Bachelor Thesis in Statistics (15 ECTS)
Lund University

Are GARCH Models Appropriate for Analysing Volatility Structures in Fundamental Valuations of the OMXS30?

February 24, 2019

Author
Gustav Furenmo

Supervisor
Krzysztof Podgórski

Abstract

This thesis investigates the volatility structures found in forward-looking fundamental valuations of the Swedish stock index OMXS30. The evaluated data constitutes daily observations of P/E ratios based on twelve months earnings estimates during the period 2009-01-02 until 2018-10-18. The analysis is conducted by applying a GARCH modelling framework to a log-return transformed return series derived from the raw data. This thesis reveals that the underlying data exhibit commonly observed properties found in financial time series with the most prominent ones being volatility clustering (i.e. heteroscedasticity) and leptokurtic behaviour. Parameter efficiency in the maximum likelihood estimation procedure is evaluated using five different distributional assumptions for the GARCH model innovations, namely: Normal distribution, Student-t distribution, skewed Student-t distribution, Generalised error distribution and the skewed Generalised error distribution. The final model choice entails a symmetric GARCH(1,1) model with innovations assumed to be generated from a skewed student-t distribution. Finally, this model proves to be sufficient in describing the volatility structure found in the return series.

Keywords: GARCH, OMXS30, financial time series, volatility, heteroscedasticity, stationarity, McLeod-Li test, normal distribution, student-t distribution, skewed student-t distribution, generalised error distribution, skewed generalised error distribution.

Table of Contents

1. Introduction

"When an investor focuses on short-term investments, he or she is observing the variability of the portfolio, not the returns - in short, being fooled by randomness." – Nassim Taleb

1.1 Background

The Stockholm Stock Exchange has become one of the primary securities exchanges in Northern Europe within equities trading. Even though over 300 companies are listed on the exchange, a majority of the total market capitalization, roughly 60 percent, originates from companies included in the OMXS30 index. The OMXS30 stock index includes thirty constituents who represent the thirty largest companies on the exchange in terms of trading volumes. An important aspect of any equity index is its valuation. This captures various fundamental variables of the economy, such as risk sentiment, corporate performance and interpretations of the general business cycle. Valuations can be quantified with numerous metrics with one of the most common valuation multiples being the price-earnings ratio (P/E). This valuation ratio is defined as price in relation to earnings, where price represents the market value of a company's equity (i.e. market capitalization) and earnings represents a company's net income (i.e. the final line item found in the income statement).

Even though fundamental prerequisites such as management, clients and structural trends for most companies might be considered static in the very short term, the market value of a company's equity is not. On the contrary, stock prices fluctuate continuously, sometimes as a consequence of market events such as earnings reports, but sometimes seemingly as a consequence of pure randomness.

Understanding and adequately describing the volatility in OMXS30 valuations can be of use in further applications such as asset pricing and risk management. For instance, the Black-Scholes option pricing formula relies on volatility estimates in calculating asset prices. Additionally, applications within risk management, such as calculations of Value at Risk (VaR) and Expected Shortfall relies on an accurate understanding of volatility structures in the underlying asset (Hull, 2011).

The analysis of financial time series often requires statistical models that can account for certain properties such as time dependent volatility, i.e. heteroscedasticity (Cryer & Chan, 2008). Models that adhere to these specifications were first introduced by Engle (1982), coined as the Autoregressive Conditional Heteroskedasticity (ARCH) model. This concept was later generalized by Bollerslev (1986) who introduced the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model.

1.2 Hypothesis

The purpose of this thesis is to investigate if the volatility structure found in fundamental valuations of the OMXS30 exhibit the general characteristics often found in financial time series and if GARCH models are appropriate in describing this structure. Furthermore, emphasis is put on investigating which distributional assumptions should be used when estimating the model in order to achieve efficient parameter estimations. Hence, this thesis aims towards investigating and answering the following:

- *Does the volatility in fundamental valuations of the OMXS30 exhibit similar properties as often seen in other financial time series?*
- *Are GARCH models appropriate for describing volatility structures in fundamental valuations of the OMXS30?*
- *Which distributional assumptions should be used for the model innovations in order to achieve accurate models for the data?*

1.3 Disposition

This thesis starts off by discussing the analysed data set and gives some intuition behind the choice of data. Relevant data transformations and data critiques are also provided. Afterwards, the theoretical frame work applied throughout the research process is described, followed by the empirical analysis. Finally, conclusions and suggested extensions of the research presented in this thesis are summarised.

2. Market Valuation Data for OMXS30

This section introduces the data set analysed in this thesis. The section discusses data collection, data transformations, data critiques and the usage of statistical software.

2.1 Raw Data

The data set has been retrieved from the Bloomberg terminal and consists of 2 462 observations (2009-01-02 until 2018-10-18). These observations constitute daily levels of P/E ratios in the stock index OMXS30. These valuation ratios are used to quantify valuations and are calculated as total market capitalization of all index constituents (P) in relation to the combined earnings estimates for all index constituents for the coming twelve months (E). This thesis investigates P/E ratios based on next twelve months (NTM) earnings estimates as oppose to earnings from trailing twelve months (TTM). There are two main reasons for this:

- Valuation multiples based on NTM estimates tend to be less volatile compared to valuation multiples based on earnings TTM.
- In theory, the value of an asset depends on its future cash flow generating capabilities. Hence, valuation multiples based on NTM estimates allows investors to compare assets based on expected corporate performance and not only historical performance.

The market capitalization of the OMXS30 index is directly observable at any given time. However, earnings estimates produced by equity analysts are reported on a discretionary basis. These estimates are commonly updated and reported in close proximity to a company's financial reporting, which usually occurs on a quarterly basis.

2.2 Data Transformations

It is often inappropriate to model financial time series directly since they can exhibit persistent trends which poses problems related to stationarity (see section 3.1 and 3.2). To combat this, several data transformations are viable. One procedure is to impose a log-return transformation to the raw data (Cryer & Chan, 2008). This transformation is defined as:

$$r_t = \log\left(\frac{p_t}{p_{t-1}}\right) = \log(p_t) - \log(p_{t-1})\,, \qquad (1)$$

where p_t is the P/E NTM ratio for OMXS30 at time *t*. Furthermore, each element in the return series, $\{r_t\}$, is multiplied by 100 so that they can be interpreted as percentage changes and to reduce numerical errors related to rounding errors.

2.3 Data Critiques

One critique of the data set that should be brought forward is the fact that forward-looking P/E ratios are based on earnings estimates projected by several equity analysts. Estimates used in calculating forward P/E ratios therefore reflect the market consensus of future earnings and not actual reported earnings. The reader should be aware of this since reported earnings can deviate significantly compared to prior estimates.

Further issues regarding the reliability of consensus earnings estimates arose in 2018 as a consequence of the implementation of new EU regulation in the form of the Markets in Financial Instruments Directive II (MiFID II). This legislation has made an impact on a wide range of aspects related to financial markets, with one being the reporting of consensus estimates. Since the implementation of MiFID II, the number of equity analysts reporting projections such as earnings estimates has decreased. Hence, the market consensus regarding these estimates originate from a smaller sample of analysts (Exane BNP Paribas & EY, 2017). With data covering the time period prior- and subsequent to the implementation of MiFID II, it is possible that these regulatory changes have made a significant impact on the data.

2.4 Statistical Software

All calculations presented in this thesis have been performed using Microsoft Excel or the statistical software R. The data was collected and structured in Excel. All statistical computations have been carried out using R. Within R, the following packages were used:

- *tseries* – Package for time series analysis and computational finance. Used for stationarity tests.
- *TSA* – Package for time series analysis. Used for calculating and plotting autocorrelation functions, partial autocorrelation functions, extended autocorrelation functions, McLeod-Li tests, kurtosis and skewness.
- *Stats* – Package for statistical calculations and normal random number generation. Used to create time series objects, density estimates, normality tests, time series plots and normal QQ plots.
- *readxl* – Package for importing excel files to R.
- *fGarch* – Package for analysing and modelling heteroscedastic time series data and non-normal random number generation. Used for model specification, model estimation, model simulations, residual analysis, empirical distribution estimation, conditional volatility plots, non-normal QQ plots and non-normal random number generation.
- *Graphics* – Package for producing various types of graphs. Used to create histograms, density plots and line segments.

3. Theory of Financial Time Series

This section aims to introduce and discuss the theoretical framework applied in this thesis. First, the concept of stationarity is discussed and then followed by an exposition of typical properties found in financial time series. Finally, the framework for heteroscedastic time series modelling is introduced.

3.1 Stationarity

One key assumption when making statistical inference about the structure of a stochastic process is that of stationarity. A stochastic process is said to be strictly stationary if its unconditional joint probability distribution does not change with respect to time. For financial time series, such an assumption can often be deemed unreasonable due to reasons discussed in section 3.2. Another similar but mathematically more relaxed definition of stationarity is that of weak stationarity. A stochastic process is said to be weakly stationary if its mean function and covariance function are not time dependent (Tsay, 2005). In this thesis, the term stationary refers to the weaker definition. Financial time series often exhibit non-stationarity behaviour due to trends in the data. To deal with this inconvenience, several data transformations are viable in making the data stationary, with a common solution being the log-return transformation (see section 2.2).

Formal statistical tests can be performed to investigate if a time series is stationary, with one procedure being the Augmented Dickey-Fuller Test. This procedure tests the null hypothesis stating that there is a unit root present in the time series sample (i.e. non-stationary) against the alternative hypothesis stating that the time series sample is stationary. The test is constructed by creating the following model:

$$r_t = \alpha r_{t-1} + X_t \quad \text{for } t = 1,2, \ldots \tag{2}$$

where $\{X_t\}$ is a stationary process and the return series, $\{r_t\}$, is a non-stationary process if $\alpha = 1$. If $|\alpha| < 1$, the return series is said to be stationary (Cryer & Chan, 2008). Hence, the hypotheses are formulated as:

$$\mathrm{H}_0: \alpha = 1$$
$$\mathrm{H}_1: |\alpha| < 1$$

3.2 Properties of Financial Time Series

After imposing a log-return data transformation to a financial time series according to equation (2), the resulting return series often exhibit properties discussed below.

3.2.1 Volatility Clustering

When analysing financial return series, a pattern often observed are alternations between calm and volatile periods where small and large values tend to be clustered together. This pattern is referred to as volatility clustering. This occurs because the conditional variance of the return series is not constant over time, i.e. heteroscedastic. When analysing heteroscedastic time series data, ARCH and GARCH models can be appropriate (Cryer & Chan 2008).

3.2.2 Leptokurtic Distribution

Distributions with leptokurtic properties assigns greater probabilities to extreme values relative to those of the normal distribution. This implies that the probability density function (PDF) for a leptokurtic distribution exhibit heavy tails and a higher peak around the mean compared to the PDF for a normal distribution. On the contrary, platykurtic distributions display a lower peak around the mean and less extreme tail behaviour compared to the normal distribution. Finally, the normal distribution is an example of a mesokurtic distribution.

These distributional properties can be measured using the excess kurtosis, where excess refers to deviations from the kurtosis of a normal distribution (i.e. 3). Kurtosis is the normalised fourth central moment where the sample kurtosis, g_2, is estimated by the following equation:

$$g_2 = \sum_{i=1}^{n} \frac{(Y_i - \bar{Y})^4}{n\hat{\sigma}^4} - 3\,, \tag{3}$$

where Y_i is the i^{th} value in the sample, $\bar{Y}$ is the sample mean and the sample variance, $\hat{\sigma}^2$, is calculated as:

$$\hat{\sigma}^2 = \sum_{i=1}^{n} \frac{(Y_i - \bar{Y})^2}{n}\,. \tag{4}$$

The level of excess kurtosis yields the following distribution classifications:

$g_2 < 0$: Platykurtic distribution
$g_2 = 0$: Mesokurtic distribution
$g_2 > 0$: Leptokurtic distribution

Financial time series often exhibit excess kurtosis, i.e. leptokurtic behaviour (Cryer & Chan, 2008).

3.2.3 Skewness

Asymmetry in probability distributions can be measured using skewness, the normalised third central moment. The sample skewness, g_1, is estimated by using (Cryer & Chan, 2008):

$$g_1 = \sum_{i=1}^{n} \frac{(Y_i - \bar{Y})^3}{n\hat{\sigma}^3}. \tag{5}$$

3.2.4 Independence and The Sample Autocorrelation Function (ACF)

Testing the return series for dependency is commonly done by calculating the sample autocorrelation function (ACF). The sample ACF, c_k, at lag k is defined as:

$$c_k = \frac{\sum_{t=k+1}^{n}(Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}. \tag{6}$$

The sample ACF, c_k, for $k = 1,2, \ldots$ is then compared to a critical value calculated as $\pm 2/\sqrt{n}$. Values of c_k that exceed the critical value indicate significant correlations. This procedure is often plotted and presented in a correlogram. If the return series does not exhibit any significant autocorrelations, serial independence is suggested and can be modelled as a white noise process.

Our experience when working with financial return series is that tests for presence of dependence using the autocorrelations, seldom show indications that are significant. If values in the return series are indeed independent, nonlinear transformations such as squaring or taking absolute values should preserve the independence. This allows for testing of higher order serial dependence by examining the autocorrelations for the squared return series. Therefore, if the returns are truly independent and identically distributed, the squared returns should also follow a white noise process.

However, return series are often expected to show heteroscedastic tendencies, indicating serial dependencies. Graphically, this would be manifested as an ACF plot showing significant correlations slowly decaying as the number of lags increase. Such a behaviour would imply serial dependence in the volatility process (Cryer & Chan, 2008).

3.3 Probability Distributions

In accordance with the discussion in section 3.2.2, probability distributions for financial time series often exhibit leptokurtic properties. To achieve efficient parameter estimates, adequate distributional assumptions for the model innovations are of importance. This thesis will focus on five different probability distributions. These are introduced below.

3.3.1 Normal Distribution

The normal distribution is a symmetric (mesokurtic) probability distribution and is widely used in several statistical models. The normal distribution was originally assumed for the model innovations by Engle (1982) in the ARCH model. The normal probability density function is defined as (Hogg et al., 2014):

$$f_X(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{7}$$

3.3.2 Symmetric Student-t Distribution

In 1987, Bollerslev proposed the use of a student-t distribution with $v > 2$ degrees of freedom. The purpose was to improve the model's recognition of conditional heteroscedasticity. It should be noted that the student-t distribution converges to the normal distribution as $v \rightarrow \infty$. The probability density function for the student-t distribution with unit variance is defined as (Hansen, 1994):

$$f_X(x; v) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi(v-2)}\,\Gamma\left(\frac{v}{2}\right)\left(1 + \frac{x^2}{(v-2)}\right)^{\frac{v+1}{2}}} \;. \tag{8}$$

3.3.3 Skewed Student-t Distribution

To allow for a more flexible probability distribution where skewness can be modelled, the skewed student-t distribution is a viable option. This distribution can account for both leptokurtic properties but also possible skewness. The probability density function for the skewed student-t distribution with mean zero and unit variance can be defined as (Hansen, 1994):

$$f_X(x; v, \lambda, a, b, c) = \begin{cases} bc\left(1 + \frac{1}{v-2}\left(\frac{bx+a}{1-\lambda}\right)^2\right)^{-\frac{v+1}{2}}, & x < -a/b \\ bc\left(1 + \frac{1}{v-2}\left(\frac{bx+a}{1+\lambda}\right)^2\right)^{-\frac{v+1}{2}}, & x \geq -a/b \end{cases} \tag{9}$$

where $2 < v < \infty$ and $-1 < \lambda < 1$. The constants *a*, *b* and *c* are defined as:

$$a = 4\lambda c\left(\frac{v-2}{v-1}\right), \tag{10}$$

$$b^2 = 1 + 3\lambda^2 - a^2 \,, \tag{11}$$

and

$$c = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi(v-2)}\Gamma\left(\frac{v}{2}\right)} \;. \tag{12}$$

The skewed student-t takes the form of its symmetric counterpart, (8), when $\lambda = 0$.

3.3.4 Symmetric Generalised Error Distribution (GED)

The generalised error distribution (GED) originates from a family of exponential distributions and can be both leptokurtic and platykurtic depending on the shape parameter, v. If $v = 2$ and λ is given by (14), the GED is a standard normal distribution, whereas $v < 2$ yields a leptokurtic leptokurtic distribution and $v > 2$ yields a platykurtic distribution. The probability density function is defined as (Karlsson, 2002):

$$f_X(x; v, \lambda) = \frac{v}{\lambda 2^{\left(\frac{v+1}{v}\right)}\Gamma\left(\frac{1}{v}\right)} e^{-\frac{1}{2}\left|\frac{x}{\lambda}\right|^v}, \tag{13}$$

where

$$\lambda = \left[\frac{2^{-\frac{2}{v}}\Gamma\left(\frac{1}{v}\right)}{\Gamma\left(\frac{3}{v}\right)}\right]^{\frac{1}{2}}. \tag{14}$$

3.3.5 Skewed Generalised Error Distribution (SGED)

The symmetric GED can be extended to allow for modelling skewness. This provides flexibility when estimating the probability distribution for empirical financial data. The probability density function for the skewed GED is defined as (Lee et al., 2008):

$$f_X(x; \kappa, \delta, \theta, \lambda, C) = Ce^{-\left(\frac{|x+\delta|^\kappa}{[1+\text{sign}(x+\delta)\lambda]^\kappa \theta^\kappa}\right)}, \tag{15}$$

where

$$C = \frac{\kappa}{2\theta}\Gamma\left(\frac{1}{\kappa}\right)^{-1}, \tag{16}$$

$$\theta = \Gamma\left(\frac{1}{\kappa}\right)^{\frac{1}{2}}\Gamma\left(\frac{3}{\kappa}\right)^{\frac{1}{2}}S(\lambda)^{-1}, \tag{17}$$

$$S(\lambda) = \sqrt{1 + 3\lambda^2 - 4A^2\lambda^2}, \tag{18}$$

$$\delta = \frac{2\lambda A}{S(\lambda)}, \tag{19}$$

and

$$A = \Gamma\left(\frac{2}{\kappa}\right)\Gamma\left(\frac{1}{\kappa}\right)^{-\left(\frac{1}{2}\right)}\Gamma\left(\frac{3}{\kappa}\right)^{-\left(\frac{1}{2}\right)}. \tag{20}$$

3.4 Heteroscedastic Time Series Models

As previously discussed, the return series for financial data, $\{r_t\}$, is often serially uncorrelated while showing volatility clustering. This indicates that the conditional volatility of r_t is not constant with respect to time, i.e. heteroscedastic. Hence, models that can account for heteroscedasticity should be employed to accurately describe the volatility structure of r_t. Such models, namely the ARCH and GARCH frameworks, are introduced below.

3.4.1 ARCH

The Autoregressive Conditional Heteroscedasticity (ARCH) model was first introduced by Engle in 1982 and models the conditional volatility, $\sigma^2_{t|t-1}$, of a time series. The ARCH(1) model generates the return series, $\{r_t\}$, in the following way:

$$r_t = \sigma_{t|t-1}\varepsilon_t \,, \tag{21}$$

$$\sigma^2_{t|t-1} = \omega + \alpha r^2_{t-1} \,, \tag{22}$$

where $0 \leq \omega$ and $0 \leq \alpha$ are unkown parameters and $\{\varepsilon_t\}$ is a sequence of independent and identically distributed random variables with mean zero, unit variance and independent of r_{t-j}, $j = 1,2, \dots$ (Cryer & Chan, 2008). These random variables are also referred to as innovations. It may be of use to expand the model by including several lagged returns for forecasting accuracy. Engle (1982) suggested a general form of equation (13), the ARCH(q) model. This model is defined as:

$$\sigma^2_{t|t-1} = \omega + \sum_{i=1}^{q} \alpha_i r^2_{t-i} \,, \tag{23}$$

where $0 \leq \omega$ and $0 \leq \alpha_i$ for $0 < i$. However, previous research has shown that a high order of ARCH(q) is needed to properly describe the conditional volatility structure (Cryer & Chan, 2008). As a remedy, Bollerslev (1986) suggested an extended version of the ARCH model, namely GARCH.

3.4.2 GARCH

The Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model is an extension of the ARCH model that includes the conditional volatility as a linear function of its own lags. The conditional volatility, using the GARCH(p,q) model, is expressed as:

$$\sigma^2_{t|t-1} = \omega + \sum_{j=1}^{p} \beta_j \sigma^2_{t-j} + \sum_{i=1}^{q} \alpha_i r^2_{t-i} \,, \tag{24}$$

where ω, β and α are constant with $0 < \omega$, $0 \leq \beta_j$ and $0 \leq \alpha_i$. The model is weakly stationary if (Cryer & Chan, 2008):

$$\sum_{i=1}^{\max(p,q)} (\beta_i + \alpha_i) < 1\,. \tag{25}$$

For the purpose of this thesis, ARCH models will be neglected with favour to GARCH models for the reasons described above.

3.5 Model Order Specification

When determining the order specification of a GARCH(p,q) model, it is convenient to express the conditional volatility in terms of the squared returns by defining:

$$\eta_t = r_t^2 - \sigma_{t|t-1}^2\,, \tag{26}$$

where $\{\eta_t\}$ is a serially uncorrelated zero mean sequence. Equation (26) is re-written as:

$$\sigma_{t|t-1}^2 = r_t^2 - \eta_t \tag{27}$$

and substituted into equation (24), yielding the following form:

$$\begin{aligned} r_t^2 = \omega + (\beta_1 + \alpha_1) r_{t-1}^2 + \cdots + \left(\beta_{\max(p,q)} + \alpha_{\max(p,q)}\right) r_{t-\max(p,q)}^2 \\ + \eta_t - \beta_1 \eta_{t-1} - \cdots - \beta_p \eta_{t-p}\,, \end{aligned} \tag{28}$$

where $\beta_k = 0$ for all integer values $p < k$ and $\alpha_k = 0$ for $q < k$. Equation (28) implies that the squared return series follows an ARMA(max(p,q),p) process. Hence, the model order identification techniques applicable for ARMA models can be applied to the squared return series (Cryer & Chan, 2008). This thesis will mainly rely on the extended autocorrelation function (EACF) for model order identification.

3.6 GARCH Parameter Estimation

Parameter estimation of the GARCH model can be carried out using maximum likelihood. The log-likelihood function will differ depending on which conditional probability distribution is assumed for the innovations. The different log-likelihood functions are presented below.

3.6.1 Maximum Likelihood Estimation Assuming Normal Innovations

Under the assumption that the GARCH innovations are normally distributed, the log-likelihood function takes the following form:

$$l_n = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{n}\left[\log(\sigma_t^2) + \frac{r_t^2}{\sigma_t^2}\right]. \quad (29)$$

The log-likelihood function cannot be maximised analytically but can be maximised using a numerical procedure (Cryer & Chan, 2008).

3.6.2 Maximum Likelihood Estimation Assuming Student-t Innovations

In accordance with previous discussions, normality in financial time series can be an unreasonable assumption. If the GARCH innovations are assumed to follow a student-t distribution, the log-likelihood function becomes (Peters, 2001):

$$\begin{aligned} l_n = \log\left[\Gamma\left(\frac{v+1}{2}\right)\right] - \log\left[\Gamma\left(\frac{v}{2}\right)\right] - \frac{1}{2}\log\big(\pi(v-2)\big) \\ -\frac{1}{2}\sum_{t=1}^{n}\left[\log(\sigma_t^2) + (1+v)\log\left(1 + \frac{r_t^2}{\sigma_t^2(v-2)}\right)\right]. \end{aligned} \quad (30)$$

3.6.3 Maximum Likelihood Estimation Assuming Skewed Student-t Innovations

If the GARCH innovations are assumed to be generated from a skewed student-t distribution, the log-likelihood function is formulated as (Peters, 2001):

$$\begin{aligned} l_n = \log\left[\Gamma\left(\frac{v+1}{2}\right)\right] - \log\left[\Gamma\left(\frac{v}{2}\right)\right] - \frac{1}{2}\log\big(\pi(v-2)\big) + \log\left(\frac{2}{\xi + \frac{1}{\xi}}\right) \\ + \log(s) - \frac{1}{2}\sum_{t=1}^{n}\left[\log(\sigma_t^2) + (1+v)\log\left(1 + \left(\frac{s r_t}{\sigma_t(v-2)} + \frac{m}{v-2}\right)\xi^{-I_t}\right)\right]. \end{aligned} \quad (31)$$

where v is the degrees of freedom, ξ is the asymmetry parameter and

$$I_t = \begin{cases} 1, & \frac{r_t}{\sigma_t} \geq -\frac{m}{s} \\ -1, & \frac{r_t}{\sigma_t} < -\frac{m}{s} \end{cases}, \quad (32)$$

$$m = \frac{\Gamma\left(\frac{v+1}{2}\right)\sqrt{v-2}}{\sqrt{\pi}\,\Gamma\left(\frac{v}{2}\right)}\left(\xi - \frac{1}{\xi}\right), \quad (33)$$

and

$$s = \sqrt{\left(\xi^2 + \frac{1}{\xi^2} - 1\right) - m^2}\,. \quad (34)$$

3.6.4 Maximum Likelihood Estimation Assuming GED Innovations

If the GARCH innovations are assumed to follow a GED distribution, the log-likelihood function is altered to the following (Karlsson, 2002):

$$l_n = \sum_{t=1}^{n} \left[\log\left(\frac{v}{\lambda}\right) - \frac{1}{2}\left|\frac{r_t}{\sigma_t \lambda}\right|^{v} - (1 + v^{-1}) \log(2) - \log\left[\Gamma\left(\frac{1}{v}\right)\right] - \frac{1}{2}\log(\sigma_t^2) \right], \tag{35}$$

where

$$\lambda = \left[\frac{2^{-\frac{2}{v}} \Gamma\left(\frac{1}{v}\right)}{\Gamma\left(\frac{3}{v}\right)} \right]^{\frac{1}{2}}. \tag{36}$$

3.6.5 Maximum Likelihood Estimation Assuming SGED Innovations

The final case where the innovations are assumed to originated from a skewed generalised error distribution, the log-likelihood function takes the following form (Altun et al., 2017):

$$l_n = -\frac{\left|\frac{r_t}{\sigma_t} + \delta\right|^{\kappa}}{\left[1 + \text{sign}\left(\frac{r_t}{\sigma_t} + \delta\right)\lambda\right]^{\kappa} \theta^{\kappa}}, \tag{37}$$

where the additional parameters are specified as in section 3.3.5.

3.7 Model Diagnostics

Model diagnostics for the final model can be done by analysing the standardised residuals. If the model is correctly specified, the standardised residuals should be approximately independently and identically distributed. The standardised residuals, $\hat{\varepsilon}_t$, are defined as (Cryer & Chan, 2008):

$$\hat{\varepsilon}_t = \frac{r_t}{\hat{\sigma}_t}. \tag{38}$$

Distribution assumptions for the innovations can be reviewed by studying QQ plots for the standardised residuals. If the plotted values significantly deviate from a straight-line pattern, the distributional assumptions can be questioned. Furthermore, the sample ACF for the squared standardised residuals are examined and the McLeod-Li test is performed to scrutinise the model assumptions. Finally, the Akaike Information Criterion (AIC) is calculated where the model yielding the lowest value is preferred. The AIC value is calculated as:

$$\text{AIC} = -2\log(\text{maximum likelihood}) + 2k\,, \tag{39}$$

where k is the number of estimated parameters. Since the AIC is estimated using maximum likelihood, adding additional model parameters can often result in better model fit, which introduces the problem of overfitting. To combat this, the AIC includes a penalty function (the second term in equation 39) with the purpose of capturing the trade-off between adding additional model parameters, risking overfitting and obtaining low AIC values (Tsay, 2005).

4. Analysis of the OMXS30 Index Data

This section is dedicated to analysing the empirical data underlying this thesis. Data transformations, model estimation and diagnostics are discussed. Procedures introduced in the previous section are applied and evaluated.

4.1 Data Transformations

The P/E NTM ratio for OMXS30 during the period 2009-01-02 until 2018-10-18 is displayed in Figure 1 (Bloomberg, 2018). There seems to be a positive trend but with large fluctuations at times with major market events such as the Euro crisis and the Brexit election. Figure 2 shows the daily logarithmic returns. There are clear indications of volatility clustering since the return pattern seems to alternate between calm and volatile time intervals.

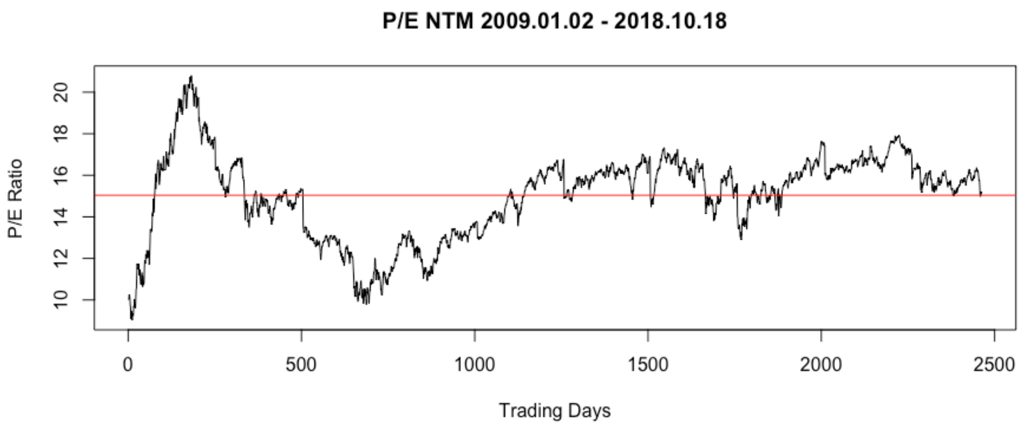


Figure 1: P/E NTM ratio for OMXS30 from 2009-01-02 until 2018-10-18 with the red line representing the mean valuation

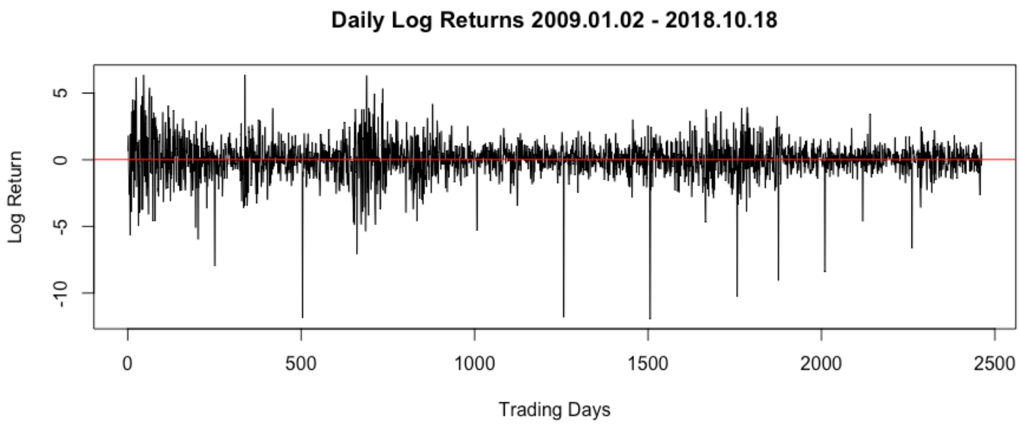


Figure 2: Daily log returns (also referred to as returns) with the mean value represented by the red line

4.2 Tests for Stationarity

The mean of the return series is close to zero (0.017) and seems to be rather constant over time, as is indicated by the red line in Figure 2. The return series also exhibit heteroscedasticity which indicates the appropriateness of applying the GARCH framework.

The Augmented Dickey-Fuller test is calculated in order to test for stationarity. The test yields a p-value of less than 0.01 with stationarity being the alternative hypothesis. This gives further indication of the return series being stationary.

The sample autocorrelation function (ACF) is calculated and investigated for the raw return series as well as for the squared return series. These are presented in Figure 3. The ACF for the raw return series is significant at four lags, indicating that there is still some autocorrelation present. However, in the context of further analysis, these dependencies are small, and the return series is deemed stationary.

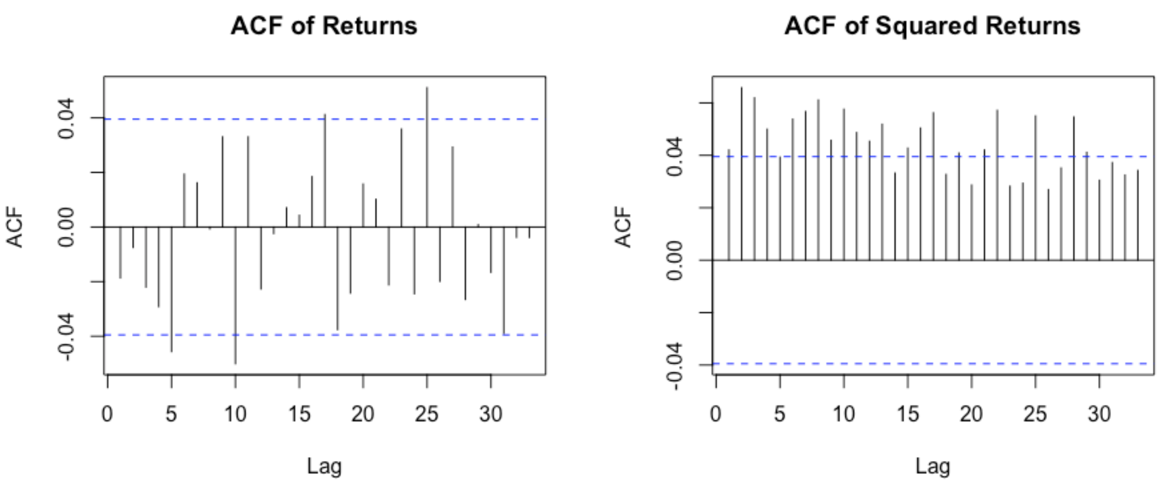


Figure 3: Sample ACF for returns (left) and sample ACF for square returns (right)

The right panel in Figure 3 presents the sample ACF for the squared return series. The ACF is significant at most lags, indicating higher order serial dependency in the volatility process (i.e. heteroscedasticity). This indicates that GARCH models are applicable for modelling the return series.

4.3 Tests for ARCH

The data are examined for ARCH by using the McLeod-Li test. This test is based on the Box-Ljung statistic for the squared returns and entails a null hypothesis stating that there are no ARCH effects in the data (see McLeod & Li (1983) and Li (2004) for further details). The test is presented graphically in Figure 4, where the red line signals the significance level ($\gamma = 5\ \%$). The test is significant at all lags, giving substance to rejecting the null hypothesis. Hence, there seems to be ARCH effects in the data which contributes further to the appropriateness for applying GARCH models.

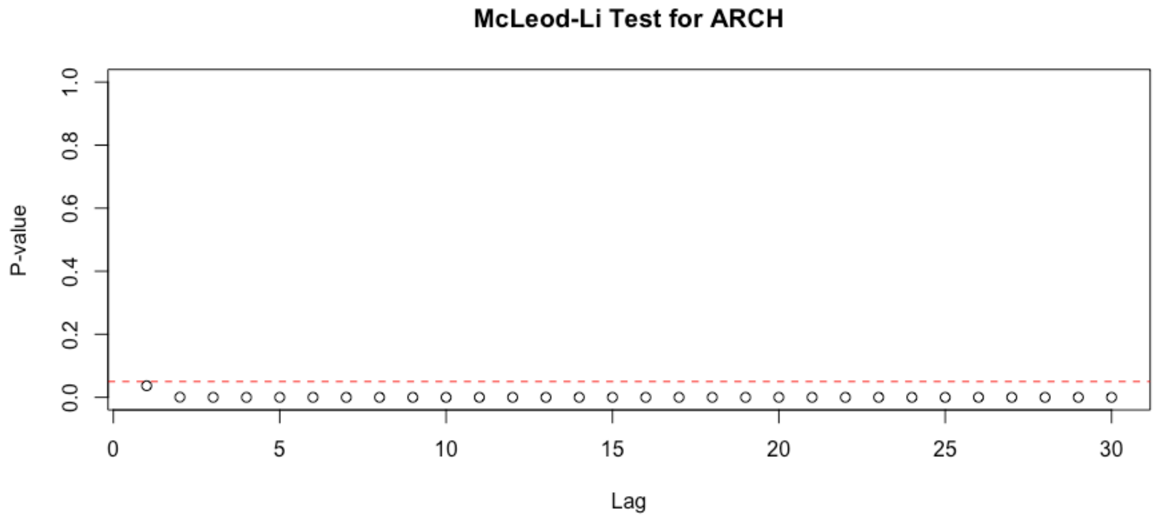


Figure 4: McLeod-Li test for ARCH effects in the return series

4.4 Empirical Distribution of Returns

An important aspect when constructing a GARCH model is accuracy in the distributional assumptions for the innovations. This is because the maximum likelihood estimation will differ depending on assumptions put on the innovations and will impact the parameter efficiency. However, innovations from the return series' data generating process are not directly observable. Therefore, the distribution of the return series is analysed empirically with the purpose of receiving an indication of which assumptions should be made for the innovations during the estimation procedure.

It should be noted that the estimated empirical distributions for the return series are unconditional and may be affected by the heteroscedastic properties found earlier. Hence, the optimal conditional distributions for the innovations used in the maximum likelihood procedure need not be the same. However, the unconditional estimates should provide guidance in finding appropriate distributional assumptions for the innovations.

4.4.1 Normal Distribution

The sample kurtosis for the return series is estimated to approximately 9.77, which is well beyond the theoretical kurtosis of a normal distribution. Furthermore, the Shapiro-Wilk test for normality yields a p-value close to zero, implying that the return series is not normally distributed. This becomes quite apparent when studying the QQ plot for the return series (see Figure 5) where the pattern deviates significantly from a straight line. Finally, the histogram for returns with a superimposed theoretical normal distribution estimated from the data clearly displays the excess kurtosis present in the data. Therefore, assuming normality for the innovations when estimating the model should be rejected.

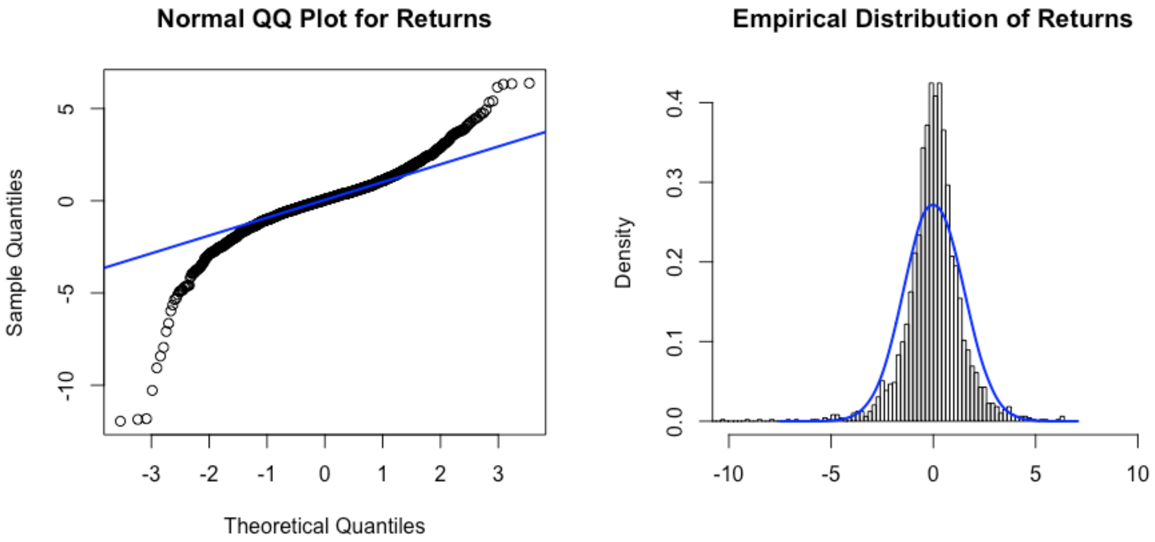


Figure 5: Normal QQ plot for returns and histogram of returns with superimposed normal curve

4.4.2 Symmetric and Skewed Student-t Distribution

The student-t distribution has leptokurtic properties under most circumstances and might provide a better fit to the data. In the symmetric case, the shape parameter (i.e. degrees of freedom, v) can be estimated empirically using maximum likelihood. The estimates, under the assumption of the symmetric student-t distribution are presented in Table 1 where the shape parameter is estimated to 2.96. The QQ plot for the return series assuming a student-t distribution with a shape parameter of 2.96 is presented in Figure 6. In addition, Figure 6 display a histogram for returns with a superimposed curve showing the density for a theoretical student-t distribution with 2.96 degrees of freedom. Both graphs indicate a significantly better fit compared to the normal distribution. This agrees with the leptokurtic properties often seen in financial time series.

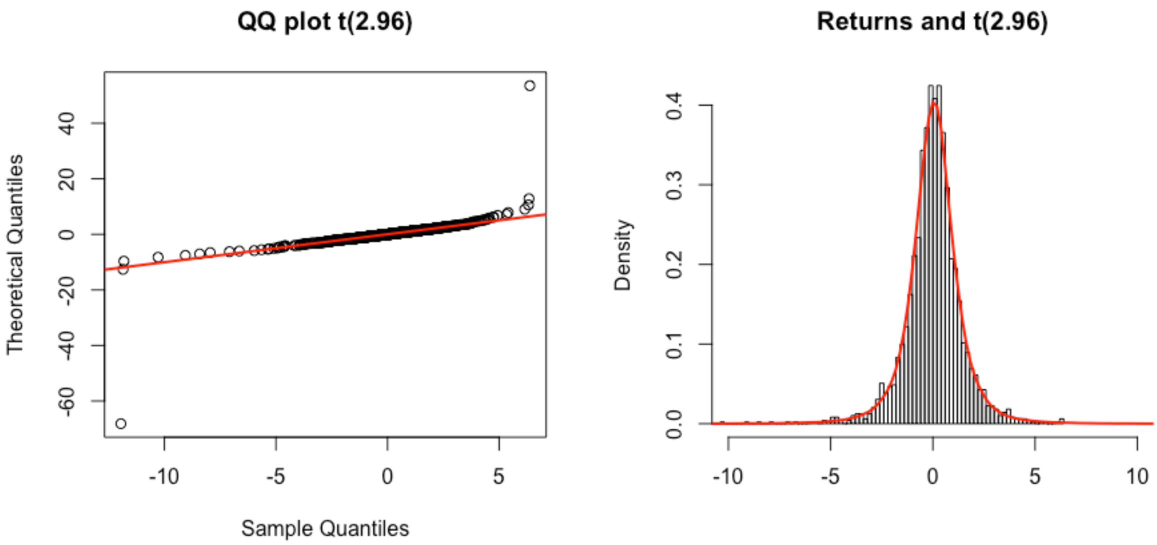


Figure 6: QQ plot for returns and histogram with superimposed theoretical t(2.96) density curve

Distribution	$\bar{X}_{ML}$	SD_{ML}	$Shape_{ML}$
Symmetric Student-t	0.05459	1.54413	2.96902

Table 1: Maximum likelihood estimates of returns assuming a symmetric student-t distribution

The sample skewness for the raw return series is estimated to approximately -1.12. This might be due to asymmetric tendencies, such as leverage effects, sometimes observed in financial time series. It is therefore interesting to investigate if introducing skewness improves model fit. The skewed student-t distribution might be of interest. This distribution has an additional skewness parameter that also can be estimated empirically. The empirical estimates under the skewed student-t assumption is presented in Table 2. Compared to the symmetric version, the shape parameter is practically identical in magnitude. A histogram with the superimposed estimated distribution and a QQ plot is presented in Figure 7. The skewed student-t distribution seems to produce almost identical results as the symmetric version.

Distribution	$\bar{X}_{ML}$	SD_{ML}	$Shape_{ML}$	$Skew_{ML}$
Skewed Student-t	0.03228	1.54296	2.97444	0.96763

Table 2: Maximum likelihood estimates of returns assuming a skewed student-t distribution

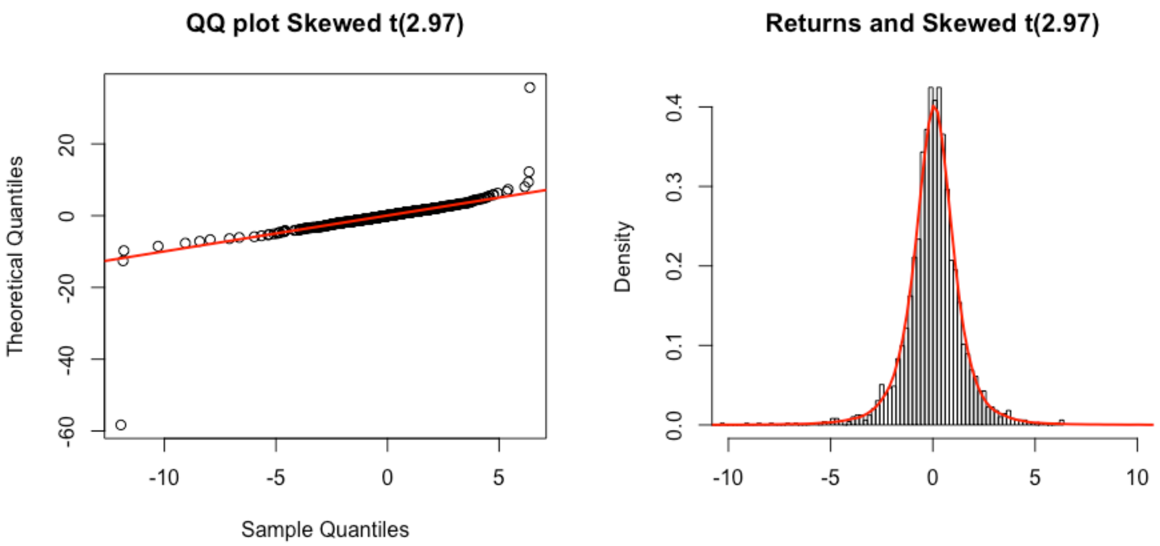


Figure 7: QQ plot for returns and histogram with superimposed theoretical skewed t(2.97) density curve

4.4.3 Symmetric and Skewed Generalised Error Distribution

If the return series belongs to a leptokurtic generalised error distribution, the shape parameter (v) should be less than two. Once again, a skewness parameter can be introduced to model asymmetric behaviour. Empirical estimates under the assumption of a symmetric and skewed generalised error distribution are presented in Table 3. In both cases, estimates for the shape parameter are almost identical. QQ plots and histograms for returns with superimposed theoretical generalised error distributions are displayed in Figures 8 and 9. Here, the GED seems to overestimate the kurtosis slightly. In addition, the QQ plots suggest that the student-t distribution provides a better fit and seems favourable compared to the normal and generalised error distributions.

Distribution	$\bar{X}_{ML}$	SD_{ML}	$Shape_{ML}$	$Skew_{ML}$
Symmetric GED	0.06681	1.38218	0.96281	0.00000
Skewed GED	0.01729	1.38065	0.96647	0.94502

Table 3: Maximum likelihood estimates of returns assuming a skewed student-t distribution

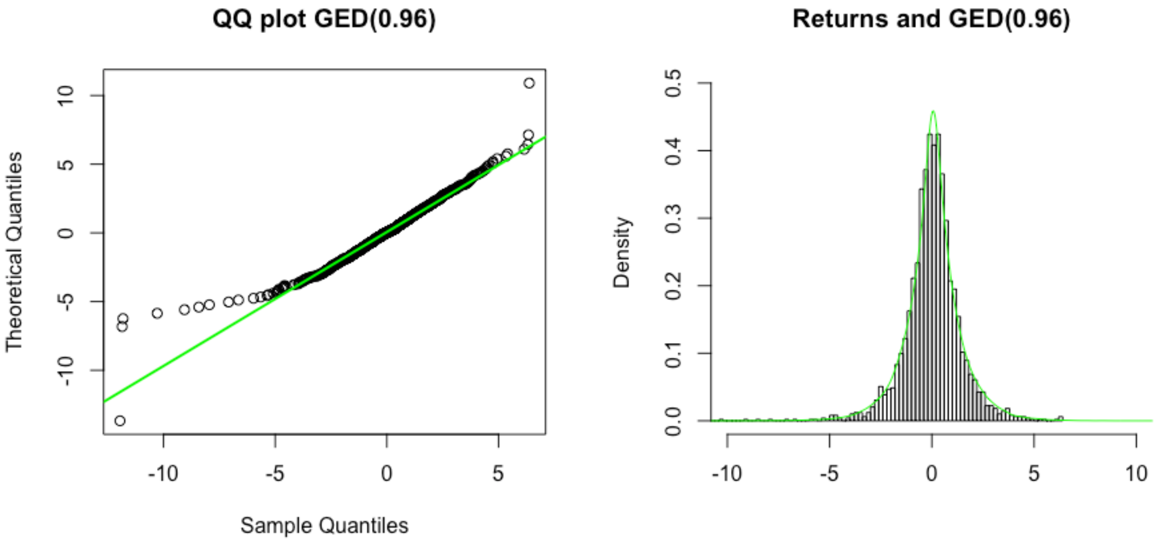


Figure 8: QQ plot for returns and histogram with superimposed theoretical GED(0.96) density curve

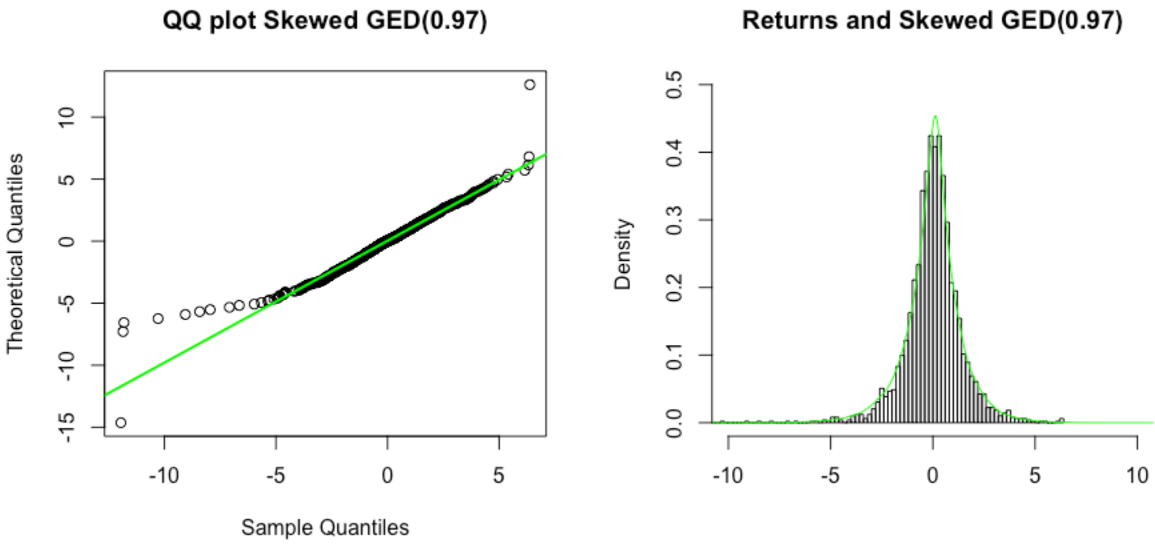


Figure 9: QQ plot for returns and histogram with superimposed theoretical skewed GED(0.97) density curve

4.5 Model Order Specification

The extended sample autocorrelation function (EACF) is calculated for the squared return series and are presented graphically in Table 4. This representation suggests that a GARCH(1,1) model is appropriate for modelling the data. Therefore, the GARCH(1,1) will be the model order specification used going forward.

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	X	X	X	X	0	X	X	X	X	X	X	X	X	0
1	X	0	0	0	0	0	0	0	0	0	0	0	0	0
2	X	X	0	0	0	0	0	0	0	0	0	0	0	0
3	X	X	0	0	0	0	0	0	0	0	0	0	0	0
4	X	0	X	X	0	0	0	0	0	0	0	0	0	0
5	x	X	X	X	X	0	0	0	0	0	0	0	0	0
6	X	0	X	X	X	X	0	0	0	0	0	0	0	0
7	X	0	X	X	0	X	0	0	0	0	0	0	0	0

Table 4: Extended autocorrelation function for squared return series

4.6 Model Estimation

Five GARCH(1,1) models have been fitted to the return series using maximum likelihood and the different distributional assumptions discussed above. The results are presented in Table 5. For all models, a constant mean function was used in the estimation procedure. Furthermore, the optimization algorithm *nlminb* was used for all models except for those relying on symmetric or skewed GED innovations. In these cases, the *lbfgsb* algorithm was used. For all models, the estimates for omega, alpha and beta are quite similar, indicating robust parameter estimates and that GARCH models are appropriate for describing the volatility structure in the return series. It should also be noted that all coefficients are significant.

Conditional Distribution	Omega	Alpha	Beta	Alpha + Beta	$Shape_{ML}$	$Skew_{ML}$	AIC
Normal	0.11751 (***)	0.07961 (***)	0.86564 (***)	0.94525	2.00000	0.00000	3.43363
Symmetric Student-t	0.03602 (**)	0.08130 (***)	0.90154 (***)	0.98284	4.51011 (***)	0.00000	3.18107
Skewed Student-t	0.03507 (**)	0.08354 (***)	0.90066 (***)	0.98420	4.56757 (***)	0.90059 (***)	**3.17484**
Symmetric GED	0.05256 (**)	0.05256 (***)	0.89418 (***)	0.97390	1.08476 (***)	0.00000	3.22319
Skewed GED	0.04532 (***)	0.07942 (***)	0.89824 (***)	0,97766	1.11091 (***)	0.87394	3.20984

Table 5: Estimates for fitted GARCH(1,1) models

At this stage, the GARCH(1,1) model assuming a skewed student-t distribution for the innovations is preferred since it has the lowest AIC value. However, further analysis concerning model fit should be performed to validate this initial conclusion. This is mainly done by analysing the standardised residuals for each model.

4.7 Model Diagnostics

In addition to the AIC value comparison, the preferred GARCH(1,1) model is determined by evaluating the standardised residuals for each model. Table 6 shows the estimated mean values and standard deviations for the standardised residuals from each model. The table also includes maximum likelihood estimates for the shape and skew parameters when applicable. Furthermore, the standardised residuals should follow the same distribution as the model is estimated from. The standardised residuals from each model is discussed further below.

Conditional Distribution	X_{ML} Standardised Residuals	SD_{ML} Standardised Residuals	$Shape_{ML}$ Standardised Residuals	$Skew_{ML}$ Standardised Residuals
Normal	-0.00276	1.00126	2.00000	N/A
Symmetric Student-t	0.04253	1.00452	4.41286	N/A
Skewed Student-t	0.01203	1.00052	4.53897	0.90733
Symmetric GED	0.04264	1.00136	1.07933	N/A
Skewed GED	-0.01074	1.00395	1.11114	0.86790

Table 6: Estimates for the standardised residuals from each fitted GARCH(1,1) model

4.7.1 Model Diagnostics Assuming Normal Distribution

Figure 10 presents a QQ plot for the standardised residuals assuming a normal distribution and a histogram of the standardised residuals with a superimposed theoretical normal distribution. The QQ plot shows quite heavy deviations from a straight-line pattern and the model does not seem to capture the leptokurtic properties. This was expected since the normality assumption for the return series was rejected prior to model fitting. Hence, a distribution with heavier tails should be better suited.

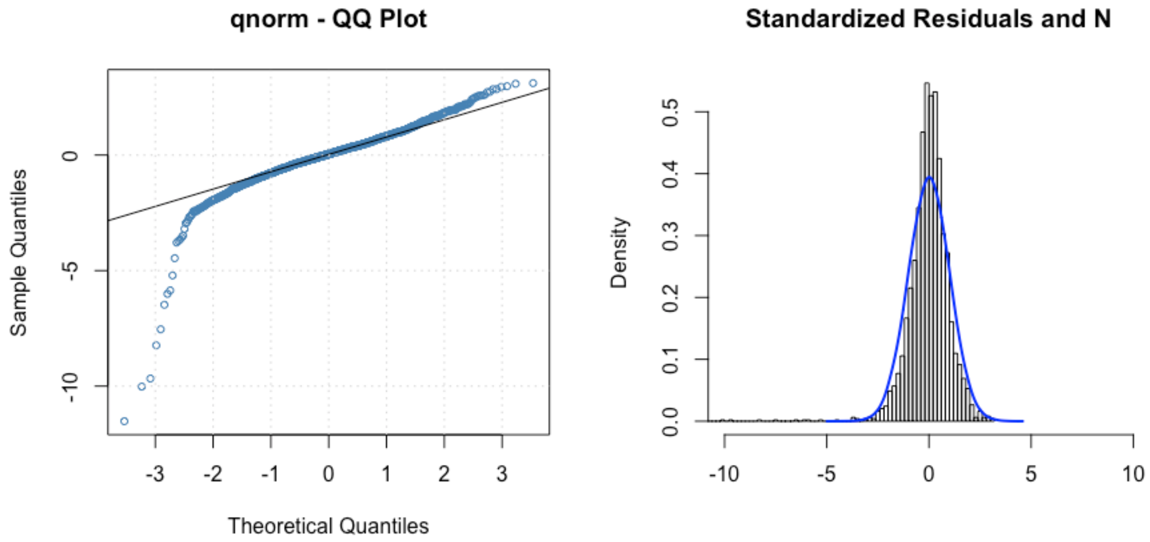


Figure 10: Normal QQ plot for standardised residuals and histogram with superimposed theoretical normal distribution

4.7.2 Model Diagnostics Assuming Symmetric and Skewed Student-t Distribution

Figure 11 displays the standardised residuals for the GARCH(1,1) model where a symmetric student-t distribution was assumed. By examining the QQ plot, we see some deviations from a straight-line pattern. However, this model seems to capture the tail behaviour better than the one analysed in section 4.7.1. It should also be mentioned that a weakness of the QQ plot is that the graphical interpretation becomes more difficult when extreme values exist in the data since these distort the results visually. Furthermore, the histogram with the superimposed theoretical counterpart display a significantly better fit compared to that of the model assuming normally distributed innovations.

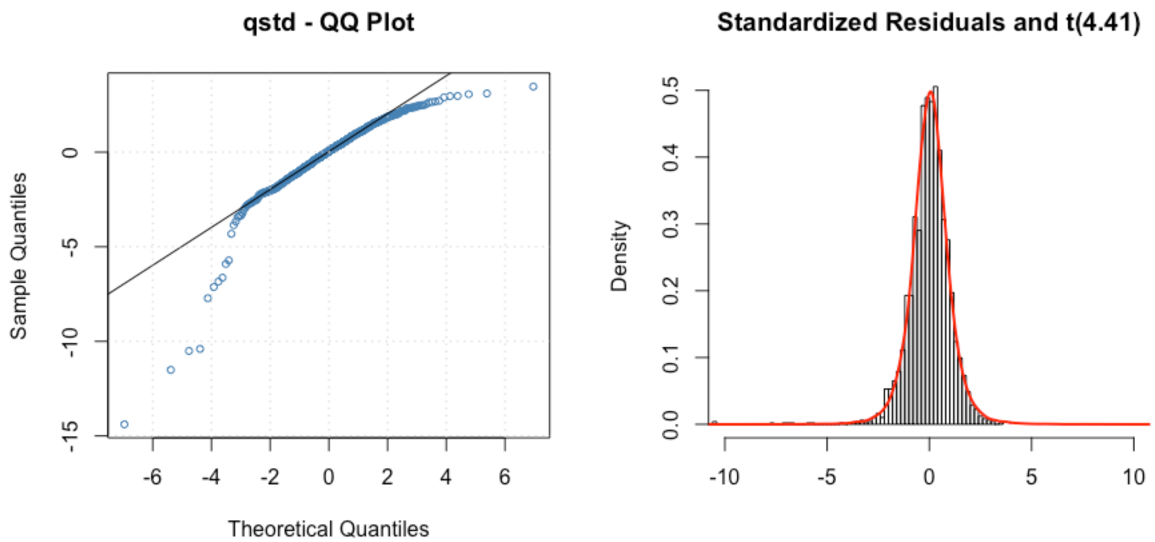


Figure 11: QQ plot for standardised residuals and histogram with superimposed theoretical t(4.41) distribution

The standardised residuals for the model assuming a skewed student-t distribution are presented in Figure 12. The results are almost identical to those in the symmetric version, perhaps with a slightly better fit. Since the model assuming a skewed student-t distribution provided a lower AIC value, this model is preferable between the two.

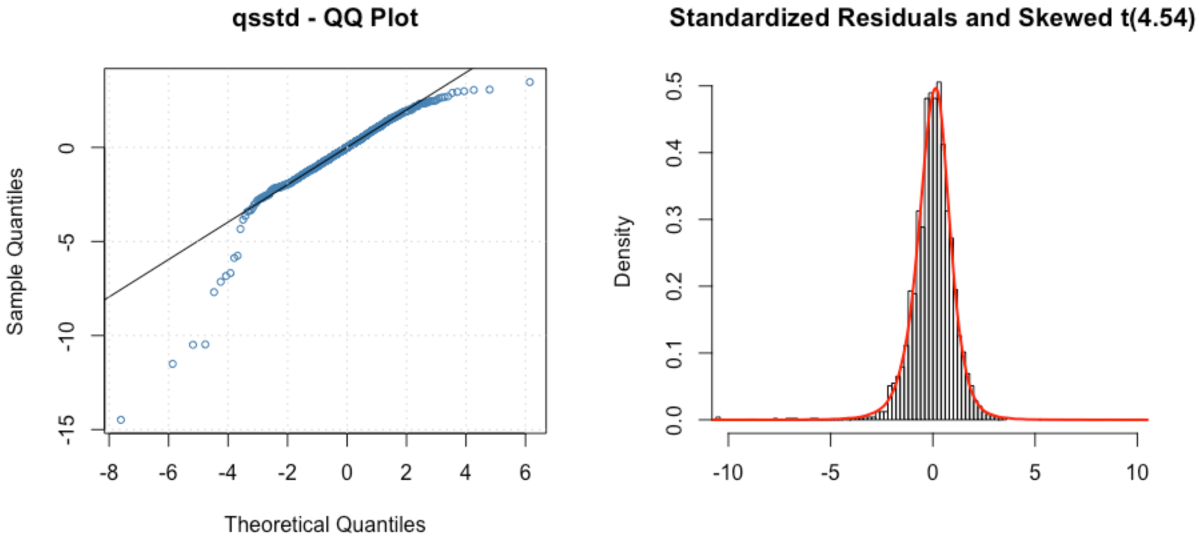


Figure 12: QQ plot for standardised residuals and histogram with superimposed theoretical t(4.41) distribution

4.7.3 Model Diagnostics Assuming Symmetric and Skewed Generalised Error Distribution

The standardised residuals for the GARCH(1,1) model estimated from a symmetric generalised error distribution are presented in Figure 13. In comparison with the model assuming normal innovations, the QQ plot indicates that the GED assumption captures the tail behaviour better. However, in comparison with the models assuming student-t innovations, the differences are quite small. Finally, the histogram in Figure 13 implies a better model fit when a symmetric GED is assumed for the innovations compared the case of the normal distribution assumption. When compared to the student-t models, smaller differences are obtained with some tendencies towards higher kurtosis.

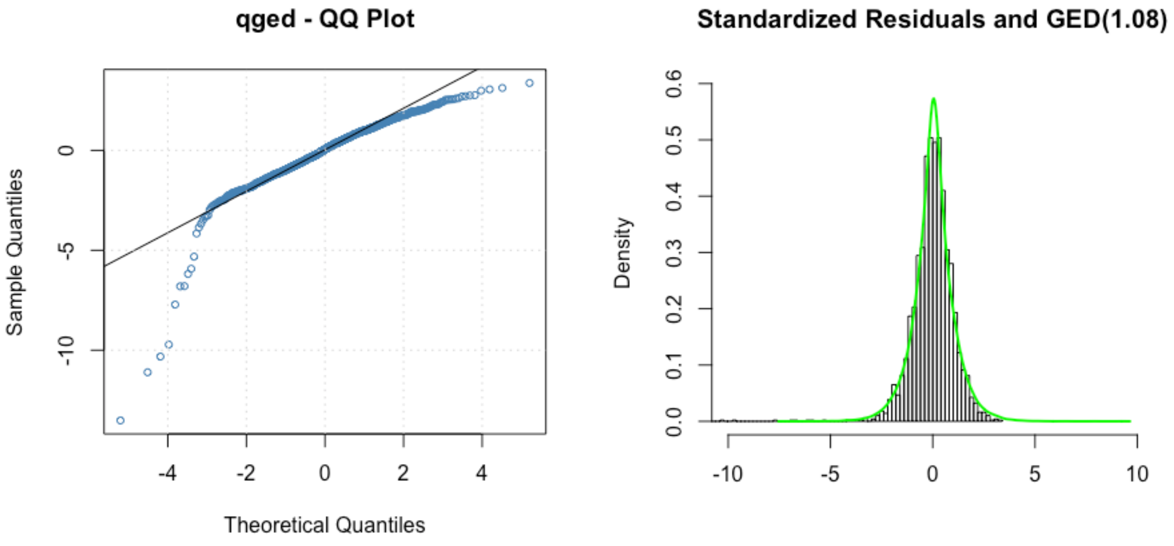


Figure 13: QQ plot for standardised residuals and histogram with superimposed theoretical GED(1.08) distribution

The last GARCH(1,1) model is estimated assuming a skewed generalised error distribution for the innovations. The standardised residuals are presented in Figure 14. Once again, the skewed version only deviates slightly from the symmetric one, showing only small differences.

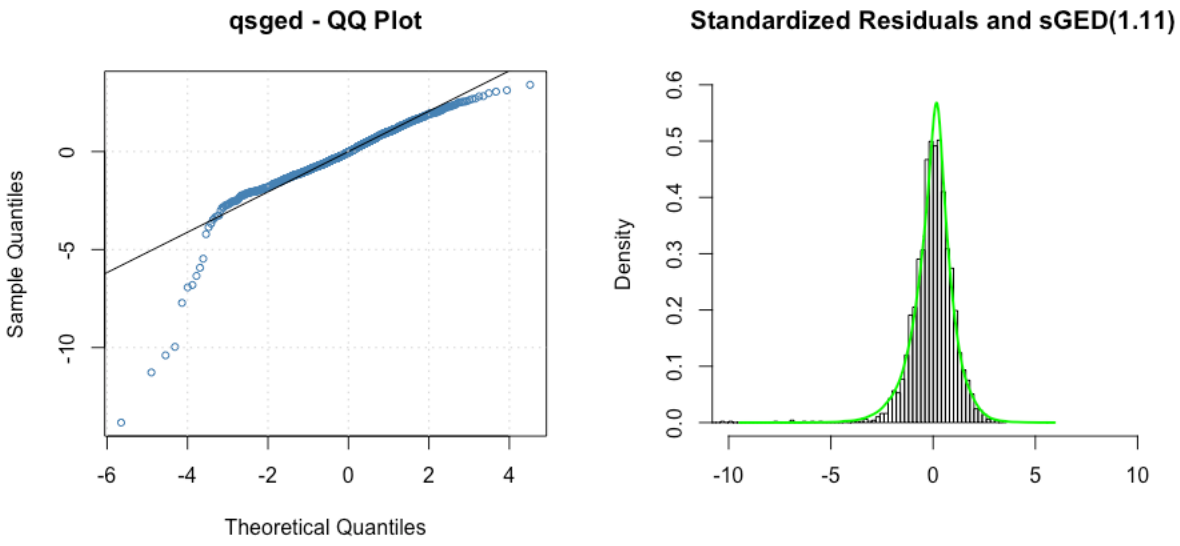


Figure 14: QQ plot for standardised residuals and histogram with superimposed theoretical skewed GED(1.11) distribution

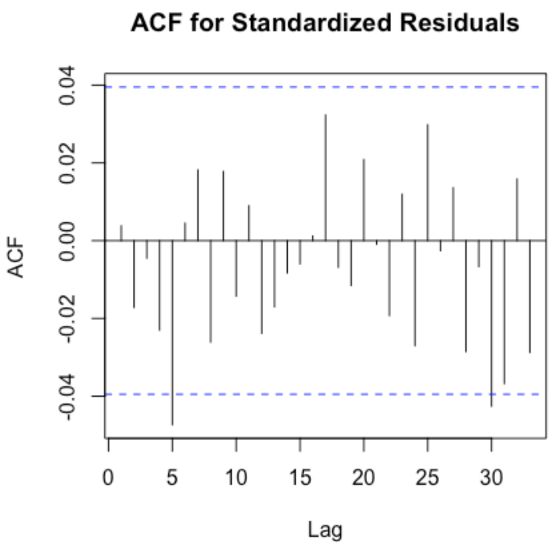
4.8 Final Model Decision

After analysing the standardised residuals for each model, it can be concluded that the GARCH models assuming leptokurtic innovations provide a better fit compared to assuming normality for the innovations. The residual analysis does not suggest major differences between these models. Therefore, the model that obtained the lowest AIC value will be chosen, in this case the GARCH(1,1) model with innovations generated from a skewed student-t distribution. The final step in validating this model is to determine if the standardised residuals are independent and identically distributed. This is done by analysing the sample ACF and EACF for the squared standardised residuals and calculating the McLeod-Li test.

The sample ACF for the standardised residuals and their squares are presented in Figure 15. These plots suggest that the standardised residuals are stationary and do not exhibit any autocorrelation in the volatility process. This result is reinforced by the EACF presented in Table 7, where no distinct pattern can be found. Finally, the visual representation of the McLeod-Li test, presented in Figure 16, indicates that there are no ARCH effects in the standardised residuals.

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	X	0	0	0	0	0	0	0	0	0	0	0	0	0
2	X	X	0	0	0	0	0	0	0	0	0	0	0	0
3	X	X	X	0	0	0	0	0	0	0	0	0	0	0
4	X	X	X	X	0	0	0	0	0	0	0	0	0	0
5	X	X	X	0	0	0	0	0	0	0	0	0	0	0
6	X	X	0	0	X	X	0	0	0	0	0	0	0	0
7	X	X	X	0	X	X	X	0	0	0	0	0	0	0

Table 7: Extenteded autocorrelation function for the squared standardised residuals



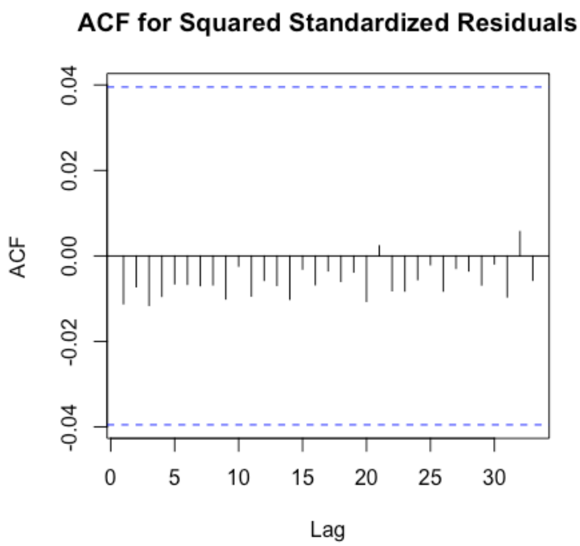


Figure 15: Sample ACF for raw and squared standardised residuals

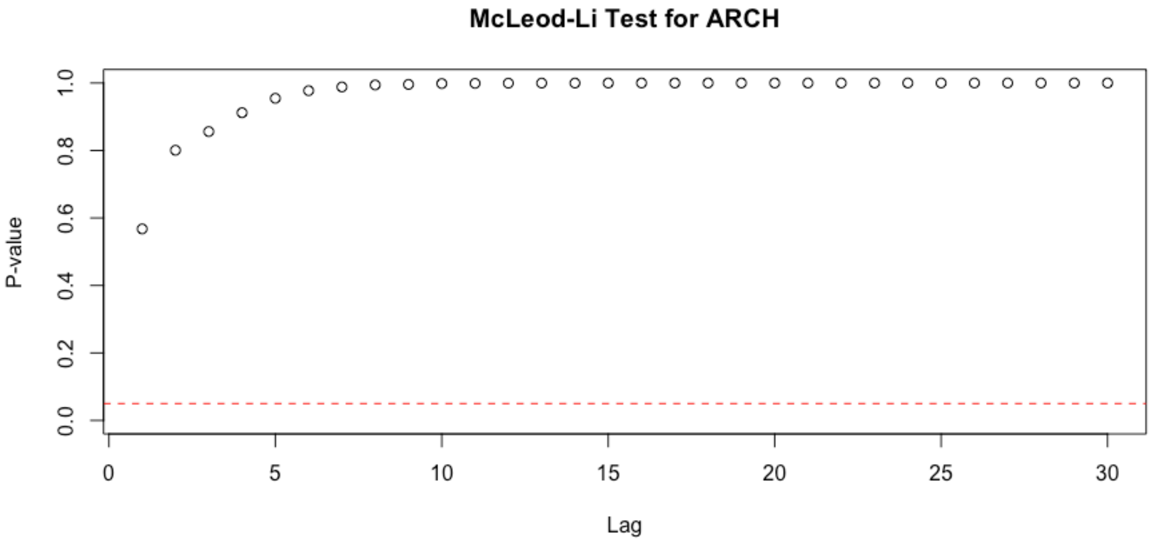


Figure 16: McLeod-Li test for ARCH effects

The final GARCH(1,1) model chosen to describe the volatility structure in fundamental valuations of the OMXS30 is summarised in Table 8.

Conditional Distribution	Omega	Alpha	Beta	Alpha + Beta	$Shape_{ML}$	$Skew_{ML}$	AIC
Skewed Student-t	0.03507 (**)	0.08354 (***)	0.90066 (***)	0.98420	4.56757 (***)	0.90059 (***)	**3.17484**

Table 8: Final GARCH(1,1) model specification

This model specification was used to simulate a new return series of the same length as the empirical return series. A histogram comparing the densities of the empirical data and the simulated series is presented in Figure 17. This plot provides evidence that the estimated GARCH model is able to produce similar data to those found the empirical return series.

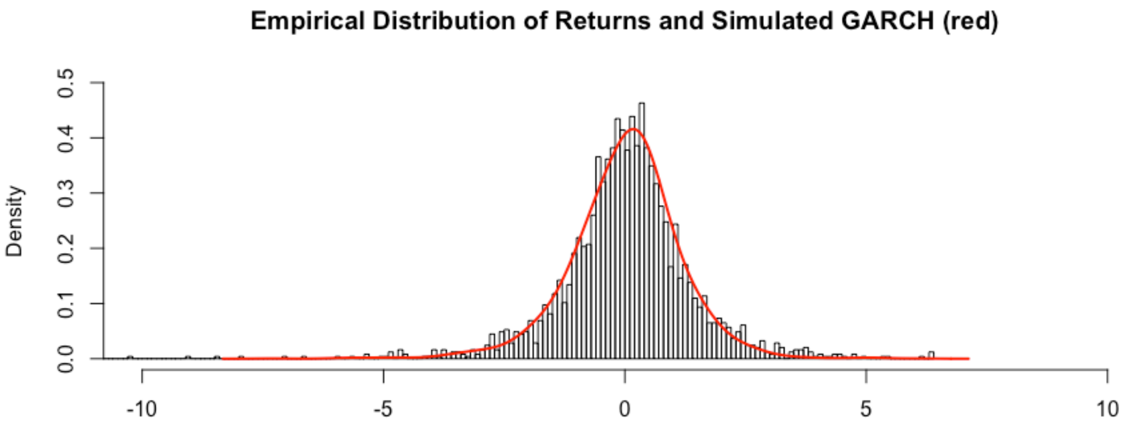


Figure 17: Histogram of returns with superimposed GARCH simulation density curve

Finally, it is imperative to investigate to what extent the model can detect volatility in the return series. This can be done by graphically comparing the conditional volatility estimated by the model with the empirical return series. The estimated conditional volatility and the return series is presented in Figure 18. One can see that the model sufficiently estimates the volatility structure in the return series as volatile time periods in the return series results in greater estimated conditional volatility. This is clearly illustrated during the time period 0 to 100 and 600 to 800.

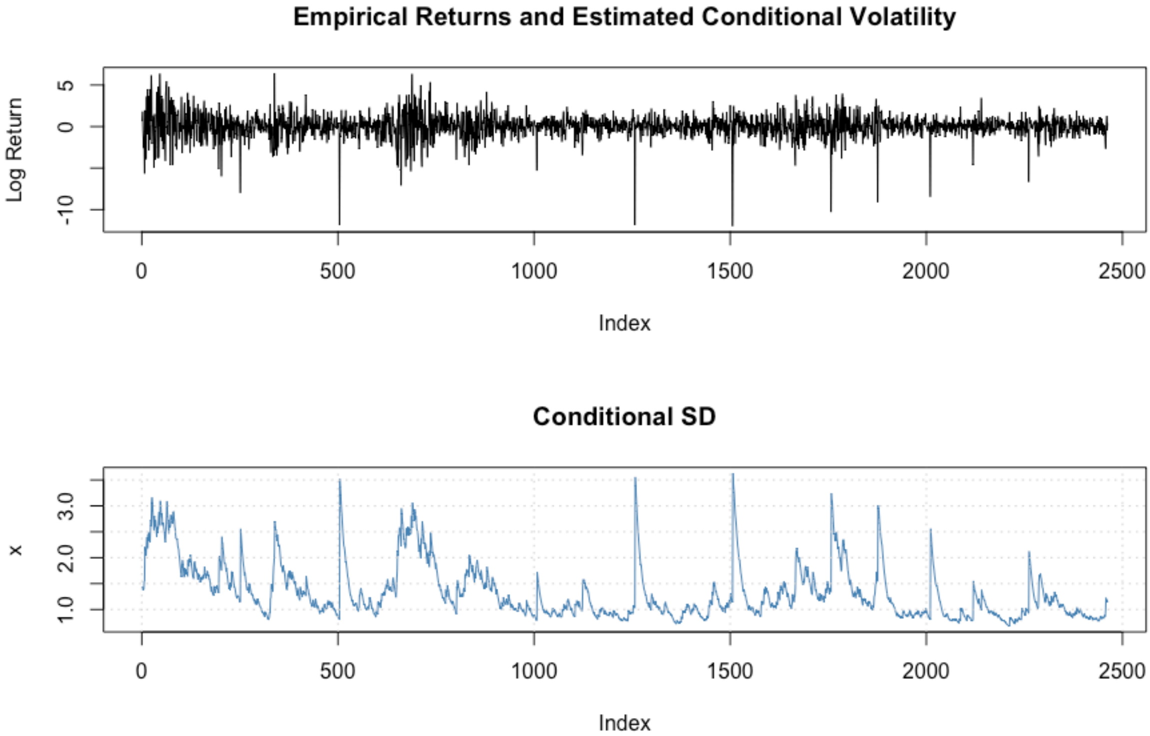


Figure 18: Return series (upper panel) and conditional volatility estimated by the final GARCH(1,1) model

5. Concluding Remarks

This section aims to answer the questions proposed in the 'Hypothesis' section. In addition, this section summarises the specific results found in the empirical analysis and draws conclusions from the estimation procedure. In addition, suggestions for future research topics that might provide further insight into the volatility structure found in fundamental valuations of stock indices are discussed.

5.1 Conclusions

The purpose of this thesis was to investigate if the volatility structures found in fundamental valuations of the OMXS30 exhibit the general characteristics often found in financial time series and if GARCH models sufficiently can describe this structure. The thesis also set out to determine which distributional assumptions should be used for the innovations during the estimation procedure in order to obtain efficient parameter estimates. Limitations were made with the purpose of focusing the analysis on a time period ranging from 2009-01-02 until 2018-10-18. This time period only reflects "bull market conditions", that is market conditions where asset prices tend to trend upwards.

It has now been established that the return series, based on forward-looking P/E ratios, exhibit the general characteristics often found in financial time series. The most prominent properties were volatility clustering and the lack of normality for the raw return series. Furthermore, it was thoroughly established that leptokurtic conditional distribution assumptions for the model innovations provides a better fit when estimating the model. Hence, for these data the normal distribution has proven to be insufficient in capturing the more extreme tail behaviour seen in the analysed financial data. Consequently, these results have given support to applying the GARCH framework with leptokurtic innovations to model the volatility structure in the return series.

The final model choice resulted in a GARCH(1,1) model estimated with the assumption of skewed student-t distributed innovations. This model achieved the lowest AIC value out of all the models and the standardised residuals acceptably resembled the conditional distribution from which they were assumed to be generated from.

In addition, the chosen model was able to sufficiently describe the volatility structure found in the return series which was showed in Figure 18. Furthermore, when the densities of the empirical return series and the simulated series based on the specified model were compared, they exhibited similar shapes (see Figure 17). This provides confidence that the specified model actually is able to produce data similar to the empirical return series.

5.2 Further Research Topics

The return series analysed in this thesis only cover a time period where bull market conditions were present. The results found may not be equivalent if the time period covers more pessimistic market conditions, for example during a recession. Such market conditions could very well yield different results and could statute an interesting extension to the research presented in this thesis. In addition, this thesis only focuses on the Swedish equity index OMXS30. A similar research procedure could be conducted for other equity indices where the results may differ.

Furthermore, this thesis only entails standard GARCH models which by nature are symmetrical. Some financial time series exhibit asymmetrical properties due to phenomena such as leverage effects. Therefore, extensions to the standard GARCH framework, such as the EGARCH could be applied and possible yield better models in terms of fitting.

Finally, general anomalies found in the stock market such as calendar effects are likely to impact the fundamental valuations of a stock index. Hence, time series models analysing these valuations might benefit from various model extensions that account for these effects.

To conclude, this thesis should be viewed as an initial attempt to bring insight into the volatility structure of fundamental valuations of the OMXS30. Extensions to this thesis suggested above could proficiently contribute to a further understanding of this interesting matter.

6. Reference List

Altun, E., Alizadeh, M., Ozel, G., Tatlidil, H., & Maksayi, N. (2017). Forecasting Value-at-Risk with Two-Step Method: Garch-Exponentiated Odd Log-Logistic Normal Model. *Romanian Journal of Economic Forecasting* 20(4), pp. 97-115.

Bloomberg L.P. (2018). *Forward PE Graph for OMXS30, 2009/01/02 to 2018/10/18.* Retrieved October 18, 2018 from Bloomberg Terminal.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics* 31(3), pp. 307-327.

Bollerslev, T. (1987). A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *The Review of Economics and Statistics* 69(3), pp. 542-547.

Cryer, J.D., & Chan, K.-S. (2008). *Time Series Analysis With Application in R*. 2nd Edition. Springer.

Engle, R.F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* 50 (4), pp. 987-1007.

Exane BNP Paribas, & EY (2017). *MIFID2 IR SURVEY*. [Online].
Accessible via: http://irsociety.org.uk/files/MiFID2_IR_survey_-_A4.pdf

Hansen, B.E. (1994). Autoregressive Conditional Density Estimation. *International Economic Review* 35(3), pp. 705-730.

Hogg, R.V., Tanis, E.A., & Zimmerman, D.L. (2014). *Probability and Statistical Inference.* 9th Edition. Pearson.

Hull, J.C. (2011). *Options, Futures and Other Derivatives: Global Edition.* 8th Edition. Pearson

Karlsson, L. (2002). *GARCH – Modeling: Theoretical Survey, Model Implementation and Robustness Analysis*. Algorithmica Research AB.

Lee, M.-C., Su, J.-B., & Liu, H.-C. (2008). Value-at-risk in US stock indices with skewed generalized error distribution. *Applied Financial Economics Letters* 4(6), pp. 425-431.

Li, W.K. (2004). *Diagnostic Checks in Time Series.* Chapman and Hall.

McLeod, A.I., & Li, W.K. (1983). Diagnostic checking ARMA time series models using squared residual autocorrelations. *Journal of Time Series Analysis* 4(4), pp. 269-273.

Peters, J.-P. (2001). *Estimating and forecasting volatility of stock indices using asymmetric GARCH models and (Skewed) Student-t densities.* University of Liège.
Accessible via:
https://pdfs.semanticscholar.org/f445/493c46b3b4d52e958f69d502ccf8f03d51c3.pdf

Tsay, R.S. (2005). *Analysis of Financial Time Series.* 2nd Edition. John Wiley & Sons.