Abstract
This essay investigates three different GARCH-models (GARCH, EGARCH and GJR-GARCH) along with two distributions (Normal and Student’s t), which are used to forecast the Value at Risk (VaR) for different return series. Seven major international equity indices are examined. The purpose of the essay is to answer which of the three models that is better at forecasting the VaR and which distribution is more appropriate. The results show that the EGARCH(1,1) $\sim t$ is preferred for all indices included in the study.

Keywords: Value at Risk, GARCH, EGARCH, GJR-GARCH, Volatility and Forecasting.
# Table of contents

1. Theory.................................................................................................................... 4
   1.1 Introduction........................................................................................................... 4
   1.2 Value at Risk (VaR) .......................................................................................... 5
   1.3 GARCH Models.................................................................................................... 7
       1.3.1 GARCH ........................................................................................................... 7
       1.3.2 EGARCH ......................................................................................................... 8
       1.3.3 GJR-GARCH .................................................................................................. 9
   1.4 Distributions......................................................................................................... 10
   1.5 Previous Research............................................................................................... 11

2. Method....................................................................................................................... 12
   2.1 Data...................................................................................................................... 12
   2.2 Data Control......................................................................................................... 15
   2.3 Forecast Procedure............................................................................................. 16

3. Result/Discussion...................................................................................................... 17
   3.1 Results.................................................................................................................. 17
       3.1.1 GARCH(1,1) .................................................................................................. 17
       3.1.2 EGARCH(1,1) ............................................................................................... 18
       3.1.3 GJR-GARCH(1,1) ......................................................................................... 18

4. Conclusion................................................................................................................ 18
   4.1. Further Research ............................................................................................. 19
   4.2. Recommendation to Investors ......................................................................... 19

5. References................................................................................................................ 20

Appendix....................................................................................................................... 22
Acknowledgements

We are grateful for the invaluable feedback and support given by our supervisor Lars Forsberg. We would also like to thank Johan Andersson, Magnus Andersson and Christer Nylander for their comments.
1. Theory

1.1 Introduction
Proper risk management is of great importance to investors. To control possible losses without missing out on profit will maximize the return. Volatility is often used to get an idea of how risky a security is. However, this asset might be characterized by upward volatility, which should not concern investors. Value at Risk (VaR) is another popular risk measure, focusing solely on the downside. VaR is formally defined as:

\[
\text{Given a probability of } \omega \text{ percent and a holding period of } t \text{ days, an entity’s VaR is the loss that is expected to be exceeded with a probability of only } x \text{ percent on the } t\text{-day holding period (Penza & Bansal 2001).}
\]

The nature of the Value at Risk measure makes it concrete and understandable for people without understanding of statistics or finance. Hence it is suitable for regulators. It is for instance used in the Basel regulations of banks (Basel III 2010).

VaR should not be seen as the sole risk measure, but should be combined with other tools and some common sense (Butler 1999).

In order to forecast VaR one has to consider the volatility patterns of financial assets. The variance tends to change over time, forming so-called volatility clusters. This means the time series will be characterized by heteroskedasticity. In 1982, Robert F. Engle elaborated on the regular AR process and created the Autoregressive Conditional Heteroskedasticity (ARCH) model, which takes the changing variance into account, and need no assumption of homoskedasticity. Engle originally created the model to describe uncertainty about British inflation. However, since the model is applicable on financial assets, substantial further research has been conducted over the years. Significant contributions to the ever-growing family of ARCH models have been made by Bollerslev (1986) and Nelson (1991) resulting in the GARCH and EGARCH models respectively.

Before applying a time series model an assumption about the error term distribution has to be made. This has also caused a discussion. Engle made the assumption of normally distributed error terms while Bollerslev (1987) preferred the student’s t-distribution. The debate has since expanded beyond these two. Earlier research shows ambiguity yet tendencies towards certain models and distributions. This paper will
provide an analysis using recent data from seven major stock indices: NASDAQ 100, S&P 500, FTSE 100, OMXS30, Euro Stoxx 50, Hang Seng and NIKKEI 225. This yields the following thesis statements:

- By investigating the conditional variance models GARCH (1,1), EGARCH (1,1) and GJR-GARCH(1,1) along with the normal- or student’s t-distribution which of these combinations is the one to prefer when forecasting VaR?

1.2 Value at Risk (VaR)
To fully understand VaR it is necessary to begin with its basics. In short, the idea of VaR is to focus on the worst losses in a return series. It can be applied for various assets for instance equities, commodities or bonds.

Based on the available data the task is to get a clear picture of what the specific loss amount corresponds to at a certain significance level. Therefore, the focus is on the left tail of the return distribution (Dowd 2010). To simplify the explanation further, assume the return distribution follows a standard normal distribution. Then at a significance level of five percent, the critical value equals -1.645. Imagine a confidence level constructed by (1-\(\alpha\)) then VaR is the upper limit of the left tail. The main objective is to establish a conclusion regarding the predicton power of the different conditional variance models. Thus VaR estimates are forecasts, which have been calculated based on a certain significance level (Lopez 1998). If the forecasted value is less than its equivalent percentile value it gives a violation. It means that one is interested in finding a model that produces violations that are close to the chosen significance level. In other words, in a scenario of a perfectly modified model it would then produce violations that equal the sum of the violations, which have been decided by the significance level (Orhan & Köksal 2012). In a situation where the data set consists of 1000 daily returns. The total sample size would then contribute to a theoretical violation amount of 50 or 10 at a five or a one percentage level respectively. The calculations are as follows: 1000*(1-0.95)=50 violations. If the VaR estimates produce for example 60 violations the model overestimate the amount since 60>50. Below is an illustration showing VaR with a 5 % significance level.
In fact, Benavides (2007) states that in general GARCH type models tend to overestimate the number of violations due to the cluster persistence volatility. An overestimation would lead to a misallocation of investments because of the restrain it brings to an investor’s capital, which in turn results in opportunity costs occurrence that takes the shape of potential profits being absent (Orhan & Köksal 2012).

To summarise, VaR with its focus on excessive losses, provides the risk analyst a loss value, which she can assume at a certain significance level (Dowd 2010).

The next question regards the decision rule of when a potential violation happens. Consider the indicator function

$$v_t = \begin{cases} 
1, & R_t < \text{VaR}_t \\ 
0, & R_t \geq \text{VaR}_t 
\end{cases} \quad (1)$$

The indicator says that when $R_t$ (Return at day t) is less than its VaR estimate it would result in a violation. If the indicator function returns the value one the predictive value ends up in the rejection region (Orhan & Köksal 2012).

The nominal size is defined as $V = \sum v/N$, which equals the $\alpha - level$. It is then compared to the size that has been found by the model.
Moving on to the utilisation of the GARCH type models when estimating VaR. Hull & White (1998) suggest to adopt the approach of modeling the forecasts based on the characteristics of the data set. The procedure involves to predict the VaR estimates based on the previous percentage changes caused by market movements. These market characteristics can then be modelled by conditional variance models such as the GARCH(1,1) model.

1.3 GARCH Models
The daily returns are calculated by

\[ r_t = \ln \left( \frac{P_t}{P_{t-1}} \right) \times 100. \]  \hspace{1cm} (2)

Where \( r_t \) is the daily return at time \( t \). The equation: \( \ln \left( \frac{P_t}{P_{t-1}} \right) \) is simply the natural logarithm of the closing price at time \( t \) divided by yesterday’s adjusted closing price \( t-1 \).

The return series is assumed to be decomposed into two parts:

\[ R_t = E(R_t | I_{t-1}) + \epsilon_t. \] \hspace{1cm} (3)

The conditional mean return \( E(R_t | I_{t-1}) \), is believed to be an Autoregressive process that captures the expected return at time \( t \) given all the available information up to and including \( t-1 \).

The second part is the so-called unpredicted part i.e. \( \epsilon_t \) that can be defined by the equation

\[ \epsilon_t = z_t \sigma_t, \] \hspace{1cm} (4)

where \( \sigma_t \) is the conditional standard deviation of \( \epsilon_t \) while the sequence of \( z_t \) is an iid with mean equal to zero and a unit variance (Angelidis et al 2004).

Moving on to the standardized student’s t-distribution. Bollerslev (1987) showed that the unpredicted part could then be written as a conditional density function for \( \epsilon_t \) that takes into account the previous period’s information and the degrees of freedom. It gives the following

\[ \epsilon_t = \sqrt{\frac{\nu}{\nu-2}} \times \frac{r_t}{\sigma_t}. \] \hspace{1cm} (5)

Here \( \nu \) represents the degrees of freedom.
1.3.1 GARCH

The unpredicted part can be estimated by conditional variance models. In 1982 Robert F. Engle presented the Autoregressive Conditional Heteroskedasticity (ARCH) model, as a way to forecast the variance of a time series. The ARCH model assumes that, just as the error terms in a regular AR process, the variance of the error term is dependent on previous error term variances.

\[ \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 \]  

Bollerslev and Taylor independently elaborated on the ARCH model and developed the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model (Bollerslev 2009). GARCH (p,q) includes \( p \) lags of the conditional variance in the conditional variance equation.

\[ \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_j \sigma_{t-j}^2 \]  

The \( \alpha_i \) coefficient captures market news. In table 1, below, there are two possible outcomes of the coefficient values. The number 1 represents a scenario where a high value of \( \alpha_i \) indicates that the volatility is sharp, spiky, and it also responds well to market movements. Whereas a low value of the \( \beta_j \) shows that the market volatility is not very persistent in the long run. In other words, the \( \beta_j \) focuses on the degree of persistence of the market news. The second outcome, number 2 in the table, is the opposite situation. A low value of \( \alpha_i \) shows that the coefficient does not respond successfully to the markets movements while the high value of \( \beta_j \) means that in the long run market news has a large persistent influence (Dowd 2010).

Table 1. Possible coefficient values.

<table>
<thead>
<tr>
<th>Coefficients:</th>
<th>( \alpha_1 )</th>
<th>( \beta_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>High values</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Low values</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

The process of a GARCH(1,1) is stationary if the condition: \( \alpha_1 + \beta_1 < 1 \) is fulfilled. Thus, in the long run the conditional variance will converge toward the unconditional variance resulting in the expression \( \frac{\alpha_0}{1-(\alpha_1+\beta_1)} \).
The forecast equation for the next period, i.e. tomorrow’s predicted value, of a GARCH (1,1) is specified by

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta_1 \sigma_t^2. \tag{8}$$

### 1.3.2 EGARCH

A modified version of GARCH is the so-called Exponential-GARCH model (EGARCH (p,q)), which uses the natural logarithmic value of the dependent variable that provides a positive value. It also enables the model to capture the asymmetric effect on the variance caused by negative and positive market news. The asymmetry can be explained by the fact that negative shocks will have a greater impact on volatility than positive shocks (Nelson 1991).

The EGARCH (p,q) model is constructed by the following equation.

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^q \alpha_i (|z_{t-i}| - E(|z|)) + \sum_{j=1}^p \beta_j \ln(\sigma_{t-j}^2) + \sum_{i=1}^q \delta_i z_{t-i}. \tag{9}$$

Where: \( E(|z|) = \sqrt{2\pi} \) if the underlying distribution follows the normal distribution while for a standardized t-distribution it can be written as

$$E(|z_t|) = \frac{2 \cdot \gamma_{v-1}^{\gamma_v}}{(v-1) \Gamma (\frac{v+1}{2}) \sqrt{\pi}} \tag{10}.$$

Furthermore, the coefficient \( \delta_i \) displays the degree of asymmetry. If \( \delta_i \) is equal to zero it is indicated that the model is perfectly symmetric whereas if it is less than zero, negative news affect the volatility more heavily than positive news. If \( \beta_j < 1 \) the process is stationary (Enders 2010). The forecasted model for tomorrow’s predicted value is given by

$$\ln(\sigma_{t+1}^2) = \alpha_0 + \alpha_1 (|z_t| - E(|z_t|)) + \beta_1 \ln(\varepsilon_t^2) + \delta_1 z_t. \tag{11}$$

### 1.3.3 GJR-GARCH

The third model is the Glosten, Jagannathan and Runkle-GARCH model. Differently from the original GARCH model it does not assume that if a shock would occur then the sign of the shock would be independent to the response variable. It would only be a function of the size of the shock (Glosten et al 1993).
The GJR-GARCH (1,1) model is defined as

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \delta_1 I(\varepsilon_{t-1} < 0)\varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 .$$  \hspace{1cm} (12)

Where $I(\varepsilon_{t-1} < 0)$ is an indicator function, which takes the value one if the corresponding lagged unconditional standard deviation is less than zero (Teräsvirta 2006). The model captures the asymmetrical nature of a time series by including the indicator. Applying the indicator function to financial return data the function would then produce a value of one if there were a loss and zero if there are profits.

The forecast model is as follows

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \varepsilon_t^2 + \delta_1 I(\varepsilon_t < 0)\varepsilon_t^2 + \beta_1 \sigma_t^2 .$$  \hspace{1cm} (13)

### 1.4 Distributions.

The maximum likelihood approach enables the three models to have the normal-or the student’s t-distribution as their underlying conditional distributions.

Firstly, consider the conditional distribution equation of $y_t$ to follow a normal distribution with mean zero and variance as $\sigma^2$

$$y_t = E(y_t|y_{t-1}) + \varepsilon_t .$$

Where $^1\varepsilon_t|y_{t-1} \sim f_\nu(\varepsilon_t|y_{t-1}) = \frac{1}{\sigma \sqrt{2\pi}}e^{-\frac{(y-\mu)^2}{2\sigma^2}}.$  \hspace{1cm} (14)

The corresponding Maximum Likelihood with normal distribution is

$$L_t = \prod_{t=1}^{T} \left(\frac{1}{\sqrt{2\pi \sigma^2}}\right) \exp\left(-\frac{\varepsilon_t^2}{2\sigma^2}\right).$$  \hspace{1cm} (15)

By taking the natural logarithm of equation (14) the log likelihood equation is:

$$\ln L_t = -\frac{T}{2} \ln(2\pi) - \frac{T}{2} \ln\sigma^2 - \frac{1}{2\sigma^2} \sum_{t=1}^{T}(\varepsilon_t)^2 .$$  \hspace{1cm} (16)

Where $\varepsilon_t = (y_t - \mu_t)$ and $\mu_t$ is the corresponding conditional mean value at time $t$ (Enders 2010).

Moving on to the student’s t-distribution. The conditional distribution equation of $y_t$ where $t=1,2,…,T$ is a standardised t-distribution with mean $y_{t|y_{t-1}}$, variance $\sigma_{t|y_{t-1}}^2$ and $\nu = \text{degrees of freedom}$. Bollerslev (1987) defines the t-distribution and the Maximum Likelihood as

---

1 Wackerly et al. (2008)
\[ y_t = E(y_t | I_{t-1}) + \varepsilon_t \]
\[ \varepsilon_t | I_{t-1} \sim f_\nu(\varepsilon_t | I_{t-1}) = \Gamma \left( \frac{\nu + 1}{2} \right) \Gamma \left( \frac{\nu}{2} \right)^{-1} \left( (\nu - 2)\sigma^2_{t|t-1} \right)^{-1/2} \left( 1 + \frac{\varepsilon^2_t}{\sigma^2_{t|t-1}} \right)^{-\frac{\nu}{2}} \]  
\[ (17) \]

Based on the condition \( \nu > 2 \) and where \( I_{t-1} \) represents all available information up through time \( t-1 \). The function \( f_\nu(\varepsilon_t | I_{t-1}) \) is the conditional density function for \( \varepsilon_t \), which gives the Maximum Likelihood\(^2\) is defined as

\[ L_T(\theta) = \sum_{t=1}^{T} \ln f_\nu(\varepsilon_t | I_{t-1}) \]

\[ (18) \]

Lastly, the \( \theta \) captures all unknown parameters that are being estimated by equation (18), including degrees of freedom.

### 1.5 Previous Research

Similar studies have been conducted investigating GARCH family model performances for several worldwide equity indices. The following studies are focused on the VaR measure. Bucevska (2012) finds the EGARCH with normal- or student’s \( t \)-distribution to be superior when making forecasts for the Macedonian stock index. A study more similar to this one, though featuring older (1987-2002) data was presented by Angelidis et al in 2003. That study focuses on bigger indices such as FTSE, NIKKEI and DAX. It as well concludes EGARCH to perform better than other GARCH-models. However they prefer the student’s \( t \)-distribution over the normal distribution. In a recent examination of three Nordic indices (OMXC20, OMXH25 and OMXS30) it is concluded that more sophisticated models such as GJR and EGARCH clearly outperform the parsimonious model: GARCH (Wennström 2014). Awartani & Corradi (2005) found that asymmetric conditional variance models such as an EGARCH outperformed the symmetric models (GARCH(1,1)) when it comes to forecasting volatility with a one-step procedure of the Standard and Poor’s 500. One of their main conclusions is that, usually, return time series often suffers of non-symmetrical shaping that for instance asymmetric models are able to include when

\(^2\) See Bollerslev (1987) for a more detailed description.
estimating return volatilities. Although there is no consensus neither regarding model nor distribution, earlier studies tend to favour the EGARCH with a non-normal distribution. Several studies include even more models, lags and distributions, drawing interesting conclusions. However, this paper will solely focus on the GARCH(1,1), EGARCH(1,1) and GJR-GARCH(1,1) with the normal- and student’s t as the underlying probability distributions.

2. Method

2.1 Data
The data sets consist of daily adjusted closing prices collected from the Yahoo Finance website and have been estimated by MATLAB (2010) software. Each data set represents one major equity index.

North American indices
- NASDAQ 100 (NDX) - USA
- Standard and Poor’s 500 (S&P) - USA

Asian indices
- NIKKEI 225 (NKY) - Japan
- Hang Seng (HSI) - Hong Kong

European indices
- FTSE 100 (FTSE) - UK
- Euro Stoxx 50 (SX5E) - Euro-zone
- OMXS30 (OMX) - Sweden

The time span corresponds to four years (07-11-2010 to 07-11-2014) where each year covers approximately 250 trading days. This gives a total sample space of roughly 1000 observations. Below, descriptive statistics are presented for the different equity indices.
Table 2. Descriptive statistics for the seven indices.

<table>
<thead>
<tr>
<th>Des. statistics</th>
<th>FTSE</th>
<th>NDX</th>
<th>S&amp;P</th>
<th>OMX</th>
<th>SX5E</th>
<th>HSI</th>
<th>NKY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Std. Error</td>
<td>0.0094</td>
<td>0.0107</td>
<td>0.0097</td>
<td>0.0122</td>
<td>0.0133</td>
<td>0.0116</td>
<td>0.0138</td>
</tr>
<tr>
<td>Min - value</td>
<td>-0.0478</td>
<td>-0.0631</td>
<td>-0.0690</td>
<td>-0.0603</td>
<td>-0.0590</td>
<td>-0.0552</td>
<td>-0.1115</td>
</tr>
<tr>
<td>Max -value</td>
<td>0.0394</td>
<td>0.0476</td>
<td>0.0463</td>
<td>0.0697</td>
<td>0.0632</td>
<td>0.00583</td>
<td>0.0552</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.2665</td>
<td>-0.3774</td>
<td>-0.5831</td>
<td>0.3424</td>
<td>0.1358</td>
<td>0.2646</td>
<td>0.8573</td>
</tr>
<tr>
<td>Sample size</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
</tr>
</tbody>
</table>

In Table 2 the time series suffers from high degrees of kurtosis meaning that there are extreme values present. The skewness values point toward asymmetrical properties of the data sets’ distribution.

Figure 2. Returns series for FTSE 100.

Figure 3. Returns series for OMXS30.
Figure 3. Returns series for Euro Stoxx 50.

Figure 4. Returns series for NASDAQ 100.

Figure 5. Returns series for S&P 500.
2.2 Data Control

The data have been checked for autocorrelation by using a Ljung-Box test statistic with fifteen-lags on the squared residuals. According to the test results\(^3\), there are generally significant lags while the latest lag is insignificant. The Engle test was also conducted by using five lags and showed similar results as the Ljung-Box test at a chosen significance level of 5\%. A conditional variance model such as a GARCH(1,1) should perform well on these time series.

\(^3\) The OMXS30, HSI, and Euro Stoxx 50 have no autocorrelations in the lags. Nikkei 225 has autocorrelation in all lags.
2.3 Forecast Procedure
The estimated models are based on the first 500 observations in the sample covering the time span from 07-11-2010 to 07-11-2012, which is the in-sample period. The next step is to forecast the out of sample space by using the rolling window method, which corresponds to the 500 last observations where the latest observation corresponds to 07-11-2014.

The method allows evaluating the prediction power of the specific model. Because the data sets are based on daily returns the window at the initial stage consists of the first 500 observations, which is then used to forecast the 501st variance. The forecasted percentile value is then compared to the actual 501st return value. The second stage involves a movement of the rolling window, which includes the 2nd value up to the 501st actual return. This gives a rolling window forecast value that corresponds to the 502nd real value. The movement of the window ends when the entire out-of-sample window, i.e. the 500 observations, has been forecasted. The forecasted values for each model are then compared to the specific return value (Hansen & Lunde 2005).
3. Result/Discussion

The estimation tables and their corresponding forecast figures for the models with normal-and student’s t-distributions are presented for NASDAQ 100 in the appendix.

3.1 Results

Below is the summarised table of the results. The most accurate model and distribution for each equity index are marked with an asterisk. The result table is followed by a separate discussion of the performance of each conditional variance model. Estimation tables and forecast figures for NASDAQ 100 are attached in the appendix.

Table 3. Results for all models and distributions.

<table>
<thead>
<tr>
<th>Model</th>
<th>GARCH(1,1)</th>
<th>EGARCH(1,1)</th>
<th>GJR-GARCH(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NASDAQ 100</td>
<td>0.082</td>
<td>0.076</td>
<td>0.081</td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>0.083</td>
<td>0.079</td>
<td>0.079</td>
</tr>
<tr>
<td>FTSE 100</td>
<td>0.078</td>
<td>0.078</td>
<td>0.077*</td>
</tr>
<tr>
<td>OMXS30</td>
<td>0.082</td>
<td>0.079</td>
<td>0.082</td>
</tr>
<tr>
<td>Euro Stoxx 50</td>
<td>0.087</td>
<td>0.088</td>
<td>0.084</td>
</tr>
<tr>
<td>Hang Seng</td>
<td>0.094</td>
<td>0.088*</td>
<td>0.089</td>
</tr>
<tr>
<td>NIKKEI 225</td>
<td>0.086</td>
<td>0.087</td>
<td>0.085</td>
</tr>
</tbody>
</table>

*Table 3 presents the violations found for each index based on the three models.*

3.1.1 GARCH(1,1)

According to table 3 the best is the GARCH (1,1) that follows the student’s t-distribution with exception to the VaR estimates for the Euro Stoxx 50 and HSI indices. A potential explanation of the outperformance is the fact that when comparing the two probability distributions, the student’s t-distribution is constructed to fit/include extreme values. In other words, this distribution includes more efficiently higher levels of kurtosis than the normal distribution by including the degrees of freedom parameter.

The largest difference is found when looking at the OMXS30 where the GARCH (1,1) ~ t outperforms the GARCH ~ N by nine violations. Hence the OMXS30 index suffers of a high degree of kurtosis, which is also supported by looking at the descriptive statistics in table 2.
3.1.2 EGARCH(1,1)
The EGARCH(1,1) VaR estimates show similar prediction patterns as the GARCH(1,1). In overall, an EGARCH model that follows a student’s t-distribution performs better compared to normal distribution, which can be described by the argument stated previously (see: GARCH(1,1) results).

However, an EGARCH(1,1) \( \sim N \) produces less violations than a GARCH(1,1) \( \sim N \) and this is due to the fact that an EGARCH model includes the leverage effect measurement. As mentioned earlier, the leverage effect captures the asymmetry caused by the negative and positive news (shocks).

Moving on to EGARCH(1,1) \( \sim t \) models it is clear that these processes outperform the EGARCH(1,1) \( \sim N \) as well as the parsimonious GARCH and give the best VaR predictions. This is because these models include the leverage effects plus the adjustment of the kurtosis problem discussed earlier.

3.1.3 GJR-GARCH(1,1)
The preferred underlying distribution is the normal distribution and the largest deviations is to be found when looking at the Standard and Poor’s 500 index where the GJR-GARCH(1,1) \( \sim N \) outperforms the student’s t-distribution. Besides the two American indices the relative violation amounts equal one another. Moreover, according to the table GJR-GARCH is, in general, the worst model out of the three models. Hence, the results show that with its indicator function it should be expected to receive a higher sum of violations due to the construction and the purpose of the indicator. Surprisingly, the fat tails dilemma seems to be fairly absent among the models, which contradict with the earlier models discussions as the previous conditional variance models should be estimated and forecasted with student’s t-distributions rather than with a normal distribution. The absence may very well be described by the high amount of degrees of freedom encountered, which contribute to the student’s t-distribution approximately begins to resemble the normal distribution.

4. Conclusion
To summarise, the results show that the most appropriate model to forecast VaR estimates is the EGARCH (1,1) model. GJR-GARCH(1,1) performs worst out of the three models. The EGARCH findings suggest using the student’s t as the underlying
distribution when forecasting VaR estimates of the examined equity indices. Moreover, the exponential conditional variance model gives the best adjustment for the heteroskedasticity in the variance.

4.1. Further Research
For future research it would be interesting to investigate the seven equity indices, during the same time span, with other conditional variance models such as the ARCH, the TARCH or EWMA models and compare the new findings with the existing results. Another suggestion is to examine the results further with for example including the Christoffersen test. The test gives the possibility to investigate whether the VaR violations are independently distributed i.e. no volatility clustering, which would suggest that the underlying model is specified correctly.

4.2. Recommendation to Investors
The results suggest that the best model to utilise for the seven equity indices (out of the three models) is the EGARCH(1,1) that follows the student’s t-distribution. For FTSE 100 and Hang Seng one can also use the EGARCH(1,1)\(\sim N\) and the GARCH(1,1)\(\sim t\) respectively.

---

4 An EGARCH model with either normal- or t-distribution can be utilised for the London based equity index (FTSE 100).
5 Exponentially Weighted Moving Average (EWMA) model.
6 FTSE 100, Euro Stoxx 50, OMXS30, NASDAQ 100, S&P 500, Hang Seng Index and NIKKEI 225.
5. References


### Appendix

Table 1. Estimation of GARCH(1,1) ~ N for NASDAQ 100.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.00080938</td>
<td>0.00045385</td>
<td>1.7834</td>
</tr>
<tr>
<td>K</td>
<td>4.7063e-06</td>
<td>2.2159e-06</td>
<td>2.1238</td>
</tr>
<tr>
<td>GARCH(1)</td>
<td>0.87122</td>
<td>0.032884</td>
<td>26.4932</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>0.097645</td>
<td>0.021343</td>
<td>4.5750</td>
</tr>
</tbody>
</table>

Log likelihood value: 1550.41

Figure 1. Forecast Figure for NASDAQ 100 with GARCH(1,1) ~ N.
Table 2: Estimation of GARCH(1,1) ~ t for NASDAQ 100.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.009247</td>
<td>0.00043668</td>
<td>2.1176</td>
</tr>
<tr>
<td>K</td>
<td>3.6376e-06</td>
<td>2.4644e-06</td>
<td>1.4761</td>
</tr>
<tr>
<td>GARCH(1)</td>
<td>0.87883</td>
<td>0.040308</td>
<td>21.8030</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>0.10189</td>
<td>0.033523</td>
<td>3.0395</td>
</tr>
</tbody>
</table>

Degrees of Freedom: 6.6176
Log likelihood value: 1557.36

Figure 2: Forecast Figure for NASDAQ 100 with GARCH(1,1) ~ t.
Table 3. Estimation of EGARCH(1,1) ~ N for NASDAQ 100.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.00025611</td>
<td>0.00043383</td>
<td>0.5904</td>
</tr>
<tr>
<td>K</td>
<td>-0.58637</td>
<td>0.1444</td>
<td>-4.0607</td>
</tr>
<tr>
<td>GARCH(1)</td>
<td>0.93437</td>
<td>0.016219</td>
<td>57.6096</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>0.13734</td>
<td>0.046927</td>
<td>2.9267</td>
</tr>
<tr>
<td>Leverage(1)</td>
<td>-0.20336</td>
<td>0.033539</td>
<td>-6.0633</td>
</tr>
</tbody>
</table>

Log likelihood value: 1567.29

Figure 3. Forecast Figure for NASDAQ 100 with EGARCH(1,1) ~ N.
Table 4. Estimation of EGARCH(1,1) ~ t for NASDAQ 100.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.00054193</td>
<td>0.00042538</td>
<td>1.2740</td>
</tr>
<tr>
<td>K</td>
<td>-0.53094</td>
<td>0.16019</td>
<td>-3.3144</td>
</tr>
<tr>
<td>GARCH(1)</td>
<td>0.94149</td>
<td>0.017842</td>
<td>52.7667</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>0.1517</td>
<td>0.0651589</td>
<td>2.4631</td>
</tr>
<tr>
<td>Leverage(1)</td>
<td>-0.21261</td>
<td>0.041707</td>
<td>-5.0977</td>
</tr>
</tbody>
</table>

Degrees of Freedom: 8.0697
Log likelihood value: 1572.3

Figure 4. Forecast Figure for NASDAQ 100 with EGARCH(1,1) ~ t.
Table 5. Estimation of GJR-GARCH(1,1) for NASDAQ 100.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.00046054</td>
<td>0.000439</td>
<td>1.0490</td>
</tr>
<tr>
<td>K</td>
<td>5.2504e-06</td>
<td>1.8076e-06</td>
<td>2.9046</td>
</tr>
<tr>
<td>GARCH(1)</td>
<td>0.862</td>
<td>0.032892</td>
<td>26.2068</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>0</td>
<td>0.033411</td>
<td>0.0000</td>
</tr>
<tr>
<td>Leverage(1)</td>
<td>0.20115</td>
<td>0.053685</td>
<td>3.7469</td>
</tr>
</tbody>
</table>

Log likelihood value: 1563.57

Figure 5. Forecast Figure for NASDAQ 100 with GJR-GARCH (1,1) ~ N.
Table 10 Estimation of GJR-GARCH(1,1) ~ t for NASDAQ 100.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>T-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.00067669</td>
<td>0.00042484</td>
<td>1.5928</td>
</tr>
<tr>
<td>K</td>
<td>4.7221e-06</td>
<td>1.9749e-06</td>
<td>2.3910</td>
</tr>
<tr>
<td>GARCH(1)</td>
<td>0.85324</td>
<td>0.044148</td>
<td>19.3270</td>
</tr>
<tr>
<td>ARCH(1)</td>
<td>0</td>
<td>0.048406</td>
<td>0.0000</td>
</tr>
<tr>
<td>Leverage(1)</td>
<td>0.23377</td>
<td>0.072608</td>
<td>2.2880</td>
</tr>
</tbody>
</table>

Degrees of Freedom: 8.0352

Log likelihood value: 1571.46

Figure 6. Forecast Figure for NASDAQ 100 for GJR-GARCH (1,1) ~ t.