Stockholms universitet

Forecasting the financial volatility of LVMH, Louis Vuitton Moët Hennessy

Pierre Serti

Kandidatuppsats i matematisk statistik
Bachelor Thesis in Mathematical Statistics

Kandidatuppsats 2013:7
Matematisk statistik
September 2013

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Mathematical Statistics
Stockholm University
Bachelor Thesis **2013:7**
http://www.math.su.se

Forecasting the financial volatility of LVMH, Louis Vuitton Moët Hennessy

Pierre Serti*

September 2013

Abstract

This paper examines the performance abilities of forecasting and modeling the financial volatility of the daily stock returns of LVMUY, belonging to Louis Vuitton Moët Hennessy. The volatility is modeled by the GARCH model and further extended into the EGARCH and the GJR-GARCH models under two distributional assumptions for the error term, i.e Gaussian distribution and Student-t distribution. We compare the performance in forecasting the volatility among the GARCH family models by calculating the AIC and the Loglikelihood for each model. The results shows that the EGARCH model with the Student-t distribution is the best model in forecasting the financial volatility. Despite the limitations of the GARCH model in comparison to the EGARCH model, the symmetric GARCH model stil provide adequate results in forecasting purposes.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail:pierre.serti@hotmail.com . Supervisor: Joanna Tyrcha.

Acknowledgements

The extent of this paper is 15 ECTS, which leads to a Bachelor's degree in Mathematical Statistics in the field of Financial Mathematics at the Department of Mathematics at Stockholm University.

I would like to thank and express my genuine gratitude to my brilliant supervisor, Joanna Tyrcha for her excellent guidance and support. I deeply appreciate Joanna's valuable suggestions and inspiring discussions.
I would also like to thank God and my family for giving me strength and faith in everything I do.

Contents

1 Introduction

Since the ability of predicting the clustering volatility has been a controversial issue for several years, many studies and methods have investigated the relationship between expected returns and volatility, producing unsuccessfully results. The dynamic volatility of an asset return seems to be very important and inspiring due to the recent financial crisis and turbulent periods. As a impact the ARCH(autoregressive conditional heteroscedasticity) model has been questioned and several different models have been developed.
Even though the volatility of asset returns is predictable and time varying, the ability of forecasting and predict the future level of volatility is difficult for several reasons. Understanding the way that the stock market volatility changes is vital to our understanding of several areas, both in financial theory and macroeconomic theory.

This paper is partly inspired by Robert Engle's contribution to the financial theory and the economic science, but in particular by Tim Peter Bollerslev's ideas and development of the generalized ARCH model, for measuring and forecasting financial market volatility.
The family of ARCH was introduced by Engle(1982) and have since then been commonly used to model time-varying volatility and the persistence of shocks to volatility.
In application, other methods in modeling and forecasting financial volatility have been presented. We will investigaste these models in their behaviour in forecasting financial volatility. These are, the GARCH model(generalized autoregressive conditional heteroscedasticity) introduced and proposed by Bollerslev(1986), the EGARCH model(exponential generalized autoregressive conditional heteroscedasticity) first introduced by Nelson(1991) and eventually the GJR-GARCH model, presented by Glosten, Jagannathan and Runkle(1993).
The problem of volatility clustering, however now can be predictable due in part to use the methodology of GARCH.

2 Background

A world leader in luxury, LVMH Moët Hennessy - Louis Vuitton have a unique portfolio of over 60 luxury and prestigious brands. The LVMH group is mostly active in following sectors[23] :

- Fashion and Leather Goods
- Wines and Spirits
- Perfumes and Cosmetics
- Watches and Jewelry
- Selective retailing

In this thesis we have chosen to examine the models of dynamic volatilities within the financial market, the closing daily price belonging to LVMUY, Louis Vuitton Moët Hennessy within the time frame of nine years and also investigate its dynamic properties. The reason why we choosed this period is because, within the above time frame this period covers different conditions on the market, both turbulent and calm conditions. By covering different conditions we get a good prediction and hence to be sufficient for our modelestimation.

Further, the different properties of the returns time series and model time varying volatility for our financial data will also be presented. This analysis is done by testing the returns distribution and dependency within time.

3 Aim of the thesis

The aim of this paper is that we will first examine the different properties of the returns time series, using the daily stock return for LVMUY and model time varying volatility for the sample data. This analysis is done by testing the returns distribution and dependency within time. Further, we will examine and hence, separately utilize the models of dynamic volatilities within the time frame of 2268 trading days and investigate its dynamic properties by using the GARCH, EGARCH and GJR-GARCH models. We will compare the performance among these GARCH extended models in order to conclude the best model in estimating and forecasting the financial volatility for the data by using two different density functions for the error term.

4 Theoretical framework

In this section we will describe the theory for the different parts of our analysis whereas this will be used in our reasoning and modeling later on. The data that we will investigate further is the daily closing price of a stock belonging to LVMH. Since the data is collected for each trading day, it is presented as time series and hence, shows how much the variance for the stock chart has changed for each and every day.

4.1 Returns and Volatility

Return on stocks is a financial ratio thats illustrates the percentage of profit a company earns in relation to their total and overall resources.
We start by defining the return of stock at time t as $R_t, t = 1, 2, 3, ..., n$

$$R_t = \frac{I_t - I_{t-1}}{I_{t-1}} \approx log(1 + \frac{I_t - I_{t-1}}{I_{t-1}}) = log(\frac{I_t}{I_{t-1}}) \tag{1}$$

where I_t is the value of the stock rate at time t and n is the number of observations.
Since it is observed in a time series of stock prices, whereas the variance of returns is both high for extended periods and then low for extended periods, i.e the variance of daily returns can be high one day and then low the next day. This property of time series of prices is called volatility clustering and we will further plot the daily return of the stock in order to see if the variance changes over time but also for the reason that it gives us the visual inspection to show that volatility changes over time and thus, it tends to cluster with periods with low volatility and periods with high volatility.

In the financial world, fluctuations in any return rate is called volatility. The volatility refers to the amount of uncertainty or more likely the size of changes in a security's value. A high volatility means that the price of security can change very dramatically over a short time period in both directions.
In Höglund T.[10], volatility for a predefined time period of one day is estimated according the formula

$$\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(R_t - \hat{\mu})^2} \tag{2}$$

where $\hat{\mu}$ is the mean value of the returns. The estimate above holds due to the assumption that volatility of returns is equal to the volatility of growth as a consequence of the relation between return and growth.

4.2 Theoretically analysis of data

If we take in consideration the Black-Scholes model in Höglund T. [10] the historical return are assumed to be presented by time series with a constant defined expexted value, μ and finite variance according to the essential theory of the financial mathematics. As part of this theory, the volatility is predictable and does not change with time.
It is assumed to be approximately normally distrubuted if the time period is sufficiently large, R_t is assumed to be a purely random process with independent increments.

Further we will investigate whether assumptions of independent increments and assumptions of the normal distribution seems to be plausible for our data. We will investigate this by using several well known tests particularly from the courses in Econometrics and Linear statistical models.
The classical linear regression model consists of seven different assumptions[3]. We will inspect two of them in depth in order to analyze the distribution of the sample data of returns, namely the assumption of No Autocorrelation and the assumption of Homoscedasticity.

4.2.1 Autocorrelation test

Autocorrelation is most common and often found in time-series data in which we would like to interpret the number of the highlighted observation, indicating the time at which the observation was made, see Andersson P. and Tyrcha J. [3]. Thus, the autocorrelation function is the measure of dependency of observations separated by defined time period of different lags.

We define the process R_t according

$$R_t = \mu + \epsilon_t \tag{3}$$

where μ is the expected value of the process R_t and ϵ_t is a independent random variable with expected value, zero and the variance σ^2.

Heteroscedasticity affects the elements on the diagonal of the covariance matrix $\sum$, but the disturbances, ϵ_j are still assumed to have zero pairwise covariances. When the disturbances are autocorrelated, the covariance matrix has the following form

$$\sum = \begin{pmatrix} \sigma^2 & \rho_{12} & . & . & \rho_{n1} \\ \rho_{12} & \sigma^2 & . & . & \rho_{n2} \\ . & . & . & . & . \\ . & . & . & . & . \\ \rho_{1n} & \rho_{n2} & . & . & \sigma^2 \end{pmatrix}$$

.

The test of Ljung and Box assesses the null hypothesis that a series of residuals exhibits no autocorrelation for a fixed number of lags, N, against the alternative hypothesis that some autocorrelation coefficient $\rho(k), k = 1, ..., N$, is nonzero.
The test statistic can be calculated according the following formula,

$$Q = T(T+2) \sum_{k=1}^{N} \Big(\frac{\hat{\rho(k)}^2}{T-k} \Big) \tag{4}$$

where T is the sample size, N is the number of autocorrelation lags, and $\hat{\rho(k)}$ is the sample autocorrelation at lag k.
Under the null hypothesis, Q is asymptotically chi-square distributed with N degrees of freedom [24].

4.2.2 Heteroscedasticity test

Recall the classical linear regression model, it is assumed that the variance of each and everyone disturbance term ϵ_j, conditional on the chosen values of the explanatory variable is constant and hence equal to σ^2. It is also assumed that the regression model is linear in the parameters, independently of the variables. If the disturbance terms have constant variance, the disturbance terms are called homoscedastic and thus, if the they do not have the same variance, they are called heteroscedastic. For the following linear model

$$Y = X\beta + \epsilon \tag{5}$$

where $Y = (Y_1, Y_2, .., Y_n)^T$, $\beta = (\beta_1, \beta_2, .., \beta_m)^T$ and X is an n x m matrix, when the disturbance terms, $\epsilon = (\epsilon_1, \epsilon_2, .., \epsilon_n)^T$ have the same variance and are uncorrelated then the ordinary least square estimator (OLS) is the best linear unbiased estimator. If this is the case then the error terms have the same variance of σ^2 of the following form

$$E[\epsilon\epsilon^T|X] = \sigma^2 I_n$$

where $E(\epsilon) = 0$.

In case of heteroscedasticity we thus get the matrix,

$$E[\epsilon\epsilon^T|X] = \begin{pmatrix} \sigma_1^2 & 0 & . & . & 0 \\ 0 & \sigma_2^2 & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & \sigma_n^2 \end{pmatrix}$$

4.2.3 Skewness and kurtosis

A further and an important characterization of our data and for the most financial data is to investigate skewness and kurtosis.
Kurtosis is a measure of whether the financial data is flat or peaked in relation to a Normal distribution. For instance, if our dataset have a high value of kurtosis then it tend to have a distinct peak near the mean, decline rather rapidly and have heavy tails. Thus, a dataset with a low kurtosis tend to have a flat top near the mean rather than a sharp peak [22]. The kurtosis for a Normal distribution is 3. Hence, values approximately close to this value should be considered as a Normal distribution.

On the other hand, skewness is a measure of the lack of a symmetry. In order to identify a dataset as symmetric we recognize it, if it looks exactly the same to the right and left of the centerpoint.

The skewness for a Normal distribution is zero and it is highly desirable that any symmetric data should have a skewness close to zero.
However, a negative value for the skewness corresponds and indicate that data is skewed left which means that the left tail is long in relation to the right tail. Thus, positive values for the skewness means that data is skewed right. Similarly for right skewness, it means that the right tail is long relative to the left tail.

4.2.4 Jarque-Bera test

The Jarque-Bera test is a goodness of fit measure of departure from Normal distribution, based on sample kurtosis and skewness. For instance, the statistic can be used to test the hypothesis that the data is from a Normal distribution. Thus, the null hypothesis

is a joint hypothesis of the skewness being zero and a kurtosis of 3. The statistic has a chi-squared distribution, $\chi^2(2)$ with 2 degrees of freedom, one for skewness and the second for kurtosis. Since we have a large sample size, we can consider this test as valid asymptotically for our financial data. The higher value of Jarque-Bera test represent the non-normality of the series data [11].
The test statistic is

$$JB = \frac{n}{6}\left(s^2 + \frac{(k-3)^2}{4}\right)$$

where n is the sample size, s is the sample skewness and respectively k is the sample kurtosis.

4.2.5 AIC

The Akaike information criterion, AIC is used to measure the relative goodness of fit of a model, in comparison to other models when both the complexity and precision of the models are taken into account. Hence, it judges a model by how close the fitted values tend to be in relation to the true values in terms of a certain expected value[2]. The criterion penalizes a model for having many parameters. We will use AIC in order to conclude which model has the best performance to forecast the financial volatility among the GARCH family models.

The Akaike information criterion is defined as follows,

$$AIC = 2k - 2ln(L)$$

where k is the number of parameters, and L is the maximum of the likelihood function for the estimated model [9]. Over the three different of GARCH models with different AIC-values, we prefer the one with the smallest AIC-value.

5 Methodology framework

Very often, returns of the most financial stocks represented as time series data are for the most of the time characterized by calm periods and turbulent periods, i.e periods of significant fluctuations. Thus, ARCH and GARCH family models are designed to deal with this issues. We will start by introducing the ARCH model in order to let you be familiar with the first and most basic model and later on proceed to three extended GARCH models, i.e GARCH(1,1), EGACRH(1,1) and GJR-GARCH(1,1) model.

The autoregressive conditional heteroscedastic (ARCH) model was first introduced by Engel, R.F [14] and it is also the first model of conditional heteroscedasticity. The ARCH model, models the conditional error as a function of the values taken by past squared error and hence, takes advantage of the commonly found presence of autocorrelations in these. Engel proposed that the conditional error variance should be conditional over the values taken by previous errors. We assume that ϵ_t is a random variable which has a mean and a variance conditionally on the realized values of the set of variables, i.e the information set $\xi_{t-1} = (y_{t-1}, x_{t-1}, y_{t-2}, x_{t-2}, ...)$ [4]. The random variable ϵ_t of the ARCH model has the following properties,

$$E(\epsilon_t|\xi_{t-1}) = 0$$

The second property is that the conditional variance,

$$\sigma_t^2 = E(\epsilon_t^2|\xi_{t-1})$$

is a positive valued parametric function of ξ_{t-1}. Further, we get that

$$\epsilon_t = x_t - \mu_t(x_t)$$

where x_t is the observed value and $\mu_t(x_t) = E(x_t|\xi_{t-1})$ is the conditional mean of x_t given ξ_{t-1}. We can thus, express the equation of ϵ_t according,

$$\epsilon_t = z_t\sigma_t$$

where z_t is a sequence of independent indentically distributed random variables with zero mean and unit variance, i.e $z_t \quad iid \sim N(0,1)$. The ARCH model of order q is then expressed by

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + ... + \alpha_q \epsilon_{t-q}^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2$$

where $\alpha_0 > 0$ and $\alpha_i \geq 0, \quad i > 0$, α_0 and α_i are coefficients, σ_t^2 is the one-step ahead forecast of the conditional variance and q is the number of the most recent squared errors included in the calculus of the conditional variance. The formula for the ARCH(q) model raises the question of what optimal size of q are given a certain financial data sample of at time series. Due to this question, there is some limitations and problems with the ARCH model. First of all, there is not any clear approach to determine the optimal size of q given certain data sample. Further, that q might need to be fairly large in order to fully capture the dynamics in the variance and since having a pretty large q, we face the problem of having many parameters to estimate in the regression and increased risk of violating the non-negativity constraints[21].

Due to these limitations and weaknesses, the ARCH(q) model are seldom used in practice today. However, the ARCH model has provided a framework for modeling the conditional variance for lags of squarred errors, which the other widely used ARCH models has been developed on. We will extend this thesis with several of these models.

With that said, it is of high importance to not forget that no matter how precisely and accurate forecasts of the future variance that ARCH class models produces, in the end they do not fully explain the true cause of the volatility[15].
In summary this is the ARCH model.

6 GARCH

Due to the limitations and drawbacks for the ARCH model, it has been extended and substituted in applications by the generalized ARCH model, GARCH first introduced by Bollerslev[19].
The GARCH(p,q) model is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + ... + \alpha_q\epsilon_{t-q}^2 + \beta_1\sigma_{t-1}^2 + ... + \beta_p\sigma_{t-p}^2 = \alpha_0 + \sum_{i=1}^{q}\alpha_i\epsilon_{t-i}^2 + \sum_{j=1}^{p}\beta_j\sigma_{t-j}^2$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$ for $i = 1,..,q$ and $\beta_j \geq 0$ for $j = 1,..,p$ and where q is the order of the ARCH terms ϵ^2 and p is the order of the GARCH terms σ^2. A sufficient condition for the conditional variance to be positive with probability one is $\alpha_0 > 0$, $\alpha_i \geq 0 \quad j = 1,...q$ and $\beta_j \geq 0, \quad i = 1,...p$. In this paper we only use the most basic and the most common GARCH model, i.e GARCH(1,1). The (1,1) is a standard notation in which the first number in the parentheses tells us how many ARCH terms appear in the equation, while the second number in the parentheses refers to how many moving average lags that is specified, which is often called the number of GARCH terms.
Hence, the GARCH(1,1) model is then given by

$$\sigma_t^2 = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2$$

where the model is a one period former estimation for the variance taken on any past information thought relevant [5].

There are some limitations in the GARCH(1,1) model. The non-negative conditions limits the estimate method because the coefficients of the model could be negative and the basic GARCH model is symmetric and does not capture the asymmetry, i.e it cannot account for leverage effect.

The GARCH model treats the influence which comes from positive and negative informations in a series equally, but still it is not reasonable in several cases. Take for instance, the negative information of stock price, it always has pronounced effect on fluctuation than the positive information. Thus, the symmetric GARCH model does not capture this kind of asymmetry performance.

Also, it is difficult to achieve all the parameters are assumed greater than zero in GARCH models. In order to solve these kind of problems, the GARCH model has been improved further.
For measuring the negative impact of leverage effect in the volatility models, Nelson(1991) proposed the EGARCH model, i.e the exponential GARCH model and Glosten, Jagannathan and Runken (1993) proposed the GJR-GARCH model. For these purposes,

we will introduce the EGARCH model and the GJR-GARCH model in following sections.

7 EGARCH

The natural logarithm of the conditional variance is assumed to be a linear function of the lagged term and allows to vary over time. Nelson[6] defined the EGARCH(p,q) model as follows :

$$log\sigma_t^2 = c + \sum_{i=1}^{p} g(Z_{t-i}) + \sum_{j=1}^{q} \beta_j log\sigma_{t-j}^2$$

where we can simplify the term $g(Z_{t-i})$ to ,

$$g(Z_t) = \gamma Z_t + \alpha(|Z_t| - E(|Z_t|))$$

where σ_t^2 is the conditional variance and c, γ, α, β are parameters to be estimated. Since the $log\sigma_t^2$ is modeled, in contrast to the GARCH model no restrictions need to be imposed on our EGARCH model estimation. This is due to that the logarithmic transformation ensures us that the forecasts of the variance is non-negative. Hence, the significant advantage of using the EGARCH model is that even though the parameters are negative, σ_t^2 will be positive.

We define $Z_{t-i} = \dfrac{\epsilon_{t-i}}{\sigma_{t-i}}$ and thus, the function can be expressed as:

$$log(\sigma_t^2) = c + \sum_{j=1}^{q} \beta_j log\sigma_{t-j}^2 + \sum_{i=1}^{p} \alpha_i \Big(|\frac{\epsilon_{t-i}}{\sigma_{t-i}}| - E(|\frac{\epsilon_{t-i}}{\sigma_{t-i}}|) \Big) + \sum_{i=1}^{p} \gamma_i \frac{\epsilon_{t-i}}{\sigma_{t-i}}$$

The term $E(|Z_{t-i}|)$ varies for different distributions for Z_{t-i}, according :

$$E(|Z_{t-i}|) = \begin{cases} \sqrt{\frac{\pi}{2}} & \text{if } Z_{t-i} \text{ is Normal distribution} \\ \frac{\sqrt{\nu}\Gamma[0.5(\nu-1)]}{\sqrt{\pi}\Gamma(0.5\nu)]} & \text{if } Z_{t-i} \text{ is Student - t distribution} \end{cases}$$

The α parameter represents a magnitude effect, i.e the symmetric effect of the model, similarly role as the ARCH effect. The β parameter measures conditional variance, the same role as the GARCH model. For instance, if β is pretty large, then volatility takes a long time to die out following a crisis in the market. The γ parameter measures the asymmetry, i.e the leverage effect. The leverage effect could be defined as, the tendency for volatility to rise more following a large price fall than following a price rise of the same magnitude. Thus, the parameter γ is an outstanding and highly important extension from the GARCH model to the EGARCH model.
When $\gamma = 0$ it means that the model does not exist any asymmetric. When $\gamma < 0$, then good news, i.e positive shocks generate less volatility than bad news, i.e negative shocks. Thus, when $\gamma > 0$ it implies that positive information is more significant than negative information.

In this thesis we use the EGARCH(1,1) model which is simplified according following,

$$log(\sigma_t^2) = c + \beta_1 log\sigma_{t-1}^2 + \alpha_1\Big(|\frac{\epsilon_{t-1}}{\sigma_{t-1}}| - E(|\frac{\epsilon_{t-1}}{\sigma_{t-1}}|)\Big) + \gamma_1\frac{\epsilon_{t-1}}{\sigma_{t-1}}$$

8 GJR-GARCH

The GJR-GARCH model also models asymmetry in the GARCH family models, see Glosten, Jagannathan and Runkle [13]. As mentioned earlier, we have $\epsilon_t = \sigma_t z_t$, where $z_t \;\; iid \sim N(0,1)$. The GJR-GARCH model is defined by

$$\sigma_t^2 = c + \sum_{j=1}^{p}\beta_j\sigma_{t-j}^2 + \sum_{i=1}^{q}\alpha_i\epsilon_{t-i}^2 + \sum_{k=1}^{r}\gamma_k\epsilon_{t-k}^2 D_{t-k}(\epsilon_{t-k} < 0)$$

where D_{t-k} is a dummy variable taking the value one if the residual is smaller than zero and respectively taking the value zero if the residuals is not smaller than zero.

$$D_t = \begin{cases} 1, & \text{if } \epsilon_t < 0 \\ 0, & \text{otherwise} \end{cases}$$

The GJR-GARCH model also captures the asymmetric impacts by the sign of the indicator term to reflect different influence between good information and bad information. In this paper we use the GJR-GARCH(1,1) model which is given by

$$\sigma_t^2 = c + \beta_1 \sigma_{t-1}^2 + \alpha_1 \epsilon_{t-1}^2 + \gamma_1 \epsilon_{t-1}^2 D_{t-1}$$

9 Distribution of the error term

The distribution of error term plays an important role in estimating the different GARCH models. In this paper we choose to introduce two common distributions.
The most common application to model and forecast the financial volatility is assumed as standard Normal distribution. But, the Normal distribution is not always good. For instance, if the error term is fat tailed, then the Normal distribution cannot capture this feature and fully explan the clustering volatility that appears in some financial data. For deal with and modeling such fat-tailed distribution, Nelson[4] proposed that a different distribution may fit the sample data better. Therefore we choose a second distribution, Student-t distribution.

9.1 Normal distribution

The probability density function of Z_t is given by

$$f(z_t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z_t-\mu)^2}{2\sigma^2}}$$

where μ is the expected value and σ^2 is the variance. From earlier courses in probability theory, it is known that the distribution is called standard Normal distribution when $\mu = 0$ and $\sigma_t^2 = 1$.

9.2 Student-t distribution

The probability density function of the Student-t distribution is given by

$$f(z_t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\nu-2)\pi}} \left(1 + \frac{z_t^2}{\nu - 2}\right)^{-\frac{1}{2}(\nu+1)}$$

where ν is the number of degree of freedom, $2 < \nu \leq \infty$, and, Γ is the Gamma function. When $\nu \to \infty$ the Student -t distribution approaches nearly equal to the Normal distribution. The lower the ν, the fatter the tails. So the Student- t distribution perhaps reflects the fat-tail of the volatility asset more precisely and may fit our data better.

10 Returns Analysis

In this section, we use the returns of the stock in order to estimate and forecast the financial volatility. Hence, we start by converting the closing daily prices of the stock return, LVMUY belonging to Louis Vuitton Moët Hennessy. We obtain the returns according formula [1], i.e

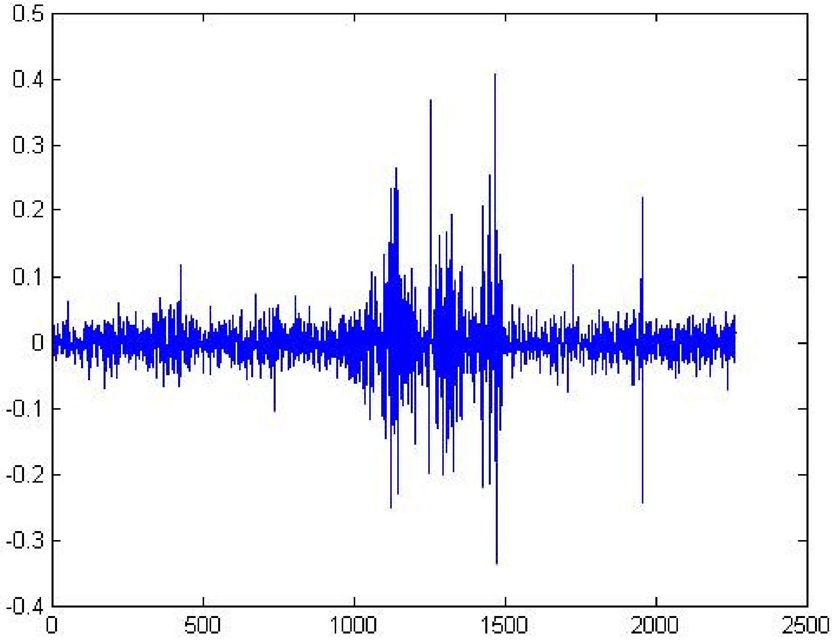
$$R_t = \frac{I_t - I_{t-1}}{I_{t-1}} \approx log(1 + \frac{I_t - I_{t-1}}{I_{t-1}}) = log(\frac{I_t}{I_{t-1}})$$

where, R_t is the return for the stock and I_t is the closing daily price for the stock at time t.

10.1 LVMUY analysis

The figure below presents the closing daily stock price development for the return series during the period of nearly nine years, i.e 2268 observations from 2004-01-02 to 2012-12-31. Where a whole year consists of 252 traiding days. The figure below shows the daily closing price changes for the return series, defined by the formula [1]. We can observe that the returns of the stock rate moves around zero basically through the whole period of time, independently of the changes in the stock price value.

If we observe the figure further, it reveals that the returns have changed within time. The amplitude of the returns has also clearly changed over time, first started as relatively low in the beginning of the period and then increased significantly. We can also see the pattern with concentration of the return movements with similiar magnitude within time. This observed pattern reveals particularly volatility clustering within the return series of the stock price which may be considered as representing a form of heteroscedasticity of the sample error terms.



10.2 Features of the sample data

After obtaining the returns of the stock price, we need to summary and list the features of our data. We choose to list the following:
Sample size, mean, standard deviation, minimum, maximum, skewness, kurtosis and Jarque-Bera test.

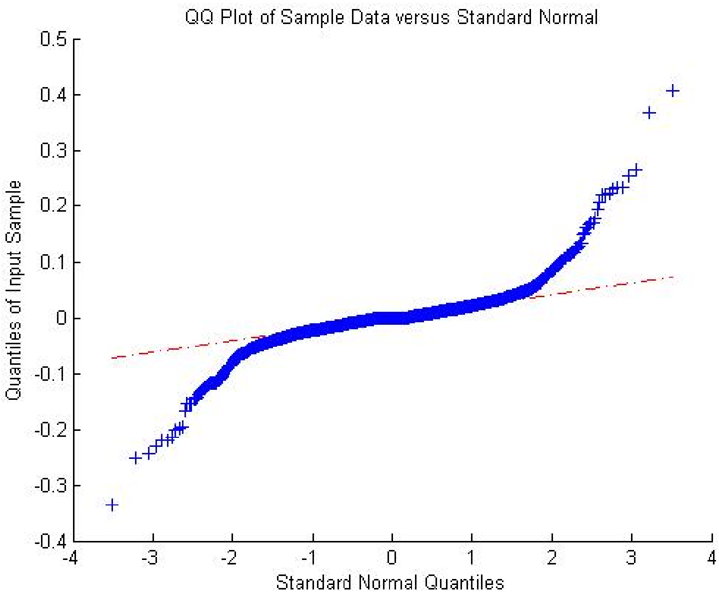
Sample size	2265
Minimum	-0.3365
Maximum	0.4055
Mean	3.2977E-004
Standard deviation	0.0403
Skewness	0.7328
Kurtosis	21.7785
Jarque-Bera	3.3482E+004

We start by examine the skewness and kurtosis since both gives us an indication of the shape of the sample distribution. The kurtosis of the LVMH daily return is 21.7785 which by far exceeds the theoretical value of the Normal distribution of 3. This implies that the financial time series data has a fat-tail characteristic. The skewness however,

is 0.7328 which means that the sample distribution is not symmetric, i.e non-normality of the series data.
The Jarque-Bera test reveals another feature, the higher value of Jarque-Bera test represent the non-normality of the time series data. Since the value of 3.3482E+004 is large enough this implies that the distribution of the returns is not normally distributed.

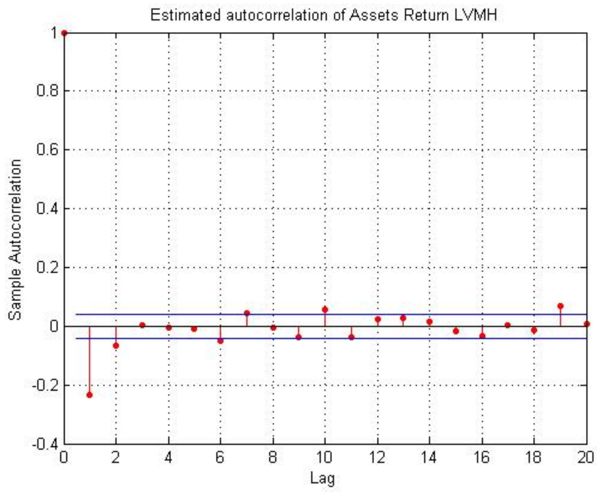
10.3 QQ-PLOT

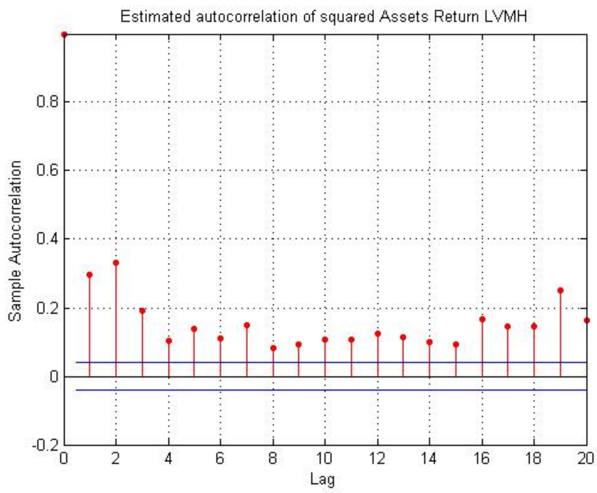
Another way to observe data and test whether the error terms of the returns are normally distributed can be done by observing it graphically. By ploting the returns with a quantile plot, we see immediately that the sample data is not Normal distributed due to the s-shape of the curve that the returns form. If the sample data is normally distributed then the observations would follow the dashed straight line marked in the figure. The more the observations diverge from this line, the less the data will be approximate to a Normal distribution. Thus, it is confirmed that the sample data is not normally distributed.



10.4 Autocorrelation

Earlier we defined a process R_t, now let us assume that the sample data could be described by the constant mean model as in R_t. Further, we calculate the autocorrelation function for the sample data and we plot it as well.





The purpose of the autocorrelation function is to measure dependency between the value in the present and the value a few days in the past. We investigate both the autocorrelation function for the stock rate returns and as well for the autocorrelation function for the squared returns.
We start first by observing the first figur. Hence, the autocorrelation of the returns could be assumed as not significant since just a few of the values easily outrun the confidence interval at the significance level of 5%.

Let us now instead consider the autocorrelation for the squared returns, by observing the plot one can easily detect a dependency in second moment of the returns. Hence, the time series data of the autocorrelation is significant.

Further, we calculate the test statistic which was defined by the formula [4] in order to test dependency within the time series.
We obtain the Q-value of 179.2 for the returns and the Q-value of 1.24E+03 for the squared returns. As we mentioned earlier, both of these test statistics are asymptotically chi squared, χ^2 distributed with $k = 20$ degrees of freedom. Thus, this confirms and reveals that both the squared returns and the returns as well are significantly autocorrelated at the significance level of 5%.

The most common and available tests for the autocorrelation are mostly based on the fact that if the true disturbances are autocorrelated, this fact will then be revealed through the autocorrelation of the least squares residuals. The most commonly used statistic is the Durbin-Watson statistic. Due to our model, we are not able to use this specified test.

10.5 Heteroscedasticity

We can determine if heteroscedasticity is likely to be present in many certain ways. A visually way of checking to see if heteroscedasticity is present is through plots. Since plots of the residuals in time series, can indicate if the assumption of the constant variance is violated.

For the linear regression model with explanatory variables one could consider a general test which does not require any form for the errors distribution. Such a test is called, White's test. This specified test proposed by White involves, including regressing ϵ_j^2 on all the explanatory variables. But since we have assumed no explanatory variables in our model, we decide to test heteroscedasticity by testing the linear model for the residuals from our previous linear model [3]

$\epsilon_t = \alpha + \beta \cdot \epsilon_{t-1} + u_t$

where α and β are parameters to be estimated and u is assumed to be white noise.

We can easily test the present of heteroscedasticity by setting up the null hypothesis of homoscedacity versus the alternative hypothesis of heteroscedasticity according :

$H_0 : \beta = 0$
$H_A : \beta \neq 0$

After using the ordinary least square method in order to estimate our unknown parameter of β, we receive the estimated value of - 0.23561. Further, we use a common test statistics that we have been encountered in previous basic courses in statistics according the formula below, in order to test the null hypothesis of homoscedasticity at the significance level of 5%,

$$t_{score} = \frac{\hat{\beta} - \beta_0}{SE_{\hat{\beta}}} \sim t_{0.05}(n-2)$$

where $SE_{\hat{\beta}}$ is the standard error of the estimated beta value. We receive a test statistics of 11.53. Hence, we can reject the null hypothesis of homoscedasticity at the significance level of 5%. Thus, the volatility of the returns is heteroscedastic which means that the variance varies for the time series. Taking the heteroscedastic of the returns in consideration, this is for instance a feature of the clustering volatility.

11 Results

In this section we will examine the GARCH family models with two kinds of different distribution of the error term to estimate and forecast the financial volatility using the stock return for the Louis Vuitton Moët Hennessy stock rate, LVMUY.

11.1 Comparison of the GARCH models

In order to find the best and most appropriate model for our data, we will separately utilize the GARCH(1,1), EGARCH(1,1) and GJR-GARCH(1,1) model including the Normal distribution and the Student-t distribution of the error term to predict and model the financial volatility with the daily stock returns belonging to LVMUY. We start by summarize the results of the different GARCH family models. Further, we compare them in the pursuit of finding which model forecasting the volatility better. In order to do so, we use the returns to fit the different GARCH family models. We are especially interested in :
the coefficients, whether they are statistically significant or not, AIC and the loglikelihood of the models.

11.2 GARCH(1,1)

	Gaussian	Student-t
c	1.78E-05***	6.24E-05***
β	0.88***	0.76819***
α	0.11053***	0.19372***
AIC	-9.695E+03	-1.0034E+04
Loglikelihood	4851.48	5021.96

Note: ***, Statistically significant
at the 10% significant level.
AIC stands for Akaike information criterion for each model.
The log-likelihood is the value which stands for
the maximum log-likelihood value for the models

11.3 EGARCH(1,1)

	Gaussian	Student-t
c	-0.029739	-0.14219
β	0.99182***	0.97903***
α	0.22109***	0.22925***
γ	-0.072569	-0.024842
AIC	-9.6133E+03	-1.0038E+04
Loglikelihood	4811.63	5025.23

Note: ***, Statistically significant
at the 10% significant level.
AIC stands for Akaike information criterion for each model.
The log-likelihood is the value which stands for
the maximum log-likelihood value for the models

11.4 GJR-GARCH(1,1)

	Gaussian	Student-t
c	1.5584E-05***	6.2683E-05***
β	0.88254***	0.76758***
α	0.063755***	0.19552***
ϕ	0.098587***	-0.0026511
AIC	-5.714E+03	-1.0032E+04
Loglikelihood	4861.98	5021.9

Note: ***, Statistically significant
at the 10% significant level.
AIC stands for Akaike information criterion for each model.
The log-likelihood is the value which stands for
the maximum log-likelihood value for the models

There are many several ways to compare the advantages of models, such like AIC, BIC, RMSE, Out of sample forecasts and Loglikelihoodvalue. However, in this paper in order to contrast the performance of the GARCH family models in forecasting the stock volatility, we will determine it by calculating the AIC and Loglikelihood.

Recall, that loglikelihood is a logarithm of the likelihoodfunction. The logarithm function is an increasing function, i.e if the value of the loglikelihood of a model is greater in comparison to another model, then this tell us that the likelihood of this model is greater and that means that this model is more likely to have a better performance in forecasting the volatility. We will consider the model with most significant coefficients together with the smallest AIC and greatest Loglikelihood as the best model in forecasting the financial volatility.

We begin with examine the GARCH(1,1) model. All of the coefficients c, α, β for this model are statistically significant at 10% significant level for both the Gaussian distribution and for the Student-t distribution. The Student-t distribution has a superior smaller AIC value of -1.0034e+04 and a Loglikelihood value of 5021.96 in comparison to the GARCH model with Gaussian distribution. Therefore we consider the GARCH model with Student-t distribution as the most appropriate model in forecasting the financial volatility.

Proceeding with the EGARCH(1,1) model, we see that the coefficient γ is not significant for both the Gaussian distribution and as well as for the Student-t distribution. Hence, the EGARCH model with Student-t distribution and with the Gaussian distribution are not asymmetric models. We also find that the EGARCH model with the Student-t distribution has the smallest value of AIC and the greatest value of Loglikelihood. Thus, it has the best performance in forecasting the volatility.

For the GJR-GARCH model, we find that all coefficient besides ϕ for the Student-t distribution are significant, therefore the GJR-GARCH model with the Student-t distribution is not an asymmetric model. We find also that the GJR-GARCH model with the Student-t distribution has a higher value of Loglikelihood and a superior smaller value of AIC in relation to the GJR-GARCH with the Gaussian distribution. Whereas, we consider the GJR-GARCH with Student-t distribution better in forecasting the financial volatility.

12 Conclusion and Discussion

This paper has considered the modeling and the forecasting of the stock returns volatility for the stock chart, LVMUY during the past nine years. The sample data has been examined in order to compare the performance in forecasting the financial volatility within three different GARCH family models, namely the GARCH, EGARCH and GJR-GARCH under two distributional assumptions for the error term, i.e the Normal distribution and the Student-t distribution. By comparing the AIC and the Loglikelihood among the different GARCH models we can easily decide which one of them is the most optimal in order to contrast the performance in forecasting the stock volatility. From the result of this paper, we get that the EGARCH model with the Student-t distribution is the best model in forecasting the financial volatility but not by much. Since the coefficient γ in the EGARCH model for both the Gaussian distribution and the Student-t distribution is not significant, we consider this model as not an asymmetric model, i.e an symmetric model. This result may show that despite that the GARCH model have some drawbacks and limitation in relation to the EGARCH model, the symmetric GARCH model may still provide useful results in forecasting purposes. However, the AIC and Loglikelihood of the different models do not tend to always provide a clear decision when to decide the best model and does not always give us the exactly decision in forecasting the volatility.

Another way to compare the several models against each other and hence, acquire the best model to forecast the financial volatility could be done by calculating the RMSE, root mean square error with a procedure called out-of-sample forecasts. But since we only have one single stock to examine and not an index, we choose not to use this method.

The choice of method to forecast the volatility may have some drawbacks which maybe lead to poor estimation for the sample data, so from this perspective, maybe we could add some extra extension to the GARCH models with higher order than (1,1) and consider an additional distribution for the error term, e.g the GED distribution. Maybe most important, we should have compared a index instead of a single stock. This may be small changes and factors but could play a crucial and distinctive role in explaining the data in a optimal and better way.

An alternative method would be that, since the EGARCH model with the Student-t distribution outperformed better forecast than the Normal GJR-GARCH and the Normal GARCH model, we could apply it to estimate the coefficients in three different time frames. Since economic factors include some cyclical changes in the financial and economic world, e.g financial crisis. The impact of a financial crisis could affect the stock price harmly and directly, especially the luxury brands belonging to Louis Vuitton Moët Hennessy. So we could, apply the modeling and fit the EGARCH model into some kind of subseries, i.e within the whole period, before the financial crisis and during the crisis. Thus, this could lead us to better estimation.

13 References

[1] Adrian R. Pagan and G. William Schwert(1990), Journal of Econometrics 45: Alternative Models for Conditional Stock Volatility, pp.267-290

[2] Alan Agresti(2002) : Categorical Data Analysis, Wiley Series in Probability and Statistics, pp. 216-217

[3] Andersson P. , Tyrcha J. (2012) : Notes in Econometrics, Department of Mathematics, Stockholm University

[4] Anil K. Bera and Matthew L. Higgins (1993), Journal of Economic Surveys: ARCH models Properties, Estimation and Testing, Vol.7 , No.4 pp. 305-362

[5] Carol Alexander and Emese Lazar(2006), Journal of Applied Econometrics 21: Normal Mixture GARCH(1,1): Applications to Exchange Rate Modelling, pp. 307-336

[6] Daniel B. Nelson(1990), Econometric Theory: Stationarity and Persistence in the GARCH(1,1) model, pp. 318-334

[7] Daniel B. Nelson(1991), Econometrica: Conditional Heteroskedasticity in Asset Returns: A New Approach, Volume 59 Issue2 pp. 347-370

[8] Hans Malmsten(2004), SSE/EFI Working Paper Series in Economics and Finance: Evaluating exponential GARCH models, No.564 , pp.1-6

[9] Hirotugu Akaike(1974), Transactions on Automatic Control: A new look at the Statistical Model Identification, Volume. AC-19, No.6, pp. 716-723

[10] Höglund T. (2008) : Mathematical Asset Management, John Wiley and Sons

[11] Jarque Carlos M. Bera, Anil K. (1980), Economics Letters 6 : Efficient tests for normality, homoscedasticity and serial independence of regression residuals, pp. 255–259

[12] Lars Forsberg and Tim Bollerslev(2002), Journal of Applied Econometrics 17: Bridging the Gap between the Distribution of Realized (ECU) Volatility and ARCH modelling(Of the EURO): The GARCH-NIG Model, pp. 535-548

[13] Lawrence R. Glosten, Ravi Jagannathan, David E. Runkle(1993), the Journal of Finance: On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, Volume.48, No.5, pp. 1779-1801

[14] Robert F. Engle(1982), Econometrica: Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, Vol. 50, No. 4, pp.

987-1007

[15] Robert F. Engle and Victor K. NG(1993), The Journal of Finance: Measuring and Testing the Impact of News on Volatility, Volume XLVIII, No.5 , pp. 1749-1778

[16] Robert Engle(2001), Journal of Economic Perspectives: The Use of ARCH/GARCH Models in Applied Econometrics, Volume 15, No. 4, pp.157-168

[17] Richard T. Baillie and Ramon P. DeGennaro (1990), Journal of Finance and Quantitative Analysis : Stock Returns and Volatility, Vol.25, No.2, pp. 1-12

[18] Stefan Lundberg and Timo Teräsvirta (1998), SSE/EFI Working Paper Series in Economics and Finance : Evaluating GARCH models, Working Paper No.292, pp. 1-25

[19] Tim Bollerslev(1986), Journal of Econometric: Generalized Autoregressive Conditional Heteroskedasticity, pp. 307-327

[20] Tim Bollerslev and Hans Ole Mikkelsen (1996), Journal of Econometrics 73: Modeling and pricing long memory in stock market volatility, pp. 151-184

[21] Timo Teräsvirta and Department of Economic Statistics (2006), SSE/EFI Working Papers in Economics and Finance : An Introduction to Univariate GARCH Models, No.646 , pp.1-30

[22] V.K.Rohatgi : An Introduction to Probability Theory and Mathematical Statistics. John Wiley and Sons, Inc.1976.

[23] http://www.lvmh.com/the-group/lvmh-companies-and-brands

[24] http://www.mathworks.se/help/econ/lbqtest.html