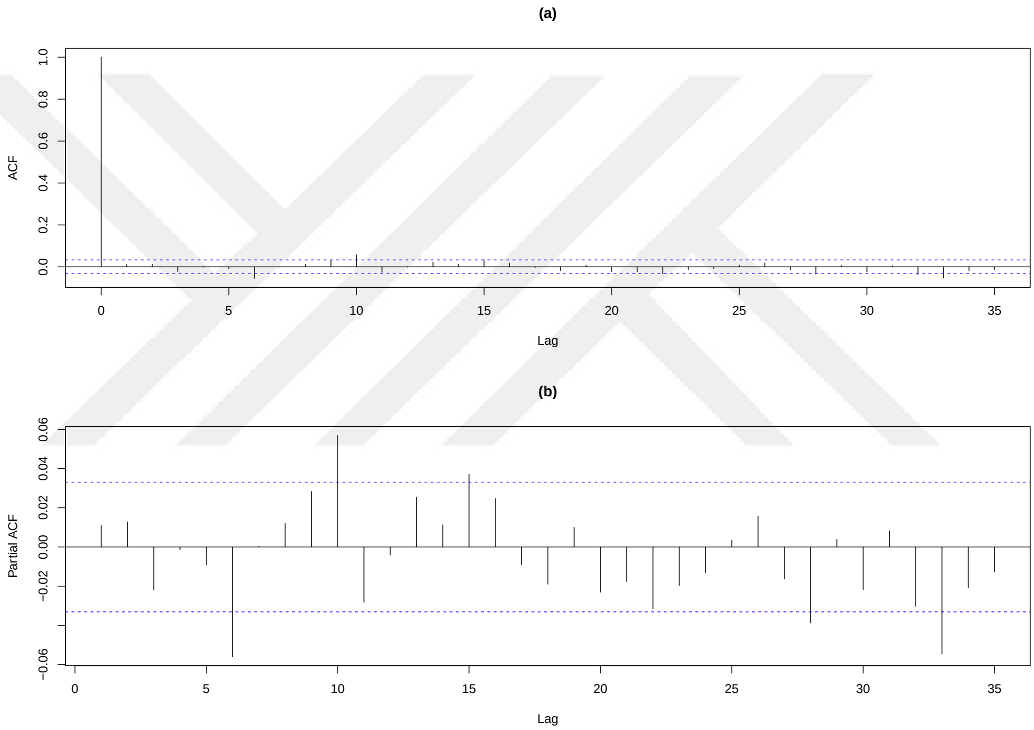

Figure 2. (a) ACF plot of Log Return of BIST100 (b) PACF plot of Log Return of BIST100.

the density function of the series is from a heavy tailed distribution. Skewness measures how symmetric the distribution is: symmetric data have (near) zero value, negative (positive) value means the distribution is skewed to the left (right) (skewness is zero for a normal distribution). The skewness of the BIST100 time series is -0.06766, indication of slight negative asymmetry - left skewed. Moreover, the Jarque-Bera Test, which is a a goodness-of-fit measure and is a function of kurtosis and skewness of a distribution to match normal distribution. The test value for our BIST100 daily returns is zero, suggesting that the data set is not normally distributed.

The return data can be further investigated for stationarity by looking at the ADF (Augmented Dickey-Fuller) [167] and Zivot & Andrews (ZA) unit root tests [206] in Table 1. As ADF test p-value is nearly zero, the null hypothesis that the data has unitroot can be rejected. A more refined unit root test, ZA test, assumes a data set can be trend stationary with a single (estimated) break in the trend. The H_0 is a unit root process that rules out exogenous break (or permanent shock such as financial crash) and H_a is a trend-stationary process that includes one-time jump (or break) in the trend [18]. The near zero value for ZA test suggest that we can reject H_0.

A widely used test for ARCH effects proposed by Engle [64]. ARCH effect test with the p-value of zero shows that there is an ARCH effect in the log returns of Istanbul stock exchange market.

From Figure 3, which shows the BIST100 squared returns, X_t^2, we are able to identify volatility clustering (as observed in [17], [57]) that there is tendency of large variations in squared-returns following large changes and small changes by small moves. Consequently, we can also deduce that the BIST100 volatility changes over time. An appropriate model for our data set is then should account for time varying volatility such as the GARCH model. The volatility in BIST100 can easily be viewed in Figure 3. The spikes are much more appearant than the Figure 1(b).

2.3.2. Distributional Properties of Log Return of BIST100

There is practical importance in deriving the distribution of asset returns: in risk management it allows a likelihood estimate of extreme price changes, i.e. tail risk; in pricing of derivatives, for example options, in assessing the performance of a given portfolio such as in Sharpe ratio, which is a function of the mean and standard deviation (volatility) of returns and quantifies the risk-reward relationship between different investment strategies.

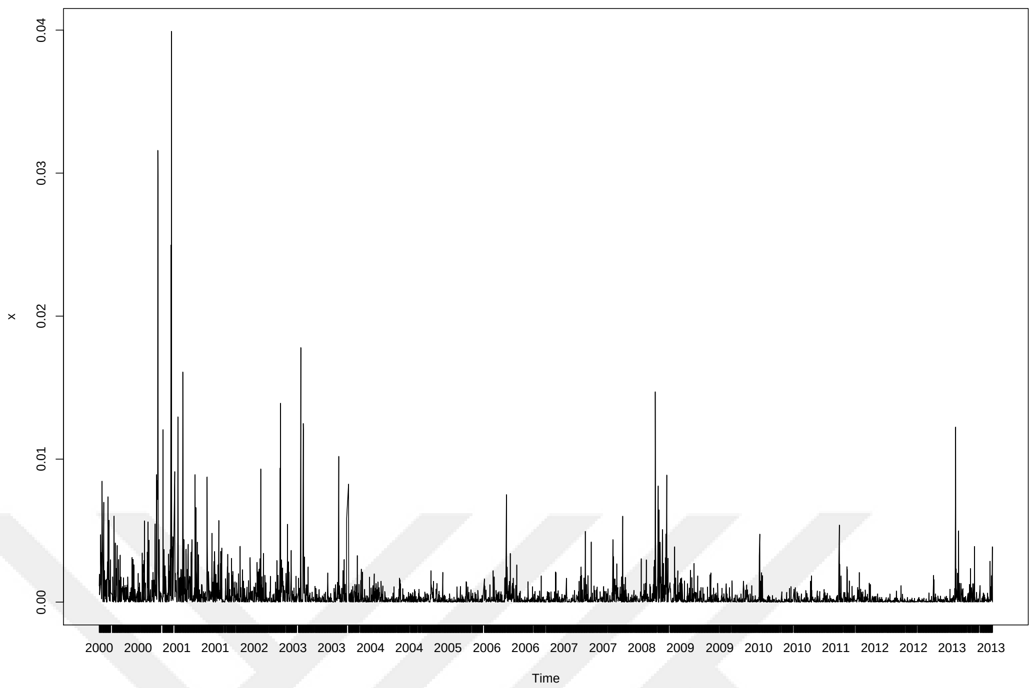

Figure 3. Squared Returns of BIST100.

As far as the assumptions in asset price distributions concerned, perhaps the most commonly used description is the Gaussian (normal) distribution, which served as a somewhat decent approximation for asset returns historically. A standard technique in finding out if a series originate from a specific statistical distribution is the Quantile-Quantile (Q-Q) plot, where the quantiles of both distributions (empirical vs theoretical) plotted against each other to check for deviations from linearity which would suggest that the two data sets do not have common distribution. Moreover, the Q-Q plot help answer whether the two data set have similar distributional profiles and tail behavior.

Q-Q plot of Normal distribution for the BIST100 data can be viewed in Figure 4. Notice that the points that fall in the middle of the plot lines up well with the normal distribution values, however, as we move towards more extreme values we find significant deviations from linearity. This suggests that large (extreme) returns in BIST100 occurs more frequently than what the normal distribution predicts. In other words, the BIST100 return data exhibit heavy tails. From this we can conclude that the log-returns in our

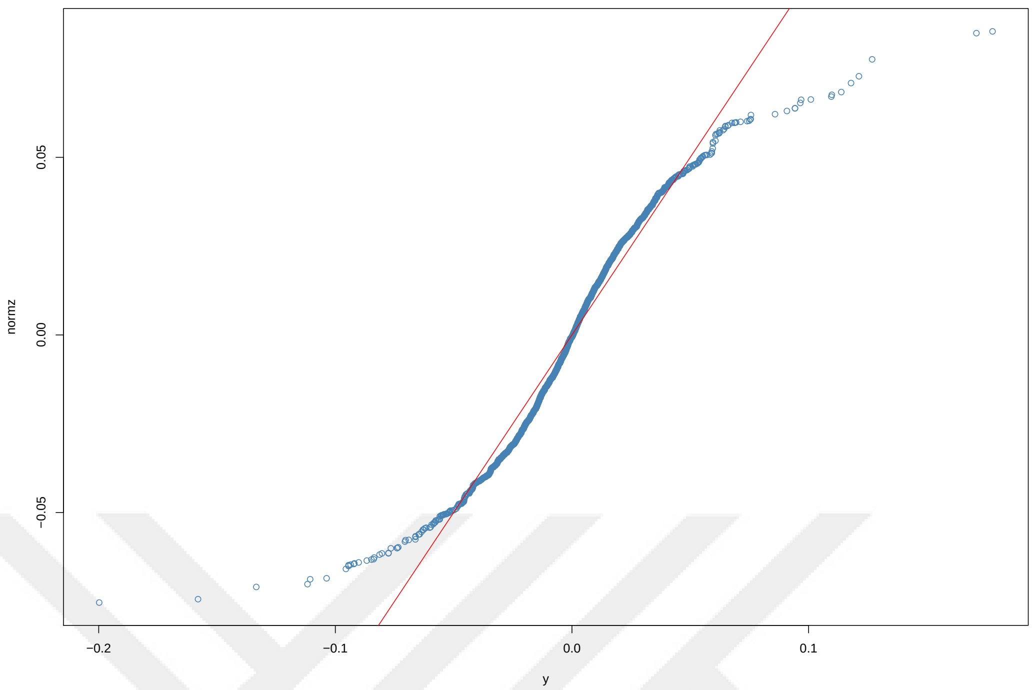

Figure 4. Q-Q Plot of Normal Distribution of Log-return of BIST100.

series does not follow normal distribution.

There are also various tests available to apply in order to specify which distribution the data comes from. In Table 2, Shapiro-Wilk [181] and Jarque-Bera [97] normality tests clearly show that daily price changes of BIST100 are not normally distributed. Furthermore histogram plot and Q-Q plot in Figures 5 and 6 show that Meixner distribution fits to the dataset much better. Meixner distribution is a favourable one because it can model distributions that exhibit symmetric or asymmetric profiles, potentially with extended tails in each direction.

Another graphical method to check the normality of a data series is to plot its density histogram, where the area under the curve is normalized to sum to 1. Figure 5 shows the BIST100 density plot together with the normal distribution - depicted with black line, and the Meixner distribution (red-circle line). It is clear that the normal distribution falls short of capturing the extreme occurrences. In contrast, the Meixner profile offers a better fit in capturing the peakedness and the heavy tailedness of the data.

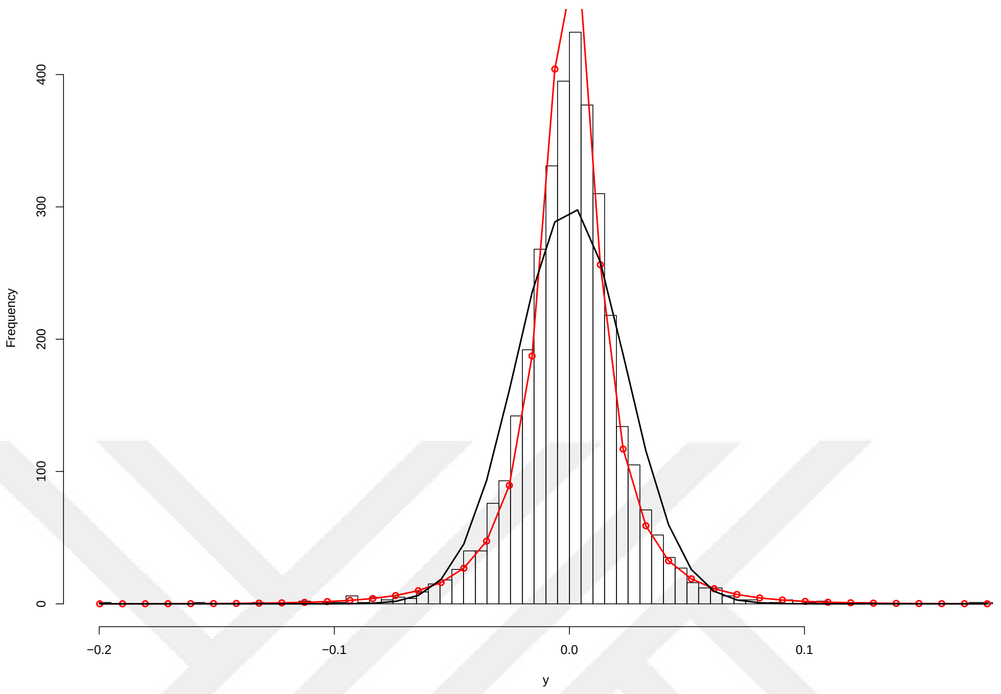

Figure 5. A density histogram of the BIST100 data, with a normal distribution (black-straight line) and Meixner distribution (red-circle line) overlaid.

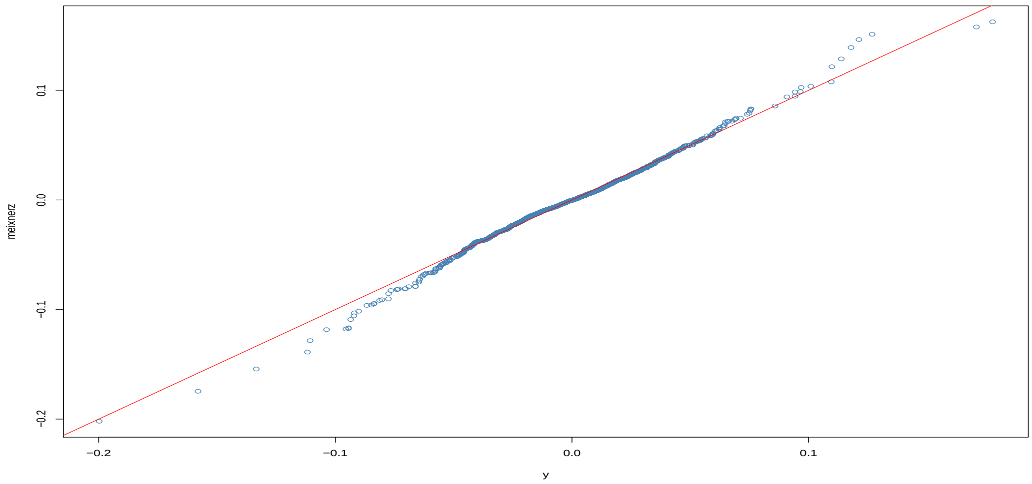

Figure 6. Q-Q Plot of Meixner Distribution of Log-return of BIST100.

Table 2 Distributional Tests of BIST100

Tests	P-values
Shapiro-Wilk Normality	0
Anderson-Darling Normality	0

2.3.3. GARCH Modelling

In the previous section; the log returns of BIST100 analysed in detail. The results showed that the log-returns is stationary, and the series does not have a unit root. Furthermore, the tests for ARCH effect proved that BIST100 needs to be modelled by a heteroskedasticity model to eliminate the ARCH effect. Also the distributional properties of the log return data showed that the log return of BIST100 has a heavy tailed distribution. In this section; the residuals of the GARCH model needs to be analysed to look into the characteristics of the residuals distribution. This result will give evidence on which distribution is appropriate to use in COGARCH modelling

The followings will be considered when selecting a good fitted model for GARCH modelling:

- Standard residuals should not exhibit any of the following structures: serial correlation, conditional heteroskedasticity, linear dependence.
- Existence of autocorrelation in residuals and the squared residuals confirmed by the Ljung-Box test.
- Engle's LM test to check for disappearance of ARCH effects in the residuals.
- One of the GARCH modelling's rule of thumb is that the sum of the model's coefficients must be less than 1 for the model to be covariance stationary. GARCH negativity constraints must be satisfied by the selected model:
  $\sum \alpha_i + \sum \beta_j < 1$

Table(3) summarizes the parameters estimates of GARCH(1,1) model. All three parameters are statistically significant and t-values are high. And the model is covariance stationary

Table 3 GARCH(1,1) Model Estimation

Parameter	Estimates	Std.Error	t-value
α_0	7.116e-06	1.072e-06	6.637 ***
α_1	1.049e-01	6.893e-03	15.215 ***
β_1	8.858e-01	6.417e-03	138.047 ***

Signif. codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

as the sum of ARCH and GARCH coefficients is less than 1. Further diagnostics tests have been carried out for GARCH residuals to see whether GARCH(1,1) is a suitable model, see Table 4. The p-value of 0.3639 of Ljung-Box suggests no autocorrelation in the residuals. Moreover, the p-value of 0.6607 of ARCH-LM rules out ARCH effects. And lastly, the Jarque-Bera p-value of near zero indicates the residuals not normally distributed.

Table 4 Diagnostic Tests for GARCH Residuals

Test	p-value
Jarque Bera Test	2.2e-16
Box-Ljung Test	0.3639
ARCH-LM Test	0.6607

GARCH(1,1) residuals have heavy tails. They are not normally distributed given the Shapiro-Wilk Normality and Anderson-Darling Normality tests in Table 5. K-S test for Normal distribution also proves that the residuals are from a heavy tailed distribution and K-S test for Meixner distribution explains the goodness of fit to the data.

Table 5 Distributional Tests for GARCH Residuals

Tests	P-values
K–S test for Meixner	0.2189641
K-S test for Normal	0.001758326
Shapiro-Wilk Normality	7.551541e-20
Anderson-Darling Normality	1.777898e-16

Given that the squared returns are a measure for the 2nd order moment, the ACF plot as shown in Figure 7, indicates that the conditional variance vary over time. In other words, BIST100 log-returns exhibit time varying conditional heteroskedasticity (volatility

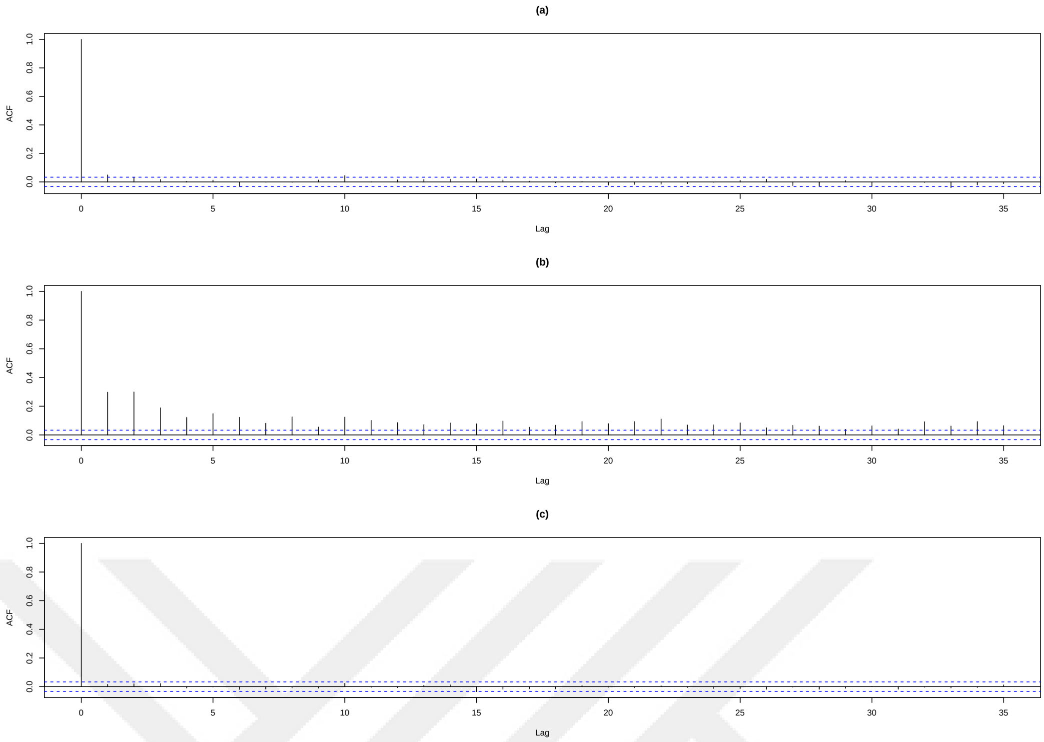

Figure 7. (a) ACF plot of GARCH residuals. (b) ACF plot of Squared Log Return of BIST100 (c) ACF plot of Squared GARCH residuals.

clustering). Figure 7(b) shows that there exists ARCH effect in the log return data and GARCH(1,1) residuals, see Figure 7(c), eliminates the ARCH effect in the time series.

2.3.4. Meixner-GARCH Model

Until now, we have showed the GARCH(1,1) fit for the log return of BIST100. And we have also explained the poor fit of Normal distribution to the GARCH(1,1) residuals. Furthermore detailed analysis proved that Meixner distribution fits well to the GARCH(1,1) residuals. In this section, considering the previous results, we will focus on GARCH(1,1) with Meixner distribution as the conditional distribution.

Coefficients of Meixner-GARCH(1,1) model are statistically significant as t-values in Table 6 are higher than 1,96. And Meixner-GARCH(1,1) model is covariance stationary as the sum of parameters are less than 1. The diagnostics tests for the residuals in Table 7 show that no significant autocorrelation exists in the residuals. The LM test,

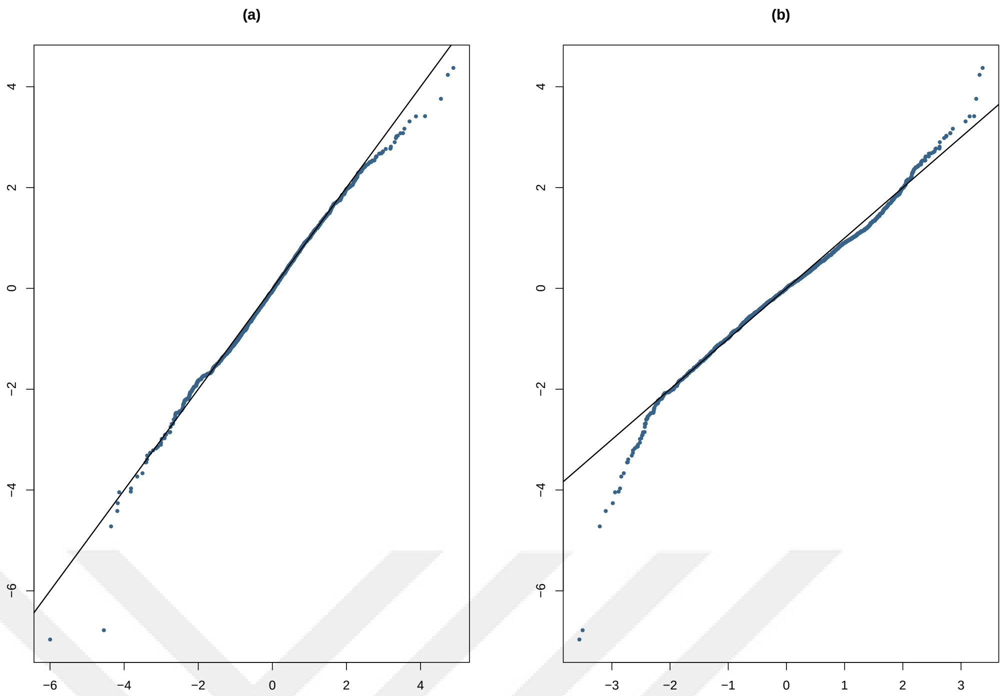

Figure 8. (a) QQ-Plot Meixner for GARCH residuals. (b) QQ-Plot Normal for GARCH residuals.

together with the ACF of squared residuals (Figure 10), supports no ARCH effect in Meixner-GARCH(1,1) residuals.

Various distributional tests in Table 8 and Q-Q plot of Meixner distribution in Figure 9 supports the good fit of Meixner distribution to the model's residuals. And Figure 11 is the histogram plot of the Meixner-GRACH(1,1) residuals with the Meixner distribution and both density plots also shows the goodness of fit.

Table 6 Meixner-GARCH(1,1) Model Estimation

Parameter	Estimates	Std.Error	t-value
α_0	5.56528e-06	1.55851e-06	3.57090 ***
α_1	8.08085e-02	1.12238e-02	7.19975 ***
β_1	8.97722e-01	1.36247e-02	65.88935 ***

Signif. codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

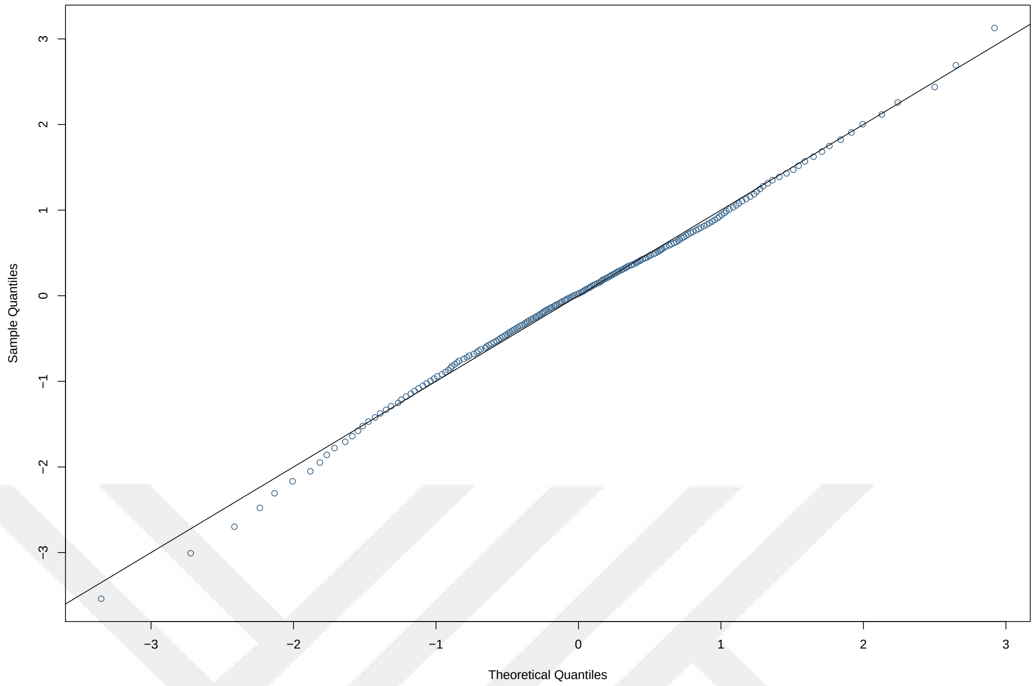

Figure 9. Q-Q Plot for Meixner-GARCH(1,1) residuals.

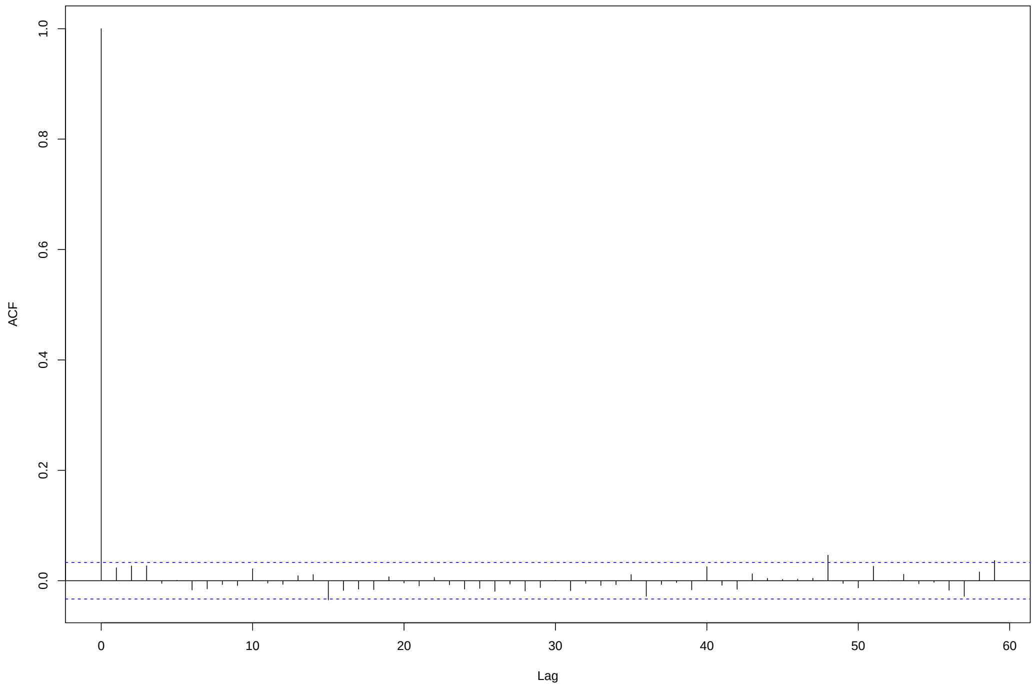

Figure 10. ACF Plot of Meixner-GARCH(1,1) squared residuals.

Table 7 Diagnostic Tests for Meixner-GARCH Residuals

Test	p-value
Jarque Bera Test	2.2e-16
Box-Ljung Test	0.06665
ARCH-LM Test	0.5592

Table 8 Distributional Tests of Meixner-GARCH(1,1) Residuals

Tests	P-values
K–S test for Meixner	0.17851
K-S test for Normal	0.0058416
Shapiro-Wilk Normality	6.1154e-20
Anderson-Darling Normality	1.3783e-16

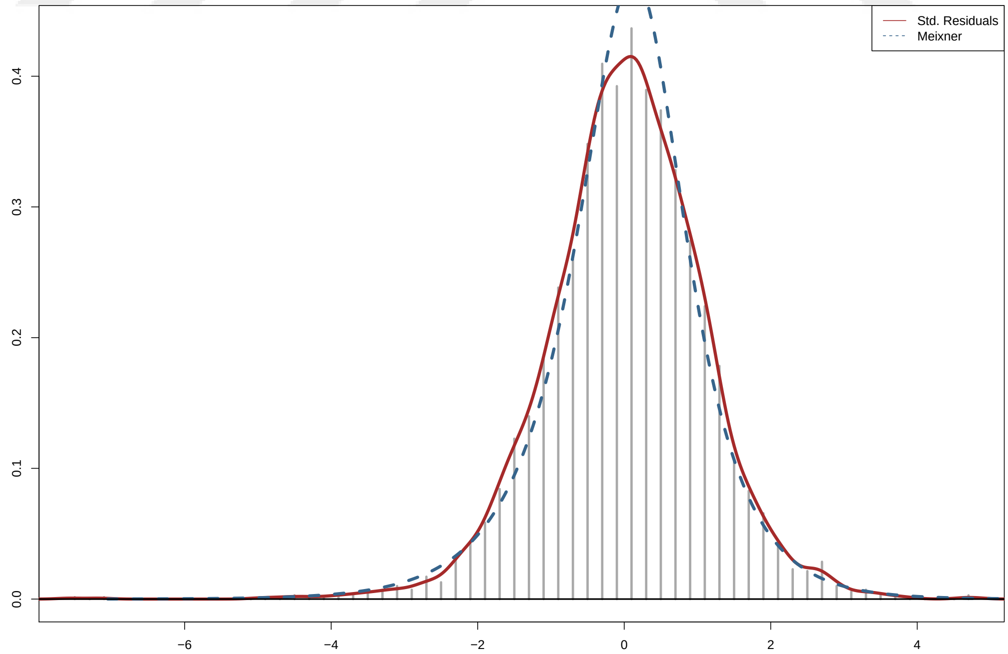

Figure 11. Density plot of Meixner-GARCH(1,1) residuals and Meixner distribution(dashed line).

2.3.5. Continuous-Time GARCH Modeling

2.3.5.1. Simulation Study

Detailed analysis showed that GARCH(1,1) residuals follow a heavy tailed distribution. During this section Meixner Lévy process will be used as the driven process for COGARCH(1,1) model.

A Lévy process can be seen as having three parts (Lévy triplet): (1) a linear deterministic part (drift) controlled by the drift coefficient (2) a Brownian part (diffusion) controlled by the diffusion coefficient (3) a pure jump part controlled by the Lévy measure which dictates how the jumps occur. A Meixner Lévy process has no Brownian component and its diffusion coefficient is also a pure jump process, see Applebaum[2], Bertoin[23], and Sato[170] for detailed results concerning Lévy processes.

To be able to estimate the parameters of COGARCH(1,1) model driven by a Meixner Lévy process, Pseudo Maximum Likelihood (PML) is used which is set out in Maller et.al. [131]. Taking the stationary value $\beta/(\eta-\varphi)$ as starting value for $\sigma^2(0)$, maximize L_M to get PML of the parameters (β,η,φ).

Using Eq.(2.21) we have

$$L_M = -\frac{\mu_z}{\rho_z}\sqrt{\sigma_t} + 2\delta\left[\log(\cos\frac{\beta}{2}) - \log(\cos(\frac{\alpha\sqrt{h_t}}{\rho_z}+\beta)/2)\right] \tag{2.25}$$

where $\theta = (\alpha,\beta,\delta)$, $u_t = \frac{Z_t-\mu_z}{\rho_z}$

Once the initial parameters (β,η,φ) are estimated via PML method, 1000 sample paths of COGARCH(1,1) are simulated using the initial parameters and various diagnostic tests are applied to the sample paths. K-S test are carried out for the simulated 1000 residuals. The estimated residuals are given by $\widehat{G}_t^{(1)}/\widehat{\sigma}_{t-1}$. 974 sample paths fit well to Meixner distribution. 940 out of 1000 simulations pass the ARCH-LM test. Figure 12 shows the density plots of COGARCH(1,1) residuals (black colored lines) with a

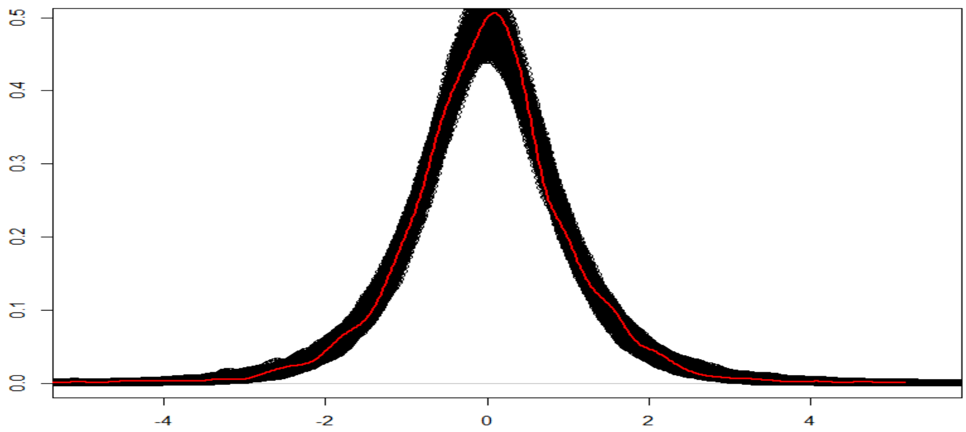

FIGURE 12: Density plots of simulated residuals (black) and density plot of a Meixner Distribution (red)

density plot of Meixner Distribution (red colored line). It is clear from the Figure 12 that the density plots of the sample residuals fit to Meixner distribution.

1000 simulations are carried out for the same length of the data which is 3504 equidistant observations. The outcome of the simulations concerning the parameters β, η, φ are summarized in Table 9 for Pseudo Likelihood Estimation Method.

Table 9 Residual analysis.

	mean($\widehat{G}_t^{(1)}/\widehat{\sigma}_{t-1}$)	std($\widehat{G}_t^{(1)}/\widehat{\sigma}_{t-1}$)
Mean	0.000216 (0.01737)	0.99965 (0.01686)
RMSE	0.030	0.02689
MAE	0.0257	0.02255

Table 10 COGARCH and GARCH Volatility Comparison.

	COGARCH(1,1)	GARCH(1,1)
BIAS	0.5161473	0.5773501
RMSE	0.9793929	1.0007466
MAE	0.7435358	0.7600032

Table 9 investigates the goodness of fit for PML estimation method by a residual analysis. The estimated residuals are given by $\widehat{G}_t^{(1)}/\widehat{\sigma}_{t-1}$. Since we assumed a symmetric jump distribution with zero mean, hence the residuals should be symmetric around zero

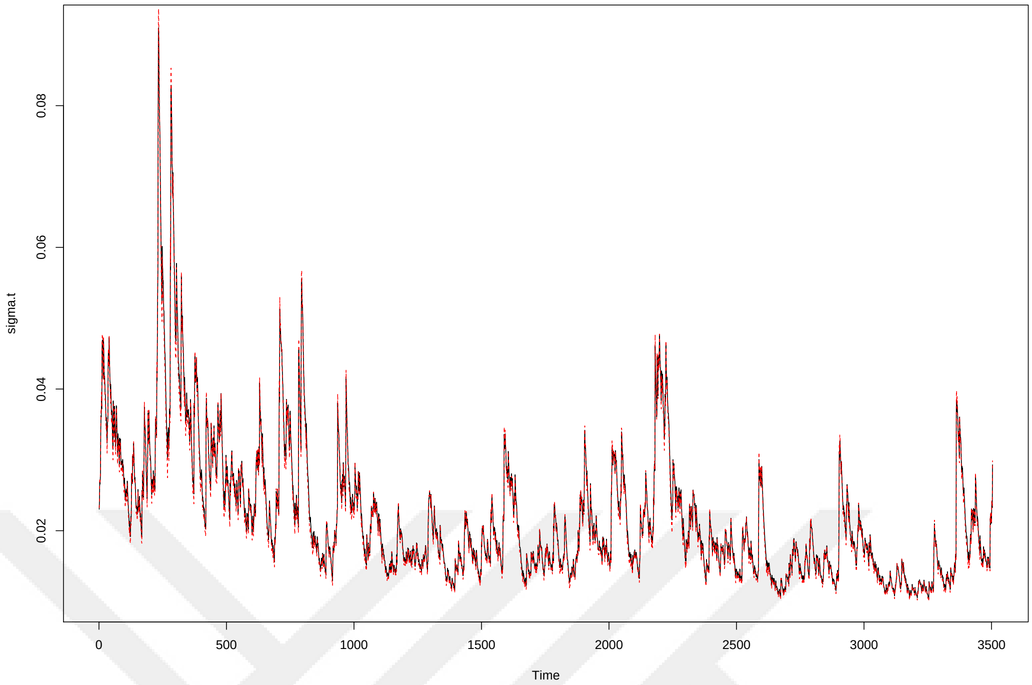

FIGURE 13: Volatility plot of COGARCH (black line) and GARCH (red dashed-line)

and their mean should be close to zero. If the volatility has been estimated correctly, we expect the standard deviation to be close to 1. We estimated mean, Root Mean Square Error (RMSE), and Mean Squared Error (MSE) for the mean and the standard deviation of the residuals based on 1000 simulations. The error statistics MAE and RMSE are given by:

$$\begin{aligned} MSE &= \frac{\sum_{t=1}^{n}(\sigma_t - \hat{\sigma}_t)^2}{n} \\ MAE &= \frac{1}{n}\sum_{t=1}^{n}|\hat{\sigma}_t - \sigma_t| \end{aligned}$$

where $\hat{\sigma}_t$ is predicted COGARCH(1,1) volatility and σ_t is the GARCH(1,1) volatility.

The estimated volatility process is calculated, [109], as recursively by

$$\hat{\sigma}_t^2 = \hat{\beta} + (1-\hat{\eta})\hat{\sigma}_{t-1}^2 + \hat{\varphi}(G_t^{(1)})^2 \tag{2.26}$$

Table 10 compares the discrete-time and continuous-time volatilities. Three popular criterion considered for the compression, which are Bias, RMSE, and MSE. GARCH column shows the GARCH volatility minus the absolute return of the real data and COGARCH column is the COGARCH volatility minus absolute return of the real data. For all the indices, COGARCH criterion is smaller than the GARCH ones. This result implies that COGARCH volatility imitates the real absolute return much better than the GARCH model's volatility does.

The empirical mean of all the estimated parameter values $\widehat{\beta}, \widehat{\varphi}, \widehat{\eta}$ is shown in the first column of Table 11. RMSE and MAE are also estimated as a guage for goodness of fit: small RMSE and MAE indication of good model in terms of describing the data set. The smaller the values of RMSE and MAE, the better the model is. The parameters $\widehat{\varphi}, \widehat{\eta}$ give important characteristic of the model concerning stationarity as $\psi(1) = -\widehat{\varphi} + \widehat{\eta}$. In Figure 14 a simulated sample path is plotted. Figure 14(a) is the driving

Table 11 Statistics of the parameters from COGARCH-Meixner-PML simulations

Parameters	Mean	St.Dev.	RMSE	MAE	Lower CI	Upper CI
$\widehat{\beta}$	9.219003e-06	9.337622e-06	8.7104e-08	4.101566e-06	8.639560e-06	9.798446e-06
$\widehat{\varphi}$	0.12267565	0.06555912	4.2937	0.029004	0.11860740	0.12674390
$\widehat{\eta}$	0.11556199	0.07425576	5.508404	0.03114683	0.11095407	0.12016991

Lévy Process (L_t), (b) COGARCH(1,1) process (G_t) , and figures (c) and (d) are differenced (G_t) and is the plot of volatility process (σ_t) respectively. (G_t) looks similar as (L_t), they only differ in their jump sizes. Also note the volatility clustering in the simulation, which is also observed in the empirical data.

The graph in Figure 15(a) shows that the increments $G_{t+1} - G_t$ exhibit no significant correlation which is the case for a GARCH process as well, see Figure 10(a). Figure 15(b) depicts the squared increments $(G_{t+1} - G_t)^2$ which shows significant correlation. Figure 15(c) is the plot of the volatility process and it is highly significant indication of persistence (long memory).

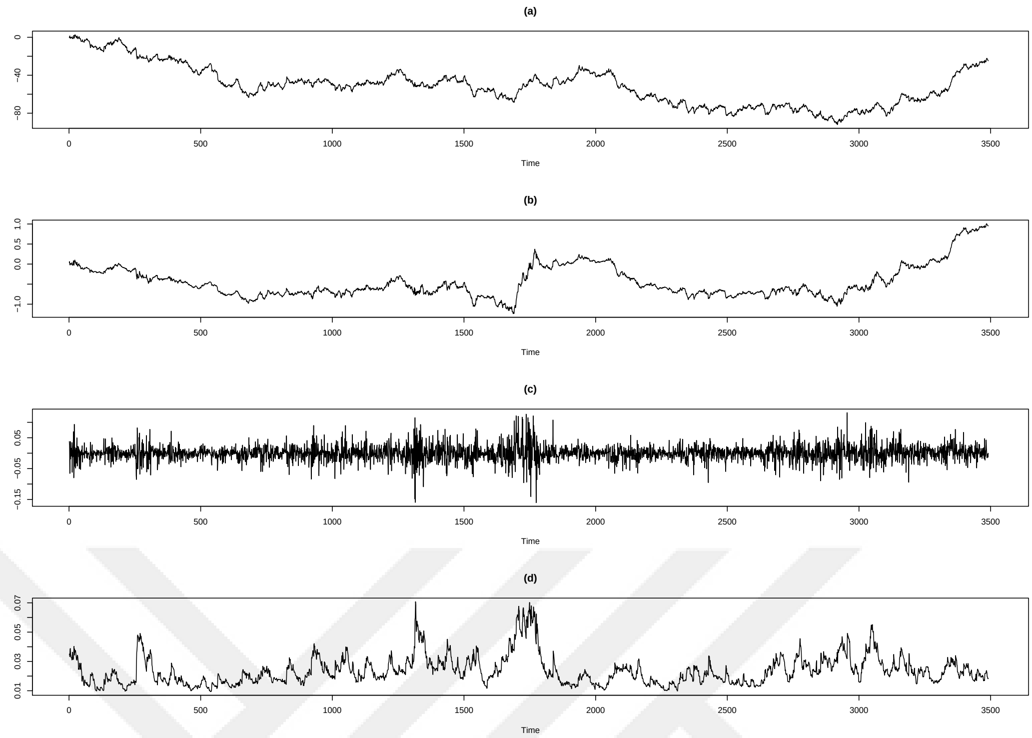

Figure 14. (a) Simulated Levy Process (L_t) (b) COGARCH process (G_t) (c) differenced COGARCH process $(G_{t+1} - G_t)$ (d) Volatility process (σ_t).

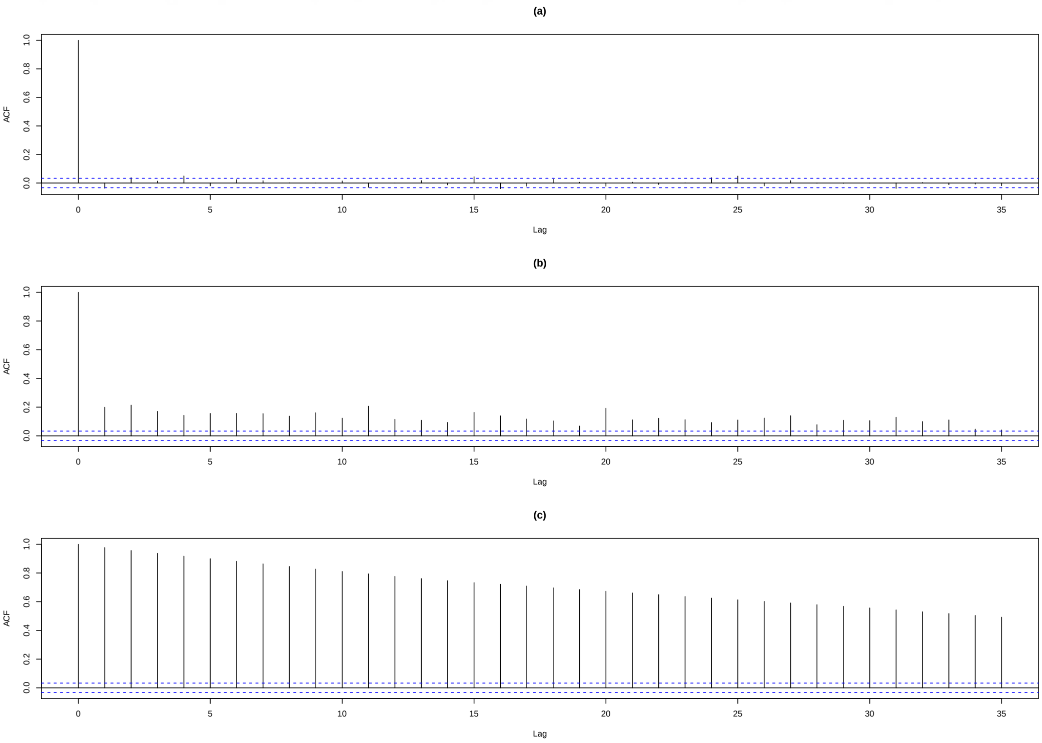

Figure 15. (a) ACF plot of the Increments $G_{t+1} - G_t$ (b) ACF plot of the Squared Increments $(G_{t+1} - G_t)^2$ (c) ACF plot of the volatility.

2.4. Conclusion

This chapter has identified a stochastic model that explains the behavioural properties of the BIS100. GARCH diagnositc tests implies a fairly good model and distributional tests show Miexner distrbution to be a better distribution to fit the GARCH residuals. We have applied COGARCH(1,1) with Meixner Lévy process to model the volatility of the series. To be able to estimate the parameters of COGARCH(1,1) model driven by a Meixner Lévy process, Pseudo Maximum Likelihood (PML) is used. To be able to test the consistency of COGARCH model over GARCH model, a simulation study carried out. According to Bias, RMSE, and MSE criterions, COGARCH volatility imitates the real absolute return much better than the GARCH model's volatility does for BIST100.

3. Long Memory Modelling of Stock Exchange Markets

3.1. Introduction

The manifestation of long memory in financial time series and economics (preceded by physical sciences such as astronomy, climatology and geophysics - see [13] for comprehensive overview) has attracted a considerable degree of attention both in academia and financial markets. In time series analysis, the term "long memory" or "long-range memory" refers to presence of significant dependence (correlation) between distant (in time) observations. For financial data this would mean that there is potentially a degree of predictability in asset returns due to correlated observations (at long lags) and therefore opportunity to generate sustained gains by exploiting trends in time series. This, by very definition, in contradiction with the weak form of EMH.

The hypothesis of long memory has become popular for stock market returns. Greene and Fielitz [78] investigated daily returns of 200 stocks trading in NYSE by utilizing a graphical method called the rescaled range (R/S) method developed by Hurst [93] to investigate long term dependence and reported presence of "persistent" statistical dependence. A couple of important considerations needs to be taken into account regarding this study. Although an effective graphical technique, the method suffers from having a well defined distribution to be used as an accurate statistical method [192]. Moreover, Lo [123] contradicted these results by suggesting that this was due to the *classic* R/S method's sensitivity to short-term correlations. Instead, Lo [123] applied a *modified* R/S method to investigate any presence of long memory in CRSP stock index returns andreported against any significant findings of long term dependence. Cheung and Lai [52] also reported lack of enough evidence in favor of long memory by applying Lo's modified method and a fractal differencing test to international stock index series. Crato [53] used exact maximum likelihood estimation and could not find any evidence of persistency either while studying G-7 stock markets. However, in their review of the modified R/S method Teverovsky et al. [192] highlights some drawbacks of Lo's method [123] and suggests that a single estimator such as the modified R/S method by itself cannot be relied upon as the ultimate method for long memory test and more diverse tests (including graphical and statistical) should be applied to rule out any long term dependence in a time series.

The characteristics profile of the ACF of time series that exhibit long memory is slowly decaying (or hyperbolic) ACF rather than exponential with increasing lags, such that their sum do not converge. In his study of multitude of time series data, Taylor [191] reported significantly higher correlation among squared returns than the returns themselves. Kariya et al. [100] reported similar results for Japanese stocks and Ding et al. [58] who reported a hyperbolic ACF decay in absolute returns of S&P 500 prices. Other notable works that found evidences of autocorrelation in asset returns are: in separate studies Fama [130] and Brock et al. [36] reported positive correlation in daily returns of publicly traded names in DJIA; Lim et al. [120] examined a number of emerging stock markets in Asia and found evidence of predictable nonlinearities in all the return series even after removing linear autocorrelation from the time series; and Serletis et al. reported mean reversion property in the US stock market.

In light of the forementioned studies, in order to be model the observed persistence in the stock market volatility Baillie et al. [12] introduced the Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity (FIGARCH) by extending the well known GARCH and allow for persistence in the conditional variance. The most important aspect of the model is its ability to explain the temporal dependencies in volatility by allowing a slow hyperbolically decaying lagged absolute or squared innovations in the conditional variance function.

Another popular method to analyze long-term temporal dependencies is the Detrended Fluctuation Analysis (DFA) method [155]. In contrast to the R/S method where the range (R) of fluctuations is determined within a given sub-period, DFA inspects the fluctuations around trend rather than the range. More details can be found at [189], [98],[117], [51].

3.2. Methodology

There are multiple ways of describing the properties of long memory processes and the most common definition is in terms of ACF. Recall that ACF measures the linear temporal dependence of a time series and therefore (potentially) indicating the degree of persistence in the data, which is evaluated by studying the rate of decay as a function of increasing lag.

Let $\gamma(k)$ be the autocovariance function and $\rho(k) = \gamma_k/\gamma_0$ be the corresponding ACF for some lag k, where $k = 1, 2, 3, \ldots$. Then the ACF indicates existence of long memory if

$$\lim_{n\to\infty} \sum_{k=-n}^{n} |\rho(k)| = \infty \tag{3.1}$$

Alternatively, a process said to have long memory if the autocovariance function decays follows a power law such as

$$\gamma(k) \sim k^{-\alpha} L(k) \quad \text{for} \quad k \to \infty \tag{3.2}$$

where $\alpha \in (0,1)$ and $L(k)$ is a slowly varying function[1]. Smaller α is indicative of longer memory [119].

[1] $\lim_{x\to\infty} L(ax)/L(x) = 1$ for all $a > 0$

The long term dependence can also be evaluated in terms of the Hurst exponent, H (also known as Hurst parameter or coefficient) [93] and related to α such that $E[R_n/S_n] \propto n^H$ as $n \to \infty$, where $H = 1 - \alpha/2$. The exponent, therefore, takes values between 0 and 1. The long memory property is characterized by $H > 0.5$. $H = 0.5$ corresponds to the standard Brownian motion, as such indicates no temporal dependence between observations, and H less than 0.5 suggests anti-persistent features i.e. mean-reversion.

Another method to identify dependence in a time series is investigating the series for periodic behavior by representing the covariance of the series by spectral density function via Fourier transform. Diverging spectral density, $f(\lambda)$ as frequency λ tends to zero is interpreted as persistence or long-memory. This may be expressed as

$$f(\lambda) \sim \lambda^{\alpha-1} L(\lambda) \quad \text{as} \quad \lambda \to 0 \tag{3.3}$$

3.2.1. Classical Rescaled Range

The idea behind the classical R/S test [[93], [134] and [136]] is dissect the data set into sub-periods and measure the widest spread in each sub-period by comparing the *min* and *max* of cumulative sum of deviations from the sample mean, $\bar{X}_n$ to obtain mean zero series. This is then normalized by the sample standard deviation, s_n. More specifically, the Hurst-Mandelbrot rescaled range statistic is defined as

$$(R_n/S_n) = \frac{1}{s_n}\left[\overset{Max}{1 \le k \le n} \sum_{j=1}^{k} \left(X_j - \bar{X}_n\right) - \overset{Min}{1 \le k \le n} \sum_{j=1}^{k} \left(X_j - \bar{X}_n\right)\right] \tag{3.4}$$

where

$$s_n = \left[\frac{1}{n}\sum_{j=1}^{n} \left(X_j - \bar{X}_n\right)^2\right]^{1/2} \tag{3.5}$$

As discussed above, the classical R/S suffers from short memory bias, such that the method may not be reliable especially when dealing with short memory process with long range component. Lo [123] proposed a modified R/S statistic.

3.2.2. Lo's Modified R/S

In the modified R/S, the denominator is redefined and is a function of a somewhat discretionary lag parameter, ξ and can be defined as

$$Q_n = \frac{R_n}{\bar{\sigma}_n(\xi)} \tag{3.6}$$

where short memory processes:

$$\tilde{\sigma}_n(\xi) = s_n^2 + \frac{2}{n}\sum_{j=1}^{\xi}\omega_j(\xi)\left[\sum_{i=j+1}^{n}\left(X_j - \bar{X}_n\right)\left(X_{i-j} - \bar{X}_n\right)\right] \tag{3.7}$$

$$= s_n^2 + 2\sum_{j=1}^{\xi}\omega_j(\xi)\tilde{\gamma}_j, \quad \omega_j(\xi) \equiv 1 - \frac{j}{\xi+1}, \quad \xi < n \tag{3.8}$$

where $\overline{\gamma}_j$ is the sample covariance up to lag ξ of X. This reduces to the classical form when $\xi = 0$.

The main practical difficulty in utilizing the modified R/S is the lack of a robust methodology in determining the optimal lag value, ξ. As noted in [63], too small ξ may lead to disregarding significant autocovariances, or lead to sample distribution deviating from its asymptotic limit if ξ is too high.

Although, as noted in [63], no known method has been proven to determine what the optimal lag ξ would be for a given time series, Andrews [3] provides the following relationship

$$\xi^* = [k_n], k_n \equiv \left(\frac{3n}{2}\right)^{1/3} \cdot \left(\frac{2\hat{\rho}}{1-\hat{\rho}^2}\right)^{2/3} \tag{3.9}$$

where $[k_n]$ is the greatest integer less than or equal to k_n and $\hat{\rho}$ is the first-order autocorrelation coefficient.

3.2.3. ARFIMA and FIGARCH

Generalized Autoregressive Conditional Heteroskedasticity, GARCH, processes are unable to adequately capture the persistency, in terms of slowly decaying ACF, in financial time series. To address this limitation, Granger and Joyeux [80] and Hosking [91] developed the ARFIMA process which stands for autoregressive fractionally integrated moving average. ARFIMA offers the flexibility to model long range dependence and expressed as

$$(1-L)^d \Phi(L)(x_i - \mu) \quad = \quad \Theta(L)\varepsilon_t \tag{3.10}$$

where the lag operator L is defined as $Lx_i = x_{i-1}$, ε_t is white noise, d is the difference parameter, $\Phi(L) = \sum_{i=1}^{p} \phi_i L^i$ and $\Theta(L) = \sum_{i=1}^{q} \theta_i L^i$ are the usual AR and MA polynomials - usual parameter restrictions apply to ensure stationarity and invertibility of the process. Regarding the non-integer integration parameter, the process is said to be nonstationary when $d \geq 1/2$. When d is in the range (0,1/2) the process is stationary and said to exhibit long range dependency. For $d = 0$ it is a short memory process corresponding to stationary and invertible ARMA, and anti-persistence for $d \in (-1/2, 0)$.

Subsequent to the introduction of the Integrated GARCH (IGARCH) [65], Baillie et al. [12] put forward the FIGARCH model, that stands for fractionally integrated GARCH, as more generalized, flexible model to be able to capture the persistence in the conditional variance. FIGARCH is an extension of ARMA type representation of GARCH(p,q) of Engle[64] and Bollerslev[28] as

$$\sigma_t^2 \quad = \quad \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \ldots + \alpha_q \varepsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \ldots + \beta_p \sigma_{t-p}^2 \tag{3.11}$$

GARCH(p,q) can also rewritten using L as lag operator

$$\sigma_t^2 \quad = \quad \alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2 \tag{3.12}$$

where $\alpha(L) = \sum_{i=1}^{q} \alpha_i L^i$, $\beta(L) = \sum_{j=1}^{p} \beta_j L^j$

Defining $v_t = \varepsilon_t^2 x \sigma_t^2$, then the GARCH process can be represented as an ARMA(m,p) process

$$[1 - \alpha(L) - \beta(L)]\varepsilon_t^2 = \alpha_0 + [1 - \beta(L)]v_t \tag{3.13}$$

where $m = max(p,q)$. Bollerslev and Engle [28] defined IGARCH(p,q) process

$$(1 - L)\phi(L)\varepsilon_t^2 = \alpha_0 + [1 - \beta(L)]v_t \tag{3.14}$$

where $\phi(L) = \sum_{i=1}^{m-1} \varphi_i L^i$ and is of order $m-1$.

The FIGARCH(p,d,m) process is then defined as

$$(1 - L)^d \phi(L)\varepsilon_t^2 = \alpha_0 + [1 - \beta(L)]v_t \tag{3.15}$$

An Exponential GARCH (EGARCH) process can also be formulated as an ARMA as a function of log of conditional variance - and therefore ensuring conditional variance remains positive. This is followed by the introduction of the fractionally integrated EGARCH or FIEGARCH process [31], which is expressed as

$$(1 - L)^d \phi(L) \ln \sigma_t^2 = \alpha_0 + \sum_{j=1}^{q} (b_j \left| x_{t-j} + \gamma_j x_{t-j} \right|) \tag{3.16}$$

where $\gamma_j \neq 0$, and x_t is the standardized residual, ε_t / σ_t. The process is stationary for $d \in (0,1)$.

3.3. Long Memory: Review of Previous Studies

There have been significant quantity of researches investigating evidences (or the lack thereof) temporal dependence (long memory) in financial markets data. Notable studies that reported evidence of persistence in financial volatility include Cunado et al. [56] who studied the S&P 500 time series extending as far back as to the Great Depression and reported long range behavior in the squared returns - higher persistence in bear markets is reported relative to bull markets. Tolvi [193] and Lillo and Farmer [119] show findings indicative of long range dependence for the stock markets of Finland and London, respectively. Presence of long memory also reported for six major global stock markets across regions in a study by [76].

On the other hand, some of the studies have reported lack of notable persistence in the US stock market data [[123],[52], [204], [200]]. Caporale and Gil-Alana [46] report little evidence of fractional integration in their study of 1700 observations of the S&P 500. Sadique and Silvapulle [166] present somewhat mixed results and report long range dependency in returns of some of the emerging Asian stock markets and New Zeland, whereas lack of persistence for Japan, Australia and the US markets. Similarly Henry [88] reported evidences of long memory for Germany, Japan, Taiwan and S. Korea and found no long range dependence from the study of the US, UK, Singapore, Hong Kong and Australia markets.

It is important to note that the forementioned studies have relied on different methodologies in their investigation, and some of the results may be effected by the choice of model parameters.

Furthermore, the studies carried out in [[5], [7], [6], [19], [105], [203]], using the FIGARCH process, the studies reported evidences of long memory in the volatility of the US market and five emerging markets: Brazil, Egypt, Kuwait, Turkey and Tunisia. However, Kilic [105] found contradictory evidence for emerging markets using parametric

FIGARCH process and non-parametric methods. Korkmaz et al. [113] reaches to similar conclusions and report no long range dependence in daily returns; however, report dynamic long memory in the conditional variance, which can be sufficiently captured by a FIGARCH model.

3.4. Results And Analysis

3.4.1. Data

Five stock exchange market index value is used to analyse the existence of the long memory in this chapter. The time series period is from Jan 2000 to Dec 2013. The variable X_t stands for the log return between time t and $t-1$. The stock market indices and the corresponding sample sizes of the time series are summarized in Table 12 alongside the moments of the distribution and Jarque-Bera probability.

Table 12 Descriptive Statistics of Log Returns

Data	Sample Size	Mean	Standard Deviation	Excess Kurtosis	Skewness	JB	ARCH-LM
BIST100	3504	0.003863	0.02327	6.669	-0.06766	0	0
DAX	3570	0.0009723	0.01577	4.179	0.02027	0	0
FTSE100	3651	-0.00007253	0.01248	6.245	-0.1532	0	0
S&P500	3520	0.0006794	0.01315	7.713	-0.1754	0	0
NIKKEI225	3440	-0.0004476	0.01576	6.263	-0.4205	0	0

Apart from DAX, the other indices' log return data exhibit negative skewness. Whereas the sample data for DAX is positively skewed with the value of 0.02027. All of the return data has positive excess kurtosis (leptokurtic) which is greater than the normal distribution, which means higher and sharper central peak and longer, heavier tails. Jarque-Bera's p-value is 0 for all of the data set which implies that none of the index series are normally distributed. And H_0 is rejected for ARCH-LM tests.

It is possible to analyze the degree of autocorrelations through different methods. One of the methods is computing the ACF and plotting it as a function of time lag

Table 13 P-values of unit root and stationarity tests of Log Returns

Data	ADF Test	PP Test	KPSS Test	Zivot-Andrews Test
BIST100	5.614e-14	0	0.0865	0.0001698
DAX	1.149e-13	0	0.2482	5.27e-09
FTSE100	1.095e-15	0	0.1973	2.261e-16
S&P500	9.334e-13	0	0.2622	5.305e-14
NIKKEI225	6.022e-11	0	0.2949	0.0002483

parameter. If the autocorrelation is not between the high and low confidence intervals then the data said to be correlated over some period of lags. The confidence interval is estimated via $(\pm)1.96/\sqrt{N}$.

The data sets that are used are; log return X_t, absolute return $|X_t|$ and squared return X_t^2 of BIST100 daily closing price index value. Table 14 has the result of the sample autocorrelations of X_t, $|X_t|$,and X_t^2 for lags $L = 1,2,3,4,5,10,20,40,70,100$. Also autocorrelogram of the datasets are plotted in Figure 16. The dotted lines show the 95% confidence interval. Barlett [21] showed that if X_t is an iid process that the sample autocorrelation can be approximated by $N(0,1/N)$.

$$(\pm)1.96/\sqrt{3504} \quad = \quad \pm 0.03311 \tag{3.17}$$

Table 14 Sample autocorrelations of Log Return X_t, Absolute Log Return $|X_t|$,and Squared Log Return X_t^2 of BIST100

	1	2	3	4	5	10	20	40	70	100
X_t	0.01099	0.01293	-0.02160	-0.00159	-0.00981	0.05780	-0.02304	-0.01883	0.02370	-0.01533
$\|X_t\|$	0.24870	0.27066	0.23831	0.21504	0.21889	0.18715	0.15659	0.14710	0.11171	0.08934
X_t^2	0.29760	0.29916	0.18905	0.12164	0.14796	0.12389	0.07834	0.10813	0.06028	0.05154

Fama [69], Taylor [191], Hamao et al. [?] reported relatively small 1st order positive autocorrelation for stock market returns. This finding is not true for BIST100 return data where the first 5 lags are within the 95% confidence interval.

The ACF plots for both $|X_t|$ and X_t^2 also shown in Figure 16. Both ACFs seem to lie outside the confidence interval, with ACF of absolute returns greater than that of squared-returns at almost every lag up to 100. This suggests that the BIST100 index return

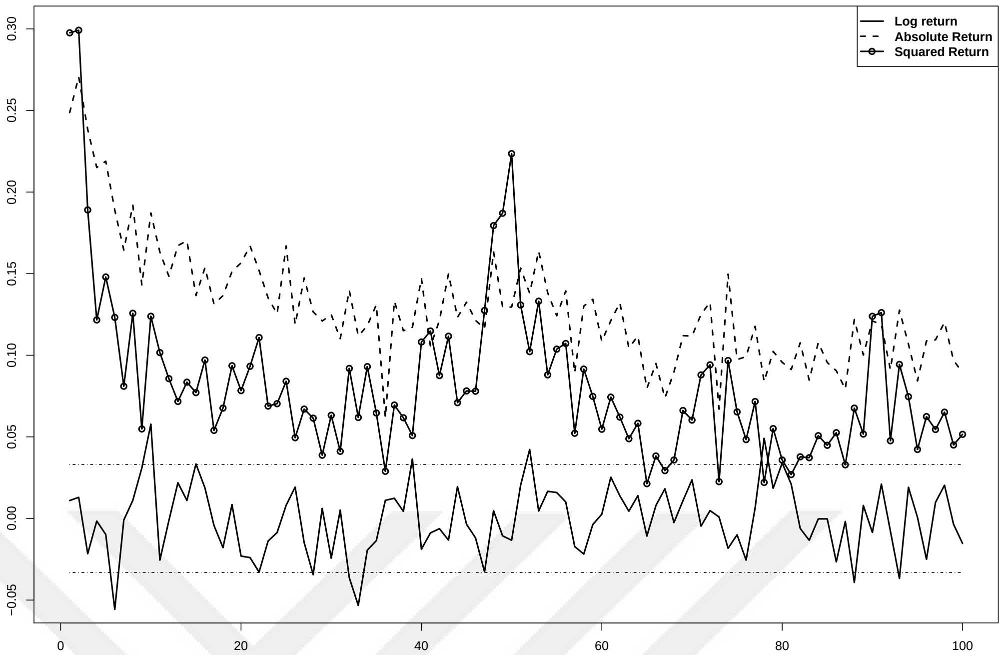

Figure 16. ACF plots of Log return, absolute return and squared return of BIST100

process not an iid process.

Absolute log return dataset $(|X_t|)^d$ is further examined with $d = 0.125, 0.25, 0.50,$ $0.75, 1,$ $1.25, 1.5, 1.75, 2, 3$ at autocorrelations lags of $1, 2, 3, 4, 5, 10, 20, 40, 70, 100$ for the major indices. From Table 15 and Figures 16, 18 and 19 support the conclusion above. The power transformations of the absolute daily returns exhibit significant positive autocorrelations at lags up to 100. Initially the ACFs rate of decay is relatively fast but slows down significantly with increasing lags. This suggests that the index returns exhibit long-range dependency. It should also be noted that the largest autocorrelations are where $d = 1$ or near 1.

The lags for which the first negative autocorrelation of absolute log returns, $|X_t|$ occur for varying d is shown Table 16. In majority of cases, $|X_t|$ has positive autocorrelations, suggesting positive autocorrelations for more than 2 years (based on 250 trading days in a year).

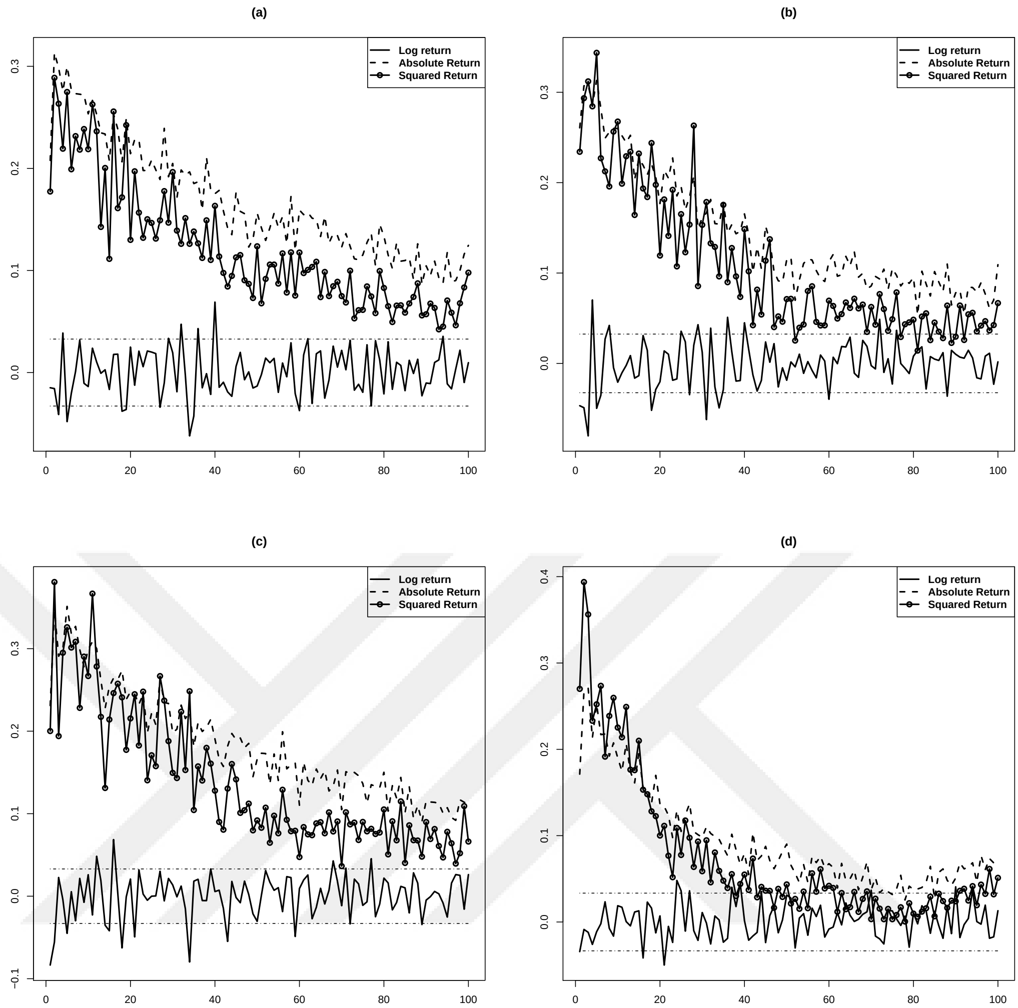

Figure 17. ACF plots of Log return, absolute return and squared return of (a)DAX, (b)FTSE100, (c)S&P500, (d)NIKKEI225.

Sample autocorrelations of $|X_t|$ up to lag 500 is plotted in Figure 20. The figure clearly shows that ACF remain outside the confidence interval, marked by the dotted lines.

Recall that in a long memory process ACF decays hyperbolic rather than exponential. The autocorrelation for $|X_t|$ and X_t^2 is almost always positive and decays rather moderately, in contrast to X_t.

Another set of tests for presence of long memory is the R/S tests of Hurst [93] and modified R/S of Lo [123], also the test proposed by Geweke and Porter-Hudak [74], which is based on the fractionally integrated process.

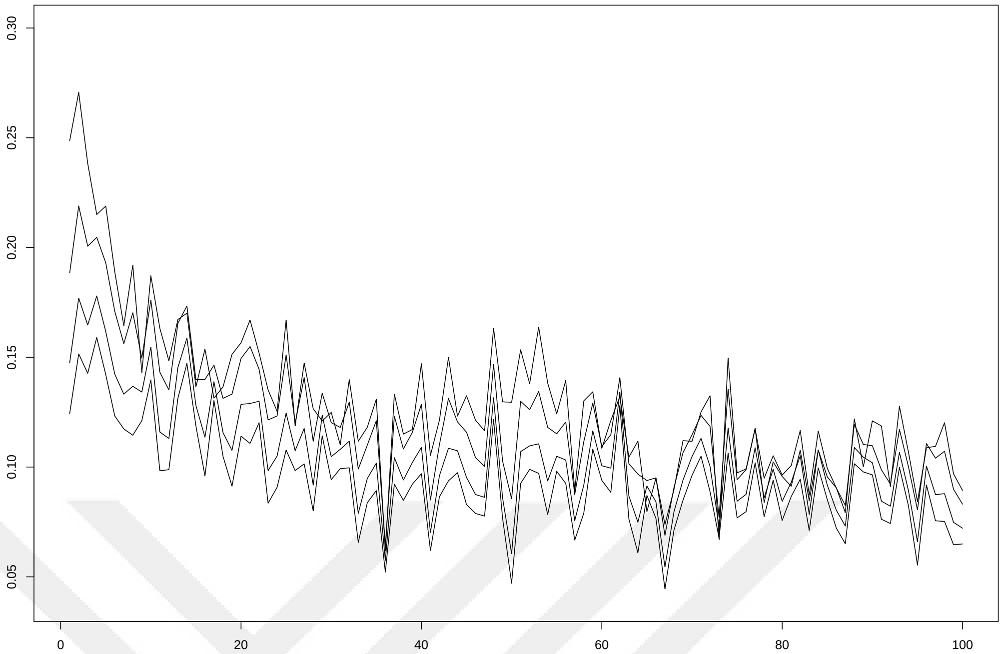

Figure 18. Autocorrelation of BIST100 absolute log return for 100 lags where d = 1, 0.5, 0.25 , 0.125

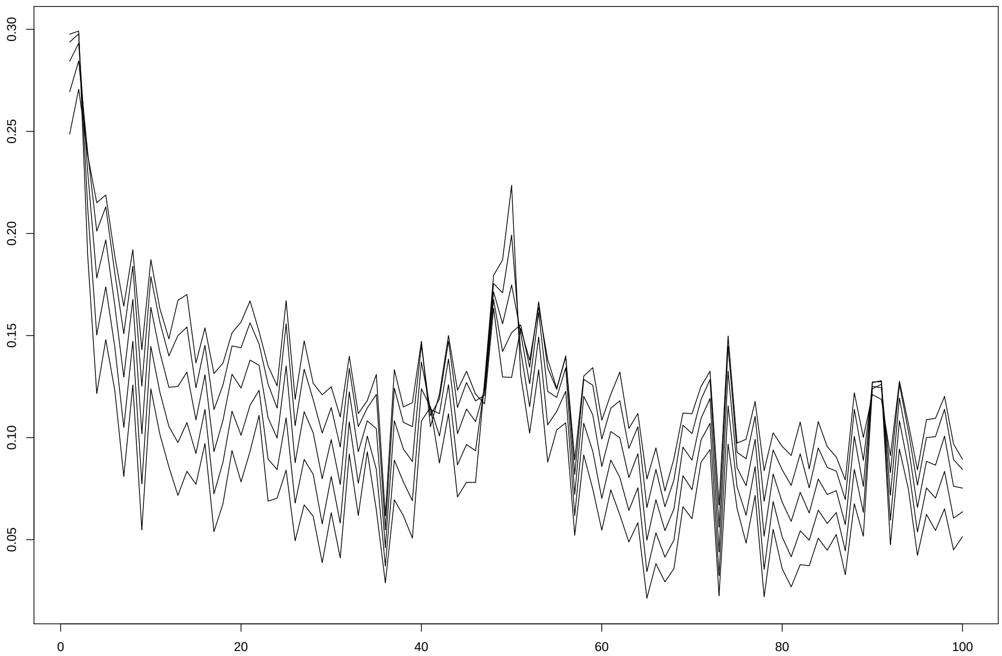

Figure 19. Autocorrelation of BIST100 absolute log return for 100 lags where d = 1, 1.25, 1.50, 1.75, 2.

Table 15 Power transformations of $|X_t|$'s Autocorrelations lags of BIST100

	1	2	3	4	5	10	20	40	70	100
0.125	0.12437	0.15153	0.14270	0.15901	0.14208	0.13972	0.11405	0.09698	0.09618	0.06496
0.25	0.14757	0.17694	0.16470	0.17794	0.16172	0.15459	0.12856	0.10899	0.10479	0.07214
0.5	0.18842	0.21892	0.20061	0.20466	0.19300	0.17609	0.14948	0.12864	0.11469	0.08309
0.75	0.22192	0.24955	0.22531	0.21677	0.21234	0.18670	0.15887	0.14156	0.11646	0.08897
1	0.24870	0.27066	0.23831	0.21504	0.21889	0.18715	0.15659	0.14710	0.11171	0.08934
1.25	0.26941	0.28454	0.23958	0.20114	0.21304	0.17883	0.14405	0.14529	0.10197	0.08440
1.5	0.28440	0.29316	0.23014	0.17818	0.19685	0.16380	0.12426	0.13699	0.08897	0.07526
1.75	0.29373	0.29779	0.21226	0.15029	0.17384	0.14463	0.10116	0.12391	0.07454	0.06368
2	0.29760	0.29916	0.18905	0.12164	0.14796	0.12389	0.07834	0.10813	0.06028	0.05154
3	0.27326	0.28054	0.09464	0.04081	0.06356	0.05568	0.02001	0.04967	0.02008	0.01689

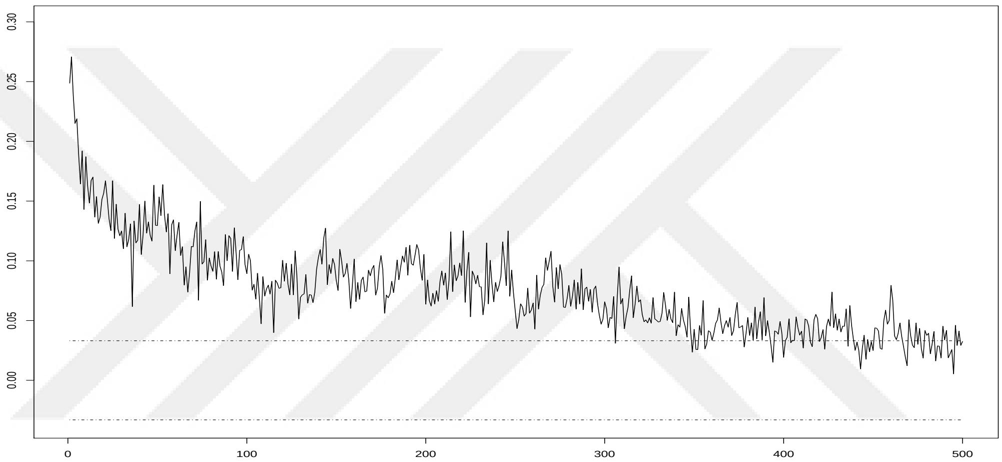

Figure 20. Autocorrologram of BIST100 absolute return data for 500 lags.

Table 16 Lags at which first negative autocorrelations of $|X_t|$ occurs of BIST100

d	0.125	0.25	0.5	0.75	1	1.25	1.5	1.75	2	3
lag	394	394	527	527	527	527	527	443	356	65

The R/S and modified R/S are significant at 1% level for squared and absolute return data and whereas there is no long-term dependence in the log return. The estimated values of d are 0.4401 and 0.4456 for BIST100, which suggests long memory. The Geweke & Porter-Hudak (1983) test statistic is 4.7019 and 5.9766, hence, H_0 is rejected at the 1% significance level for the squared and absolute return data. Although the d values are within the stationary range, they are nevertheless closer to the nonstationary value of 0.5. The opposite is true for long return data as GPH test statistic is not significant.

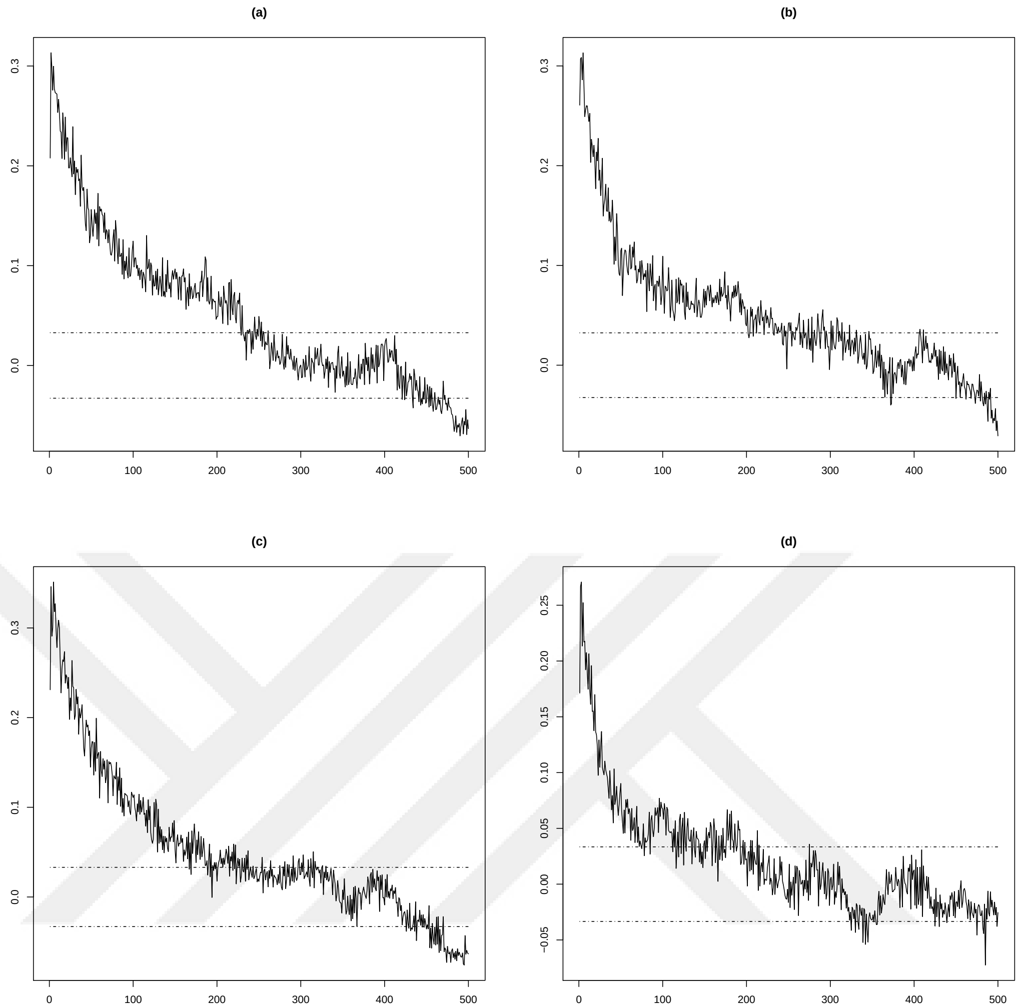

Figure 21. ACF plots of Log return, Absolute return for 500 lags (a)DAX, (b)FTSE100, (c)S&P500, (d)NIKKEI225.

We are also interested in calculating the Hurst exponent, H using the R/S method. Moreover, we will also use the periodogram and Whittle's methods.

For a stationary long memory process, Mandelbrot [136] showed that R/S converges to a r.v. at rate n^H, where H is the Hurst exponent. If the rate of convergence is $n^{1/2}$, this would indicate no long memory. And on a log-log plot of R/S as a function of sample size, the series exhibit long memory for slope $H > 1/2$.

Table 17 Long Memory Tests for Log Return X_t, Absolute Log Return $|X_t|$,and Squared Log Return X_t^2

Data	M-R/S	R/S	GPH
BIST100			
X_t	1.1014	1.0814	-0.9413(d=-0.0881)
$\|X_t\|$	4.5405**	7.925**	4.7019**(d=0.4401)
X_t^2	3.5837**	5.9766**	4.7608**(d=0.4456)
DAX			
X_t	1.5486	1.456	0.8115(d=0.076)
$\|X_t\|$	3.3816**	6.3016**	5.3074**(d=0.4967)
X_t^2	2.743**	4.8461**	4.4388**(d=0.4154)
FTSE100			
X_t	1.1419	0.9977	0.1558(d=0.0144)
$\|X_t\|$	3.4502**	6.5267**	5.2651**(d=0.488)
X_t^2	2.4954**	4.6362**	5.0308**(d=0.4663)
S&P500			
X_t	1.4451	1.2399	-0.5532(d=-0.0518)
$\|X_t\|$	3.342**	6.4406**	6.2476**(d=0.5847)
X_t^2	2.7718**	5.1833**	5.8277**(d=0.5454)
NIKKEI225			
X_t	1.1413	1.0654	0.4863(d=0.046)
$\|X_t\|$	2.6152**	4.5639**	3.8083**(d=0.36)
X_t^2	2.1945**	4.1799**	2.6581**(d=0.2513)

Following the R/S definition in section 3.2.1, an estimate for H can be obtained using log-log plots [207] :

- Calculate R/S using k_1 observations in succession.
- Increase number of sample size by some factor f such that $k_i = f k_{i-1}$ for $i = 2, \ldots, s$
- Fit a straight line of R/S vs k_i on the log-log scale.

Table 18 lists the R/S estimates of d and H. The Hurst value of log return X_t is near 0.5 which suggests a random walk. Whereas the Hurst estimates for absolute and squared return is 0.8085 and 0.8008 respectively, which are much higher then 0.5, showing that the absolute and squared returns are more persistent than the return data.

The weakness of the log-log plot of R/S is that it is not obvious which k_1 is correct for a chosen sample. Nevertheless, it is useful as an indication of presence of long memory.

Table 18 R/S estimate of long memory parameters of Log Return X_t, Absolute Log Return $|X_t|$,and Squared Log Return X_t^2

Data	d	H
BIST100		
X_t	0.02306515	0.5230651
$\|X_t\|$	0.3085443	0.8085443
X_t^2	0.3008888	0.8008888
DAX		
X_t	0.01826846	0.5182685
$\|X_t\|$	0.3853712	0.8853712
X_t^2	0.4041862	0.9041862
FTSE100		
X_t	-0.02702745	0.4729725
$\|X_t\|$	0.3277623	0.8277623
X_t^2	0.3522171	0.8522171
S&P500		
X_t	0.03847053	0.5384705
$\|X_t\|$	0.3864095	0.8864095
X_t^2	0.3264169	0.8264169
NIKKEI225		
X_t	0.04058932	0.5405893
$\|X_t\|$	0.2876189	0.7876189
X_t^2	0.2699192	0.7699192

One other approach to estimating H is the periodogram method. Figure 22, Figure 23 and Figure 24 show the log-log plot of periodogram against the frequency λ of spectral density function as defined in Equation 3.2. From Figure 22, $H \sim 0.5$ suggests random walk for log return series, however, absolute and squared return series exhibit persistency as suggested by H values higher than 0.5.

Table 19 Coefficients of FIGARCH(1,d,1) for Stock Exchange Markets

	α_0	α_1	β_1	d
BIST100	0.03796	0.09055	0.37852	0.39729
	(2.956)	(1.501)	(4.736)	(11.246)
DAX	0.03956	0.01926	0.58219	0.59205
	(3.452)	(0.571)	(9.484)	(10.523)
FTSE100	0.03351	0.06041	0.54976	0.54976
	(3.072)	(1.762)	(9.551)	(9.551)
S&P500	0.031706	0.003462	0.714944	0.724081
	(2.7449)	(0.1316)	(15.2587)	(11.6983)
NIKKEI225	0.03201	0.10502	0.62506	0.58965
	(2.388)	(3.230)	(10.671)	(8.837)

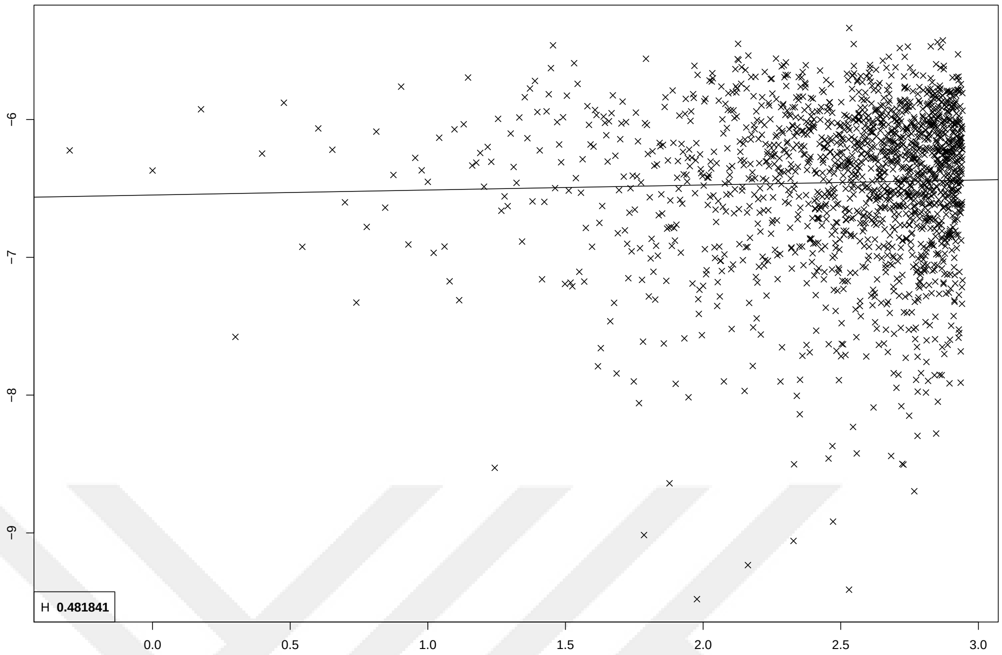

Figure 22. Log-Log Periodogram of Log Return

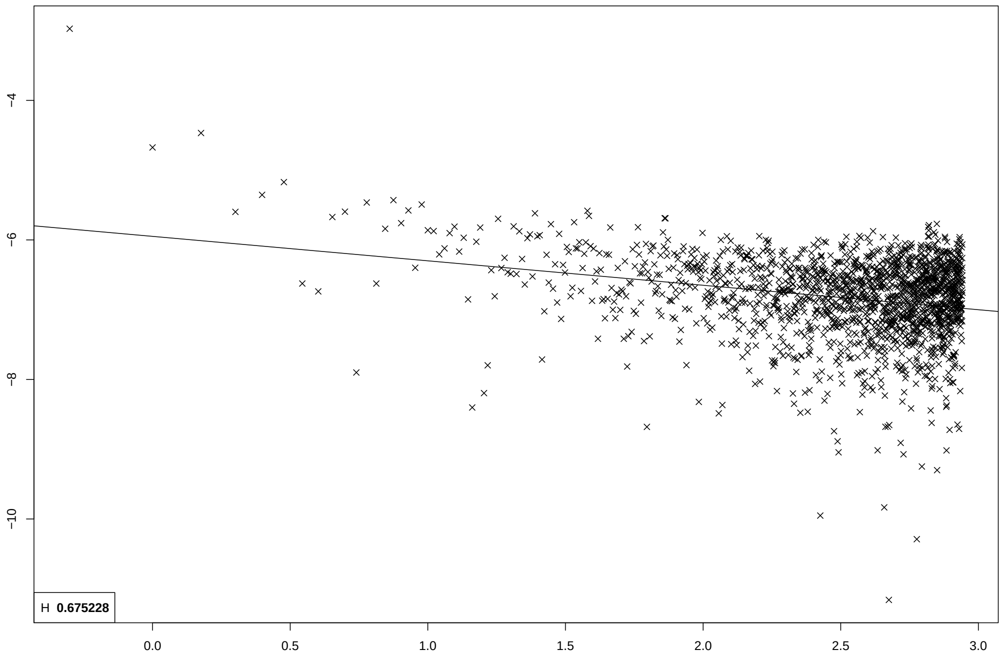

Figure 23. Log-Log Periodogram of Absolute Return

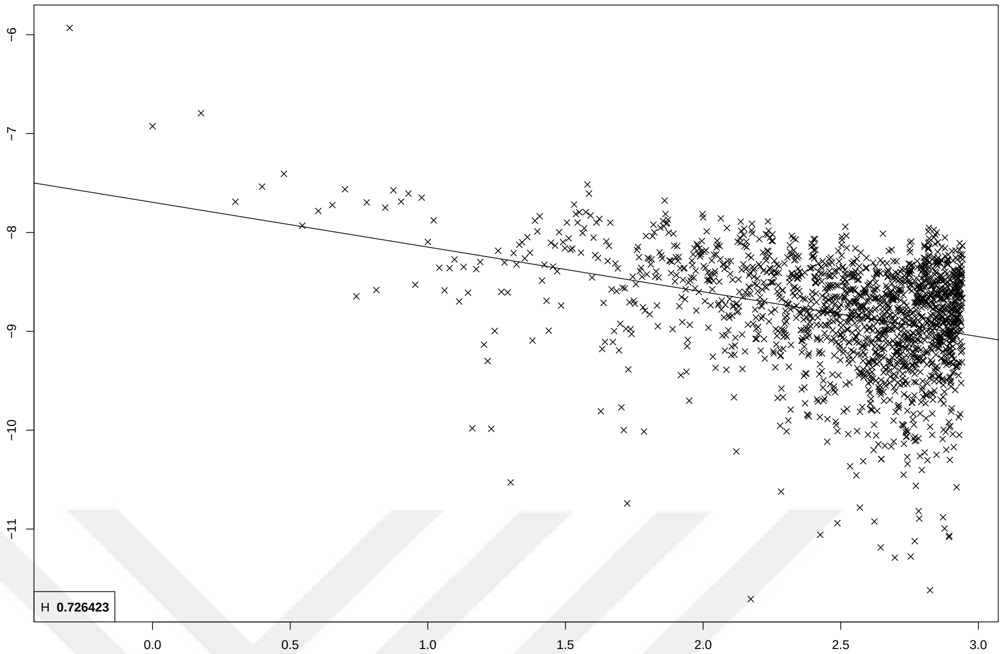

Figure 24. Log-Log Periodogram of Squared Return

Table 20 Coefficients of FIEGARCH(1,d,1) for Stock Exchange Markets

	α_0	α_1	β_1	d	Leverage
BIST100	0.02099	0.21530	0.54770	0.52340	-0.07956
	(1.789)	(11.727)	(8.701)	(15.983)	(-8.770)
DAX	0.01812	0.11744	0.67468	0.53532	-0.11671
	(1.610)	(7.694)	(12.795)	(16.428)	(-10.071)
FTSE100	0.01440	9.551	0.69394	0.57002	-0.10218
	(1.392)	(8.168)	(13.370)	(15.74)	(-9.853)
S&P500	0.009828	0.083863	0.808399	0.463368	-0.114650
	(0.9767)	(7.1191)	(25.7146)	(12.8054)	(-12.1877)
NIKKEI225	0.004211	0.179685	0.949967	0.094472	-0.070832
	(0.314)	(11.1800)	(60.4185)	(1.4756)	(-10.1306)

3.5. Conclusion

In this study, we investigated the presence of long memory in five major stock indices. The descriptive statistics on the log returns suggested none of the indices follow a normal distribution. Studying the ACF of the series we find strong correlations in absolute and squared returns but not in the log-return measure. The relatively slow rate of decay of ACF suggest persistence for absolute and squared returns. This is also

Table 21 Diagnostic tests for FIGARCH(1,d,1) & FIEGARCH(1,d,1) residuals

	FIGARCH		FIEGARCH	
	Ljung-Box test	ARCH LM Test	Ljung-Box test	ARCH LM Test
BIST100	13.32 (0.3465)	28.8050 (0.7606)	20.82 (0.05311)	42.1963 (0.1878)
DAX	12.25 (0.4259)	35.1687 (0.4602)	25.36 (0.01323)	71.0744 (0.0003)
FTSE100	12.17 (0.4319)	37.4398 (0.3578)	10.75 (0.5505)	49.8887 (0.0492)
S&P500	10.87 (0.54)	29.1021 (0.7479)	23.11 (0.02678)	49.4482 (0.0536)
NIKKEI225	12.82 (0.3825)	36.1242 (0.4158)	16.64 (0.1638)	33.7397 (0.5289)

Table 22 Comparison of FIGARCH and FIEGARCH for Stock Exchange Markets' Volatility

	FIGARCH			FIEGARCH		
	AIC	BIC	Likelihood	AIC	BIC	Likelihood
BIST100	8824	8855	-4407	8808	8845	-4398
DAX	8752	8783	-4371	8623	8660	-4305
FTSE100	8796	8827	-4393	8669	8707	-4329
S&P500	8348	8379	-4179	8202	8239	-4095
NIKKEI225	8901	8932	-4446	8855	8892	-4421

Table 23 Diagnostics; in-sample forecast of FIGARCH(1,d,1), FIEGARCH(1,d,1) & ARFIMA models

	FIGARCH		FIEGARCH		ARFIMA	
	MSE	MAE	MSE	MAE	MSE	MAE
BIST100	0.4767409	0.514694	0.464641	0.5109722	0.3811597	0.4281277
DAX	0.4461451	0.5139675	0.4152824	0.4953952	0.3395098	0.4177897
FTSE100	0.4680225	0.5159568	0.4322233	0.4963509	0.3370364	0.4077076
S&P500	0.4640043	0.5116978	0.4221534	0.4850083	0.3395626	0.4057265
NIKKEI225	0.4610738	0.52935	0.4400202	0.5179966	0.3361	0.4248185

supported by the R/S analysis we carried out as well as the Hurst exponent estimates obtained from the log-log periodogram plots.

In the light of dependency analysis, FIGARCH, FIEGARCH and ARFIMA models considered to find the most suitable model for the indices. By looking at the values of AIC and BIC, FIEGARCH model is a better model compared to FIGARCH. Whereas once the in-sample forecast is carried out, it has been easily seen that ARFIMA model outweighs the other two considered model. MSE and MAE values for ARFIMA model

Table 24 Parameter estimation of ARFIMA model for absolute standardized returns of BIST100

	Value	Std. Error	t value
d	0.3326	0.0182	18.2815
MA(1)	0.2305	0.0226	10.2041
log-likelihood	-3461.539		
BIC	6939.398		

Table 25 Parameter estimation of ARFIMA model for absolute standardized returns of DAX

	Value	Std. Error	t value
d	0.4825	0.0157	30.7807
MA(1)	0.5056	0.0173	29.2950
log-likelihood	-3381.462		
BIC	6779.282		

Table 26 Parameter estimation of ARFIMA model for absolute standardized returns of FTSE

	Value	Std. Error	t value
d 0.5387	0.0503	10.6994	
AR(1)	-1.4416	0.0257	-56.0144
AR(2)	-1.5025	0.0182	-82.4735
AR(3)	-0.7264	0.0253	-28.6566
MA(1)	-0.9492	0.0593	-16.0148
MA(2)	-0.7762	0.0483	-16.0641
MA(3)	0.0403	0.0554	0.7281
MA(4)	0.4049	0.0328	12.3364
log-likelihood	-3537.180		
BIC	7139.971		

give a much more promising results. Also the parameter estimation of ARFIMA model supports a good fit to the data sets.

Table 27 Parameter estimation of ARFIMA model for absolute standardized returns of S&P500

	Value	Std. Error	t value
d	-0.3058	0.0517	-5.9161
AR(1)	1.7395	0.0101	171.5700
AR(2)	-1.6449	0.0153	-107.5588
AR(3)	0.8981	0.0091	98.6059
MA(1)	1.4695	0.0617	23.8298
MA(2)	-1.4841	0.0992	-14.9565
MA(3)	0.7335	0.0494	14.8501
MA(4)	-0.0626	0.0607	-1.0320
log-likelihood	-3360.498		
BIC	6786.314		

Table 28 Parameter estimation of ARFIMA model for absolute standardized returns of NIKKEI225

	Value	Std. Error	t value
d	0.4511	0.0164	27.4389
MA(1)	0.4414	0.0188	23.4474
log-likelihood	-3308.396		
BIC	6633.076		

4. Wavelet and Hurst Analysis of DAX and NIKKEI Stock Markets

4.1. Introduction

Wavelet coherence and Hurst analysis of two major stock market's daily log returns and volatility is the main focus in this chapter. The presence of long memory in the data sets are analysed by applying autocorrelation and long memory tests. It has been seen that both stock markets' volatility has a long memory property. Furthermore the volatility analysed by multifractal model. The multifractal process of DAX and NIKKEI volatility is apparent by considering the Hurst exponent and multifractal spectrum. MD-DFA method is utilized to estimate the Hurst exponent. The results of wavelet coherence demonstrates that two stock markets are quiet coherent at the frequency of 64 and 128 days and strongly coherent at 256 days. Given the selected period and values of Hurst, VARIMA model is applied to forecast the next 30 days for the stock indices.

Multifractals (or multi-scaling - in terms of scaling exponents or powers) process (Kantelhardt [103], Eisler and Kertész [62], Hentschel and Procaccia [89], Halsey et al. [86]) is becoming a popular concept for financial time series. First applied by Mandelbrot [133], multifractals provide a more comprehensive framework for modelling financial time series in comparison to *benchmark* methods including GARCH based processes. The multifractal models offer generalized framework to capture the financial markets stylized facts including heavy tailedness, persistency, correlation, volatility clustering as well as correlations across assets [179]. One of the main advantages multifractals offer is their ability to produce varying levels of long memory in distinct powers of asset returns, which is the property of nearly all financial prices. The usefulness of this approach is reflected by the following recent works; [4], [70], [33], [186], [141].

This topic has been attempted in many different papers such as Vassilicos, Demos and Tata [198], Ghasghaie et al. [75] studied the scaling behavior of the distributions of foreign exchange price volatilities, which would not be captured by the more traditional GARCH type models. The multifractal property of various financial data were also reported in [178], ([196], and [197]). Mandelbrot et al. [139] presented the multifractal model of asset returns (MMAR) that incorporates the empirical features of asset returns such as heavy tailedness and long memory, and proposes the model as a substitute to ARCH-type processes with advantages of scale-consistency (when dealing with different sampling frequencies) [139] that simulated multifractal processes able to recreate the majority of the observational characteristics of asset returns.

The key point of this section is to forecast Hurst exponent values using vector ARIMA model and analyze the results to see if this model can be considered a useful tool for investors. As discussed in chapter 3, the Hurst coefficient is a quantitative measure indicating whether a given time series exhibit persistence or not. The exponent in effect is an indication of the structure of the correlation in the asset returns and used as a statistical tool to gauge the level of deviation of time series from Brownian motion. This may be of some interest in deciding investment strategies: if, for a given time scale, the exponent changes, it may signal a regime shift.

Moreover, to determine monofractal scaling characteristics and to identify persistence in stationary time series, another extensively used method is the DFA method [189], [98], [117], [51]).

Another methodology that has been gathering traction in the study of time series applications is the wave function framework called wavelet analysis. This involves transforming sequential observations or signals into a different, more useful representation that would otherwise might not be observable using the more traditional methodologies. In financial time series, wavelet analysis can be used to reveal multiple structures that the time series might exhibit at varying scales of time. Perhaps the most beneficial aspect of wavelet methods over similar procedures, for example Fourier transform is that the latter assumes stationarity by imposing a scale and therefore not an efficient method to investigate localized time-frequency information [194], [85], whereas wavelet is able to conserve time-frequency and suitable to apply to non-stationary data set.

Multiscale is not an unnatural concept in the context of time series analysis, meaning the data can exhibit multiple structures that occur at varying scales of time. The method of wavelets is able to break the series down into smaller series each defined with a specific wave function (and hence the expression wave-let), where each time-window may be associated with a specific scale. Once the wavelet method is employed to compartmentalize these different time-scales, each compartment of serial observations can then be analyzed with ordinary time series methods.

Commonly used wavelets are grouped into two types: the continuous wavelet transform (CWT), which is preferred for scale analysis and therefore better in analyzing the time series features; and orthogonal or discrete wavelets which are better to utilize for data compression and noise reduction purposes. More detailed discussions on types of wavelets can be found in [85], [115]. A wide range of utilization of wavelet analysis in financial time series, economics, as well as econometrics has been reported in [161], [162], [163], [106], [94], [116].

In this study we will use CWTs for the two financial time series, DAX and NIKKEI to construct what is called a cross wavelet transform (XWT) [194] to measure the coherence or cross-correlations between the return series of the two indexes. Firstly the data sets will be analysed for long memory property by using autocorrelation techniques and various tests. Then by looking at the Hurst exponent and multifractal spectrum, multifractality will be detected for the two stock markets. We will apply wavelet coherence to study the cross-correlation of the two data sets with each other. As a final step in this study, forecast will be carried out using FARIMA model.

4.2. Methodology

In the interest of clarity, the two properties of a time series, persistence and multifractality, are closely linked but independent measures: long memory is a characteristics of temporal dependence and correlation withing the structure of the whole time series whereas fractals may be understood in the sense of localized memory of the time series. Fractals are typically referred or measured in terms of their dimension, represented by the letter D. This can be thought as a measure of 'complexity' or 'roughness' in terms of how detail in a given data structure changes with scale with higher values indicating increased complexity. In contrasts, the Hurst coefficient, represented by H, measures long-range dependecy in the time series of interest and associated with power law correlations [77]. A more detailed discussion on long memory and the review of previous studies are discussed in sections 3.2 and 3.3.

The notion of multifractal process was first introduced by Mandelbrot [135]. Analogous to the measure of Hurst coefficient H, which quantifies persistency in a time series, similarly, an equivalent fundamental concept for multifractal processes exists. This is called the (local) Hölder coefficient. Hölder exponent is used to quantify the scaling rate that governs an interval of time. Multifractality can then be viewed as the process to identify a range of Hölder exponents in comparison to a single value, namely the Hurst parameter - that applies to unifractal processes (for more detailed discussion see Calvet & Fisher [43]).

To study the distribution of Hölder exponents, multifractal spectrum is used as a suitable representation, which is a function of the Hölder exponent and measures the 'fractal dimension' or shows how regular a signal varies locally over time: unifractal signal exhibit the same regularity throughout as a function of time whereas multifractal signal show variations with time. Several analyses have already been carried out by utilizing the multifractal spectrum for stock indices and foreign exchange markets [[140], [22], [187], [42], [188], [45], [8], [96])] to study the financial time series structures.

However, another approach that is founded on the detrended fluctuation method and is referred to as multifractal DFA proposed by Peng et al. [155] and more details can be found in that study. We will not focus in this method in our study. We will be looking at the Hölder exponent and multifractal spectrum for multifractal analysis.

In this study, Grinsted et al. [85] wavelet package is used.A popular wavelet definition that is used for nonstationary time series is Morlet Wavelet, which we will consider in our analysis. We use CWT for each stock index and then construct XWT from the two. XWT permits the detection of cross-magnitude, phase differences, nonstationarity, and coherency between signals. And finally for the coherency of the data sets, Wavelet Transform Coherence (WTC) method will be used to analyze the coherence and any phase lags linking the two return series in temporal and frequency dimensions. /*

4.2.1. Detrended Fluctuation Analysis

The data $Y(t)$ is first integrated by

$$Y(t) = \sum_{i=1}^{N}(X_i - \bar{X}) \tag{4.1}$$

The de-meaned data is then dividen into equal time intervals and a trend line is generated using least squares method for each interval. The local trend, $y_v(i)$ of length of v in each time window is then subtracted from the integrated time series to obtain the integrated & detrended series. The RMS fluctuation is the given by

$$F(s,v)^2 = \frac{1}{s}\sum_{i=1}^{s}[Y((v-1)s+i) - y_v(i)]^2 \tag{4.2}$$

For each time scale v, the above step is repeated to generate a mapping between the average fluctuation, $F(s,v)$ and the scale v. The slope of log-log plot of $F(s,v)$ against v, gives the scaling exponent Δ. A completely uncorrelated data, i.e. white noise has $\Delta = 0.5$. And long range persistence is indicated when $\Delta \in (0.5, 1)$. */

4.2.2. Hurst Exponent

Let X_i be the observations or asset returns over N periods, the total deviation over the period is given by

$$X_{t,N} = \sum_{i=1}^{t} (X_i - \bar{X}_N) \tag{4.3}$$

The range R is then given by

$$R = Max(X_{t,N}) - Min(X_{t,N}) \tag{4.4}$$

R is then normalized by the std. dev. S to obtain the rescaled range R/S. Hurst gives the following relationship for the exponent H

$$R/S = aN^H \tag{4.5}$$

In the log-linear form, the formula can be re-written as

$$log(R/S) = log(a) + Hlog(N) \tag{4.6}$$

Calculating the gradient of the log-log plot provides an estimate for the coefficient H. Whereas the value of H cannot hint the shape of the distribution.

For the range of values the Hurst H exponent can take and their interpretations, the reader should refer to section 3.2.

4.2.3. Continuous Wavelet Transform (CWT)

A wave-like function, wavelet has limited duration (as opposed to continuous form such as sine waves) whose average is zero. A wavelet can be characterized by time locality δt and frequency $\delta \omega$. However, the choice of locality and frequency is limited according to Heisenberg's principle, therefore the user seeks to find the optimal resolution when defining the two parameters.

The CWT of a function $f(t)$ at scale $a > 0$ and position (translation) b is expressed as

$$C(a,b;f(t),\psi(t)) \simeq \int_{-\infty}^{\infty} f(t)\frac{1}{\sqrt{a}}\psi^{*}\left(\frac{t-b}{a}\right)dt \tag{4.7}$$

where ψ is a continuous (wave) function and ψ^{*} denotes the complex conjugate. The choice of ψ (mother wavelet) effectively provides the source function to generate the wavelet coefficients. The coefficients $C(a,b)$ can be solved for by varying a and b parameters.

A choice of suitable mother wavelet is required, which defines how the transformation between spatial and frequency domains are performed. A popular choice is the Morlet wavelet for its simplicity and it is well suited to analyse cyclical behaviour.

$$\psi_0 \simeq \pi^{-\frac{1}{4}} e^{i\omega_0 \eta} e^{-\frac{1}{2}\eta^2} \tag{4.8}$$

where ω_0 and η are dimensionless frequency and time, respectively.

Through graphical inspection of the map of signal dilation, characterized by long scale (stretched wavelength - low frequency), and compression, characterized by small scale (compressed wavelength - high frequency), the CWT is an effective method for examining how series changes with time and scale.

4.2.4. Cross Wavelet Transform (XWT)

Consider two different time series, $x(t)$ and y(t), their XWT may be expressed as [92].

$$W_{xy} \simeq W_x W_y^* \tag{4.9}$$

where W_x and W_y^* are wavelet transforms of x and y respectively, and W^* refers to the complex conjugate of function W. Between the two time series, as a measure of localized covariance, the cross-wavelet power is given by

$$P_{xy} \simeq |W_{xy}| \tag{4.10}$$

Torrence and Compo [194] give the following equality

$$D\left(\frac{\left|W_x W_y^*(s)\right|}{\sigma_x \sigma_y} < p \simeq \frac{Z_v(p)}{v}\sqrt{P_{x,k}P_{y,k}}\right) \tag{4.11}$$

where $P_{x,k}$ stands for background power spectrum of series x and frequency index k, the d.o.f denoted by v, σ stands for std. dev. and $Z_v(p)$ is the level of significance of prob. p.

4.2.5. Wavelet Transform Coherence (WTC)

The strength of relationship in time-frequency space between two different series is depicted by WTC plot. In a typical WTC heatmap chart, the value of 1 (red) indicates that the two signals are strongly correlated and 0 (blue) indicates no correlation.

Another measure of coherency between two time series is introduced by Torrence and Webster [195] for a given local time index and wavelet scale of one, and smoothing operator S may be expressed as

$$R^2 \simeq \frac{\left|S(W_{xy})\right|^2}{S(|W_x|^2)S(|W_y|^2)} \tag{4.12}$$

The above equation measures the strength of association and is analogous to the classic correlation coefficient definition. The operator S is defined by

$$S(W) \simeq S_s(S_t(W)) \tag{4.13}$$

where the terms S_s and S_t represent smoothing in the scale and in time dimensions, respectively. The smoothing applied to the Morlet wavelet is [195]

$$\begin{aligned} S_t(W) &\simeq W c_1 e^{\frac{-t^2}{2}} \\ S_s(W) &\simeq W c_2 \omega(0.6) \end{aligned} \tag{4.14}$$

where we used the scale parameter $s = 1$, c_1 and c_2 are constants and $\omega(.)$ the rectangle function, and decorrelation of 0.6 from empirical observation for Morlet [194].

4.3. Results And Analysis

There exists a number of ways to investigate the serial correlation in a time series. One of the most popular methods is through computing the ACF of the return series. If the autocorrelation is not between the high and low confidence intervals then the data said to be correlated over some period of lags. The confidence interval is estimated via $(\pm)1.96\sqrt{N}$.

Tables 32 and 33 has the result of sample autocorrelations of $X_t = ln(P_t/P_{t-1})$, $|X_t|$, and X_t^2 for lags $1,2,3,4,5,10,20,40,70,100$. Also autocorrelogram of the datasets are plotted in Figure 25 shown the ACF as a function of lags of X_t (line without marker) and X_t^2 (line with marker) for DAX and NIKKEI. The dotted lines show the 95% confidence interval. It is clear that there is no strong evidence to suggest that any of the log-return series exhibit autocorrelation, suggesting that X_t may have come from an iid process.

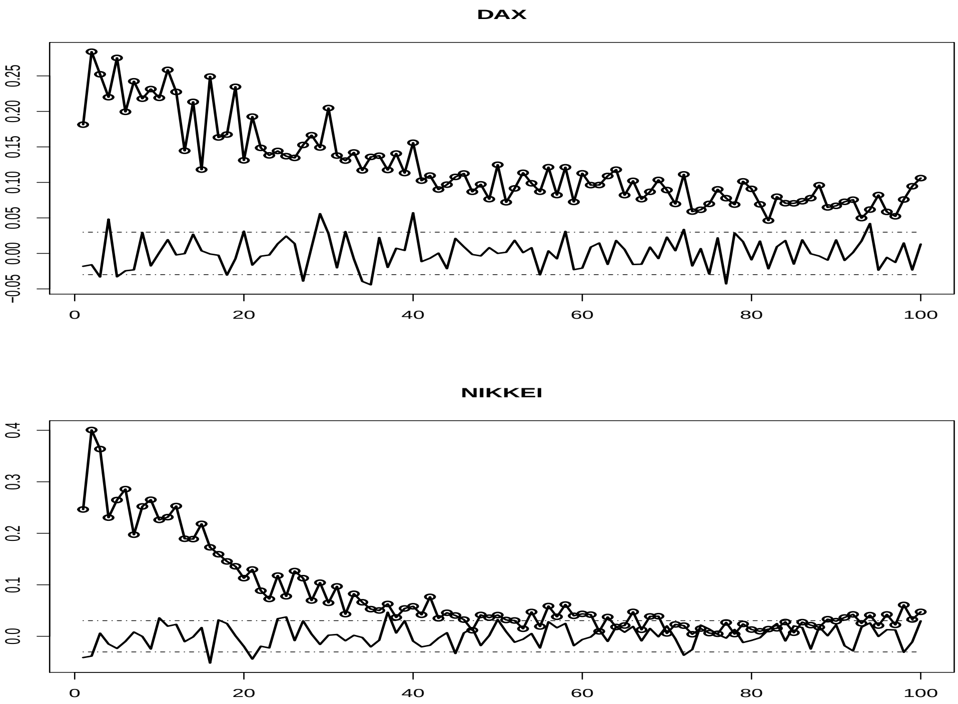

FIGURE 25: ACF plot of log returns and squared returns of DAX and NIKKEI.

According to Fama [69], Taylor [191], Hamao et.al. [**?**] there exists a relatively small positive autocorrelation in the first order in most stock index series. This finding is not supported neither with DAX nor NIKKEI return data's first and second lag autocorrelation. The negative autocorrelation supports the 'mean-reversion' thesis suggesting that the two return series not realizations of iid process.

Recall that if a process is an iid process then any transformation of that also an iid process. From Figure 25 we can see that X_t^2 ACF values are all positive, mostly outside the confidence region, particularly at lags 40 or less, and this is more true for DAX then NIKKEI. This strongly suggests that there is some higher order correlation and that the return processes not an iid.

Squared log returns dataset $(X_t^2)^d$ is further examined with $d = 0.125, 0.25, 0.50,$ $0.75, 1, 1.25, 1.5, 1.75, 2, 3$ at autocorrelations lags of $1, 2, 3, 4, 5, 10, 20, 40, 70, 100$. From Tables 34 and 35, and Figures 27 and 28 supports the conclusion above. The results supports the assertion of presence of persistence in stock index returns. This

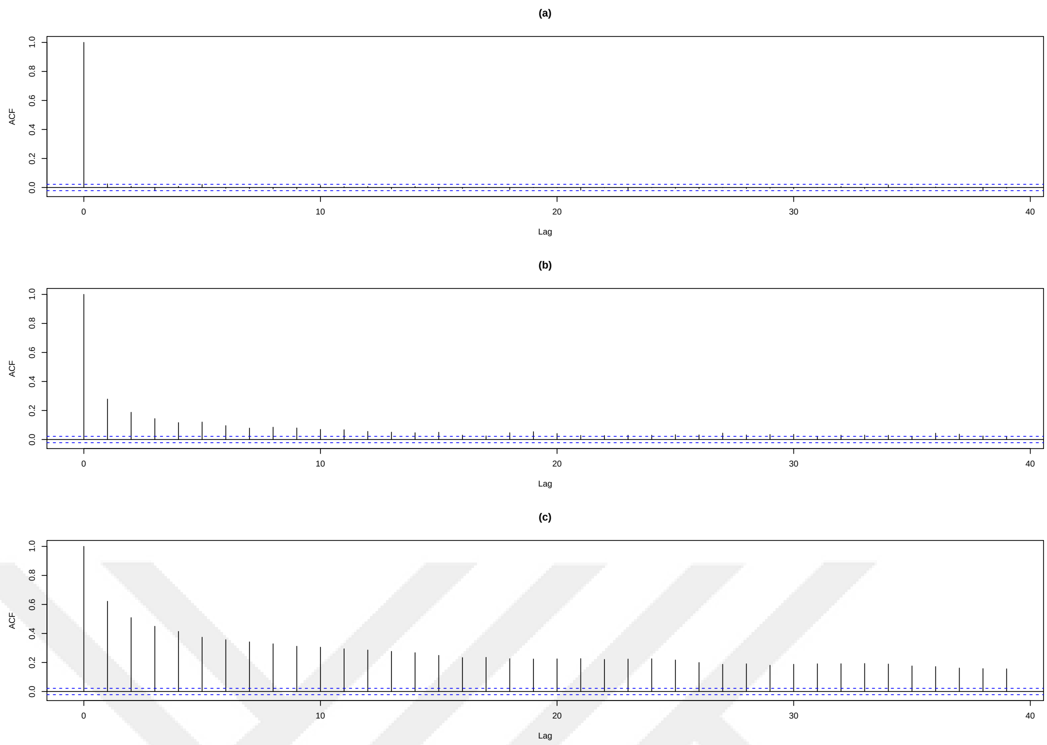

FIGURE 26: ACF plot of (a) White Noise; (b) Monofractal (c) Multifractal.

is indicated by the pronounced positive serial correlations observed to relatively long lags in the power transformations of squared log-returns of both indices as shown in Tables 34 and 35,

The property of long memory and multifractal process may be also graphically spotted. As a reference, Figure 26 shows theoretical ACF plot of different fractal time series. It is clear that white noise and monofractal data have an exponentially declining ACF plot whereas power law type behavior would be expected when series exhibit long memory as depicted in the multifractal ACF plot in Figure 26(c). It can be graphically seen from the Figures 27 and 28, squared return of DAX and NIKKEI are look alike multifractal time series: more like a power law profile than exponential decay, and ACFs almost always significant and positive up to 40 lags.

Another commonly used memory test to check for long-range dependence is the R/S statistic, which provides an estimate for the Hurst coefficient H. The strength of persistency

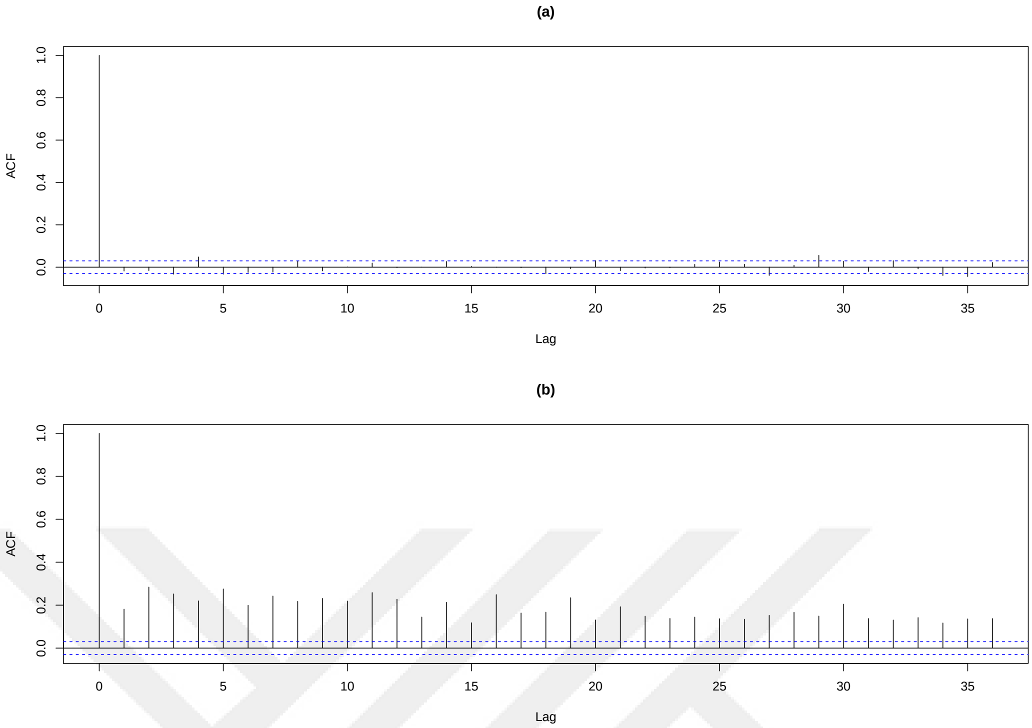

FIGURE 27: ACF plot of DAX; (a) Log return (b) Squared Log Return.

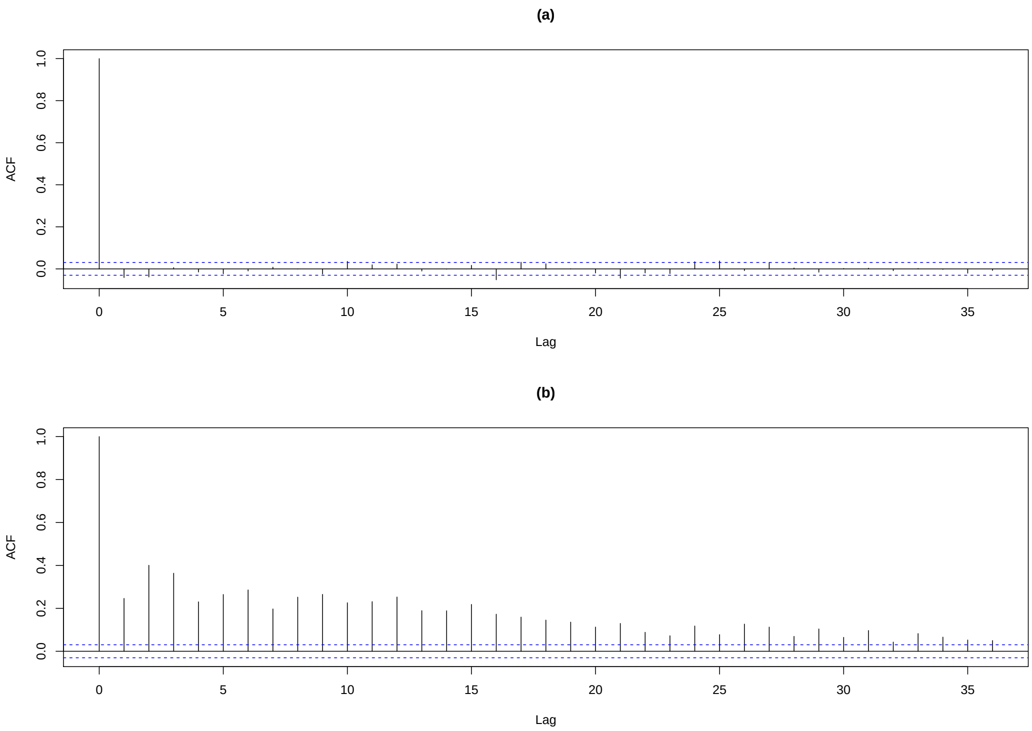

FIGURE 28: ACF plot of NIKKEI; (a) Log return (b) Squared Log Return.

TABLE 29: Long Memory Tests-DAX

Data	Modified R/S	R/S	GPH
Log Return	1.5486	1.456	0.8115 (d=0.076)
Squared Return	2.743**	4.8461**	4.4388**(d=0.4154)

** significant at 1% level, * significant at 5% level

TABLE 30: Long Memory Tests-NIKKEI

Data	Modified R/S	R/S	GPH
Log Return	1.1413	1.0654	0.4863 (d=0.046)
Squared Return	2.1945**	4.1799**	2.6581**(d=0.2513)

** significant at 1% level, * significant at 5% level

can also be studied by the differencing parameter, d which is related to the scaling H by $d = H - 1/2$.

From Table 29 and Table 30, the R/S and modified R/S of squared returns are significant at one percent suggesting persistence in the data for both DAX and NIKKEI. The same could not be said for log returns, supporting our findings from the ACF analysis. The squared-return d values are 0.4154 and 0.2513 for DAX and NIKKEI, respectively. This indicates that the series is stationary (though DAX value closer to the nonstationary range) and exhibit long memory. Moreover, based on the GPH test statistic for squared-returns, H_0 of no long memory can be rejected at 1% level. The opposite is true for long return data as GPH test statistic is not significant for the two financial series.

Another method for estimating H is the periodogram method. The slope of the log-log scatter plot of periodogram as a function of frequency gives an estimate for $1 - 2H$, where H is the Hurst coefficient. Figure 29 and 30 depict the results for DAX and NIKKEI indexes, respectively. Recall that when $H = 0.5$ the data is said to be uncorrelated, whereas for $H > 0.5$ data said to exhibit persistence. From the periodogram analysis we have values about 0.6 and 0.7 for DAX and NIKKEI squared returns supporting our earlier results regarding presence of long memory.

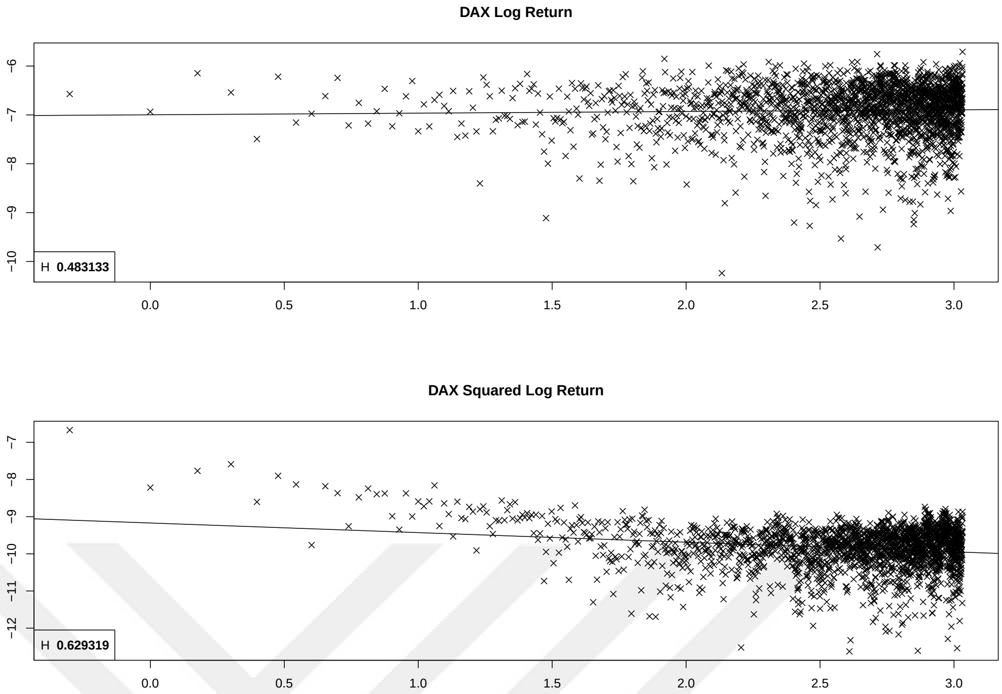

FIGURE 29: Log-log Periodogram-DAX; Log return H:0.4831 and Squared Log Return H:0.6293.

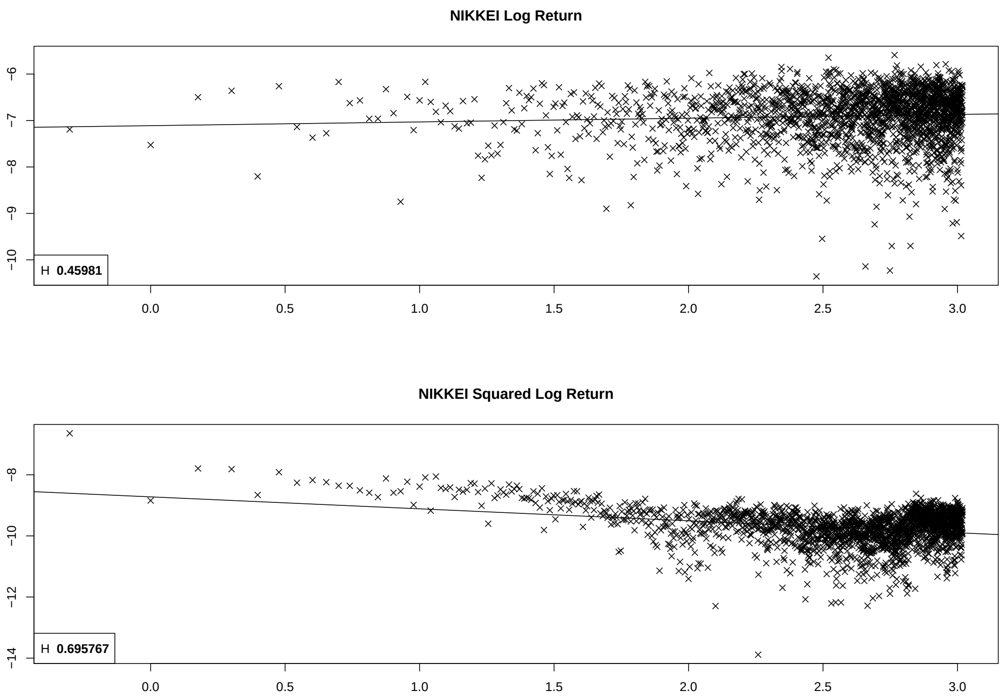

FIGURE 30: Log-log Periodogram-NIKKEI; Log return H:0.4598 and Squared Log Return H:0.6957.

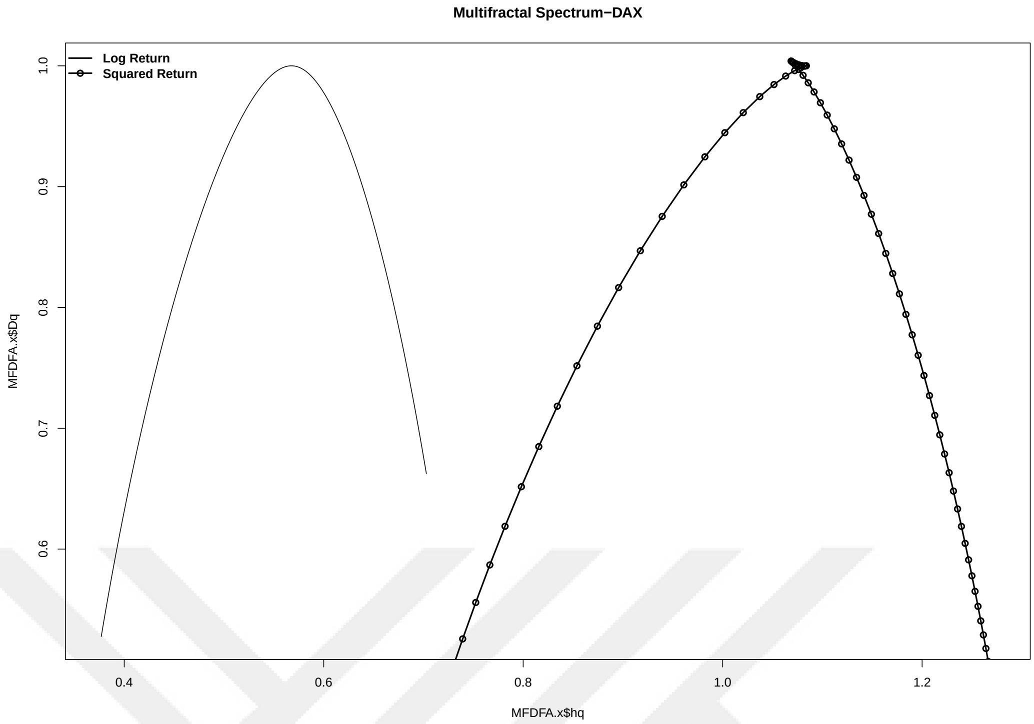

FIGURE 31: Multifractal Spectrum-DAX; Log return and Squared Log Return.

The multifractal (MF) spectrum (Stanley and Nature [185]) reflects the n-point correlations (Barabasi and Vicksek [14]) which gives information regarding the temporal structure of variations in the price. Multifractal spectrum is considered to be a useful tool to detect a multifractal effect in data. The wider the spectrum the stronger multifractal effect there is. Figures 31 and 32 show that squared return of the data has multifractal properties.

4.3.1. Wavelet Analysis

The aim behind applying wavelet coherence is to study correlations in terms of frequency and time interval (scale). Typically, this is visualized with a 3D contour plot and the heatmap indicates the dependece level with the color spectrum ranging from maximum dependency indicated by red and minimum dependecy by blue. Frequency parameter is typically shown on the horizontal and time on the vertical axis. 5% level of significance is used and this is depicted as thick contour lines. The edge effects, where there is insufficient data count, are shown (drawn) with a line in black. The direction arrows

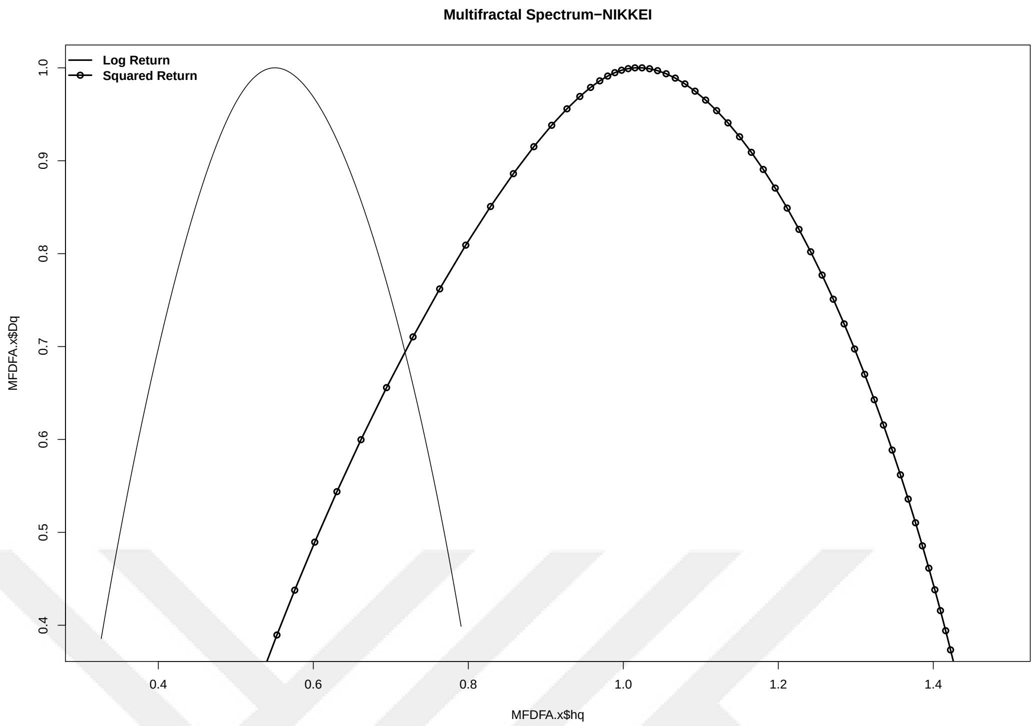

FIGURE 32: Multifractal Spectrum-NIKKEI; Log return and Squared Log Return.

indicate the alignment of phases: when there is positive correlation between two series, this is indicated by arrow pointing to right; out of phase indicated by arrow pointed to left. Downward pointing arrow is indication of the first series leading the second. The wavelet transforms for both indexes are shown in Figure 33.

The cross-wavelet transform contour plot highlights time-frequency domains of high common power. In Figure 34. it is clear to note that the high power area is in the years 2002-2003, and 2006-2007. During those years, the relationship gathers around at the frequencies 64 and 256. The thick contours showing 5% sig. level is much more apparent during those years.

By looking at the Figure 35, the coherence contour plot highlights co-varying (and possibly of high power but not a requirement) time-frequency domains. It is obvious to note that there is a strong coherence for the frequencies over 64 period. For the years 2000 and 2004, the arrows are pointing right which explains the positive correlation

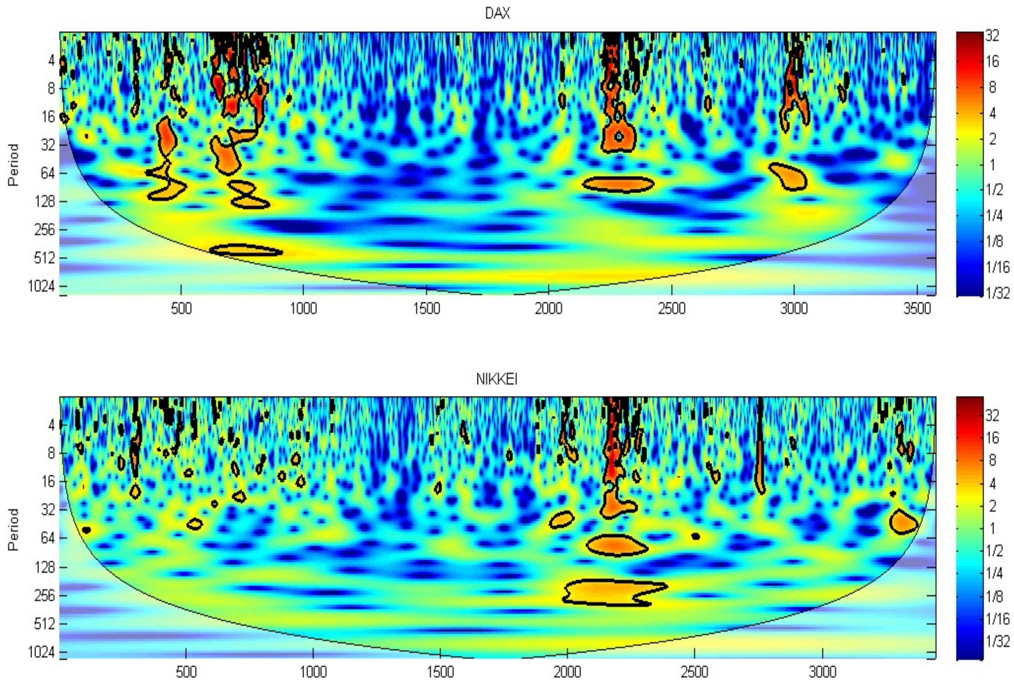

FIGURE 33: Continuous Wavelet Transform of DAX and NIKKEI Log return.

of the data sets. Beyond 2005 the arrows heads to the left which shows that time series are not in phase.

4.4. Hurst Analysis and Forecast

In the above sections we looked at the long-range dependency nature of DAX and NIKKEI and signs of any cross-correlations between the two indexes. After specifying the most coherent period between DAX and NIKKEI stock exchange markets, the Hurst values will be looked into details by applying Continuous Wavelet and Inverse Continuous Wavelet Transformations. A suitable and a popular model for linear forecasting is a class of vector ARMA or VARMA processes. VARMA is an application of ARMA process to multivariate time series [126]. Forecasting the future time step of a given series is achieved by using the past observations and any other time series that are

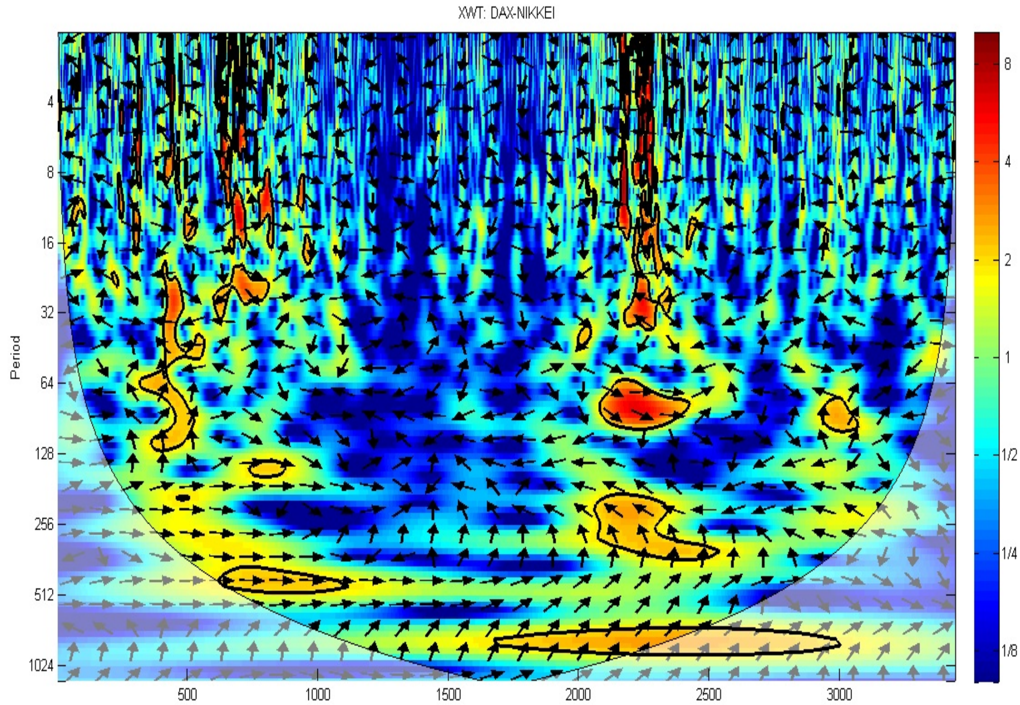

FIGURE 34: Cross Wavelet Transform of DAX and NIKKEI Log return.

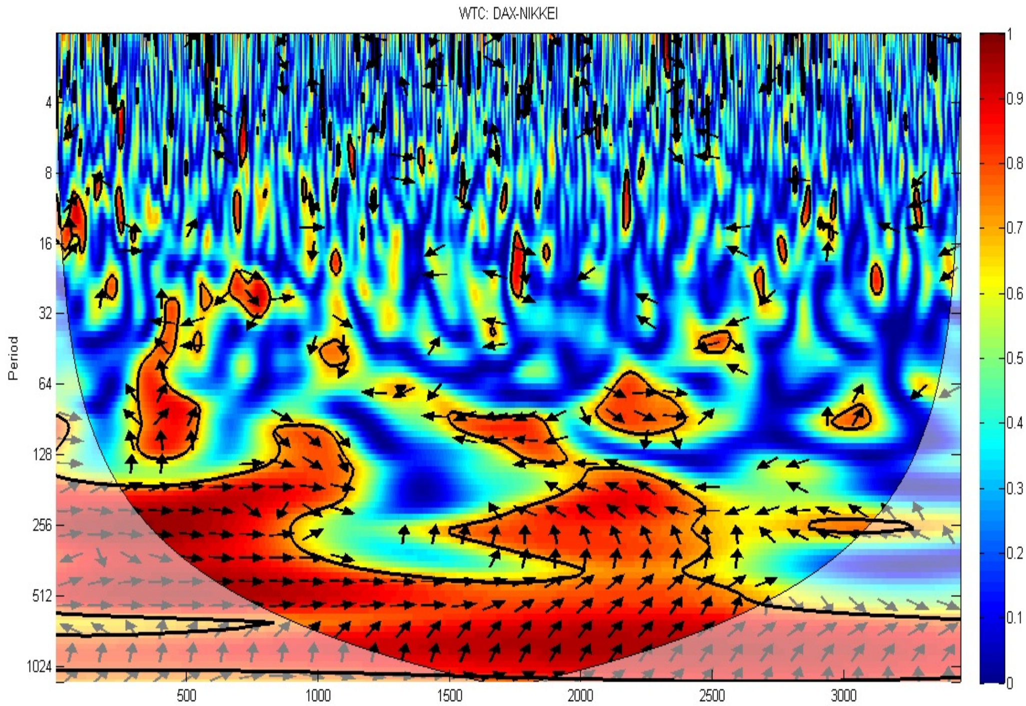

FIGURE 35: Wavelet Coherence of DAX and NIKKEI Log return.

TABLE 31: VARMA(1,1) Coefficients

Estimate	Std. Error	t-value	Pr($>|t|$)
0.960002	0.007130	134.645	$<$ 2e-16 ***
0.032303	0.007155	4.515	6.34e-06 ***
0.044718	0.009292	4.813	1.49e-06 ***
0.939859	0.009347	100.554	$<$ 2e-16 ***
0.174175	0.026905	6.474	9.57e-11 ***
0.003863	0.020567	0.188	0.851
-0.033595	0.031388	-1.070	0.284
0.223215	0.027078	8.244	2.22e-16 ***
***significant	at 0.001 level		
** significant	at 0.01 level		
aic=19.87169	bic=19.89927		

correlated. VARMA (1,1) model is applied to the Hurst values and forecast analysis carried out for the Hurst values of DAX and NIKKEI.

Figures 36 and 37 are the plots of DAX and NIKKEI most coherent period's Hurst values' (blue) and Inverse CWT (brown).

Figures 38 and 39 show the forecast plot against the real 30 days data and can be seen clearly that the forecast data of VARMA(1,1) mimics the trend of real data.

TABLE 32: Estimated Autocorrelation of DAX

	1	2	3	4	5	10	20	40	70	100
Log(X)	-0.01812	-0.01597	-0.03315	0.04829	-0.03292	0.00113	0.03128	0.05722	0.02289	0.01311
X^2	0.18135	0.28402	0.25226	0.21998	0.27534	0.21906	0.13106	0.15583	0.08922	0.10605

TABLE 33: Estimated Autocorrelation of NIKKEI

	1	2	3	4	5	10	20	40	70	100
Log(X)	-0.04131	-0.03846	0.00604	-0.01499	-0.02358	0.03575	-0.01962	-0.00940	0.02021	0.02979
X^2	0.24628	0.40059	0.36362	0.23039	0.26459	0.22603	0.11288	0.05828	0.00559	0.04749

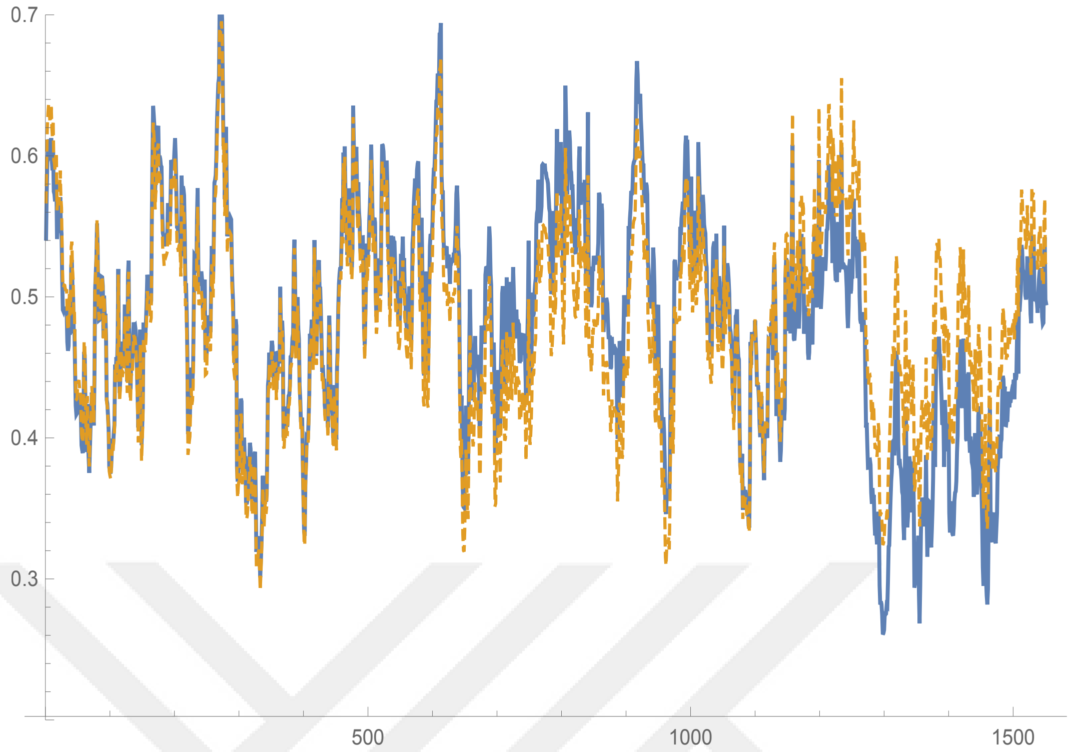

FIGURE 36: DAX's Most Coherent Period's Hurst Values CWT(blue) and ICWT (brown).

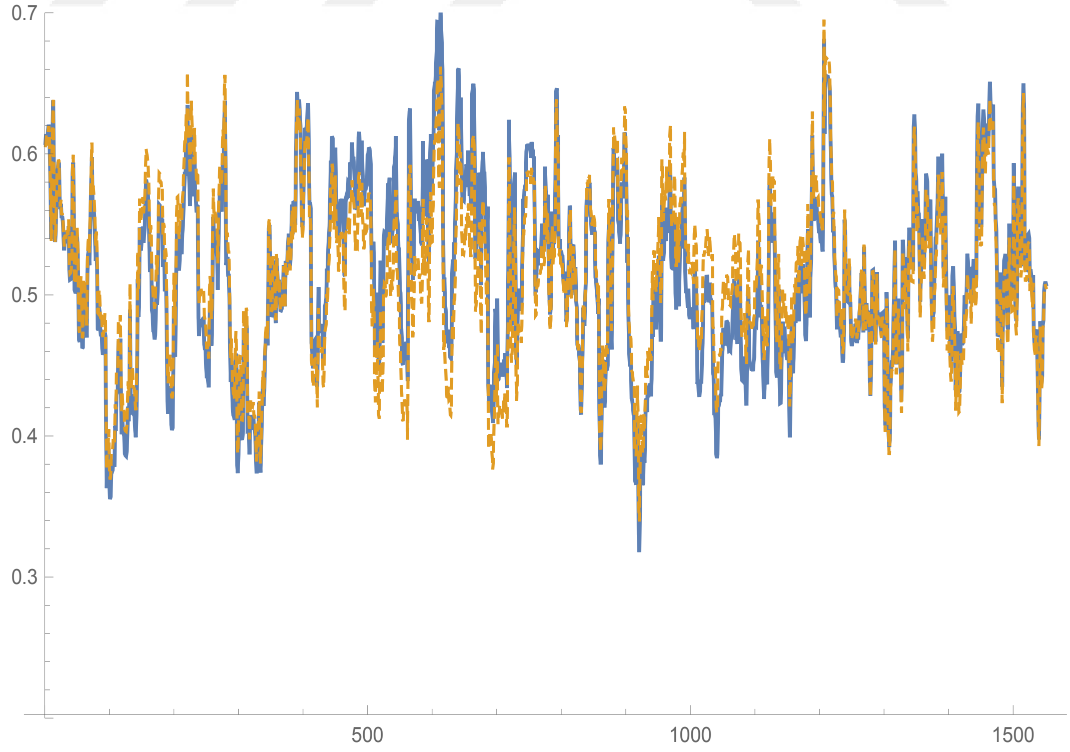

FIGURE 37: NIKKEI's Most Coherent Period's Hurst Values CWT(blue) and ICWT (brown).

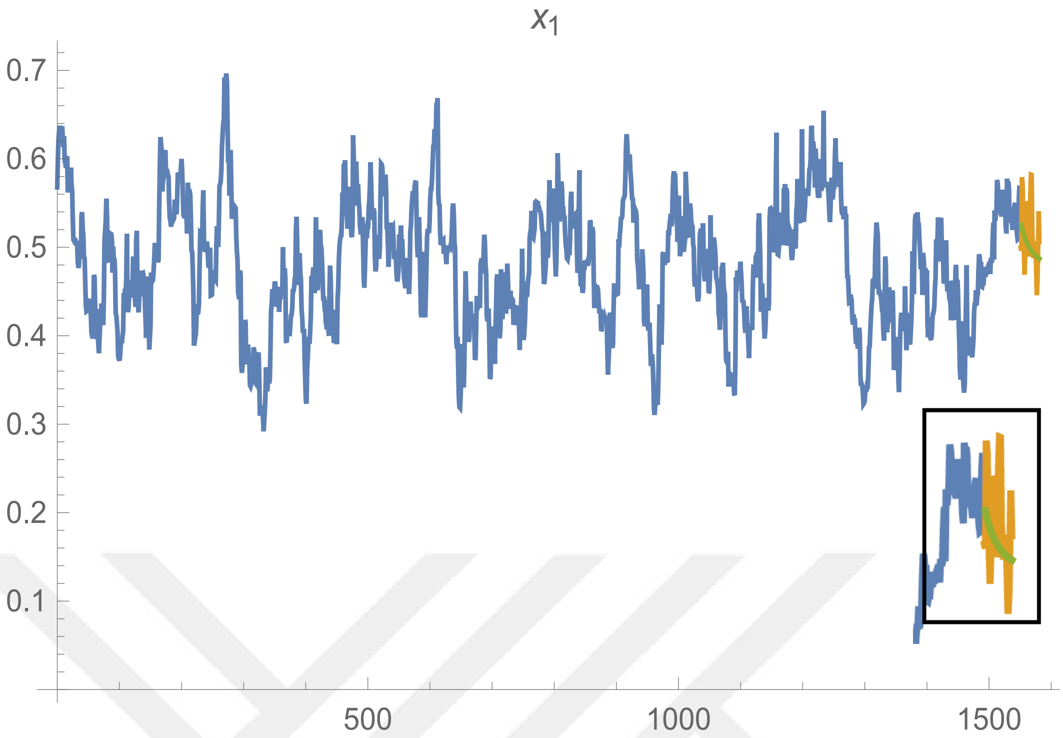

FIGURE 38: DAX's Hurst Values CWT(blue), Real 30days data (brown) and Forecast Values (green).

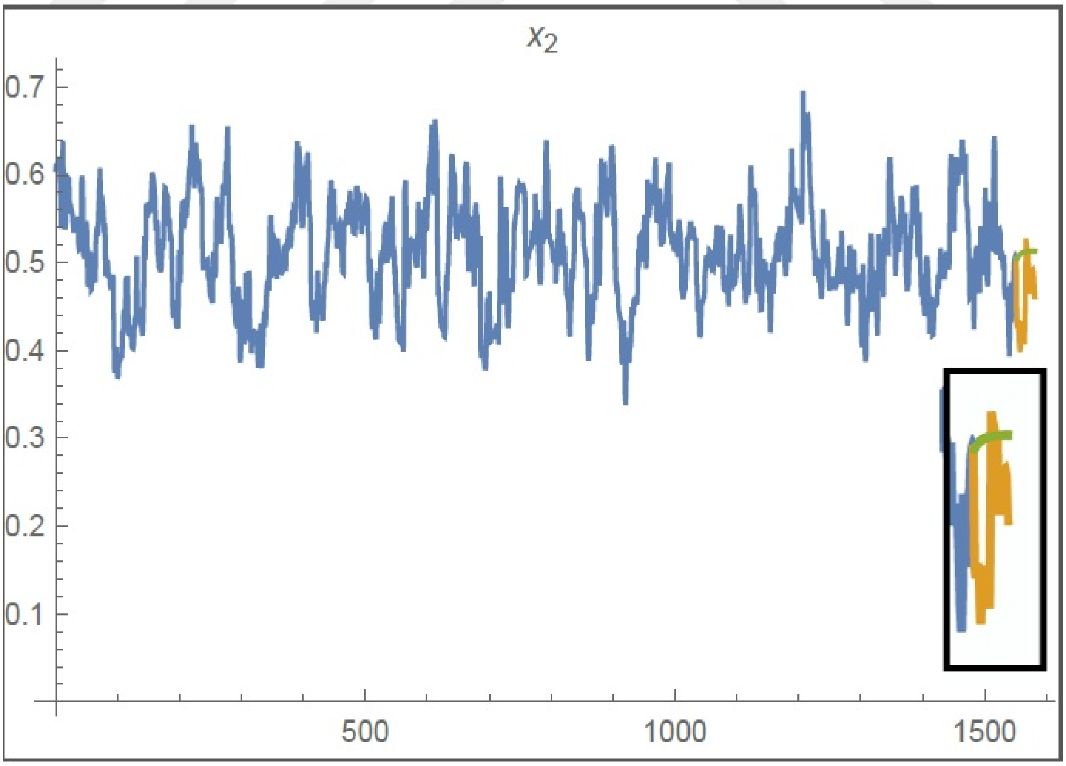

FIGURE 39: NIKKEI's Hurst Values CWT(blue), Real 30days data (brown) and Forecast Values (green)

TABLE 34: Power transformation of autocorrelation lags for DAX

	1	2	3	4	5	10	20	40	70	100
0.125	0.12184	0.20466	0.20115	0.18238	0.21311	0.18365	0.16254	0.11639	0.09461	0.10454
0.25	0.15737	0.25270	0.24376	0.22947	0.25629	0.21851	0.19553	0.14227	0.11881	0.12537
0.5	0.19847	0.30496	0.28469	0.27186	0.29937	0.24987	0.21424	0.16793	0.13647	0.13949
0.75	0.20183	0.30990	0.28261	0.26228	0.29983	0.24460	0.18362	0.16881	0.12064	0.12937
1	0.18135	0.28402	0.25226	0.21998	0.27534	0.21906	0.13106	0.15583	0.08922	0.10605
1.25	0.15288	0.24298	0.20898	0.16658	0.24033	0.18611	0.08130	0.13839	0.05794	0.07902
1.5	0.12574	0.19744	0.16392	0.11710	0.20297	0.15260	0.04521	0.12075	0.03400	0.05436
1.75	0.10304	0.15431	0.12357	0.07813	0.16747	0.12185	0.02297	0.10414	0.01827	0.03495
2	0.08484	0.11743	0.09067	0.05034	0.13603	0.09548	0.01066	0.08893	0.00894	0.02122
3	0.04052	0.03652	0.02370	0.00734	0.05662	0.03359	-0.00074	0.04455	-0.00077	0.00144

TABLE 35: Power transformation of autocorrelation lags for NIKKEI

	1	2	3	4	5	10	20	40	70	100
0.125	0.06756	0.13127	0.14497	0.12245	0.13630	0.10834	0.09449	0.06217	0.02685	0.00818
0.25	0.10308	0.17110	0.17934	0.14948	0.17566	0.13499	0.12123	0.07128	0.03181	0.02323
0.5	0.16996	0.25179	0.25153	0.19891	0.23887	0.18118	0.14801	0.07941	0.03267	0.04737
0.75	0.22133	0.33479	0.32020	0.23191	0.27168	0.21431	0.14221	0.07400	0.02004	0.05495
1	0.24628	0.40059	0.36362	0.23039	0.26459	0.22603	0.11288	0.05828	0.00559	0.04749
1.25	0.24737	0.43649	0.37404	0.19973	0.23151	0.21576	0.07746	0.04031	-0.00231	0.03447
1.5	0.23572	0.44959	0.36159	0.15951	0.19334	0.19233	0.04847	0.02570	-0.00449	0.02293
1.75	0.21950	0.45139	0.33795	0.12242	0.16025	0.16477	0.02874	0.01559	-0.00421	0.01466
2	0.20234	0.44876	0.30994	0.09232	0.13379	0.13818	0.01645	0.00911	-0.00335	0.00915
3	0.13951	0.42659	0.19961	0.02834	0.06917	0.06366	0.00105	0.00041	-0.00128	0.00085

4.5. Conclusion

In this chapter, we have shown various data properties of two major stock exchange markets, DAX and NIKKEI time series and the relationship between them. The volatility of both stock indices has long memory which were proved by autocorrelation analysis and tests such as modified rescaled test and GPH test. There were signs of multifractal process in the square returns from the ACF plots and Hurst exponent values. So the data sets further analysed for multifractal process. And as the initial findings suggested, the volatility of the time series were multifractal for the studied period. By looking at the ACF plots decay slope, the DAX mimics stronger multifractality than NIKKEI which is also supported by GPH test significance values. As for our final analysis for this study, wavelet coherence investigated between two stock indices. Even if there were not strong coherency at the shorter frequency level, there were at some level of correlation between frequency 32 and 64, and they are also multifractal. Stronger coherency happens to be at frequency between 256 and 512.

Our main purpose was to look for any existent correlation between chosen stock markets. We have proven that even if there did not seem to have a strong relation between two of the markets, there are some level of correlation and coherency between them.

We have to test our model by forecasting 30 days period and the result of the VARMA(1,1) shows consistency in the model as the trend of the forecasted data follows the trend of the real data as seen in Figures 38 and 39.

5. Concluding Remarks

This thesis is comprised of three complementary studies on financial time series analysis of BIST100 and some of the major stock indexes across the globe: DAX, FTSE100, NIKKEI225 and S&P500. Our studies mainly focused on modeling the volatility of stock market returns and attempting to capture the empirically observed structures in the time series.

The normality tests we performed including the Jarque-Bera test indicated that the return distribution for all five indexes do not come from a normal distribution. The distributions are leptokurtic and skewed to the left, except for DAX, for the time series studied. We found evidence of ARCH effects on time series of all the indexes i.e. squared residuals are autocorrelated, supporting the adoption of GARCH framework as a model of volatility.

In our first study we used both discrete-time and continuous time GARCH driven by Meixner process for BIST100 and showed that COGARCH(1,1) to be a good fit for the return series. We have also shown presence of long range dependency in the squared and absolute returns of all the stock market indexes in our study, and estimated the predictive power of fractionally integrated GARCH processes for the indexes. In our third study we performed wavelet analysis on DAX and NIKKEI series to study the cross-correlations between the two stock markets. Although we did not find strong relation between the two, we observed some level of correlation and coherency. We also introduced an application of VARMA(1,1) to forecast Hurst values and the in-sample forecast appeared to agree with the trend of the realized data.

The trading data we have used covers the period between 2000 and 2013, which is a reasonably long time frame to study the structures of the stock market returns as it covers at least two stock market crises: bursting of the dot-com bubble and the global credit-crunch of 2007-08. Different results than what we reported here may be obtained by either utilizing different volatility models or using different data sets. Therefore, for future work, the analysis we carried out here may be performed on a longer period, extending time series to the end of 2020, which would include the market stresses experienced during the initial outbreak of Covid-19. Also, although we performed our analyses on daily stock market returns, as this typically pertains to investors and institutional portfolio managers, COGARCH framework on modeling intra-day returns can be performed as an extension to our study as this may be of interest to other market participants. Lastly, another future work can be an extension of our wavelet analysis to other stock market indices to investigate the cross-correlations between different markets or, potentially, between different asset classes, which may be a particular interest to portfolio managers in terms of studying risk diversification or concentration.

References

[1] Albrecher, H., & Predota, M. (2004). On Asian option pricing for NIG Lévy processes. Journal of computational and applied mathematics, 172(1), 153-168.

[2] Applebaum, D. (2004). Lévy Processes and Stochastic Calculus, Cambridge Studies in Advanced Mathematics 93, Cambridge University Press Cambridge.

[3] Andrews, D. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, Econometrica, Vol. 59

[4] Arnedo, A., Bacry, E., & Muzy, J. F. (1995). The Thermodynamics of Fractals Revisited with Wavelets, Physica A 213-232.

[5] Assaf, A. (2004). A FIGARCH modeling of the emerging equity market of Egypt. International Journal of Applied Business and Economic Research, 2(1), 67-80.

[6] Assaf, A., & Cavalcante, J. (2005). Long range dependence in the returns and volatility of the Brazilian stock market. European Review of Economics and Finance, 4(2), 5-20.

[7] Assaf, A. (2006). Persistence and long-range dependence in the emerging stock market of Kuwait. Middle East Business and Economic Review, 18(1), 1-17.

[8] Ausloos, M., & Ivanova, K. (2002). Multifractal nature of stock exchange prices, North-Holland, Comput. Phys. Commun. 147, 582-585.

[9] Aydogan, K., & Booth, G. G. (1998). Are there long cycles in common stock returns?, Southern Economic Journal 55(1), 141-149.

[10] Bachelier, L. (1900). Théorie de la Speculation (Thesis), Annales Scientifiques de l'École Normale Superieure III 17, 21-86. English Translation;Cootner (ed.), (1964) Random Character of Stock Market Prices, Massachusetts Institute of Technology pp 17-78

[11] Bacry, E., Delour, J., & Muzy, J. F. (2001). Multifractal random walk Phys. Rev. E 64, 026103, American Physical Society.

[12] Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractioanlly Integrated Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, 74, 3-30.

[13] Baillie, R. T., (1996). Long memory processes and fractional integration in econometrics, journal of Econometrics

[14] Barabasi, A. L., & Vicsek, T. (1991). Multifractality of self-affine fractals Phys. Rev. A 44, 2730-2733, American Physical Society.

[15] Barkoulas, J. T., Baum, C. F., & Travlos, N. (2000). Long memory in the Greek stock market. Applied Financial Economics, 10(2), 177-184.

[16] Barndorff-Nielsen, 0. E. (1995). Normal Inverse Gaussian Processes and the Modelling of Stock Returns, Research Report 300, Department of Theoretical Statistics, Institute of Mathematics, University of Aarhus.

[17] Barndorff-Nielsen, O. E., & Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics, (with discussion) Journal of the Royal Statistical Society, Series B, 63, 167-241.

[18] Barndorff-Nielsen, O. E., Mikosch, T., & Resnick, S. I. (Eds.). (2012). Lévy processes: theory and applications. Springer Science & Business Media.

[19] Bellalah, M., Aloui, C., & Abaoub, E. (2005). Long-range dependence in daily volatility on Tunisian stock market. International Journal of Business, 10(3), 191-216.Banerjee, A., & Sahadeb, S. (2006). Long memory property of stock returns:evidence from India. Working Paper.589. Indian Institute of Management.

[20] Bera, A. K., & Higgins, M. L. (1993). ARCH Models: Properties, Estimation and Testing, Journal of Economic Surveys, Vol.7, No.4, 307-366.

[21] Bartlett, M. S. (1946). On the theoretical specification and sampling properties of auto-correlated time series. Supplement to the Journal of the Royal Statistical Society 8 27-41.

[22] Bershadskii, A. (2001). Multifractal diffusion in NASDAQ Journal of Physics A: Mathematical and General, Volume 34, Number 10.

[23] Bertoin, J. (1996). L'evy Processes, Cambridge University Press Cambridge.

[24] Bhansali, R., Kokoszka, P. S. (2003). In Processes with Long-Range Correlations, Lecture Notes in Physics vol. 621

[25] Biane, P., Pitman, J., & Yor, M. (2001). Probability laws related to the Jacobi theta and Riemann zeta functions, and Brownian excursions. Bull. Amer. Math. Soc. 38, 435465.

[26] Björk, T. (2009). Arbitrage theory in continuous time. Oxford University Press, third edition.

[27] Blaesild, P. (1995). Lecture given at workshop on stochastic processes and financial markets, Personal communication. Macmillan college publishing Co.

[28] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics 31, 307-327.

[29] Bollerslev, T., Chou, R. Y., & Kroner, K. F. (1992). Journal Journal of Econometrics Elsevier 52,1-2 pp5-59.

[30] Bollerslev, T., Engle, R. F., & Nelson, D. B. (1994). GARCH Models, in: Engle, R.F., and McFadden, D.L. (eds.) Handbook of Econometrics, Vol. 4, Elsevier, Amsterdam, 2961-3038.

[31] Bollerslev, T., & Mikkelsen, H. O. (1996). Modeling and pricing long memory in stock market volatility, Journal of Econometrics 73, 1551-184.

[32] Bouchaud, J. P., & Sornette, D. (1994). The Black-Scholes option pricing problem in mathematical finance: generalization and extensions for a large class of stochastic processes. Journal de Physique I, 4(6), 863-881.

[33] Boveroux, N. V. P., Minguet, A., & Ausloos, M. (1998). The Crash of October 1997 Seen as a Phase Transition Amplitude and Universality, Physica A 255: 201-210.

[34] Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: forecasting and control, John Wiley & Sons.

[35] Bożejko, M., & Lytvynov, E. (2009). Meixner class of non-commutative generalized stochastic processes with freely independent values I. A characterization. arXiv:0812.0895v2, MR2540072

[36] Brock, W., Lakonishok, J., & LeBaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns, The journal of Finance Economics 41(3), 441-464.

[37] Brockwell, P. J. (2004). Representations of continuous-time ARMA processes. J. Appl. Probab. 41A 375-382.

[38] Brockwell, P. J., & Marquardt, T. (2005). Lévy driven and fractionally Integrated ARMA process with continuous tome parameter, Statist. Sinica 15(2), 477-494.

[39] Brockwell, P. J., Chadraa, E., & Lindner, A. (2006). Continuous-time GARCH Processes, Ann.Appld.Probab. Volume 16, Number 2, 790-826.

[40] Brockwell, P., Chadraa, E., & Lindner, A. (2006). Continuous-time GARCH processes. The Annals of Applied Probability, 16(2), 790-826.

[41] Calvet, L., & Fisher, A. (1999). Forecasting Multifractal Volatility. New York University, Leonard N. Stern School Finance Department Working Paper Series 99-017, New York University, Leonard N. Stern School of Business.

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

[43] Calvet, L. E., & Fisher, A. J. (2008). Multifractal volatility: theory, forecasting, and pricing. Academic Press.

[44] Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). The Econometrics of Financial Markets, Princeton, New Jersey: Princeton University Press.

[45] Canessa, E. (2000). Multifractality in time series ,J. Phys. A: Math. Gen. 33, 3637-3651.

[46] Caporale, G. M., & Gil-Alana, L. A. (2004). Long-range dependence in daily stock returns. Applied Financial Economics 14. 375-383.

[47] Caporin, M. (2003). Identification of long memory in GARCH models. Statistical Methods and Applications, 12, 133-1510.

[48] Chakravart I. M., Laha, R. G., & Roy, J. (1967). Handbook of Methods of Applied Statistics, Volume I, pp. 392-394, John Wiley.

[49] Chang, G. (2010). Time-frequency dynamics of resting-state brain connectivity measured with fMRI NeuroImage 50(1) 81-98.

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

[51] Chen, Z., Ivanov, P. C., Hu, K., & Stanley, H. E. (2002). Effect of nonstationarities on detrended fluctuation analysis. Physical review E, 65(4), 041107.

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

[53] Crato, N. (1994). Some international evidence regarding the stochastic behaviour of stock returns, Applied Financial Economics, 4, 33-9.

[54] Crato, N., & de Lima, P. J. F. (1994). Long range dependence in the conditional variance of stock returns, Economics Letters 45(3), 281-285.

[55] Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues, Quantitative Finance, Vol.1, No.2.

[56] Cunado, J., Gil-Alana, L., & Perez De Gracia, F. (2008). Fractional integration and structural breaks: Evidence from international monthly arrivals in the USA. Tourism Economics. 14. 13-24. 10.5367/000000008783554884.

[57] Dacorogna, M. M., Müller, U. A., Nagler, R. J., Olsen, R. B., & Pictet, O. V. (1993). A Geographical Model for the Daily and Weekly Seasonal Volatility in the Foreign Exchange Market, J. Int'l Money and Finance 12, 413-438.

[58] 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, 83-106.

[59] Ding, Z., & Granger, C. W. J. (1996). Modeling volatility persistence of speculative returns: A new approach, Journal of Econometrics 73(1), 185-215.

[60] Eberlein, E., & Keller, U. (1995). Hyperbolic distributions in finance. Bernoulli 1, 281299

[61] Einstein, A. (1905). Über die von molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. Ann. Physik, 17:549–560.

[62] Eisler, Z., & Kertész, J. (2004). Multifractal model of asset returns with leverage. Physica A: Statistical Mechanics and its Applications. 343. 603-622. 10.1016/j.physa.2004.05.061.

[63] Eitelman, P. S., Vitanza, J. T. (2008). A Non-Random Walk Revisited: Short- and Long-Term Memory in Asset Prices, Board of Governors of the Federal Reserve System, International Finance Discussion Papers, No. 956.

[64] Engle, R.F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation, Econometrica 50, 987-1008.

[65] Engle, R. F., and Bollerslev, T. (1986). Modelling the persistence of conditional variances. Econometric Reviews, 5, [1-500].

[66] Esche, F., & Schweizer, M. (2005). Minimal entropy preserves the Lévy property:how and why. Stoch. Proc. Appl, 115(2):299–327, MR2111196

[67] Fama, E. F. (1963). Mandelbrot and the stable Paretian hypothesis. The Journal of Business, 36(4), 420-429.

[68] Fama, E. F. (1965). The Behavior of Stock-Market Prices,The Journal of Business Vol. 38, No.1, pp.34-105

[69] Fama, E. F. (1976). Foundations of finance: Portfolio decision and security prices, New York: Basic Books Inc.

[70] Fisher, A., Calvet, L., & Mandelbrot, B. B. (1997). Multifractality of Deutschemark & US Dollar Exchange Rates, Cowles Foundation Discussion Paper, No 1165.

[71] Gallant, A. R., & Tauchen, G. (1996). Which moments to match?. Econometric theory, 12(4), 657-681.

[72] Geman, H., Madan, D. B., & Yor, M. (2001). Asset prices are Brownian motion: only in business time. In Quantitative Analysis In Financial Markets: Collected Papers of the New York University Mathematical Finance Seminar (Volume II) (pp. 103-146).

[73] Gencay R., Selcuk, F., & Whitcher, B. (2002). An Introduction to Wavelet and Other Filtering Methods in Finance and Economics, San Diego, Academic Press.

[74] Geweke, J., & Porter-Hudak, S. (1983). The estimation and application of long memory time series models. Journal of time series analysis, 4(4), 221-238.

[75] Ghashghaie, S., Breymann, W., Peinke, J., Talkner, P., & Dodge, Y. (1996). Turbulent cascades in foreign exchange markets. Nature, 381(6585), 767-770.

[76] Gil-Alana, L. A. (2006). Fractional integration in daily stock market indexes. Review of Financial Economics 15, 28-48.

[77] Gneiting, T., & Schlather, M. (2001). Stochastic models wich separate fractral dimension and hurst effect, Technical report, Technical Report Series,NRCSE-TRS.

[78] Greene, M. T., & Fielitz, B. D. (1997). Long-term dependence in common stock returns,Journal of Financial Economics 4(3),339-349.

[79] Gouriéroux, C., Monfort, A., & Renault, E. (1993). Indirect Inference, Journal of Applied Econometrics 8, S85-S118.

[80] Granger, C. W. J., & Joyeux, R. (1980). An Introduction to Long-Memory TİmeSeries Models and Fractional Differencing, Journal of TİmeSeries Analysis, 1, 15-29.

[81] Grant, P. M. (2008). Option Pricing: A GARCH Model With Lévy Process Innovations, Dissertation at University of Johannesburg.

[82] Grau-Carles, P. (2005). Tests of long memory: A bootstrap approach, Computational Economics 25(1-2), 103-113.

[83] Grigelionis, B. (1999). Processes of Meixner type, Lithuaninan Mathematical Journal

[84] Grigelionis, B. (2000). Generalized z-Distributions and related Stochastic Processes. Matematikos Ir Informatikos Institutas Preprintas Nr.22, Vilnius.

[85] Grinsted, A., Moore, J. C., & Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. Non-linear Processes in Geophysics 11, 561-566.

[86] Halsey, T. C., Jensen, M. H., Kadanoff, L. P., Procaccia, I., & Shraiman, B. I. (1986). Fractal measures and their singularities: the characterization ofstrange sets. Phys. Rev. A 33, 1141–1151. bibitemHamao1990 Hamao, Y., Masulis, R. W., & Ng, V. (1990). Correlations in price changes and volatility across international stock markets. The review of financial studies, 3(2), 281-307.

[87] Haug, S. (2006). Exponential COGARCH and other continuous time models with applications to high frequency data, Ph.D. dissertation at Munich University of Technology.

[88] Henry, O. T. (2002). Long memory in stock returns: some international evidence. Applied Financial Economics 12, 725-712.

[89] Hentschel, H. G. E., & Procaccia, I. (1983). The infinite number of generalized dimensions of fractals and strange attractors, Elsevier, Physica D: Nonlinear Phenomena Volume 8, Issue 3, September 1983, Pages 435-444.

[90] Heston, S.L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, The Review of Financial Studies, Volume 6-2,327-343.

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

[92] Hudgins, L., Friehe, C. A., & Mayer, M. E. (1993). Wavelet transforms and atmopsheric turbulence. Physical Review Letters, 71(20), 3279.

[93] Hurst, H. E. (1951). Long-term storage capacity of reservoirs, Trans. Amer. Soc. Civil Eng. 116: 770-808.

[94] In, F., & Kim, S. (2003). Hedge Ratio and Correlation between the Stock and Futures Markets: Evidence from the Wavelet Analysis, Journal of Business

[95] Ito, K. (1957). Stochastic processes.

[96] Ivanova, K., & Ausloos, M. (1999). Low q-moment multifractal analysis of Gold price, Springer, Dow Jones Industrial Average and BGL-USD exchange rate Eur. Phys. J. B 8, 665-669.

[97] Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Economics letters, 6(3), 255-259.

[98] Kantelhardt, J. W., Koscielny-Bunde, E., Rego, H. H., Havlin, S., & Bunde, A. (2001). Detecting long-range correlations with detrended fluctuation analysis. Physica A: Statistical Mechanics and its Applications, 295(3-4), 441-454.

[99] Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., & Stanley, H. E. (2002). Multifractal detrended fluctuation analysis of

nonstationary time series. Physica A: Statistical Mechanics and its Applications, 316(1-4), 87-114.

[100] Kariya, T., Tsukuda, Y., & Maru, J. (1990). Testing the random walk hypothesis for Japanese stock price in S. Taylor's Model, Working paper, University of Chicago.

[101] Karlis, D. (2002). An EM type algorithm for maximum likelihood estimation of the normal–inverse Gaussian distribution. Statistics & probability letters, 57(1), 43-52.

[102] Kendall, M. S. (1953). "The analysis of economic time series - Part I: Prices," Journal of the Royal Statistical Society, Series A 66, 11-25; 1953.

[103] Kantelhardt, J. H. (2008). Fractal and Multifractal Time Series, arXiv:0804.0747v1 [physics.data-an].

[104] Khan, M. S., & Ross, K. Z. (1977). The functional form of the aggregate import demand equation. Journal of International Economics, 7(2), 149-160.

[105] Kilic, R. (2004). On the long memory properties of emerging capital markets: evidence from Istanbul stock exchange. Applied Financial Economics, 14(13), 915-922.

[106] Kim, S., & In, F. H. (2003). The Relationship Between Financial Variables and Real Economic Activity: Evidence From Spectral and Wavelet Analyses, Studies in Nonlinear Dynamics and Econometrics, 7, No. 4.

[107] Kirichenko, L., Radivilova, T., & Deineko, Z. (2011). Comparative analysis for estimating of the Hurst exponent for stationary and nonstationary time series. Information Technologies & Knowledge, 5(1), 371-388.

[108] Kitchen, C. (2009). Normal Inverse Gaussian Process With Applications in Mathematical Finance, Talk at University of Calgary

[109] Klüppelberg, C., Lindner, A., & Maller, R. (2004). A continuous time GARCH process driven by a Lévy process: stationarity and second order behaviour. J. Appl. Probab.41, (3), 601-622.

[110] Koekoek, R., & Swarttouw, R. F. (1998). The Askey-scheme of hyper-geometric orthogonal polynomials and its q-analogue. Report 98-17, Delft University of Technology.

[111] Kolmogorov, A.N. (1931). Uberdie analytischen Methoden in der Wahrschein lichkeits vech-nung,Mathematische Annalen.

[112] Kristoufek, L., & Vosvrda, M. (2013). Measuring capital market efficiency:Global and local correlations structure, Physica A: Statistical Mechanics and its Applications 392(1): 184-193.

[113] Kurkmaz, T., Cevic, E. I., & Özatac, N. (2009). Testing for long memory in ISE using ARFIMA-FIGARCH Model and Structural Break Test. International Research Journal of Finance and Economics 26, 196-192.

[114] Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. Journal of econometrics, 54(1-3), 159-178.

[115] Lau, K. M., & Weng, H., (1995). Climate Signall Detection Using Wavelet Transform: How to Make a Time Series Sing, American Meteorological Society Bulletin.

[116] Lee H. S. (2004). International transmission of stock market movements: a wavelet analysis, Applied Economics Letters, 11, 3/20, 197-201.

[117] Liang, Z. M., Hu, Y. M., & Wang, J. (2011). Advances in hydrological frequency analysis of non-stationary time series. Advances in Water Science, 22(6), 864-871.

[118] Lillestöl, J. (2000). Risk Analysis and the NIG Distribution,J.Risk, 4,pp.41-55.

[119] Lillo, F., & Farmer, J. D. (2004). The long memory of the efficient market. Studies in nonlinear dynamics & econometrics, 8(3).

[120] Lim, S. Y., Brooks, R. D., & Hinich, M. J. (2008). Nonlinear serial dependence and the weak-form efficiency of Asian emerging stock markets, Journal of International Financial Markets, Institiutions and Money 18(5), 527-544.

[121] Liu, Y., Cizeau, P., Meyer, M., Peng, C. K., & Stanley, H. E. (1997). Correlations in economic time series. Physica A: Statistical Mechanics and its Applications, 245(3-4), 437-440.

[122] Lo, A. W., (1989). Long-term memory in stock market prices, Technical report, National Bureau of Economic Research.

[123] Lo, A. W., (1991). Long-term memory in stock market prices, Econometrica, Vol.59, No.5.

[124] Lobato, I. N., & Savin, N. E. (1998). Real and spurious long memory porperties of stock market data, Journal of Business & Economic Statistics 16(3), 410-427.

[125] Lux, T. (1996). Long-term stochastic dependence in financial prices: evidence from the German stock market, Applied Economics Letters, 3, 701-6.

[126] Lütkepohl, H. (2002). Forecasting cointegrated VARMA processes in M.P. Clements & D. F. Hendry (eds), A Companion to Economic Forecasting, Blackwell, Oxford, pp.179-205.

[127] Madan, D. B., & Seneta, E. (1990). The VG model for share market returns. Journal Business, Vol.63, No.4, pp.511-524

[128] Madan, D. B., & Milne, F. (1991). Option pricing with vg martingale components 1. Mathematical finance, 1(4), 39-55.

[129] Madan, D., Carr, P., & Chang, E. (1998). The variance gamma process and option pricing, European Finance Review, 2 pp. 79-105.

[130] Malkiel, B. G., & Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. The journal of Finance, 25(2), 383-417.

[131] Maller, R. A., Müller, G., & Szimayer, A. (2008). GARCH modelling in continuous time for irregularly spaced time series data. Bernoulli, 14(2), 519-542.

[132] Mandelbrot, B. (1963). The Variation of Certain Speculative Prices, The University of Chicago Press, The Journal of Business, Vol. 36, No. 4 (Oct., 1963), pp. 394-419.

[133] Mandelbrot, B. van Ness, J. W. (1968). Fractional Brownian Motions, Fractional Noises and Applications, SAIM, Review 10, 422.

[134] Mandelbrot, B. (1972). Statistical methodology for non-periodic cycles: from the covariance to R/S analysis, Annals of Economic and Social Measurement 1,259-290.

[135] Mandelbrot, B. (1972). Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence, ”In Statistical models and turbulence,”, Springer, 333-351.

[136] Mandelbrot, B. (1975). Limit theorems on the self-normalized range for weakly and strongly dependent processes. Zeitschrift ftr. Wahrscheinlichkeitstheorie. und Verwandte Gebiete, 31, 271-285.

[137] Mandelbrot, B. (1977). Fractals: form, chance, and dimension (Free Press, New York).

[138] Mandelbrot, B. (1983). The fractal geometry of nature. New York : W.H. Freeman

[139] Mandelbrot, B. Fisher, A., & Calvet, L. (1997). A Multifractal Model of Asset Returns, mimeo: Cowles Foundation for Research in Economics.

[140] Mandelbrot, B. (1999). A Multifractal Walk down Wall Street, Scientific American, February, 50-53.

[141] Mantegna, R. N., & Stanley, H. E. (2000). An Introduction to Econophysics. Cambridge University Press.

[142] Markowitz, Harry M. (1952). Portfolio Selection, Journal of Finance, C.7 No.1 pp.77–91

[143] Matia, K., Amaral, L. A. N., Goodwin, S., & Stanley, H. E. (2002). Different scaling behaviors of commodity spot and future prices Phys. Rev. E 66, 045103, American Physical Society.

[144] McNees, S.K., (1976a). An evaluation of economic forecasts: Extension and update. New England. Economic Review, 30–44.

[145] Merton, R. (1973). The theory of rational option pricing, Bell Journal of Economics and Management Science 4, 141-183.

[146] Mills, T. C. (1993). Is there long-term memory in UK stock returns?, Applied Economic Research, 3(4), 303-306.

[147] Mills, T. C. (1997). Stylized Facts of the Temporal and Distributional Properties of Daily FTSE Returns, Applied Financial Economics, 7, 599-604.

[148] Nath, G. C. (2001). Long memory and Indian stock market - an emprical evidence, National Stock Exchange of India Ltd., Mumbai.

[149] Nawrocki, D. (1995). R/S analysis and long term dependence in stock market indices, managerial Finance, 21(7), 78-91.

[150] Nelson, D.B. (1990). ARCH models as diffusion approximations. J. Econometrics 45, 7-38.

[151] Nelson, D. B., & Cao, C. Q. (1992). Inequality constraints in the univariate GARCH model. Journal of Business & Economic Statistics, 10(2), 229-235.

[152] Osborne, M. F. M. (1959). "'Brownian motion in the stock market"', Operations Research, 7, 145 – 173.

[153] Osborne, M. F. M. (1964). "'Brownian motion in the stock market"', in The random character of stock market prices, P. H. Cootner (Editor). Pages 100 – 128. Cambridge, MA: MIT Press.

[154] Panas, E. (2001). Estimating fractal dimension using stable distributions and exploring long memory through ARFIMA models in the Athens Stock Exchange, Applied Financial Economics 11(4), 395-402.

[155] 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.

[156] Peters, E. E. (1996). Chaos and Order in the Capital Markets (2nd ed.), John Wiley & Sons, Inc.

[157] Peters, E. E. (1994). Fractal Market Analysis: Applying Chaos Theory to Investment and Economics, John Wiley & Sons, Inc.

[158] Pitman, J., & Yor, M. (2003). Infifinitely divisible laws associated with hyperbolic functions. Prébluciations du Laboratoire de Probabilitiés et Modéles Aléatoires 616, Universités de Paris 6 & Paris 7, Paris.

[159] Pochart, B., & Bouchaud, J. P. (2002). The skewed multifractal random walk with applications to option smiles arXiv:cond-mat/0204047v2.

[160] Prause, K. (1999). The generalized hyperbolic model: Estimation, financial derivatives and risk measures (Doctoral dissertation, Verlag nicht ermittelbar).

[161] Ramsey, J. B., & Lampart, C. (1998a). The Decomposition of Economic Relationship by Time Scale Using Wavelets: Money and Income, Macroeconomic Dynamics, 2, 49-71.

[162] Ramsey, J. B., & Lampart, C. (1998b). The Decomposition of Economic Relationship by Time Scale Using Wavelets: Expenditure and Income, Studies in Nonlinear Dynamics and Econometrics, 3, No. 4, 23-42.

[163] Ramsey, J. B. (2002). Wavelets in Economics and Finance: Past and Future, C.V. Starr Center for Applied Economics, No. 2002-02.

[164] Roberts, H. V. (1959). "'Stock market patterns and financial analysis: Methodological suggestions,"' Journal of Finance 14, 1-10.

[165] Rydberg, T. H. (1997). The normal inverse Gaussian Lévy process: simulation and approximation. Communications in statistics. Stochastic models, 13(4), 887-910.

[166] Sadique, S., & Silvapulle, P. (2001). Long.term memory in stock returns, internacional evidence. International Journal of Economics and Finance 2001, 6, 50-67.

[167] Said, S. E., & Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. Biometrika, 71(3), 599-607.

[168] Samuelson, P. (1965). Proof That Properly Anticipated Prices Fluctuate Randomly.

[169] Samuelson, P. A. (1965). Rational theory of warrant pricing, Industrial Management Review, 6(2):13–31.

[170] Sato, K. I. (1999). Lévy Processes and Infinitely Divisible Distributions. Cambridge University Press

[171] Sato, K. (2000). Lévy Processes and Infinitely Divisible Distributions. Cambridge Studies in Advanced Mathematics 68. Cambridge University Press, Cambridge.

[172] Scholes, M., & Black, F. (1973). The Pricing of options and corporate liabilities. Journal of Political Economy.

[173] Schoutens, W., & Teugels, J. L. (1998). Lévy processes, polynomials and martingales. Communications in statistics:Stochastic models,Vol.14, pp.335-349.

[174] Schoutens, W. (2000), Stochastic Processes and Orthogonal Polynomials, Lecture Notes in Statistics 146, Springer-Verlag, New York.

[175] Schoutens, W., & Teugels, J. L. (2001). Meixner processes in finance.

[176] Schoutens, W., (2002). The Meixner Processes: Theory and Applications in Finance, Eurandom

[177] Schoutens, W. (2003). Lévy Processes in Finance:Pricing Financial Derivatives, New York: John Wiley & Sons.

[178] Schmitt, F., Schertzer, D., & Lovejoy, S. (1999). Multifractal Analysis of Foreign Exchange Data, AppliedStochastic Models and Data Analysis, in press.

[179] Segnon, M., Lux, T. (2013). Multifractal Models in Finance: Their Origin, Properties and Applications, Kiel Working Paper, No. 1860.

[180] Serletis, A., & Rosenberg, A. A. (2009). Mean reversion in the US stock market, Choas, Solitons & Fractals 40(4), 2007-2015.

[181] Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrika, 52(3/4), 591-611.

[182] Smith, Jr. A .A. (1990). Three Essays on the Solution and Estimation of Dynamic Macroeconomic Models, Ph.D. dissertation (Duke University).

[183] Smith, Jr. A .A. (1993). Estimating Nonlinear Time-series Models Using Simulated Vector Autoregressions, Journal of Applied Econometrics 8, S63-S84.

[184] Snedecor, G. W., & Cochran, W. G. (1989). Statistical Methods, Eighth Edition, Iowa State University Press.

[185] Stanley, H. E., & Meakin, P. (1988). Multifractal phenomena in physics and chemistry Nature 335, 405-409.

[186] Stanley, H. E., Amaral, L. A. N., Canning, D., Gopikrishnan, P., Lee, Y., & Liu, Y. (1999). Econophysics: Can physicists contribute to the science of economics?, Physica A: Statistical Mechanics and its Applications Vol. 269(1) 156-169.

[187] Struzik, Z. R. (2001). Wavelet methods in (financial) time-series processing Physica A: Statistical Mechanics and its Applications 296 (1), 307-319.

[188] Sun, X., Chen, H., Wu, Z., & Yuan, Y. (2001). Multifractal analysis of Hang Seng index in Hong Kong stock market, Physica A: Statistical Mechanics and its Applications Volume 291, Issues 1–4, 1 March, 553-562.

[189] Taqqu, M., Teverosky, W., & Willinger, W. (1995). Estimators for long-range dependence: an empirical study. Fractals 3(4)

[190] Taquu, M. S., Teverovsky, V., in: Alder, R., Feldman, R., & Taquu, M. S. (Eds.) (1998). A practical guide to heavy tails: Statistical techniques and applications, Birkhauser, Boston. Pages 177-217.

[191] Taylor, S. (1986). Modelling financial time series, New York, John Wiley & Sons.

[192] Teverovsky, V., Taqqu, M. S., Willinger , W. (1999). A critical look at Lo's modified R/S statistic, Journal of Statistical Planning and Inference, vol. 80, issues 1-2.

[193] Tolvi, J. (2003). Long memory in a small stock market, Economic Bulletin 7(3), 1-13.

[194] Torrence, C., & Compo, G. P. (1998). A Practical Guide to Wavelet Analysis. Bull. Amer. Meteor. Soc., 79, 61-78.

[195] Torrence, C., & Webster, P. (1999). Interdecadal Changes in the ESNO-Monsoon System, J. Clim., 12, 2679–2690.

[196] Vandewalle, N., & Ausloos, M. (1998a). Sparseness and Roughness of Foreign Exchange Rates, International Journal of Modern Physics C, 9, 711-719.

[197] Vandewalle, N., & Ausloos, M. (1998b). Multi-affine Analysis of Typical Currency Exchange Rates, European Physical Journal B, 4, 257-261.

[198] Vassilicos, J. C., Demos, A., & Tata, F. (1993). No Evidence of Chaos But Some Evidence of Multifractals in the Foreign Exchange and the Stock Market, in: Crilly, A.J., R.A. Earnshaw and H. Jones, eds., Applications of Fractals and Chaos. Berlin: Springer.

[199] Von Seggern, D. H. (1993). CRC Standard Curves and Surfaces. CRC Press.

[200] Wang, F., Yamasaki, K., Havlin, S., & Stanley, H. E. (2006). Scaling and memory intraday volatility return intervals in stock markets. Physical Review 73, Issue 2.

[201] Wang, Y., Wu, C., & Pan. (2011). Multifractal detrending moving average analysis on the us dollar exchange rates, Physica A: Statistical Mechanics and its Applications 390(20): 3512-3523.

[202] Wiener, N. (1949). Extrapolation, Interpolation and Smoothing of Stationary Time Series.

[203] Wright, J. H. (2002). Log-periodogram estimation of long memory volatility dependencies with conditionally heavy tailed returns. Econometric Reviews, 21(4), 397-417.

[204] Yamasaki, K., Munchnik, L., Havlin, S., Bunde, A., & Stanley, H. E. (1995). Scaling and volatility in memory return intervals in financial markets- Proceedings of the National Academy of Sciences of the United States of America 102. 9424-9428.

[205] Zhuang, Y., Green, C. J., & Maggioni, P. (2000). The great rebound, the great crash, and persistence in British stock prices. Economic Research Paper, Loughborough University.

[206] Zivot, E., & Andrews, D. W. K. (1992). Further Evidence on the Great Crash,the Oil-Price Shock and the Unit-Root Hypothesis. Journal of Business and Economics Statistics , Vol. 10, pp. 251-270.

[207] Zivot, E. (2009). Financial Econometrics Lecture notes, University of Washington

[208] Zumbach, G. O., Dacorogna, M. M., Olsen, J. L., & Olsen, R. B. (2000). Measuring shock in financial markets. International Journal of Theoretical and Applied Finance, 3(03), 347-355.

Appendix A. List of Symbols

S_t	stock price at time t
μ	drift term
σ	stock price volatility
σ^2	variance rate
W_t	Weiner process
V	variance term
B^1	Brownian motion
B^2	Brownian motion
G_t	COGARCH process
L	quadratic variation process
y_t	return time series
ε_t	error term
u_t	white noise process
T	sample size
F	probability space
v	measure
$\phi(u)$	characteristic function
$\kappa(u)$	cumulant generating function
$\gamma(k)$	ACF function
$\tilde{Q}_n$	rescaled range statistic
s_n	maximum likelihood standard deviation estimator
X_t	log return
$\|X_t\|$	absolute return
X_t^2	squared return
Y_t	demeaned time series

Appendix B. Abbreviations

GBM	Geometric Brownian Motion
BIST	The stock market index for Borsa Istanbul
FTSE	The stock market index for London Stock Exchange
NIKKEI	The stock market index for Tokyo Stock Exchange
DAX	The stock market index for Frankfurt Stock Exchange
S&P	Standard & Poor's stock market index
ARIMA	Autoregressive Integrated Moving Average
ARCH	Autoregressive Conditional Heteroskedasticity
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
EGARCH	Exponential GARCH
COGARCH	Continuous Generalized Autoregressive Conditional Heteroskedasticity
FIGARCH	Fractionally Integrated GARCH
FIEGARCH	Fractionally Integrated Exponential GARCH
ARFIMA	Autoregressive Fractionally Integrated Moving Average
BIC	Bayesian Information Criterion
AIC	Akaike Information Criterion
PML	Psuedo Maximum Likelihood
MM	Method of Moments
IIM	Indirect Inference Method
CP	Compound Poisson
VG	Variance-Gamma
NIG	Normal Inverse Gaussian
IG	Inverse Gaussian
H	Hurst Exponent
DFA	Detrended Fluctuation Analysis
MF-DFA	Multifractal Detrended Fluctuation Analysis

CWT	Continuous Wavelet Transform
XWT	Cross Wavelet Transform
WTC	Wavelet Transform Coherence

Appendix C. Meixner Cumulant generating Function

$$
\begin{aligned}
E[e^{uX}] &= \int_{\Re} e^{ux}\frac{(2\cos(\beta/2))^{2\delta}}{2\alpha\pi\Gamma(2\delta)}\exp(\frac{\beta x}{\alpha})\left|\Gamma(\delta+\frac{ix}{\alpha})\right|^2 dx \\
&= \int_{\Re}\frac{(2\cos(\beta/2))^{2\delta}}{2\alpha\pi\Gamma(2\delta)}\exp(\frac{\beta x+\alpha ux}{\alpha})\left|\Gamma(\delta+\frac{ix}{\alpha})\right|^2 dx \\
&= (2\cos(\beta/2))^{2\delta}\int_{\Re}\frac{1}{2\alpha\pi\Gamma(2\delta)}\exp(\frac{(\beta+\alpha u)x}{\alpha})\left|\Gamma(\delta+\frac{ix}{\alpha})\right|^2 dx \\
&= \frac{(2\cos(\beta/2))^{2\delta}}{(2\cos((\beta+\alpha u)/2))^{2\delta}}\int_{\Re}\frac{(2\cos((\beta+\alpha u)/2))^{2\delta}}{2\alpha\pi\Gamma(2\delta)}\exp(\frac{(\beta+\alpha u)x}{\alpha})\left|\Gamma(\delta+\frac{ix}{\alpha})\right|^2 dx \\
&= \frac{(2\cos(\beta/2))^{2\delta}}{(2\cos((\beta+\alpha u)/2))^{2\delta}}\int_{\Re} f(x;\alpha,\beta+\alpha u,\delta)dx \\
&= (\frac{\cos(\beta/2)}{(\cos((\beta+\alpha u)/2))})^{2\delta} \qquad \text{(C.1)}
\end{aligned}
$$

Hence,

$$\kappa(u) = \log E[e^{uX}] = 2\delta[\log(\cos(\beta/2)) - \log(\cos((\beta+\alpha u)/2)))] \qquad \text{(C.2)}$$

Appendix D. Meixner Characteristic Function

Usuing the relationship, $\varphi(u) = exp\kappa(iu) = E[e^{iuX}]$ and the following relationship between complex trigonometric and hyperbolic functions

$$\cos(iz) = cosh(z) \qquad \text{(D.1)}$$

we have

$$\begin{aligned}\varphi(u) &= (\frac{\cos(\beta/2)}{\cos((\beta+\alpha iu)/2)})^{2\delta} \\ &= (\frac{\cos(\beta/2)}{\cos(i(\alpha u - i\beta)/2)})^{2\delta} \\ &= (\frac{\cos(\beta/2)}{\cosh((\alpha u - i\beta)/2)})^{2\delta} \qquad \text{(D.2)}\end{aligned}$$

(D.3)

Appendix E. Estimation of Method of Moments (MoM)

The model parameters estimated by equating ACF and sample moments to their theoretical paremeters.

The algorithm is as follows

(i) Calculate the moment estimators

$$\begin{aligned}\hat{m}_1 &:= \frac{1}{N}\sum_{n=1}^{N}(G_n^{(1)})^2 \\ \hat{m}_2 &:= \frac{1}{N}\sum_{n=1}^{N}(G_n^{(1)})^4 \\ \hat{\mu}_i &= \frac{1}{N}\sum_{n=1}^{N}(G_n^{(1)} - \hat{m}_1)^i, i = 2,3,4 \qquad \text{(E.1)}\end{aligned}$$

where $(G_n^{(1)})_{n\in N}$ denotes COGARCH process as a function of $\beta, \eta, \varphi > 0$, stationary increment process. And for lag $h \geq 2$ the observed autocovariances $\hat{p}(h) := (\hat{p}(1)....\hat{p}(h))$

$$\hat{p}(h) := \frac{1}{N}\sum_{n=1}^{N-s}((G_n^{(1)})^2 - \hat{\mu}_2)(G_{n+s}^{(1)})^2 - \hat{\mu}_2) \qquad \text{(E.2)}$$

For $s = 1, ...h$

(ii) For fixed $h \geq 2$, minimize

$$\sum_{i=1}^{h} (\hat{p}(i) - k_p e^{-pi})^2 \tag{E.3}$$

where k_p, $p > 0$ are constants. This gives estimators $\hat{k}_p$, $\hat{p}$ Which is given by

$$\begin{aligned} \hat{p} &= -\frac{\sum_{i=1}^{h} (\log \hat{p}(i) - (\frac{1}{h}\sum_{j=1}^{h} \log \hat{p}(i)))(i - \frac{h+1}{2})}{\sum_{i=1}^{h}(i - \frac{h+1}{2})^2} \\ \hat{k} &= \exp((\frac{1}{h}\sum_{j=1}^{h} \log \hat{p}(i)) + \frac{h+1}{2}\hat{p}) \end{aligned} \tag{E.4}$$

(iii) Calculate $\hat{k} = \hat{k}_p/(\hat{m}_2 - \hat{m}_1^2)$ and then the following estimators

$$\begin{aligned} \hat{\beta} &= \hat{p}\hat{m}_1 \\ \hat{\varphi} &= \hat{p}\sqrt{1 + \hat{M}_2 - \hat{p}} \\ \hat{\eta} &= \hat{p}\sqrt{1 + \hat{M}_2(1 - \tau_L^2)} + \hat{p}\tau_L^2 = \hat{p} + \hat{\varphi}(1 - \tau_L^2) \end{aligned} \tag{E.5}$$

where τ_L^2 is the variance of the Brownian component of the Lévy process, L and

$$\begin{aligned} \hat{M}_2 &:= \frac{2\hat{k}_p}{\hat{M}_1(e^{\hat{p}} - 1)(1 - e^{-\hat{p}})} \\ \hat{M}_1 &:= \hat{m}_2 - 3\hat{m}_1^2 - 6\frac{1 - \hat{p} - e^{-\hat{p}}}{(e^{\hat{p}} - 1)(1 - e^{-\hat{p}})}\hat{k} \end{aligned} \tag{E.6}$$

Alternatively, for step (ii), the autocovariance function $\hat{\gamma}(i)$ can also be used in the least square estimation. Hence the function would be

$$\sum_{i=1}^{h} (\hat{\gamma}(i) - k_p e^{-pi})^2 \tag{E.7}$$

According to Klüppelberg et. al. [109], the estimation on the autocorrelaiton function is more accurate as k_p is independent of β.

An estimate of the volatility process $(\sigma_t^2)_{t\geq 0}$ is given as

$$\hat{\sigma}_i^2 = \hat{\beta} + (1-\hat{\eta})\hat{\sigma}_{i-1}^2 + \hat{\varphi}(G_i^{(1)})^2 \tag{E.8}$$

Appendix F. Indirect Inference Method (IIM)

To estimate the COGARCH with NIG-Lévy process parameters, IIM is used. IIM is a simulation-based estimator to estimate parameters of a given model. More details can be found in [182], [183], [79] and [71]. As NIG process is an infinite process it is not possible to use Haug's [87] method of moment estimation that is used for Compound Possion and Variance Gamma processes. In IIM estimation GARCH model is used as an auxiliary model as COGARCH is an analogue of GARCH model and also it captures the aspects of the data. See [79] for a detailed explanation of the IIM.

The IMM optimization is based on minimizing the estimates of parameters between the underlying model and the auxiliary model, such as

$$\begin{aligned}\min f = \sqrt{(mean(G_n^{(1)}) - mean(\hat{G}_n^{(1)}))^2 + (Var(G_n^{(1)}) - Var(\hat{G}_n^{(1)}))^2} \\ + (skew(G_n^{(1)}) - skew(\hat{G}_n^{(1)}))^2 + (kurtosis(G_n^{(1)}) - kurtosis(\hat{G}_n^{(1)}))^2\end{aligned} \tag{F.1}$$

Appendix G. Kolmogorov–Smirnov (K–S) Test

The K-S is a nonparametric, goodness-of-fit test to measure how good a sample distribution describes a reference probability distribution [48].

There are two types of K-S test. One sample K-S is based on max deviation between the observed and the theoretical distributions given as

$$D = \max_{1\leq i\leq N}\left|F(Y_i) - \frac{i}{N}\right| \tag{G.1}$$

where Y_i is a data point within N data points and $F(.)$ is the theoretical cdf.

In the two-sample K-S test, there are no reference distribution, instead the empirical distributions E_1 and E_2 are compared against one another,

$$D = |E_1(i) - E_2(i)| \tag{G.2}$$

The hypotheses of the test are stated as

- H_0: Both samples come from the same distribution,
- H_a: The samples do not come from the same distribution.

Appendix H. Chi-Square Test

This is also a goodness-of-fit measure, to test if sample data points comes from a population with a specific distribution. The chi-square test is an alternative to the Anderson-Darling and Kolmogorov-Smirnov goodness-of-fit tests.

The test statistic is defined as

$$\chi^2 = \sum \frac{(X_i - E_i)^2}{E_i} \tag{H.1}$$

where X_i and E_i are the observed and expected values, respectively.