Stockholms universitet

The volatility of tomorrow - Comparison of GARCH and EGARCH models applied to Texas Instruments stock returns

Sanna Kronman

Kandidatuppsats i matematisk statistik
Bachelor Thesis in Mathematical Statistics

Kandidatuppsats 2015:14
Matematisk statistik
Juni 2015

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Mathematical Statistics
Stockholm University
Bachelor Thesis **2015:14**
http://www.math.su.se

The volatility of tomorrow - Comparison of GARCH and EGARCH models applied to Texas Instruments stock returns

Sanna Kronman*

June 2015

Abstract

In this thesis we will apply and compare two autoregressive conditional heteroscedasticity models, the GARCH$(1, 1)$ and EGARCH$(1, 1)$, to see which one is the better to use for predicting future volatility in time series data. In statistics these models are used when we encounter time series data that is heteroscedastic, i.e. has a non-constant variance. Both models use information about previous values and volatility to determine future volatility, but EGARCH$(1, 1)$ includes properties that takes in to consideration that volatility respond asymmetrically to positive and negative shocks. In this thesis we will apply both the assumption of standard normal distribution and Student's t distribution to the stock returns of Texas Instruments and then apply the models to this time series. The models and distributions will be compared by constructing confidence intervals based on predicted variances, computed AIC and MeSSIE values and Ljung-Box tests. The results will show us that the models and assumed distributions were quite alike, but after discussion we would rather assume a Student's t distribution and use a EGARCH$(1, 1)$ model to predict the volatility of tomorrow of Texas Instrument stock returns.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: sannakronman@gmail.com. Supervisor: Mathias Lindholm, Karl Rökaeus, Joanna Tyrcha.

Acknowledgement

My sincerest gratitude to my supervisors; Mathias Lindholm, Karl Rökaeus and Joanna Tyrcha, for your inspiration, dedication and feedback.

I would also like to thank my fellow students; Caroline Jernström, Andrea Klemming, Martina Sandberg and Amanda Wiman for your support and feedback.

Contents

Abstract **i**

Acknowledgement **ii**

1 Introduction **3**

2 Background **4**

3 Theory **5**

3.1 ARCH and GARCH 5

3.1.1 ARCH(q) 5

3.1.2 ARCH(1) 5

3.1.3 GARCH(p, q) 6

3.1.4 GARCH($1, 1$) 6

3.2 EARCH and EGARCH 7

3.2.1 EARCH 7

3.2.2 EGARCH(p, q) 8

3.2.3 EGARCH($1, 1$) 9

3.3 Summary of model theory 9

3.4 Estimate parameters 9

3.4.1 Maximum likelihood estimator, MLE 10

3.4.2 Normal distribution 10

3.4.3 Student's t-distribution 11

3.5 Return 12

3.6 Variance, standard deviation and volatility 12

3.7 t-test 12

3.8 ACF and PACF 12

3.9 Ljung-Box test 13

3.10 Mean sum of squares of interval error, MeSSIE 13

3.11 Akaike Information Criteria, AIC 14

4 Data **14**

4.1 Texas Instruments 14

4.2 Data analysis 14

4.2.1 Texas Instruments stock 14

4.2.2 Histogram and normal Q-Q plot 15

4.2.3 Autocorrelation 16

4.2.4 Heteroscedasticity 17

5 Modeling **18**

5.1 Model of returns 18

5.2 Forecasting 18

5.3 Confidence intervals 19

6 Result **20**

7 Discussion and conclusion **26**

8 Further research **28**

9 Appendix **29**

9.1 Appendix 1: Conditional and unconditional expected value and variance of a_t in different models 29

9.1.1 ARCH(q) and ARCH(1) 29

9.1.2 GARCH(p, q) and GARCH(1, 1) 29

9.2 Appendix 2: `rugarch` in `R` 30

References **32**

1 Introduction

The financial market is one area where it is common to see values which over time have a changing variation. Return of assets as stocks, exchange rates or changes of inflation rates have the tendency to vary differently over time. These are values which we would like to be able to model so that we can predict how they will behave in the future. If we know how much the value of an asset will vary tomorrow, we can decide if we would like to invest in it today or not. Or if a central bank knew how much the inflation rate would vary in the future, they could make changes to regulate it. But when they vary differently over time this becomes more difficult. One of the assets in which we can see that there is a difference in variation over time is in the daily stock returns of Texas Instruments.

Two models that have the ability to predict volatility is the GARCH$(1,1)$ and EGARCH$(1,1)$. These have different properties that are constructed to model reality and these are just a few of the autoregressive conditional heteroscedasticity models. Since they have different properties but both are invented to be able to predict the same thing, it is interesting to compare them so that we in the future are sure to be using the one with the best predictive ability. The aim of this thesis is to analyse which of these two models that are the best to use when we want to predict the future volatility of Texas Instruments stock returns.

This thesis will in section 3 start with a presentation of the theory that will be used to analyse and compare these models. The theory will give a short overview of how the two models are constructed. Since a few assumptions about the time series data (the stock returns) has to be fulfilled for the models to work properly, we will in section 4 continue with an analysis of the data to see if these underlying assumptions hold. In this section we will analyse which of two distributions that is suitable to assume that the returns have, the standard normal distribution or Student's t distribution.

In section 5 the modeling and forecasting of volatility will be described and results of which the comparison of the two models predictive ability will be based on is presented in section 6.

In the discussion, section 7, we will compare the results and see if they are enough to separate the models ability from each other and choose the one that is the best.

2 Background

In most statistical models we have to assume that data has constant volatility for the models to work properly. The phenomenon of constant volatility is called homoscedasticity while non-constant is called heteroscedasticity. The assumption of constant volatility is often violated in the financial market. We refer to it as clustering volatility when there are periods of high volatility and periods of low volatility in time series data and this is a common behaviour amongst returns of assets.

In 1982 Robert F. Engle introduced a model which uses the fact that data is heteroscedastic: *the ARCH model* (AutoRegressive Conditional Heteroscedasticity model). He received The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2003 for his financial analyses with the model. The ARCH model uses information about past values of data to determine future volatility.

An extension of the ARCH model was introduced by Tim Bollerslev in 1986: *the GARCH model* (Generalized AutoRegressive Conditional Heteroscedasticity model). The model does not only include information about past values to determine future volatility, but also the past conditional variances that are known.

The ARCH and GARCH model both use a shock term as one of the components to predict the future volatility, however they only consider the magnitude of the shock term. It is known that the financial market responds asymmetrically to positive and negative shocks. The variance of an asset seems to increase when the market is hit by a negative shock, while it does not increase as much when it is hit by a positive shock. This property is something that Daniel B. Nelson's extension, the EGARCH model, allows for.

The EARCH model (Exponential AutoRegressive Conditional Heteroscedasticity model) was introduced by Daniel B. Nelson in 1991. He suggests a model which predicts the natural logarithm of the conditional variance and include properties which also changes the variance due to what algebraic sign the shock term has. The question is if the asymmetrical response to shock have to be included to make a better prediction of the volatility of tomorrow.

There are many extensions of the ARCH and GARCH models, such as integrated GARCH, GJR-GARCH, asymmetric power ARCH and so on. They all have different properties and are said to handle different phenomenons within the data better than the other. In this thesis, the GARCH and EGARCH models will be applied and compared.

3 Theory

This section will present definitions of models, tests and terms used to analyse data and compare the results of applied GARCH and EGARCH models.

3.1 ARCH and GARCH

3.1.1 ARCH(q)

The ARCH model (AutoRegressive Conditional Heteroscedasticity model) was introduced by Robert F. Engle in 1982. The ARCH model should be used when time series data are heteroscedastic and with information about past values the model can help us determine future volatility, σ_t. Engle suggests the following model for the random variable a_t:

$$a_t = \sigma_t \varepsilon_t, \tag{3.1}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + ... + \alpha_q a_{t-q}^2. \tag{3.2}$$

In the model ε_t is white noise with $Var(\varepsilon_t) = 1$ and we will refer to it as the shock term. It is often assumed to be standard normal distributed or standardized Student's t distributed. The coefficients $\alpha_0 > 0$ and $\alpha_i \geq 0$ for $i = 1, ..., q$ are scalar parameters (Engle, 1982). The random variable a_t is stationary, which means that its distribution will not change over time (Tsay, 2005, pp. 105).

The ARCH(q) model suggests that the value of a_t depends upon knowledge about previous values $\{a_{t-1}, ..., a_{t-q}\}$, also called the lagged values. The parameters that are to be estimated in the model are α_0 and α_i, $i = 1, ..., q$. The latter ones will represent the weight that previous values will have on determining σ_t^2. We call this previous information Ψ_{t-1}, the information known at time $t-1$ (Engle, 1982).

For the interested reader, the conditional and unconditional expected values and variances of a_t will be derived in Appendix 1.

3.1.2 ARCH(1)

Letting the lag, q, in an ARCH model (equation 3.2) be one give us the ARCH(1) model:

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2. \tag{3.3}$$

In the model $\alpha_0 > 0$ and $\alpha_1 \geq 0$. There are now only two parameters, α_0 and α_1, that are to be estimated in the model, where α_1 represent the impact that a_{t-1}^2 have on σ_t^2. See Appendix 1 for derivations of the expected value and variance of a_t in the model ARCH(1).

3.1.3 GARCH(p, q)

The GARCH model (Generalized AutoRegressive Conditional Heteroscedasticity model) is an extension of the ARCH model. It was introduced by Tim Bollerslev in 1986. The model does not only use the lagged values to determine the conditional variance of a_t (as in the ARCH model), it also includes the lagged conditional variances to the known information. Bollerslev argues that the GARCH extension gives a more flexible lag structure and allows for a longer memory. We use the same expression and assumptions for a_t (equation 3.1), and Bollerslev's extension follows:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i a_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2, \tag{3.4}$$

$$p \geq 0,\ q > 0,$$

$$\alpha_0 > 0,\ \alpha_i \geq 0 \text{ for } i = 1, ..., q,$$

$$\beta_j \geq 0, \text{ for } i = j, ..., p,$$

see Bollerslev, 1986. The parameters α_i are called the ARCH parameters and β_j the GARCH parameters and they are both scalar and non-stochastic. They will be estimated to determine the impact that either the previous known squared values, a_{t-i}^2, $i = 1, ..., q$, or previous conditional variances, σ_{t-j}^2, $j = 1, ..., p$, will have on σ_t^2. If $p = 0$ the model reduces to an ARCH(q).

See Appendix 1 for derivations of the conditional and unconditional expected values and variances of a_t in a GARCH(p, q) model.

3.1.4 GARCH$(1, 1)$

The simplest version of a GARCH model is the GARCH(1,1). Remember that we define a_t as a $a_t = \sigma_t \varepsilon_t$, where ε_t is white noise with variance 1. The conditional variance of a_t is then expressed as

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \tag{3.5}$$

See equation 3.4 for conditions for the coefficients. As we see, this variance depends upon the squared one lag value of a and the one lag conditional variance and their effect on σ_t^2 is determined by the estimated α_0, α_1 and β_1.

See Appendix 1 for derivations of the conditional and unconditional expected values and variances of a_t in a GARCH$(1, 1)$ model.

3.2 EARCH and EGARCH

3.2.1 EARCH

EARCH, exponential ARCH, was introduced by Daniel B. Nelson in 1991. In ARCH and GARCH models we use a constraint of non-negative parameters in the expression for the conditional variance, σ_t^2, to ensure that the implied variance does not become negative. This restricts the modeling since the parameter estimates do become negative at times. This constraint is not included in the EARCH model. Instead Nelson introduces a model to estimate $\ln(\sigma_t^2)$ and thereby ensure a non-negative variance and negative parameters can be used. The EARCH model also handle a shock to the market, represented by the ε term, differently from the previously studied models. Some theories imply that the volatility of a financial asset does not have a symmetric response to a negative and positive shock. While ARCH and GARCH only use the magnitude of a shock to predict future volatility, EARCH also includes the sign of the shock.

Nelson suggests the following model, called EARCH, to define the conditional variance, σ_t^2, of $a_t = \sigma_t \varepsilon_t$. The shock term, ε_t, is once again white noise with variance equal to 1.

$$ln(\sigma_t^2) = \alpha_0 + \sum_{k=1}^{\infty} \alpha_k g(\varepsilon_{t-k}), \qquad \alpha_1 = 1. \tag{3.6}$$

In the model α_0 and α_k are scalar, non-stochastic parameters and g is a function defined as

$$g(\varepsilon_t) \equiv \theta\varepsilon_t + \gamma[|\varepsilon_t| - E|\varepsilon_t|],$$

see Nelson, 1991. The random sequence of the white noise variable, $\{g(\varepsilon_t)\}_{t=(-\infty,\infty)}$, is independent and identically distributed. The unknown parameters in EARCH are $\alpha_0, \alpha_k, \theta, \gamma$. This function gives the conditional variance, σ_t^2, the properties that are special for EARCH and EGARCH. The two terms in the g function ($\theta\varepsilon_t$ and $\gamma[|\varepsilon_t| - E|\varepsilon_t|]$) represent the effects of the algebraic sign of ε_t and a magnitude effect respectively.

We begin by studying the magnitude effect ($\gamma[|\varepsilon_t| - E|\varepsilon_t|]$) of ε_t on $\ln(\sigma_t^2)$. Assume that $\theta = 0$ and $\gamma > 0$, then

$$\text{if } |\varepsilon_t| < E|\varepsilon_t| \Longrightarrow g(\varepsilon_t) < 0,$$
$$\text{if } |\varepsilon_t| > E|\varepsilon_t| \Longrightarrow g(\varepsilon_t) > 0.$$

That is, if the magnitude of the white noise, ε_t, is smaller (larger) than expected, then $g(\varepsilon_t)$ will be less (larger) than zero and have a negative (positive) effect on $\ln(\sigma_t^2)$. The assumption of $\gamma > 0$ assumes that the conditional variance of, say an asset, will be larger when a large shock hits and smaller when a small shock hits, which often is how we assume that an asset varies due to shocks.

If we now study the "sign effect" of ε_t, we assume that $\gamma = 0$ and $\theta < 0$, then

$$\text{if } \varepsilon_t < 0 \Longrightarrow g(\varepsilon_t) > 0,$$
$$\text{if } \varepsilon_t > 0 \Longrightarrow g(\varepsilon_t) < 0.$$

That is, if there is a negative (positive) shock accuring, then $g(\varepsilon_t)$ will have a positive (negative) effect on $\ln(\sigma_t^2)$ (Nelson, 1991, 351). Now the assumption of $\theta < 0$ indicates that the volatility in a market is higher when affected of negative shocks, which also often is what is assumed in financial markets. The volatility in a market will increase more with negative shocks than with positive shocks. Therefore we would expect θ to be estimated to a negative value.

Note that $g(\varepsilon_t)$ has expected value zero since $E[\varepsilon_t] = 0$ and

$$\begin{aligned} E[g(\varepsilon_t)] &= E[\theta\varepsilon_t + \gamma[|\varepsilon_t| - E|\varepsilon_t|]] = \\ &= E[\theta\varepsilon_t] + E[\gamma[|\varepsilon_t| - E|\varepsilon_t|]] = \\ &= \theta E[\varepsilon_t] + \gamma(E|\varepsilon_t| - E|\varepsilon_t|) = 0, \end{aligned}$$

see Nelson, 1991.

3.2.2 EGARCH(p, q)

Equation 3.6 extends to a EGARCH(p, q) model where we include the logarithmic value of the lagged conditional variances, σ_{t-j}^2, of $a_{t-j} = \varepsilon_{t-j}\sigma_{t-j}$. We get

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{q} \alpha_i g(\varepsilon_{t-i}) + \sum_{j=1}^{p} \beta_j \ln(\sigma_{t-j}^2), \quad \alpha_1 = 1. \tag{3.7}$$

Inserting the expression for $g(\varepsilon_{t-i})$ in $\ln(\sigma_t^2)$ gives us

$$\ln(\sigma_t^2) = \alpha_0 + \textstyle\sum_{i=1}^{q} \alpha_i\theta\varepsilon_{t-i} + \gamma[|\varepsilon_{t-i}| - E|\varepsilon_{t-i}|] + \sum_{j=1}^{p} \beta_j \ln(\sigma_{t-j}^2),$$

and using that $\varepsilon_t = \frac{a_t}{\sigma_t}$ we get

$$\ln(\sigma_t^2) = \alpha_0 + \textstyle\sum_{i=1}^{q} \alpha_i\theta\frac{a_{t-i}}{\sigma_{t-i}} + \sum_{i=1}^{q} \alpha_i\gamma[\frac{|a_{t-i}|}{\sigma_{t-i}} - E[\frac{|a_{t-i}|}{\sigma_{t-i}}]] + \sum_{j=1}^{p} \beta_j \ln(\sigma_{t-j}^2),$$

see Tsay, 2005, page 125. The unknown parameters that are to be estimated in this model are $\alpha_0, \alpha_i, i = 2, ..., q, \beta_j, j = 1, ..., p, \theta$ and γ. The EGARCH model does not only consider the magnitude and algebraic sign of the earlier white noise variables, ε_{t-k}, but also includes the knowledge of earlier conditional variances σ^2_{t-j} and use the logarithm of these. This extension can be compared to the extension from ARCH to GARCH where previous conditional variances also was included in the extension.

3.2.3 EGARCH$(1, 1)$

The simplest model of EGARCH is the one with orders $(1, 1)$, given by

$$\ln(\sigma_t^2) = \alpha_0 + \alpha_1 g(\varepsilon_{t-1}) + \beta_1 \ln(\sigma^2_{t-1}). \tag{3.8}$$

This model can also be expressed as

$$\ln(\sigma_t^2) = \alpha_0 + \alpha_1 \theta \tfrac{a_{t-1}}{\sigma_{t-1}} + \alpha_1 \gamma [\tfrac{|a_{t-1}|}{\sigma_{t-1}} - E[\tfrac{|a_{t-1}|}{\sigma_{t-1}}]] + \beta_1 \ln(\sigma^2_{t-1}).$$

or we can take the exponential of the first expression to get it in terms of σ_t^2:

$$\sigma_t^2 = (\sigma^2_{t-1})^{\beta_1} e^{\alpha_0 + \alpha_1 g(\varepsilon_{t-1})}.$$

We then can determine the conditional variance, σ_t^2, by the one lagged conditional variance, standard deviation and one lagged value of a. The definition $\alpha_1 = 1$ leaves us with the unknown parameters α_0, θ, γ and β_1 whom will represent impacts on $\ln(\sigma_t^2)$ by the one lag a value, conditional standard deviation and logarithmic variance.

3.3 Summary of model theory

To easier see the difference between the two models that will be compared and analysed they are here presented together. Remember that $a_t = \sigma_t \varepsilon_t$.

GARCH$(1, 1)$: $\sigma_t^2 = \alpha_0 + \alpha_1 a^2_{t-1} + \beta_1 \sigma^2_{t-1}, \quad \alpha_0 > 0,\ \alpha_1 \geq 0,\ \beta_1 \geq 0.$

EGARCH$(1, 1)$: $\ln(\sigma_t^2) = \alpha_0 + \alpha_1 g(\varepsilon_{t-1}) + \beta_1 \ln(\sigma^2_{t-1})$

$$\text{where } g(\varepsilon_{t-1}) \equiv \theta \varepsilon_{t-1} + \gamma[|\varepsilon_{t-1}| - E|\varepsilon_{t-1}|].$$

3.4 Estimate parameters

As mentioned in section 3.1.1. the white noise term ε_t in the variable $a_t = \sigma_t \varepsilon_t$ is often assumed to be either standard normal distributed or standardized Student's t distributed. The density function of the conditional a_t will differ depending on which distribution we assume and thereby also the likelihood function which we use to obtain the maximum likelihood estimators of unknown parameters. In this thesis both distributions will be applied to ε_t and the results will be compared.

3.4.1 Maximum likelihood estimator, MLE

A maximum likelihood estimator is a value of a parameter that maximizes the maximum likelihood function. $\hat{\theta}_{ML}$ is an maximum likelihood estimator of θ if

$$L(\hat{\theta}_{ML}; x) = \text{argmax}_{\theta \in \Theta} L(\theta, x)$$

for all x in its sample space and Θ is the parameter space. The MLE is one of the most widely used estimators and under certain regularity conditions it is consistent, asymptotically normal and efficient (Liero & Zwanzig, 2012, 79, 116).

3.4.2 Normal distribution

If we assume that ε_t is standard normal distributed we know that the elements of a vector $\boldsymbol{a'} = (a_t, a_{t-1}, ..., a_{t-T'})$, conditionally with respect to the known information, has a marginal normal distribution with variance σ_t^2.

$$a_t | \Psi_{t-1} \sim N(0, \sigma_t^2)$$

Therefore we can estimate the α_i parameters in the conditional variance by using the likelihood function for every element in $\boldsymbol{a}$. Let $f(a_t|\boldsymbol{\alpha}, \Psi_{t-1})$ denote the conditional density function of a_t given the known information and parameters $\boldsymbol{\alpha'} = (\alpha_0, \alpha_1, ..., \alpha_i)$.

$$f(a_t|\boldsymbol{\alpha}, \Psi_{t-1}) = \frac{1}{\sqrt{2\sigma_t^2}} e^{-\frac{a_t^2}{2\sigma_t^2}}$$

We obtain the log likelihood of the t^{th} a as

$$l(\boldsymbol{\alpha}; a_t, \Psi_{t-1}) = C - \frac{1}{2}\ln(\sigma_t^2) - \frac{a_t^2}{2\sigma_t^2}$$

where C is a constant independent of $\boldsymbol{\alpha}$.

Taking the derivative with respect to α_i we can obtain the maximum likelihood estimate of the parameter when setting the derivative to zero and solve for α_i (Engle, 1982, 990).

The log likelihood function will look differently depending upon which model we assume, GARCH or EGARCH:

$$\textbf{GARCH}(1,1)\textbf{:}\ \boldsymbol{\alpha'} = (\alpha_0, \alpha_1, \beta_1),$$

$$l(a_t|\overline{\alpha}, \Psi_{t-1}) = C - \tfrac{1}{2}\ln(\alpha_0+\alpha_1 a_{t-1}^2+\beta_1\sigma_{t-1}^2) - a_t^2/(2(\alpha_0+\alpha_1 a_{t-1}^2+\beta_1\sigma_{t-1}^2)),$$

$$\textbf{EGARCH}(1,1)\textbf{:}\ \boldsymbol{\alpha'} = (\alpha_0, \alpha_1, \theta, \gamma, \beta_1),$$

$$l(a_t|\overline{\alpha}, \Psi_{t-1}) = C - \tfrac{1}{2}(\alpha_0 + \alpha_1 g(\varepsilon_{t-1}) + \beta_1 \ln(\sigma_{t-1}^2)) -$$

$$a_t^2/(2(exp\{\alpha_0 + \alpha_1 g(\varepsilon_{t-1}) + \beta_1 \ln(\sigma_{t-1}^2)\})).$$

3.4.3 Student's t-distribution

The Student's t distribution can have the property of larger tails than a normal distribution have, depending upon which degree of freedom that is used. Stock returns often have larger tails than a normal distribution and the Student's t distribution is therefore one of the most commonly assumed distributions i GARCH and EGARCH modeling. If we assume that ε_t has a standardized Student's t distribution, then our a_t gets the following conditional density function

$$f(a_t|\boldsymbol{\alpha}, \Psi_{t-1}) = \tfrac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{(v-2)\pi}} \tfrac{1}{\sigma_t}(1 + \tfrac{a_t^2}{(v-2)\sigma_t^2})^{-(v+1)/2}, \qquad v > 2,$$

where v are the degrees of freedom and the gamma function $\Gamma(x)$ is defined as $\Gamma(x) = \int_0^\infty y^{x-1}e^{-y}dy$ (Tsay, 2005, 108).

As for the standard normal distribution we can obtain the maximum likelihood estimates in the same way, but now with a different log likelihood function. We can also estimate the degrees of freedom, v, with the maximum likelihood function and differentiate with respect to v. The log likelihood function will be

$$l(\boldsymbol{\alpha}; a_t, \Psi_{t-1}) = C - \ln(\sigma_t) - \tfrac{v+1}{2}\ln(1 + \tfrac{a_t^2}{(v-2)\sigma_t^2}), \qquad v > 2,$$

where C is a constant which will vanish when we differentiate with respect to any of the parameters we want to estimate. The log likelihood will be the following in the GARCH and EGARCH models.

$$\textbf{GARCH}(1,1)\textbf{:}\ \boldsymbol{\alpha'} = (\alpha_0, \alpha_1, \beta_1)$$

$$l(\boldsymbol{\alpha}; a_t, \Psi_{t-1}) = C - \ln(\sqrt{\alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1\sigma_{t-1}^2}) -$$

$$\tfrac{v+1}{2}\ln(1 + \tfrac{a_t^2}{(v-2)(\alpha_0+\alpha_1 a_{t-1}^2+\beta_1\sigma_{t-1}^2)})$$

$$\textbf{EGARCH}(1,1)\textbf{:}\ \boldsymbol{\alpha'} = (\alpha_0, \alpha_1, \theta, \gamma, \beta_1)$$

$$l(\boldsymbol{\alpha}; a_t, \Psi_{t-1}) = C - \ln(\sqrt{exp\{\alpha_0 + \alpha_1 g(\varepsilon_{t-1}) + \beta_1 \ln(\sigma_{t-1}^2)\}}) -$$

$$\tfrac{v+1}{2}\ln(1 + \tfrac{a_t^2}{(v-2)(exp\{\alpha_0+\alpha_1 g(\varepsilon_{t-1})+\beta_1\ln(\sigma_{t-1}^2)\})})$$

3.5 Return

The arithmetic return of an asset at time t is defined as

$$r_t = \frac{S_t - S_{t-1}}{S_{t-1}}$$

where S_t is the value of the asset at time t and S_{t-1} the value of the asset at time $t-1$, one time unit before t. In this thesis the time units will be in days.

3.6 Variance, standard deviation and volatility

The sample variance of $X = (x_1, x_2, ..., x_n)$ is defined as

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

where n is the number of observations in the sample and $\bar{x}$ is the sample mean. We call $\sigma = \sqrt{\sigma^2}$ the standard deviation of a sample.

Volatility is the standard deviation of a time series (Capinski & Zastawniak, 2011, 200).

3.7 t-test

Assuming that $X = (x_1, x_2, ..., x_n)$ is a sample from a normal distribution with expected value μ and standard deviation σ we can test the hypothesis $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$ with the test statistic

$$T_\mu = \frac{\overline{X} - \mu}{s(X)/\sqrt{n}} \sim t(n-1)$$

which is Student's t-distributed with $(n-1)$ degrees of freedom. The sample standard deviation is $s(X)$. We reject the null hypothesis if $|T_\mu| > t_{\alpha/2}(n-1)$, where $t_{\alpha/2}(n-1)$ is the $\alpha/2$ t-distributed quantile (Alm & Britton, 2008, 328).

3.8 ACF and PACF

ACF (AutoCorrelation Function) and PACF (Partial AutoCorrelation Function) can be used to study the correlation between values in a time series.

We start by defining the estimated k^{th} lag autocovariance, c_k

$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} (r_t - \bar{r})(r_{t+k} - \bar{r}) \qquad k = 0, 1, 2, ..., K$$

where N is the number of observations in the time series r and $\bar{r}$ is the sample mean of the time series. Then the estimated autocorrelation function is defined as

$$ACF_k = \hat{\rho}_k = \frac{c_k}{c_0}$$

This is the estimated autocorrelation of lag k. If the process r_t is stationary, this is a function only depending on the lag, k (Box *et al.* , 2008, 31).

The $PACF_k$ is a measure of autocorrelation between r_t and r_{t-k} when we have adjusted for the effects that the time series values between r_t and r_{t-k} have, i.e. the values $\{r_{t-1}, r_{r-2}, ..., r_{t-k+1}\}$ (Box *et al.* , 2008, 68).

3.9 Ljung-Box test

The Ljung-Box test tests the hypothesis that autocorrelations are zero.

$$H_0 : \rho_1 = \rho_2 = ... = \rho_m = 0$$

The Ljung-Box test statistic is defined as

$$Q(\rho) = N(N+2) \sum_{k=1}^{m} \frac{\widehat{\rho}_k^2}{N-k} \sim \mathcal{X}_m^2$$

where N is the number of observations in the time series and $\hat{\rho}_k$ is the estimated autocorrelation. This statistic is $\mathcal{X}^2$ distributed with m degrees of freedom (Ljung & Box, 1978).

3.10 Mean sum of squares of interval error, MeSSIE

Sometimes it can be hard to visually see which of several confidence intervals that are the most fitting to the data if you don't plot them in the same figure. When confidence intervals will be analysed in this thesis, we will measure the length from the value of a return to the percentiles in that time. To do this, we here propose the measure Mean Sum of Squares of Interval Error, MeSSIE, which is defined as

$$\text{MeSSIE}(\alpha) = \frac{1}{N} \sum_{i=1}^{N} [(r_i - q_{\alpha/2})^2 + (r_i - q_{1-\alpha/2})^2]$$

where r_i are the returns and q_α is the $100 * \alpha\%$-percentile when $1 - \alpha$ is the confidence level of the interval. N is the number of returns the interval is covering. The MeSSIE value will always be larger or equal to zero. The smaller the value, the closer the values of the confidence interval will be to the return values. The lowest MeSSIE value should not alone be regarded as the best fitted confidence interval, since this does not say anything about if the confidence level is obtained or not.

3.11 Akaike Information Criteria, AIC

The Akaike Information Criterion (AIC) is a measurement of goodness of fit of a model. AIC is defined as

$$\text{AIC} = -2(\text{maximized log likelihood} - \text{number of parameters in model}).$$

When compared, the model with the best fit is the one who has the lowest AIC (Agresti, 2012, 212).

4 Data

This section will describe and analyse the properties of Texas Instrument stock returns which GARCH$(1,1)$ and EGARCH$(1,1)$ will be applied to.

4.1 Texas Instruments

Texas Instruments is an American company which for more than 80 years have developed new innovations in technology. They are today located in 35 countries world wide and their production of chips help costumers in industrial, energy, media and medical industries to improve their products and equipments. Texas Instruments also provide software and development tools for different markets (Texas Instruments, 2009). One of their most famous products are their calculators which was developed in the 1990's.

Texas Instruments have been traded in The New York Stock Exchange since 1953 to 2011 and is listed in NASDAQ since 2012.

The data analyzied in this thesis is the closing price of the Texas Instruments Incorporated stock (TXN), collected from NASDAQ Historical Quote. The data is from 10 years, covering the time 2005/01/26 - 2015/01/23. This is 2516 observed closing prices. This will result in 2515 return values r_t, $t = 2, 3, ...2516$.

4.2 Data analysis

4.2.1 Texas Instruments stock

All the 2515 r_t values will not be used to estimate GARCH$(1,1)$ and EGARCH$(1,1)$ models at the same time. The first half of values, which will cover the first five years (2005/01/26 to 2010/01/25), will be used to estimate the first models (see section 5.2 Forecasting for more about this procedure). This first half will be analysed with histograms, normal QQ-plots and tests to see whether it seems that GARCH and EGARCH models are a proper models to estimate future conditional volatility with.

Figure 1 show plots of the closing price and calculated returns of the period 2005/01/26 - 2010/01/25. The plot of returns indicate that there may be heteroscedasticity, its variance is not constant. It also illustrates what is called clustering volatility, there are periods of higher volatility and lower volatility. The volatility is especially higher in the end of year 2008, which could be due to the financial crisis in 2008, and then seems to decrease through year 2009.

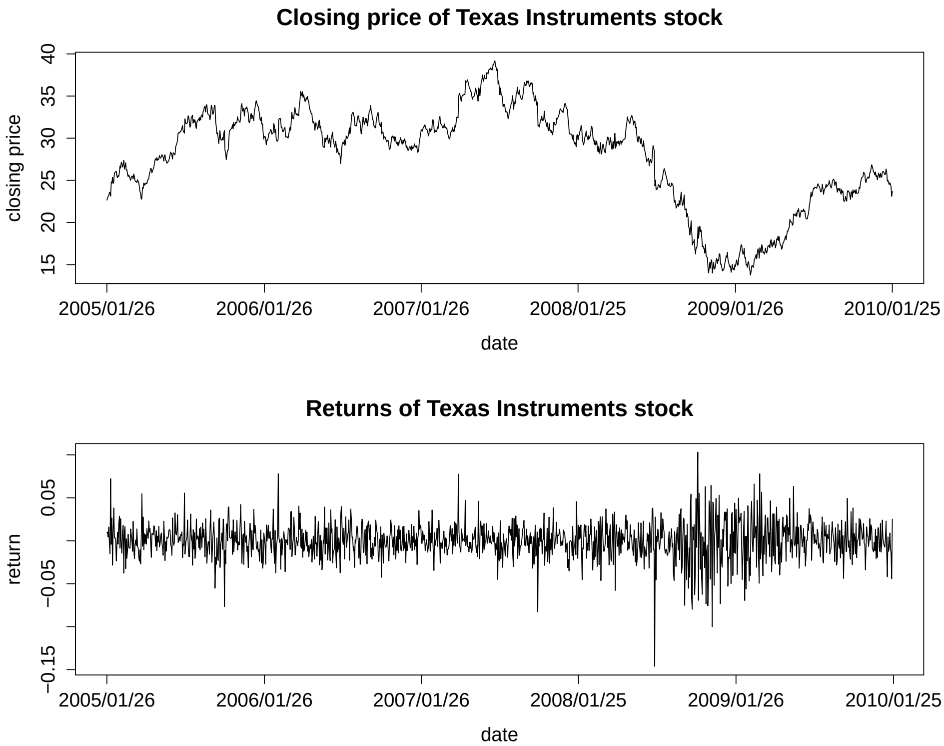

Figure 1: Closing price and returns of Texas Instruments stock

4.2.2 Histogram and normal Q-Q plot

Figure 2 is a histogram of the return probabilities and an estimated density function (the dashed line) and a fitted density function of a normal distribution with mean and standard deviation computed from the returns. The histogram and density function seems to almost follow a normal distribution. But the normal Q-Q plot in the same figure deviates some from the line and indicates that the returns may not follow a normal distribution.

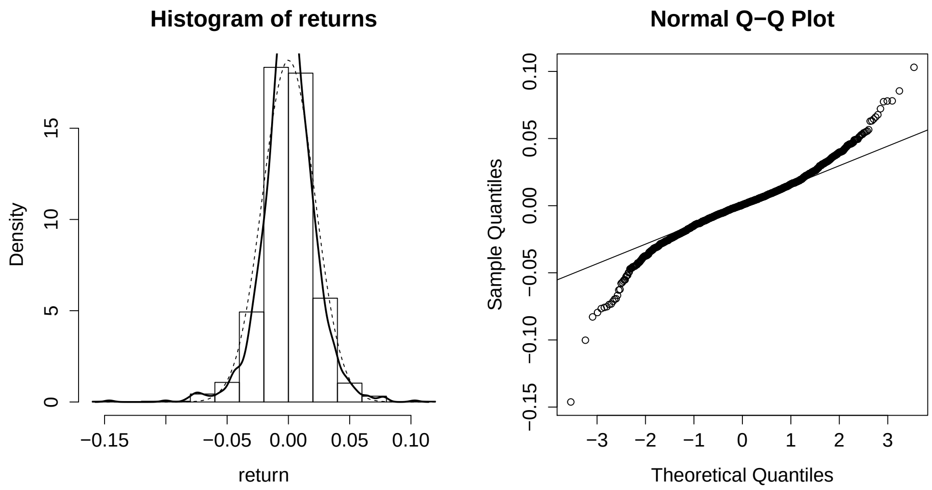

Figure 2: Histogram of returns and normal Q-Q plot. The solid line in the histogram is an estimated kernel density function of the returns and the dashed line is the density function of a fitted normal distribution with the mean and standard deviation of the returns.

4.2.3 Autocorrelation

To study if there occurs autocorrelation in the returns we compute the ACF and PACF, seen in Figure 3. For the autoregressive conditional heteroscedasticity models to work properly we want to apply the models to data that is serially uncorrelated, but still dependent (Tsay, 2005, 99). Unfortunately we can see in the ACF of returns that there seems to be several significant correlations in the return data. The occurrence of autocorrelations can be a problem when constructing GARCH and EGARCH models and for the forecasting of volatility. The high values of PACF of squared returns tell us that the return series seems to be serially dependent.

We also test if the autocorrelations can be assumed to be zero with the Ljung-Box test. When we test the hypothesis that the autocorrelations for 10 lags are zero ($m = 10$ in the test) we obtain the value of the test statistic to 34.24. The statistic is $\mathcal{X}^2_{10}$ distributed and we get a p-value of 0.0002. This means that we reject the null hypothesis of zero autocorrelations on a 95% level, which then again validates what was seen in the ACF figure. There seems to be serial autocorrelation in the returns.

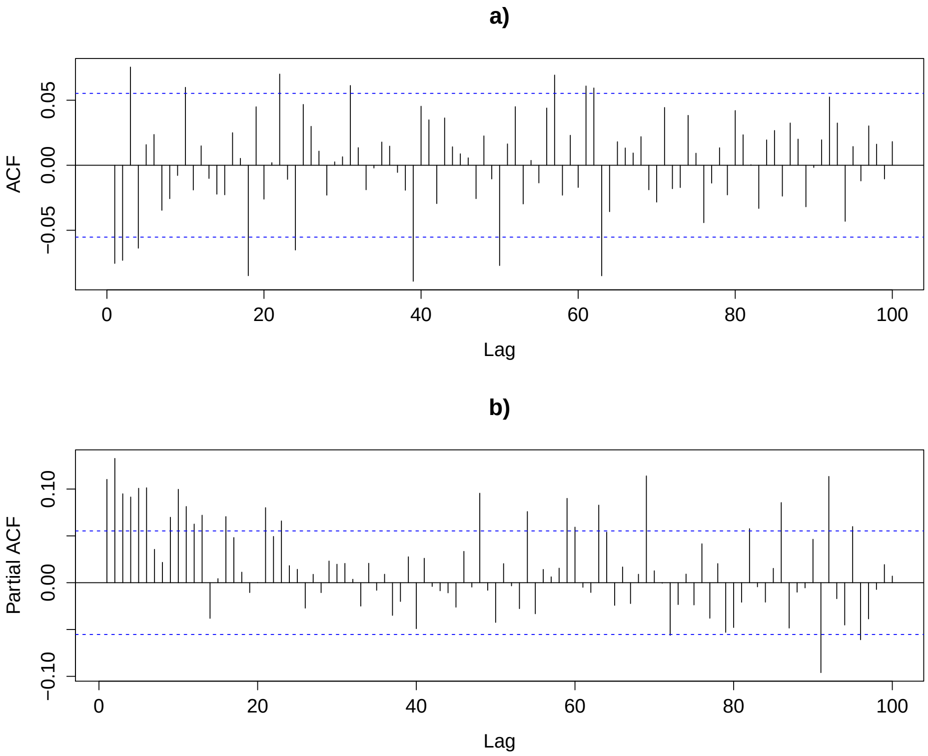

Figure 3: Autocorrelation function (a) for returns and partial autocorrelation function (b) for squared returns with 100 lags.

4.2.4 Heteroscedasticity

To test for heteroscedasticity in the returns we may again use the Ljung-Box test. As we will assume that the returns are equal to the stationary process which will be modelled with GARCH and EGARCH, $r_t = a_t$ (se section 5.1), we may test if there are any autocorrelation between squared a_t's, or equivalently if there are any autocorrelations between the squared r_t's. If autocorrelations are present, it means that we can assume heteroscedasticity in the data.

We get the Ljung Box test statistic when we use the lag 10 to 437.75, which is very high and the p-value is less than 2.2e-16. We may reject the null hypothesis of zero autocorrelations on a 95% level and assume that the data is heteroscedastic. This was what we thought was seen in figure 1 and is now stated.

5 Modeling

This section will present some assumptions, steps and tools that are used in this thesis to compare the result when we apply GARCH$(1,1)$ and EGARCH$(1,1)$ to Texas Instruments stock returns. All the modeling and calculations will be done in `R` (R Core Team, 2015).

5.1 Model of returns

To model returns, r_t, we assume a simple time series model

$$r_t = \mu_t + a_t = \mu + \sigma_t \varepsilon_t \tag{5.1}$$

where $\mu_t = E[r_t|\Psi_{t-1}]$. Returns are a sum of expected value and the random variable a_t. The expected value, μ_t, can be expressed in several ways, for example it could be constant or be an ARMA(p,q) process (this can be read more about in chapter 3 in Tsay, 2005). To see if we can simplify our model we will test the assumption that the expected value is zero.

This is tested with a t-test. We assume that the returns are normally distributed and the null hypothesis, $H_0 : \mu = 0$ is tested against $H_a : \mu \neq 0$ with the first half of our returns (period 2005/01/26 to 2010/01/25) as the observed returns. We get the test statistic to 0.44 and the p-value to 0.6612, which means we can not reject the null hypothesis on a 95% level.

We assume that the expected value of the returns are zero and therefore reduce the model of returns to

$$r_t = a_t = \sigma_t \varepsilon_t. \tag{5.2}$$

5.2 Forecasting

To investigate which of the GARCH and EGARCH models that are most suitable to use for forecasting future volatility we will construct several of these models. We will always predict one day ahead and use observations from five years back to estimate the models, approximately 1257 observations. The models that will be estimated are GARCH$(1,1)$ and EGARCH$(1,1)$, which will use the one lagged value r_t and the one lagged conditional variance σ_t^2 to forecast σ_{t+1}^2.

The first forecast will be done to predict σ_{1259}^2, which is the conditional variance of the return at date 2010/01/26. This will be predicted with models estimated with observations 2005/01/26-2010/01/25. When this is done we will predict σ_{1260}^2, the conditional variance of the return at date 2010/01/27, with models which are estimated with observations from

2005/01/27-2010/01/26, and so on. This will be done 1258 times over and give us 1258 predicted conditional variances covering the time 2010/01/26-2015/01/23. We call this process a “moving window”, since we always will be using the same amount of observations backwards to estimate the models and predict one day ahead. Since we have the real historical values of returns for that five year time period, we will have the opportunity to compare and analyse the predicted values to those of reality to see which model is most suitable.

The “moving window” procedure, which cover 1257 observations at a time, will only move one day a head for every new model. This means that only one observation will be dropped and one new will enter the windows. Since 1257 observations are quite many for just predicting the volatility one day a head and the difference between successive windows will be so small, this might indicate that the parameter estimations will not change as much for all the models. This might be something to consider if further analyses of the parameters are to be done.

Remember that we have only analysed the distribution, mean, autocorrelation and heteroscedasticity of the returns from the first five years, which will be the ones used to estimate the first models. The process of analysing all the 1258 “windows” of each 1257 observations will be to extensive and therefore we assume that all the windows have approximatley the same distribution, mean and properties of autocorrelation and heteroscedasticity as the first window.

In totalt there will be four groups of predicted variances, one for every combination of GARCH$(1,1)$ or EGARCH$(1,1)$ and assumed standard normal distribution or Student’s t distribution.

The predicted variances can then be used to construct confidence intervals by calculating percentiles for every return. One indication of a good model is one where the confidence interval will cover the real returns closely. If it does not, it means that the predicted variances were misleading.

To do this we will use a package in `R` called `rugarch`. Details of how this package work can be seen in Appendix 2.

5.3 Confidence intervals

We will construct 95% confidence intervals based on the estimated standard deviations from the four models with combinations of GARCH/EGARCH and standard normal distribution/Student’s t distribution. We use percentiles of the standardized distributions and multiply them with the estimated standard deviations to thereby get a percentile for a distribution with that standard deviation. The procedure follows.

If we have a stochastic variable, Z, with $E[Z] = \mu$ and $Var(Z) = \theta^2$ we get the $100 * \alpha\%$-percentile, z_α, of that distribution from

$$P(Z \leq z_\alpha) = \alpha.$$

To standardize this distribution and get a stochastic variable, $\tilde{Z}$, with $E[\tilde{Z}] = 0$ and $Var(\tilde{Z}) = 1$ we subtract the mean and divide by the standard deviation and get the standardized percentile, $\tilde{z}_\alpha$.

$$P(\tfrac{Z-\mu}{\theta} \leq \tfrac{z_\alpha-\mu}{\theta}) = P(\tilde{Z} \leq \tilde{z}_\alpha) = \alpha.$$

We want to find the percentiles from the distributions that have the estimated standard deviation which were calculated with our models, $\hat{\sigma}$, and therefore multiply the percentiles with it.

$$P(\tilde{Z} \leq \tilde{z}_\alpha) = P(\tilde{Z}\hat{\sigma} \leq \tilde{z}_\alpha\hat{\sigma}) = P(Q \leq q_\alpha) = \alpha.$$

Now the stochastic variable Q has $E[Q] = 0$ and $Var(Q) = \hat{\sigma}^2$, as we wanted. The percentile q_α is the ones which are shown in the plots of the confidence intervals (see Result).

6 Result

In Figure 4 to 7 we can see the returns of Texas Instruments stocks for the period 2010/01/26 to 2015/01/23 and the blue dashed lines are the estimated confidence intervals for the four models. They all seem to cover the returns pretty well and grasp areas of clustering volatility, however none of them catches the high and low peaks, see especially the beginning of 2012 and end of 2014. The GARCH$(1,1)$ models give quite similar results for standard normal distribution (Figure 4) and Student's t distribution (Figure 5) and the same holds for the two distributions when we have used EGARCH$(1,1)$ models (Figure 6 and 7). However, the confidence intervals based on EGARCH$(1,1)$ seems to be narrower in the end of 2010 as well as in the end of 2011. The EGARCH$(1,1)$ models also give us confidence intervals that are a bit smoother from the end if 2012 to the beginning of 2013 than what the confidence intervals from GARCH$(1,1)$ are.

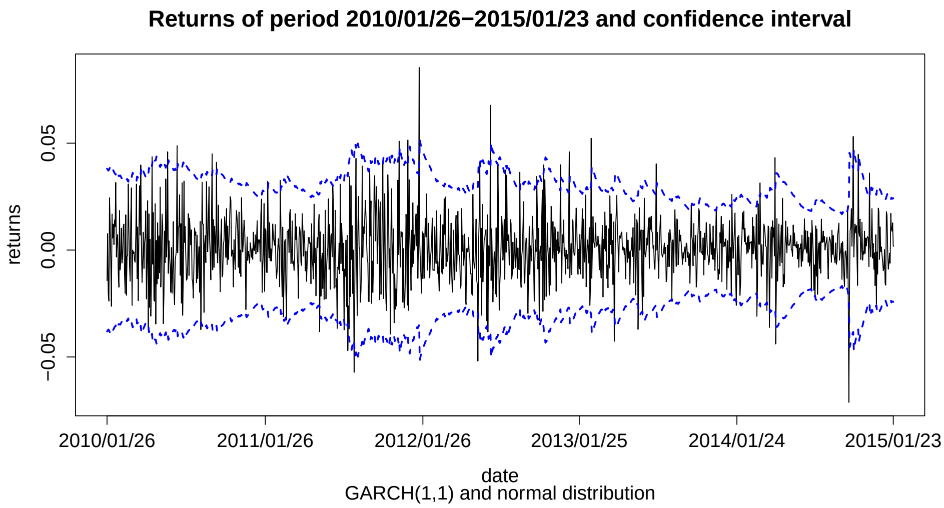

Figure 4: Return of period 2010/01/26-2015/01/23 and GARCH$(1,1)$ forecasted 95% confidence intervals from a normal distribution

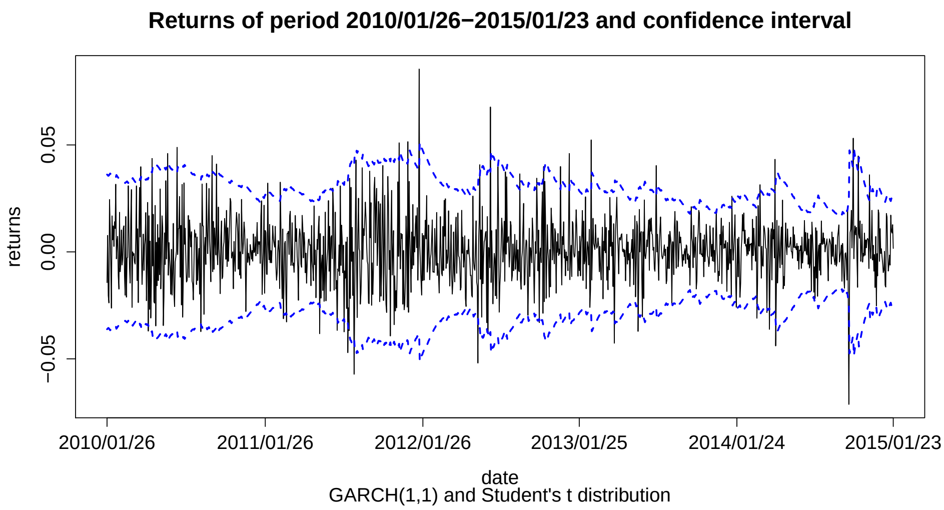

Figure 5: Return of period 2010/01/26-2015/01/23 and GARCH$(1,1)$ forecasted 95% confidence intervals from a Student's t distribution

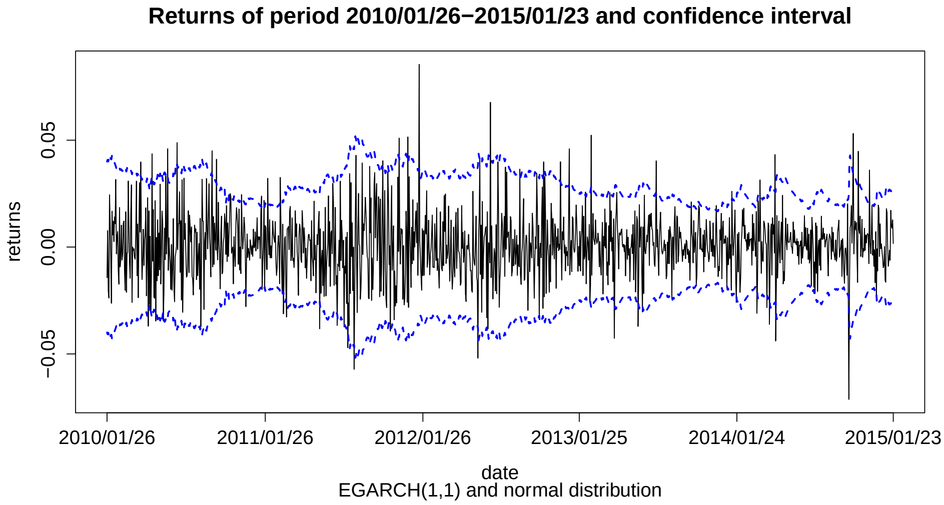

Figure 6: Return of period 2010/01/26-2015/01/23 and EGARCH(1, 1) forecasted 95% confidence intervals from a normal distribution

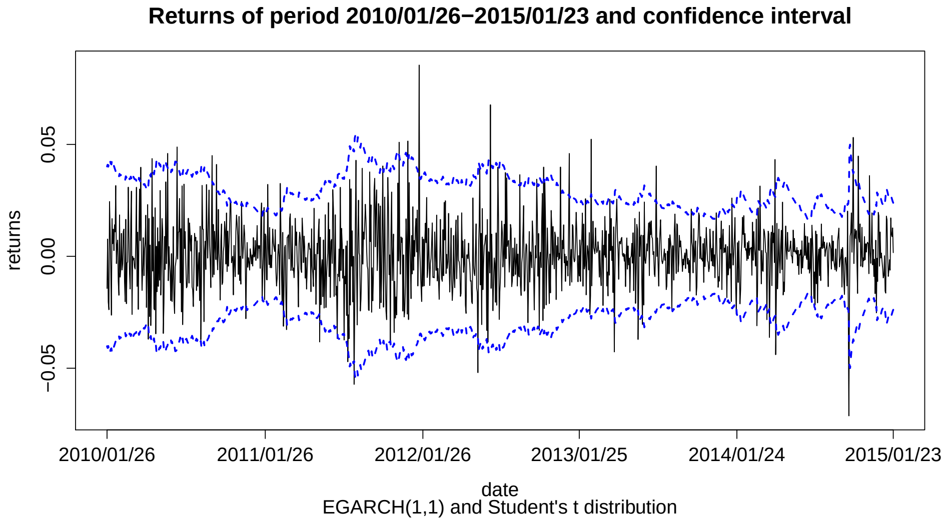

Figure 7: Return of period 2010/01/26-2015/01/23 and EGARCH(1, 1) forecasted confidence intervals from a Student's t distribution

In Table 1 we can se calculated per cent of returns that violate the estimated confidence intervals and MeSSIE for all the models. The confidence intervals that are estimated should be of the level 95%, which means that we assume that 5% will not be within the interval. All the models have approximately that amount of returns outside their intervals, although EGARCH(1, 1) and Student's t distribution are closest to the wanted level and both GARCH$(1, 1)$ models have almost a too low level of returns outside their confidence intervals.

The MeSSIE values tell us how narrow or wide the fit of the confidence interval is. An indication of a good confidence interval should have a low MeSSIE value, since this means that the confidence interval is close to the real values. The model with the lowest MeSSIE value is the EGARCH$(1, 1)$ with normal distribution, while the GARCH$(1, 1)$ models have the same MeSSIE values for both distributions.

Table 1: Per cent of returns that violates the estimated confidence intervals and MeSSIE(0.05) values

Model	**Distribution**	**Per cent of returns outside the CI**	**MeSSIE**(0.05)
GARCH$(1, 1)$	Normal	0.045	0.0025
GARCH$(1, 1)$	Student's t	0.043	0.0025
EGARCH$(1, 1)$	Normal	0.053	0.0023
EGARCH$(1, 1)$	Student's t	0.051	0.0024

In Figure 8 and 9 we see all p-values for Ljung-box tests of standardized residuals (small dotted line) and squared standardized residuals (dashed line). With the standardized residuals we are testing the null hypothesis that there are zero autocorrelation in the standardized residuals serie, which should not be rejected (the small dotted line should not be less than 0.05) if the autoregressive heteroscedasticity models have done their job. When we test the squared standardized residuals we test the null hypothesis of homoscedasticity, which we wish to have for there to be good models (the dashed line should not be less than 0.05) (Tsay, 2005, 119).

As we can see in Figure 8 and 9 both the null hypothesis of zero autocorrelations and homoscedasticity in the standardized residuals is not rejected in the 95% level in most of the models, especially the first models of GARCH$(1, 1)$ (Figure 8). But we can also se that the null hypothesis of zero autocorrelations will be rejected for all the models from the beginning of 2013 to the end of 2014. The test of homoscedasticity will never be rejected for either of the models.

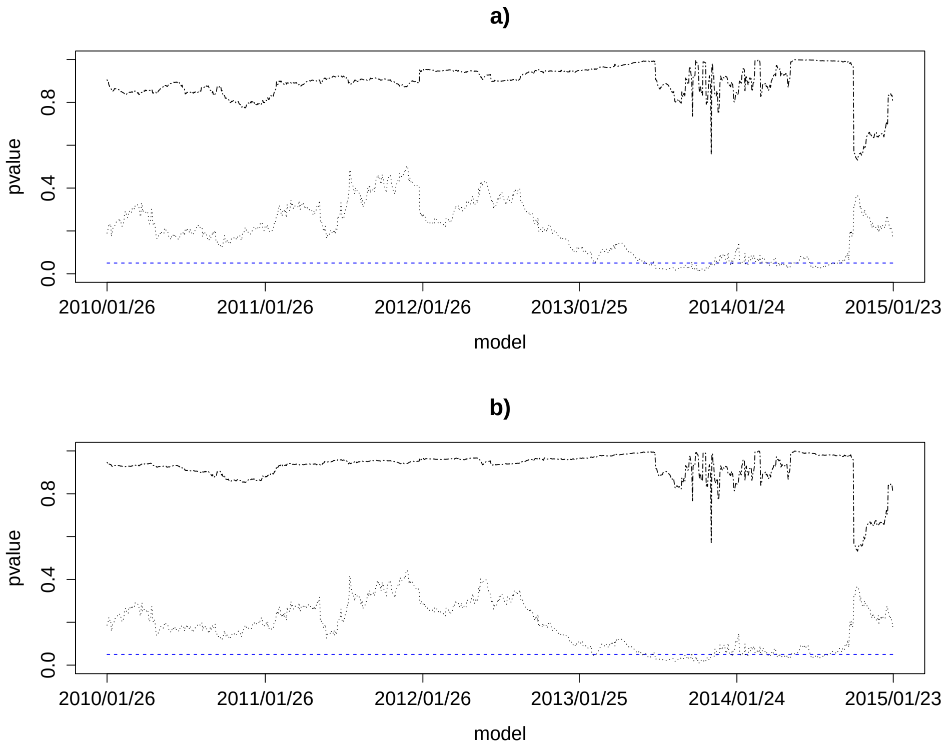

Figure 8: Ljung-Box test of standardized residuals (small dotted line) and squared standardized residuals (dashed line). a) is GARCH$(1, 1)$ and normal distribution, b) is GARCH$(1, 1)$ and Student's t distribution. The straight line marks 0.05.

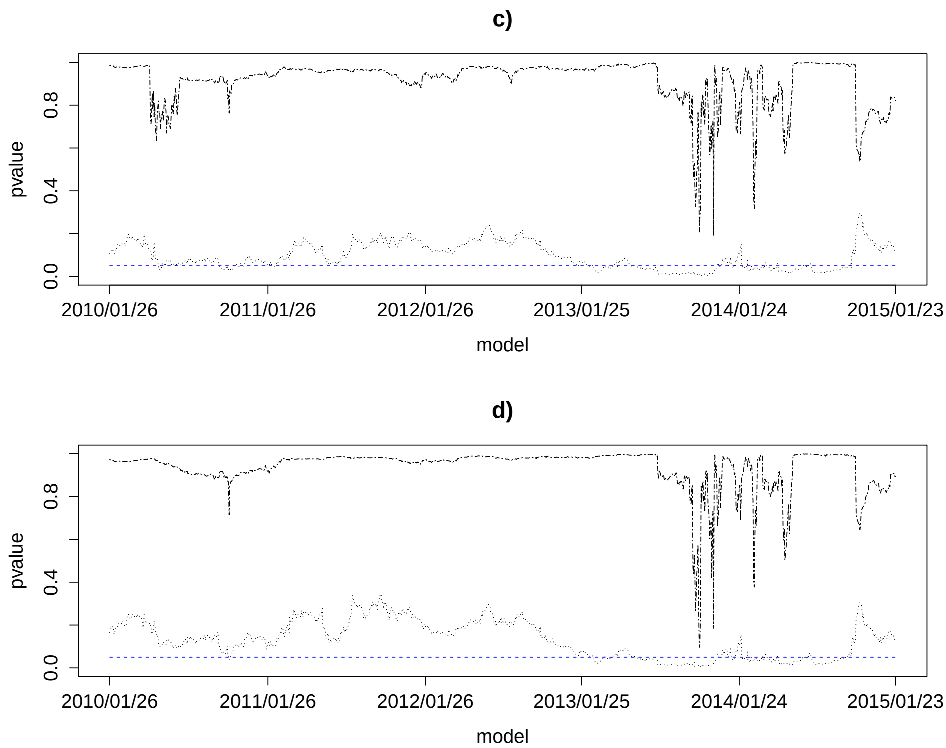

Figure 9: Ljung-Box test of standardized residuals (small dotted line) and squared standardized residuals (dashed line). c) is EGARCH$(1,1)$ and normal distribution, d) is EGARCH$(1,1)$ and Student's t distribution. The straight line marks 0.05.

In Figure 10 we can see calculated AIC for all the models. For the first models, GARCH$(1,1)$ and Student's t distribution seems to have the smallest AIC, while both EGARCH$(1,1)$ models after the beginning of 2013 then have the smallest AIC (there AIC are so alike that their curves are approximately the same all the time, that's why we see the bold line). We can calculate the average AIC for all models, which are -5.16 (GARCH$(1,1)$, standard normal distribution), -5.2 (GARCH$(1,1)$, Student's t distribution), -5.19 (EGARCH$(1,1)$ and standard normal distribution) and -5.22 (EGARCH$(1,1)$, Student's t distribution). We here see that EGARCH$(1,1)$ and Student's t distribution has the lowest avarge AIC for this period.

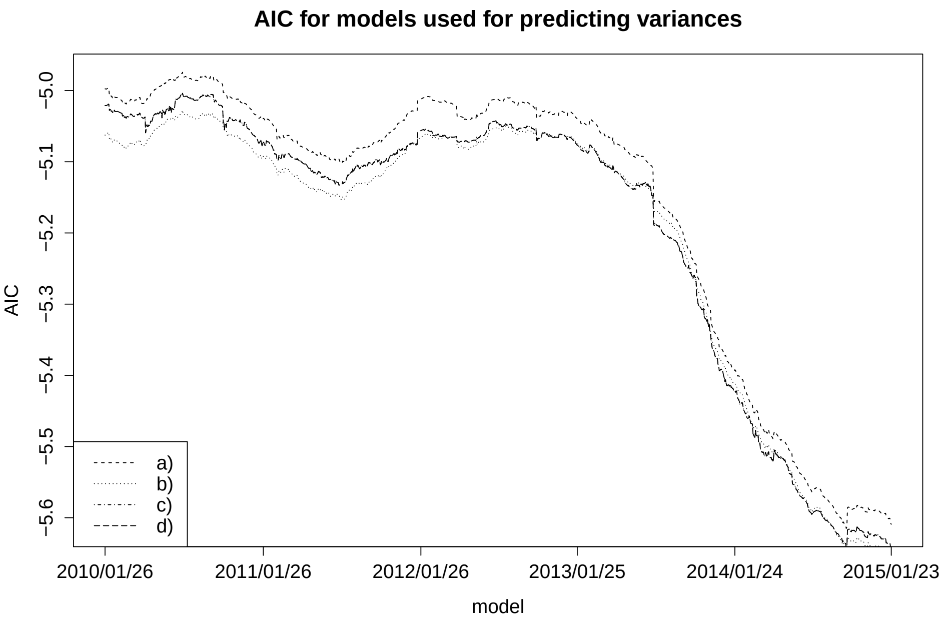

Figure 10: AIC for all models. a) is GARCH$(1, 1)$ and normal distribution, b) is GARCH$(1, 1)$ and Student's t distribution, c) is EGARCH$(1, 1)$ and normal distribution, d) is EGARCH$(1, 1)$ and Student's t distribution. c) and d) is the bold line, since they are almost exactly the same for all models.

7 Discussion and conclusion

The purpose of this thesis was to analys which of the models GARCH$(1, 1)$ and EGARCH$(1, 1)$ that are most suitable to use for predicting volatility of Texas Instruments stock returns. Since we have to assume a distribution for the returns and could not determine which one was the most suitable just by analysing the data, we applied the standard normal distribution and the Student's t distribution to the two models. This gave us four models to compare when it comes to predictive ability.

We constructed confidence intervals based on the predicted volatility given from 1258 models for every one of the four models (seen in Figure 4 to 7). The visual analysis showed us that the difference in distributions among GARCH$(1, 1)$ was not that big and this was also confirmed when per cent of returns outside the confidence interval and MeSSIE was calculated in Table 1, they are almost exactly the same. The similarity is also confirmed in the Ljung-Box tests of autocorrelation and homoscedasticity in the models (Figure 8). But when it comes to AIC, the GARCH$(1, 1)$ with assumed Student's t distribution has the lower AIC values in the time period observed.

But this is the only advantage shown by assuming Student's t distribution in the GARCH$(1,1)$ model. When we in section 4 of this thesis studied which distribution was appropriate to assume that the returns had, it seemed that assuming normal distribution was not the best choice. But the little difference in all these results also seem to tell us that assumed Student's t distribution did not give us a big advantage.

When it comes to the difference due to the distributions in the EGARCH$(1,1)$ models, neither they are as extensive. The narrower confidence interval is the one where we have assumed normal distribution, but this can hardly be seen visually when comparing Figure 6 and 7, and is only noticed when calculating the MeSSIE values. But the narrow confidence interval of assumed normal distribution also have a higher level of returns that violates the confidence interval. The confidence intervals are calculated to have a 95% confidence level, but in the case of EGARCH$(1,1)$ and normal distribution we have a bit smaller confidence level than that. This means that the predictions of volatility calculated by these models are a bit to small at times, which can make us miss clusters of higher volatility. We can see some differences in the p-values of the Ljung-box tests in Figure 9. We can especially see that there seems to be autocorrelation in the standardized residuals for a period of 2010 when we assume normal distribution, while this does not occur when we assume Student's t distribution. The AIC of the EGARCH$(1,1)$ models are so alike that we can barely see a difference. So the two distributions both had a disadvantage of the other, one seems to under estimate the volatility at times making the confidence interval narrower while the other one could not produce models in which there were no autocorrelation.

If we compare GARCH$(1,1)$ and EGARCH$(1,1)$ over all we see that EGARCH$(1,1)$ in both cases had the narrower confidence interval, however the confidence level was violated more times than what was wanted. The EGARCH$(1,1)$ model with assumed Student's t distribution was the closest to the wanted confidence level and the GARCH$(1,1)$ models have a bit to high levels. When the standardized residuals are tested for zero autocorrelation and homoscedasticity there seem to be several cases where these properties are not rejected in the GARCH$(1,1)$ models, while the EGARCH$(1,1)$ models do not only violate these properties from the period 2013 to end of 2014 (which all models does) but also some times in the time of 2010. If we look at the AIC's, the EGARCH$(1,1)$ models have the smallest average AIC if we only compare the ones with the same distributions (if we compare a) and c) and compare b) and d) in Figure 10) and therefore may be the better model to use.

Something that we may consider is the fact that the returns was not serially uncorrelated, which should not have been the case if the model estimation should work properly. It could be due to this that the Ljung-Box tests of the properties of standardized residuals was rejected at some times. Also, the analysis of data was only made to the first five years of data. The "moving window" procedure make us use all the returns, but at different times, in the five year windows. To simplify our analysis, we assumed that the properties of the first five years can be assumed for all the windows. This may not be the case and several assumption may not hold.

There is something that is happening to the estimated models in the beginning of 2013 to the end of 2014. During this period we can not in any of the models assume that there are zero autocorrelation in the standardized residuals of the models, but at the same time AIC is decreasing, telling us there is a better fit in the models than before.

There is no model and distribution that has a big advantage to another in this study. Although, the QQ-plot in our data analysis show us that a normal distribution is not appropriate to assume and this upholds me to say that the models with this assumption are the best. If we then compare the GARCH$(1, 1)$ and EGARCH$(1, 1)$ with Student's t distribution, the EGARCH$(1, 1)$'s accuracy of the confidence level is to its advantage. The GARCH$(1, 1)$ underestimated the volatility, and although the hypothesis of zero autocorrelation in the Ljung-Box tests seems to have gotten the result we wanted more times for the GARCH$(1, 1)$ the difference is small when compared to EGARCH$(1, 1)$. Therefore, the better choice of model and distribution for predicting the volatility of Texas Instruments stock returns are a EGARCH$(1, 1)$ model with assumed Student's t distribution.

8 Further research

As mentioned in the discussion, there is something that is happening to the models when we predict the volatility from the beginning of 2013 to the end of 2014. It would be interesting to look further into this and see what it was that made the models behave as they did. Since it occurs in all the models, a guess would be that it has something to do with the data. What happened there in the Texas Instruments stock returns?

It would also be interesting to have studied the parameters in the models further, since this could have given us some information of how much either previous conditional variances or previous values of the returns effected the future values. Which of these two values had the large impact on the volatility of tomorrow?

9 Appendix

9.1 Appendix 1: Conditional and unconditional expected value and variance of a_t in different models

9.1.1 ARCH(q) and ARCH(1)

The shock term, ε_t has expected value zero. Thereby, both the unconditional expected value of a_t and the conditional expected values will be zero since

$$E[a_t] = E[E[a_t|\Psi_{t-1}]] = E[E[\sigma_t\varepsilon_t]] = E[\sigma_t E[\varepsilon_t]] = 0.$$

The conditional variance of a_t is σ_t^2 since

$$Var[a_t|\Psi_{t-1}] = Var(\sigma_t\varepsilon_t) = \sigma_t^2 Var(\varepsilon_t) = \sigma_t^2.$$

These results hold for both ARCH(q) and ARCH(1).

If we assume that a_t is a stationary process, which means that $Var(a_t) = Var(a_{t-1})$, then the unconditional variance for a ARCH(q) model is

$$Var(a_t) = E[a_t^2] = E[E[a_t^2|\Psi_{t-i}]] = E[\sigma_t^2 E[\varepsilon_t^2]] = E[\sigma_t^2] =$$
$$= E[\alpha_0 + \alpha_1 a_{t-1}^2 + ... + \alpha_q a_{t-p}^2] = \alpha_0 + \alpha_1 E[a_{t-1}^2] + ... + \alpha_q E[a_{t-q}^2] =$$
$$= \alpha_0 + \alpha_1 Var(a_{t-1}) + ... + \alpha_p Var(a_{t-q}) = [\text{stationarity}] =$$
$$= \alpha_0 + \alpha_1 Var(a_t) + ... + \alpha_q Var(a_t) \Longrightarrow$$
$$\Longrightarrow Var(a_t) = \frac{\alpha_0}{1-(\sum_{i=1}^{q}\alpha_i)}.$$

If the unconditional variance should be finite, some regularity conditions have to be fulfilled for α_i (Tsay, 2005, pp. 103).

For a ARCH(1) model the conditional variance of a_t only depends upon the squared value of a one time unit earlier. The unconditional variance becomes

$$Var(a_t) = \alpha_0 + \alpha_1 Var[a_t] \Longrightarrow Var(a_t) = \frac{\alpha_0}{1-\alpha_1}.$$

9.1.2 GARCH(p,q) and GARCH(1,1)

The unconditional and conditional expected value of a_t in a GARCH(p,q) and GARCH(1,1) model will (as in ARCH) be zero since

$$E[a_t] = E[E[a_t|\Psi_{t-1}]] = E[E[\sigma_t\varepsilon_t]] = E[\sigma_t E[\varepsilon_t]] = 0.$$

And the conditional variance is

$$Var[a_t|\Psi_{t-1}] = Var(\varepsilon_t\sigma_t) = \sigma_t^2 Var(\varepsilon_t) = \sigma_t^2.$$

The expression for the unconditional variance of a_t in a GARCH(p,q) can be derived as

$$\begin{gathered}
Var(a_t) = E[a_t^2] = E[E[a_t^2|\Psi_{t-i}]] = E[\sigma_t^2 E[\varepsilon_t^2]] = E[\sigma_t^2] = \\
= E[\alpha_0 + \textstyle\sum_{i=1}^{q} \alpha_i a_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2] = \\
= \alpha_0 + \textstyle\sum_{i=1}^{q} \alpha_i E[a_{t-i}^2] + \sum_{j=1}^{p} \beta_j E[\sigma_{t-j}^2] = \\
= \alpha_0 + \textstyle\sum_{i=1}^{q} \alpha_i Var(a_{t-i}) + \sum_{j=1}^{p} \beta_j E[\sigma_{t-j}^2] = [\text{stationarity of } a_t] \\
= \alpha_0 + \textstyle\sum_{i=1}^{q} \alpha_i Var(a_t) + \sum_{j=1}^{p} \beta_j E[\sigma_{t-j}^2] \Longrightarrow \\
\Longrightarrow Var(a_t) = \frac{\alpha_0 + \sum_{j=1}^{p} \beta_j E[\sigma_{t-j}^2]}{1 - (\sum_{i=1}^{q} \alpha_i)}.
\end{gathered}$$

For a GARCH$(1,1)$ the unconditional variance reduces to

$$Var(a_t) = \frac{\alpha_0 + \beta_1 E[\sigma_{t-1}^2]}{1 - \alpha_1}.$$

9.2 Appendix 2: `rugarch` in `R`

The `rugarch` package (Ghalanos, 2014b) can help us model several of the GARCH models that is used in statistics and econometrics. The package properties are extensive and here only the ones used for this thesis will be presented. For more information and descriptions of the commands you can find reference manuals at
`http://cran.r-project.org/web/packages/rugarch` and Alexio Ghalanos have written a helpful introduction to the package (Ghalanos, 2014a).

To construct a model we first have to specify which model we want. This is done with `ugarchspec`.

```
ugarchspec <- ugarchspec(variance.model = list(model = "sGARCH",
garchOrder = c(1, 1)), mean.model = list(include.mean = FALSE),
distribution.model = "norm")
```

Here we have specifed which GARCH model we want, “sGarch”, which give us the standard GARCH. If we wanted the EGARCH we would have used “eGARCH”. Then we specify the orders of the model and also if we want to include a mean in the model, which we do not in this thesis. We also specify which distribution we assume for the process, where we will use the standard normal distribution, “norm”, and Student’s t distribution, “std”.

When the model is specified we can fit a model to the data we have. This is done with `ugarchfit`. The estimations of parameters in the specified model will be maximum likelihood estimates.

```
modelfit <- ugarchfit(spec=garchspec, data=return)
```

`garchspec` is the specified model and `return` are the returns that are used to estimate the model. Since we will use the "moving window" approach to estimate all the models, these returns will be different for every model.

When this is done we want to forecast the volatility one day ahead, and therefore use `ugarchforecast`.

```
forecast <- ugarchforecast(spec=garchspec, data = return, n.ahead
= 1, n.roll = 0, out.sample =0)
```

With this command we then use the specified model and the data for which the model was constructed to forecast the volatility one day ahead (Hybel Pedersen, 2013).

References

Agresti, Alan. 2012. *Categorical Data Analysis*. 3 edn. Hoboken, New Jersey: John Wiley and Sons, Inc.

Alm, Sven Erick, & Britton, Tom. 2008. *Stokastik, Sannolikhetsteori och statistikteori med tillampningar*. 1 edn. Stockholm: Liber.

Box, George E. P. Box, Jenkins, Gwilym M., & Reinsel, Gregory C. 2008. *Time Series Analysis, Forcasting and Control*. 4 edn. Hoboken, New Jersey: John Wiley and Sons, Inc.

Capinski, Marek, & Zastawniak, Tomasz. 2011. *Mathematics for Finance: An Introduction to Financial Engineering*. 2 edn. Springer-Verlag London.

Engle, Robert F. 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, **50**(4), 987–1007.

Ghalanos, Alexios. 2014a. *Introduction to the rugarch package. (Version 1.3-1)*. 2015-02-20. CRAN. http://cran.r-project.org/web/packages/rugarch/vignettes/Introduction_to_the_rugarch_package.pdf.

Ghalanos, Alexios. 2014b. *rugarch: Univariate GARCH models*. R package version 1.3-4.

Hybel Pedersen, Jesper. 2013. *ARMA(1,1)-GARCH(1,1). Estimation and forecast using rugarch 1.2-2*. 2015-02-20. unstarched. http://unstarched.net/wp-content/uploads/2013/06/an-example-in-rugarch.pdf.

Liero, Hannelore, & Zwanzig, Silvelyn. 2012. *Introduction to the theory of Statistical Inference*. Boca Raton: Taylor & Francis Group.

Ljung, G. M., & Box, G. E. P. 1978. On a measure of lack of fit in time series models. *Biometrika*, **65**(2), 297–303.

Nelson, Daniel B. 1991. Conditional Heteroskedasticity in Aseet Returns: A New Approach. *Econometrica*, **59**(2), 347–370.

R Core Team. 2015. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Texas Instruments. 2009. *Company, Our Product and technology*. 2015-02-13. http://www.ti.com/corp/docs/company/tibus.shtml.

Tsay, Ruey S. 2005. *Analysis of Financial Time Series*. 2 edn. Hoboken, New Jersey: John Wiley and Sons, Inc.