Value-at-Risk Estimation Under Shifting Volatility

Ola Skånberg

Supervisor: Hossein Asgharian

Abstract

Due to the Basel III regulations, Value-at-Risk (VaR) as a risk measure has become increasingly important in Europe for financial institutions. But even though it has become an important risk measure, both internally within company reporting and externally due to legislation, there is no one single way to forecast VaR that has yet proven to be superior. The aim of this paper is to examine different models of VaR estimation on the OMXS30 and FTSE100 indices. I divided the in-sample time periods into one period of low volatility and one period of high volatility. From there, I have calculated VaR with different underlying GARCH models, both symmetrical and asymmetrical. To evaluate the different Value-at-Risk models, the Christoffersen test was used.

For the two time series where the in-sample period had high volatility and the out-of-sample period has low volatility, the asymmetric GARCH models seemed to perform best at estimating Value-at-Risk. The converse relationship was found for the time series where the in-sample had low volatility. Furthermore, an assumption of t-distributed returns worked better than the normal distribution.

Contents

1. Introduction

Due to the Basel III regulations, Value-at-Risk (VaR) as a risk measure has become increasingly important in Europe for financial institutions. But even though it has become an important risk measure, both internally within company reporting and externally due to legislation, there is no one single way to forecast VaR that has yet proven to be superior. Different distributional assumptions can be made, and calculations of volatilities can differ, since there is no generally agreed upon consensus on what the best estimation method is. As Angelidis et al. (2004) puts it: "the choice of an appropriate historical sample size as well as of the adequate model for forecasting volatility [for forecasting VaR] should be considered far from resolved".

Furthermore, the different approaches of calculating VaR for an asset or portfolio depend heavily on the volatility, which is calculated from the previous period. I.e. to calculate the Value-at-Risk for the next time period, volatilities from the previous time periods are required. If the market is experiencing fluctuating periods of volatility, one might have to calculate VaR with high historical volatilities on a future period when the volatility has decreased. Therefore, an in-sample period of low (high) volatility may underestimate (overestimate) the forecasted VaR one period ahead. As the latest financial crisis has proven, volatility on the stock market can be quite different from one year to another. Firstly, the question this paper focuses on is what happens to forecasted VaR that is used as a risk measure in many financial institutions, when the forecasted VaR is based on market data with different volatility than the current market conditions? Secondly, is there a model that works better than the others to model VaR duringunder shifting volatilities? For that reason, I will calculate one day ahead estimates of VaR using two different samples for two different stock indices. The first sample has an in-sample period with high volatility and an out-of-sample period with low volatility. The second has an in-sample period with low volatility and an out-of-sample period with high volatility. The purpose is to study how the models behave in times when the volatility unexpectedly changes and whether they can adapt to changing market conditions to predict accurate measures for VaR.

To compare and evaluate the different volatility models, the Christoffersen (1998) conditional coverage test is used. It tests the accuracy of the VaR estimations as well as the independence of the estimates.

The aim of this paper is to examine different models of VaR estimation on the OMXS30 and FTSE100 indices. I will divide the in-sample time periods into one period of low volatility and one period of high volatility. From there, I will calculate VaR with different underlying GARCH models, both symmetrical and asymmetrical. The models I will use are GARCH, EGARCH and TARCH. GARCH stands for Generalized Autoregressive Conditional Heteroskedasticity and is a model that uses past volatilities to forecast future ones. It is a symmetrical model, meaning that both negative and positive deviations are dealt with in the same manner. EGARCH is an exponential version of the standard GARCH model, which takes into effect whether the shocks are positive or negative, therefore resulting in different forecasted volatilities. TGARCH, or Threshold GARCH, also captures the effects that can arise from asymmetric shocks (like EGARCH). It is, however, even more flexible than the EGARCH model. These three models will be tested under the assumption of both the normal and t-distribution, resulting in 6 different time series of VaR. My resulting VaR estimates will then be evaluated using the Christoffersen test to see which volatility model produces the most accurate VaR estimates. The purpose is to find out which Value-at-Risk measure is better during periods of shifting volatility.

2. Value-at-Risk

Value-at-Risk is defined as the smallest loss, L, on an asset or instrument such that the probability of a future loss larger than L is less than or equal to $1 - \alpha$, where α is the chosen significance level. Since this paper will analyze VaR under the assumption of different continuous distributions, VaR can be defined as:

$$\Pr\big(L > VaR_\alpha(L)\big) = 1 - \alpha$$

The probability that a loss incurs that is larger than the VaR estimate is equal to 1 minus the significance level. Another way to put it is that VaR is the quantile of a loss distribution, where a

time series of losses (L) is a stochastic variable. Under the assumption of a normal distribution, VaR with a significance level of 95% can be illustrated as the 5% of the left tail, i.e. the losses (Dowd, 2007).
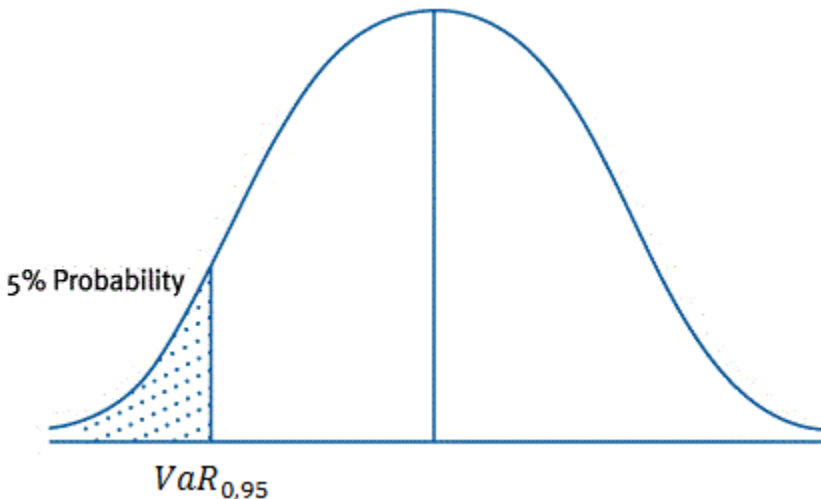

Figure 1.The 95% Value-at-Risk under the assumption of a normal distribution.

Value-at-Risk has become the most used measure for downside financial risk. A large contribution to the popularity of VaR derived from JP Morgan in the late 1980's and their "RiskMetrics" model. Out of many different measures that were developed under the name of RiskMetrics, one was the one day ahead 95% VaR. This was used internally by JP Morgan and replaced an older and more complicated risk measurement system. Due to the influence of JP Morgan in the finance business and the relative simplicity of VaR, it spread externally to other financial institutions, such as the clients of JP Morgan. Following the popularity of the VaR measure, the Basel Committee introduced VaR in regulatory text in 1993. In 1995, it was amended so that banks could use their own VaR models instead of a simplistic VaR version proposed by the Basel Committee. The only provision for using its own VaR measure was that the banks had to get the approval from the regulators (Holton, 2002). There are, however, some drawbacks to VaR as a risk measure. Firstly, it is not sub-additive. The effect of a non-sub additive risk measure is that it does not take diversification into account. In practice, it means that the total VaR of a portfolio can be higher than the sum of the individual VaR's from the assets in the portfolio. Secondly, the Value-at-Risk only looks at the smallest loss at a certain

significance level. The size of the losses in the tail is not considered. Consequently, the potential losses can be underestimated. Because of this, a potential future challenger to VaR as the most popular risk measure might be Expected Shortfall, which takes the average of the tail losses instead of the smallest one.

2.1 Value-at-Risk under normal distribution

The probability density function for the normal distribution with a mean μ and standard deviation σ is:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

By assuming that the losses are normally distributed with a mean μ and a standard deviation σ and using the definition of Value-at-Risk above, an expression for VaR under the normal distribution can be derived:

$$\Pr\big(L > VaR_\alpha(L)\big) = \Pr\left(\frac{L-\mu}{\sigma} > \frac{VaR_\alpha(L)-\mu}{\sigma}\right)$$

$$= \Pr\left(z > \frac{VaR_\alpha(L)-\mu}{\sigma}\right) = 1 - \alpha$$

Therefore:

$$\frac{VaR_\alpha(L)-\mu}{\sigma} = z_\alpha \leftrightarrow VaR_\alpha(L) = \mu + \sigma z_\alpha$$

So, under the assumption of normally distributed losses, the Value-at-Risk can be calculated by using the mean and standard deviation of the stochastic variable at hand. In this case, the stochastic variable is a time series of losses. The z_α is the quantile for the relevant significance level α. For a significance level of 95%, for example, the quantile under the normal distribution is 1,65. When dealing with short time horizons of VaR, such as 1 day ahead or 10 day ahead VaR, the mean is usually assumed to be zero, leaving the VaR to be calculated as:

$$VaR_\alpha(L) = \sigma_{T+1} z_\alpha$$

Where the standard deviation one period ahead, σ_{T+1}, will be estimated with different specifications of GARCH models.

2.2 Value-at-Risk under Student t-distribution

The student t-distribution can be useful when dealing with financial return data, since it accommodates for fatter tails, or larger kurtosis than the normal distribution. The normal distribution has a fixed kurtosis of 3, while the t distribution can have larger kurtosis. The probability density function for the t distribution is:

$$f(x) = \frac{\Gamma[(v+1)/2]}{\sigma\sqrt{(v-2)\pi}\Gamma(\frac{v}{2})}\left[1 + \frac{1}{v-2}\left(\frac{x-\mu}{\sigma}\right)^2\right]^{-(v+1)/2}$$

Similar to the normal distribution, σ is the standard deviation, μ is the sample mean and v represents the degrees of freedom. This specification requires the degrees of freedom to be larger than 2, i.e. for the variance to exist. As a result, the VaR under the student t-distribution is:

$$VaR_\alpha(L) = \sqrt{\frac{v-2}{v}}\,\sigma t_{\alpha,v}$$

Where $t_{\alpha,v}$ is the relevant α-quantile of the t-distribution.

3. Basel III

The Basel III rules are a consequence of the latest financial crisis. As the title of the accord implies, it is the third accord of its kind. The goal of Basel III is to "strengthen global capital and liquidity rules with the goal of promoting a more resilient banking sector" (Basel, 2010). In essence, it means stricter rules for the banking sector, amongst other things a higher capital ratio and larger liquidity buffers. These measures, put in place to minimize the probability of future financial crises, come with a cost. The higher demands on the financial institutions mean

that they take on larger responsibilities and costs, resulting in lower profitability in the banking system. There is also research suggesting that there is a macroeconomic impact of the new rules, reducing GDP growth by about 0.1 per cent annually (Slovik, 2011). It is, in other words, a regulation aimed at stabilizing the financial system at the expense of profits in the financial sector and, to a small degree, economic growth.

There are three types of risk in the Basel framework, named Operational, Credit and Market risk. The parts of the Basel accord that deal with Value-at-risk are related to the capital requirements put on market risk. The capital reserves required by Basel III are 8% of the banks risk weighted assets. For market risk, the risk weighted assets (RWA) are calculated in part by the Value-at-Risk:

$$K = max\left(VaR_{t-1}, m_c * VaR_{avg}\right) + max\left(sVaR_{t-1}, m_s sVaR_{avg}\right)$$

$$RWA = K * 12.5$$

VaR_{t-1} is the VaR for the previous trading day and VaR_{avg} is the average VaR for the previous 60 trading days. sVaR is the stressed VaR and m_c/m_s are multiplicative factors set by the regulators for each specific bank. For less stable banks, the multiplicative factors are higher. When it comes to Value-at-Risk, the previous Basel accords have already stated a requirement of VaR being calculated daily on a 99% level. In addition to this, Basel III introduces a new VaR measure called Stressed VaR. This VaR measure is to be calculated by a one year (250 days) in-sample period of "significant financial stress". There are different ways to choose this period of significant stress; either a judgment-based or a formulaic way. The formulaic approach could, for example, be that the financial institution uses quantitative methods to identify a one year period with the highest volatility. The reason for introducing such a concept is that the banks tend to decrease their coverage during periods of low volatility. The stressed VaR is meant to prevent this from happening in the future, so that the capital reserves of the financial institutions remain on a higher level even in times of low market risk, or volatility.

The banks perform their own calculations of their standard and stressed VaR, and therefore use different methods of calculating their VaR's. The newly introduced concept of the stressed VaR

serves as an inspiration for using an in-sample period with high volatility to estimate VaR for a subsequent period with lower volatility.

4. GARCH Models

4.1 Symmetrical Model

To be able to forecast VaR under the different distributional assumptions, the volatilities of the returns are required. In order to obtain the volatilities, the GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model can be used. Unlike many other time series models, the GARCH model allows for the conditional volatility to change over time, depending on past error terms and conditional variances. There are several variations of the GARCH model, and I will estimate the volatilities using a variety of different GARCH specifications in order to see which one best serves the purpose of estimating Value-at-Risk.

In a paper by Bollerslev (1986), the original model was introduced and formally GARCH(p,q) it can be expressed as:

$$\sigma_t^2 = a_0 + \sum_{i=1}^{q} a_i e_{t-i}^2 + \sum_{i=1}^{p} b_i \sigma_{t-i}^2$$

Where σ_t^2 is the the variance at time t, e_{t-i}^2 represents the squared residuals from previous time periods and σ_{t-i}^2 is the variance from the previous time periods. Furthermore, $a_0 > 0$, $a_1 \geq 0$ and $a_2 \geq 0$. The **p** and **q** determine how many GARCH and ARCH terms respectively that are included in the model specification. The specification states that the variance of time period t is determined by a combination of past variances and past squared residuals of variances. It is a symmetrical model, meaning that the impact of positive and negative movements in the returns is equivalent. This is due to the use of quadratic past residuals, which is a constraint on the model that has been proved to be troublesome empirically. There is evidence that there is an asymmetry in how volatilities react to different types of "shocks". I will therefore compare the traditional GARCH model to a couple of asymmetrical models.

For the GARCH(1,1) process, the specification is given by

$$\sigma_t^2 = a_0 + a_1 e_{t-1}^2 + a_2 \sigma_{t-1}^2$$

The GARCH model can, therefore, be used to provide one period ahead forecasts of volatility, which in turn can be used to estimate the one period ahead Value-at-Risk.

4.2 Asymmetrical Models

4.2.1 Exponential GARCH

A drawback of the GARCH model proposed by Bollerselev (1986) is that only the magnitude of residuals matter, not whether these residuals are positive or negative. This is due to the fact that the residuals are squared. The exponential GARCH (EGARCH) model, proposed by Nelson (1991), does not have this property. Due to its formulation, the variance can react differently to shocks of different signs, i.e. good or bad news. Since there might be an asymmetry in how the volatility reacts to positive and negative shocks, the EGARCH might produce more reliable volatility forecasts. Support of such an asymmetry has been found in the past by, for example, Nelson (1991) and Schwert (1989). These authors noted that when the asset price rises, the volatility tends to decrease and vice versa. I.e., volatility increases more as a result of negative shocks compared to positive shocks. According to Brooks & Persand (2003), there are two different widely used explanations for this. The first explanation is the so called leverage effect, which states that when equity prices fall, the debt to equity ratio increases. Due to the increase, the asset will be perceived to be riskier and therefore becomes more volatile. The second explanation is the volatility feedback hypothesis which states that an increase in volatility leads to a decrease in expected returns. The original EGARCH model proposed by Nelson (1991) is:

$$\ln(\sigma_t^2) = a_0 + \sum_{i=1}^{q} a_i g_i(Z_{t-i}) + \sum_{i=1}^{p} b_i \ln(\sigma_{t-i}^2)$$

The function $g_i(Z_{t-i})$ is defined as:

$$g_i(Z_{t-i}) = a_i Z_{t-i} + \Phi_i(|Z_{t-i}| - E|Z_{t-i}|), \qquad i = 1, \dots, q$$

Where Z_t is a sequence of independent and identically distributed (iid) random variables. The distribution of the sequence therefore has a mean of zero and a variance of one. Due to the specification of $g(Z_t)$, it becomes a function of both the size and sign of Z_t. This is the difference between EGARCH and GARCH. Another difference is that the variance is modeled in logarithmic form. This ensures that the variance will be positive, so no constraints on the parameters in the model are necessary (Brooks, 2008).

4.2.2 Threshold GARCH

Similar to the EGARCH model, the primary difference between threshold GARCH (T-GARCH) and the traditional GARCH model is that it captures the asymmetric effect of different signs on the past residuals on the volatility. There are, however, differences to the EGARCH. One of the differences is that the threshold GARCH has a more variable structure in the lags. In the threshold GARCH model, the asymmetric effect can be different at each lag. This is not the case for the exponential GARCH (Zakoian, 1991). The threshold GARCH model, put forth by Zakoian (1991), uses a specification with a standard deviation instead of variance. Since there are no squared terms, the past residuals can be both positive and negative, thereby capturing the asymmetric effects on volatility. The T-GARCH process is given by:

$$e_t = Z_t \sigma_t$$

$$\sigma_t = a_0 + \sum_{i=1}^{q} a_i^+ e_{t-i}^+ - a_i^- e_{t-i}^- + \sum_{i=1}^{p} b_i \sigma_{t-j}$$

Where e_t denotes a discrete time process, $e_{t-i}^+ = \max(e_t, 0)$ and $e_{t-i}^- = \min(e_t, 0)$. Z_t is an independent and identically distributed variable with mean zero and variance equal to unity.

5. Christoffersen Test

In order to evaluate and compare the performance of the different VaR estimates, some sort of back testing is required. The conditional coverage framework by Christoffersen (1998) is a well-used test, used in similar settings to this paper by Angelidis et al (2003) and Jansky et al (2011). The Christoffersen test is a frequency based test that has an unconditional and a conditional part.

5.1 Unconditional Coverage

The unconditional coverage part of the test measures the frequency of actual VaR violations with the expected frequency of violations. In this case, a VaR violation is defined as when, for any day in our sample, the actual loss is larger than the VaR estimate of the loss. For example, if VaR at a 99% significance level is calculated, the expected number of VaR violations would be 1% of the sample size. If the actual number of VaR violations significantly diverges from 1%, the VaR model is rejected.

This test can be used in a likelihood ratio framework, suggested by Christoffersen (1998). The unconditional coverage part of the test is defined as:

$$LR_{UC} = -2ln\left(\frac{L_0}{L_1}\right) = -2(ln(p^x(1-p)^{N-x}) - ln(\pi^x(1-\pi)^{N-x})\ \sim\chi^2(1)$$

Where N is the number of observations, x is the number of observed violations, p is the expected frequency of violations dictated by the significance level of the VaR estimate and π is the observed frequency of violations.

5.2 Independence test

The independence part of the test deals with independence of the VaR violations. The point of the testing for independence is to rule out models that cluster volatilities. The reasoning behind

the conditional test is that volatility clustering leads to VaR violations being clustered as well. Therefore, the conditional coverage tests whether the VaR violations are independently distributed over the length of the sample. This second part of the Christoffersen test is defined as:

$$LR_{IND} = -2ln\left(\frac{L_0}{L_1}\right) = -2\left[ln\left(\pi_0^{n_0}\pi_1^{n_1}\right) - \ln\left(\pi_{00}^{n_{00}}\pi_{01}^{n_{01}}\pi_{10}^{n_{10}}\pi_{11}^{n_{11}}\right)\right] \sim \chi^2(1)$$

The null hypothesis is that violations and non-violations are independent, and the alternative hypothesis is that the violations follow a two-state Markov chain over time. π_{xx} denotes the probabilities of transitioning states. For example, π_{10} is the probability of moving from state 1 at time t (violation) to state 0 (non-violation) at time t+1. Independence is defined as when the probabilities of moving to state 0 in time period t+1 are equal, independently of what state (1 or 0) we are in at time t. That is; the violations are independent when $\pi_{10} = \pi_{00}$ and $\pi_{01} = \pi_{11}$. If they are not independent, for example if $\pi_{10} \neq \pi_{00}$, then the probability of a non-violation tomorrow is not independent of whether there was a violation today.

The unconditional coverage and independence test are the two components of the Christoffersen test. The conditional coverage test is therefore:

$$LR_{CC} = LR_{UC} + LR_{IND} \sim \chi^2(2)$$

6. Previous Studies

Due to the wide spread popularity of VaR as a risk measure, a lot of research has been done over the last decade. Different authors have tested various ways of estimating variances, over different time periods and data. The results from these studies arrive in diverse conclusions, but with some similarities as well.

Angelidis et al (2004) has tested different ARCH models (GARCH, TARCH, EGARCH) with different distributions and sample sizes to model daily VaR for five stock indices. The indices that provide the daily data are S&P 500, NIKKEI 225, FTSE 100, CAC 40 and DAX 30. This enables the analysis to compare the ARCH models between different geographical markets. Furthermore, the authors test different distributions, namely the normal, t-distribution and the GED distribution. Their conclusions are that leptokurtic distributions outperform the normal distribution in their samples, especially the t-distribution. The most accurate volatility forecasting model in their sample was EGARCH. Finally, they point out that the significance of sample sizes is hard to establish, since different sample sizes yield different results over the data sets.

Jansky et al (2011) conduct similar tests to Angelidis et al, but they test whether volatility models with an in-sample period of lower volatility can be used for forecasting VaR during a period of higher volatility. Six different world stock indices were used during the time period 2004 to 2009. They evaluate the performance of the different models by employing the conditional coverage test by Christoffersen. Jansky et al's conclusion was that the EGARCH or TARCH with a t-distribution or GED distribution fit the in-sample data best. However, when it came to forecasting volatilities, the symmetrical GARCH model performed better. This was along the line of the authors' hypothesis that the symmetrical GARCH would perform better in times of high volatility. One of the major conclusions, therefore, was that the in-sample significance tests can be misleading when it comes to the performance of the models out-of-sample. It also marks the largest deviation from the results of Angelidis et al, who got the result that EGARCH was the preferred model. However, both articles find that the t-distribution is the most valid.

Brooks & Persand (2003) attempted a similar analysis on five Southeast Asian stock indices. Their conclusions were that asymmetric models (unsurprisingly) led to more stable VaR estimates than their symmetric counterparts. They found that symmetric models tended to underestimate VaR on average.

Berkowitz & O'Brien (2002) compared the trading risk models of six large U.S. commercial banks with the outcome of their portfolios to evaluate how well their Value-at-Risk estimations were performing. In addition, the authors also compared the VaR estimates from the commercial banks with their own estimates using a GARCH model. Their findings indicate that the risk models employed by the banks were on the conservative side and overestimating the daily VaR. The GARCH model that the authors used to compare to the internal bank models produced better VaR estimates. This seemed to be because the GARCH model was more flexible in the estimation of volatilities. The authors ascribe these shortcomings of the internal models to limitations due to legislation. The commercial banks need to adhere to legislation regarding risk reporting, and therefore the VaR models suffer since the models become less flexible. For example, time-varying volatilities are not assumed in the structural VaR modeling in the commercial banks. So & Yu (2006) test seven different GARCH models on exchange rates and market indices from the time period 1980 through 1998 and find that the standard GARCH model with the assumption of a t-distribution produces the most accurate estimates of VaR when it comes to long positions. Short positions in market indices, however, are quite surprisingly better estimated when employing the normal distribution. This phenomenon is only visible when it comes to market indices, not exchange rates.

7. Data

7.1 OMX 30

In order to test the different GARCH models and distributions, I have chosen to use data from the Swedish stock market, more specifically from the OMX 30 index, and the FTSE 100, based on the London Stock Exchange. The OMX 30 is comprised of the 30 most traded stocks on the Swedish stock market. The weights of the different companies are decided through capital weighting, which means that the market capitalization of each company in the OMX 30 is used as a relative weight to determine its size in the index portfolio. The base date for the index is September 30, 1986. All the data employed in this study is daily and comprised of closing prices.

Due to the purpose of this paper, two different samples are required; one where the in-sample period has high volatility and one where the in-sample period has low volatility. The sample period I have chosen for the in-sample period with low volatility and out-of-sample period with high volatility (from now on called low-to-high) ranges from January 1st 2004 through January 12th 2010. The number of in-sample observations is 949. The other sample with a high in-sample period volatility (high-to-low), ranges from the 1st of November 2006 through the 28th of March 2013. The in-sample period in this sample consists of 1045 observations. The table below illustrates some descriptive statistics, such as the distributional properties, of the log return series.

	Sample size	Mean	St. Dev.	Skewness	Kurtosis	Jarque-Bera
High-to-Low	1610	7.78E-05	0.0166	0.10	6.58	861.9
-In-sample	1045	8.28E-05	0.0177	0.22	6.57	564.0
-Out-of-sample	565	6.85E-05	0.0144	-0.32	5.53	160.3
Low-to-High	1512	0.00308	0.0155	0.12	7.93	1537.9
-In-sample	949	7.41E-04	0.0106	-0.49	5.75	336.8
-Out-of-sample	563	-4.23E-04	0.0213	0.31	5.40	143.5

Table 1: Descriptive statistics for OMXS 30 sample returns in logarithmic form.

The skewness, kurtosis and, by consequence, the Jarque-Bera statistics imply that the two return series do not seem to be normally distributed. This gives further reason to think that distributions with thicker tails such as the student's t-or the GED distribution might be more appropriate in the estimation of the volatilities. Each forecasted volatility for time t+1 will be estimated at time t with the same in-sample size over time. For each time period, the first in-sample observation will be dropped, and the t+1 observation will be added. Therefore, the rolling regression technique means that the data will produce 565 and 563 volatility forecasts respectively for the two samples. This, in turn, will lead to the same number of one-day-ahead Value-at-Risk estimates.
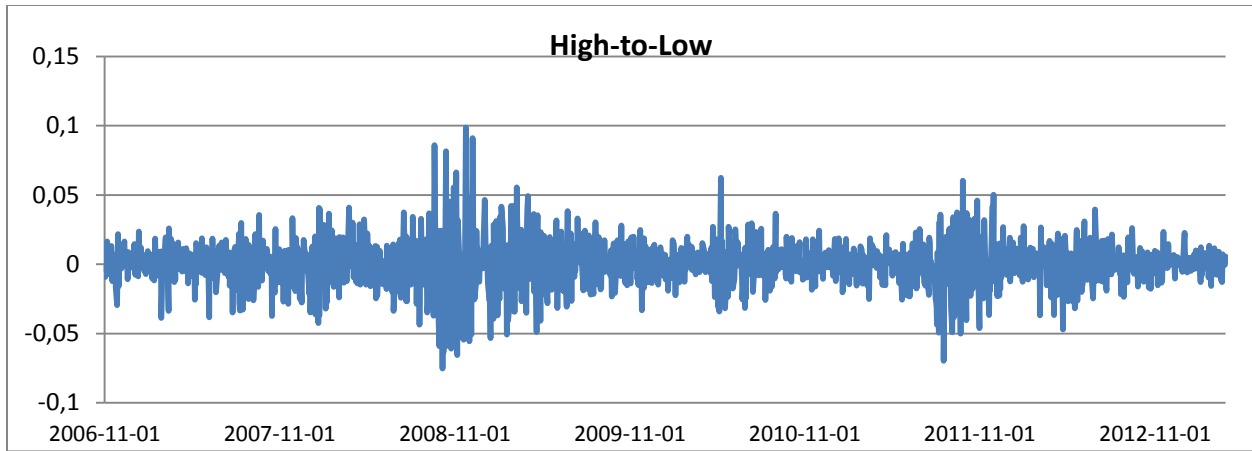
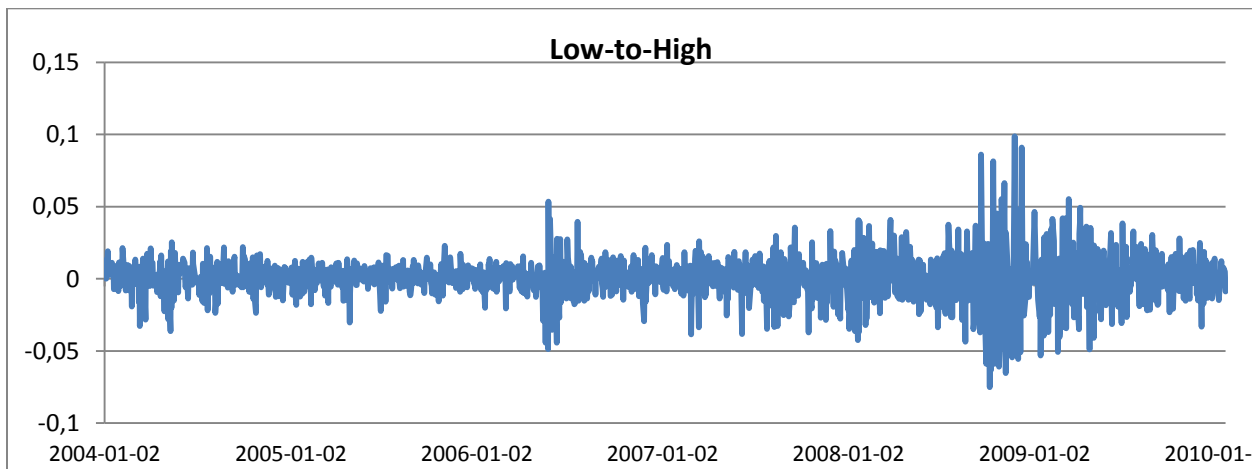*Figure 2: The High-to-Low time series of log returns for the OMX 30 sample, from November 1st 2006 through March 28th 2013.*

 Figure 3: The Low-to-High time series of log returns for the OMX 30 sample, from January 1st 2004 through January 12th 2013.

The two figures above illustrate the two different sample returns. The returns are in logarithmic form. Differences in the volatility between the in-sample and out-of-sample periods are quite visible by looking at the time series.

7.2 FTSE 100

The FTSE 100 is an index containing a weighted share of the 100 largest companies on the London Stock exchange (measured by market capitalization). Just as the OMX 30 index, it is weighted using the size of market capitalization, so that the larger companies in the index are

given a larger weight. The base date for the index is January 3, 1984 with a starting level of 1000. All the data employed in this study is daily and is comprised of closing prices.

Just as for the OMX30-sample, I have constructed two different samples. The first sample (high-to-low) has an in-sample period with high volatility and an out-of-sample period with a low volatility. This sample has a total of 1535 observations, and the time period is from March 1st 2007 through March 28th 2013. The second sample (low-to-high) has the opposite properties, i.e. an in-sample period of low volatility and an out-of-sample period with high volatility. The total number of observations in this sample is 1306 and the time period is from February 2nd 2004 through March 31st 2009.

	Sample size	Mean	St. Dev.	Skewness	Kurtosis	Jarque-Bera
High-to-Low	1535	2.49E-05	0.0145	-0.1	9.2	2459.1
-In-sample	906	-1.17E-04	0.0166	-0.05	8.44	1115.6
-Out-of-sample	629	2.30E-04	0.0107	-0.25	4.95	106.3
Low-to-High	1306	-8.56E-05	0.0133	-0.14	13.5	6047.3
-In-sample	801	4.52E-04	0.0069	-0.42	4.7	119.7
-Out-of-sample	505	-9.39E-04	0.0196	0.03	7.41	410.6

Table 2: Descriptive statistics for FTSE 100 sample returns in logarithmic form.

Just as the OMX 30 sample, the different time series for the FTSE index do not seem to be normally distributed when looking at the Jarque-Bera statistics. The method for calculating one-day ahead VaR is the same here as for the OMX30 index, namely to do a rolling regression with the in-sample size constant. In total, 14 118 regressions will be carried out to get the corresponding number of volatilities.

The two figures below illustrate the two time series log returns for the FTSE 100 index for the two samples, high-to-low and low-to-high.
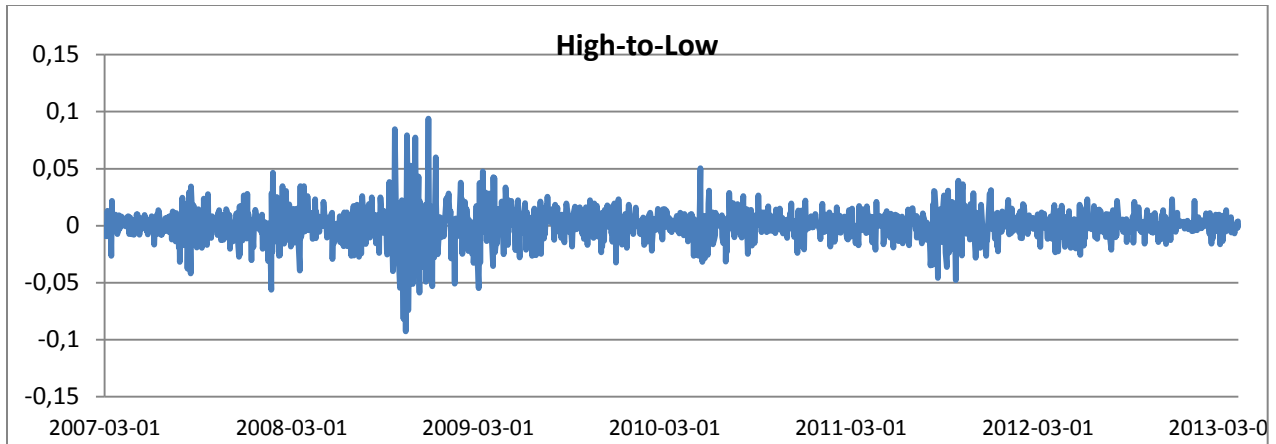
Figure 4: The High-to-Low time series of log returns for the FTSE 100 sample, from March 1st 2007 through March 28th 2013.
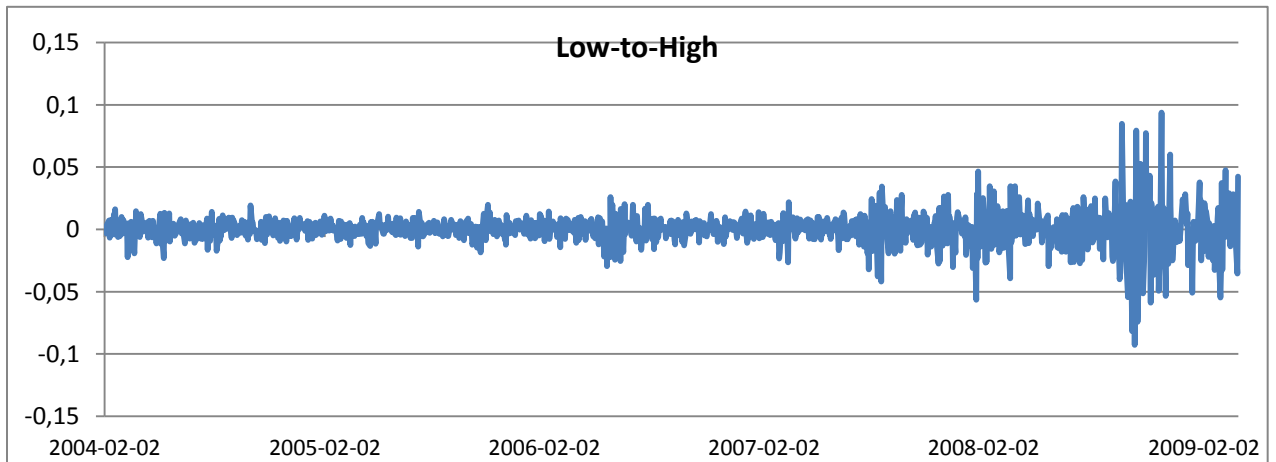

Figure 5: The Low-to-High time series of log returns for the FTSE 100 sample, from February 2nd 2004 through March 31st 2009.

The figures above are daily log returns of the FTSE 100 index. The differences in volatility between the early half of the sample and the later part are visible by simply looking at the time series.

8. Results

8.1 GARCH Estimates

The figures below illustrate the conditional one-day-ahead standard deviations produced by the different GARCH specifications for all samples.

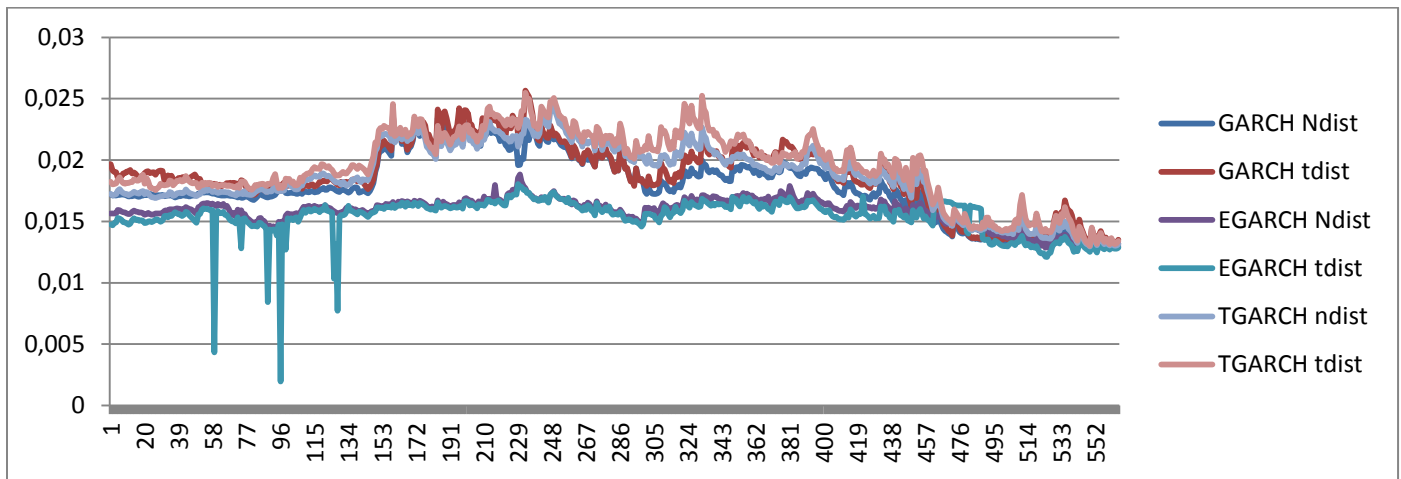### 8.1.1 High-to-Low samples

Figure 6: Estimated standard deviations (High-to-Low) for the OMX 30 sample. Volatilities estimated with EGARCH are lower than the other models.
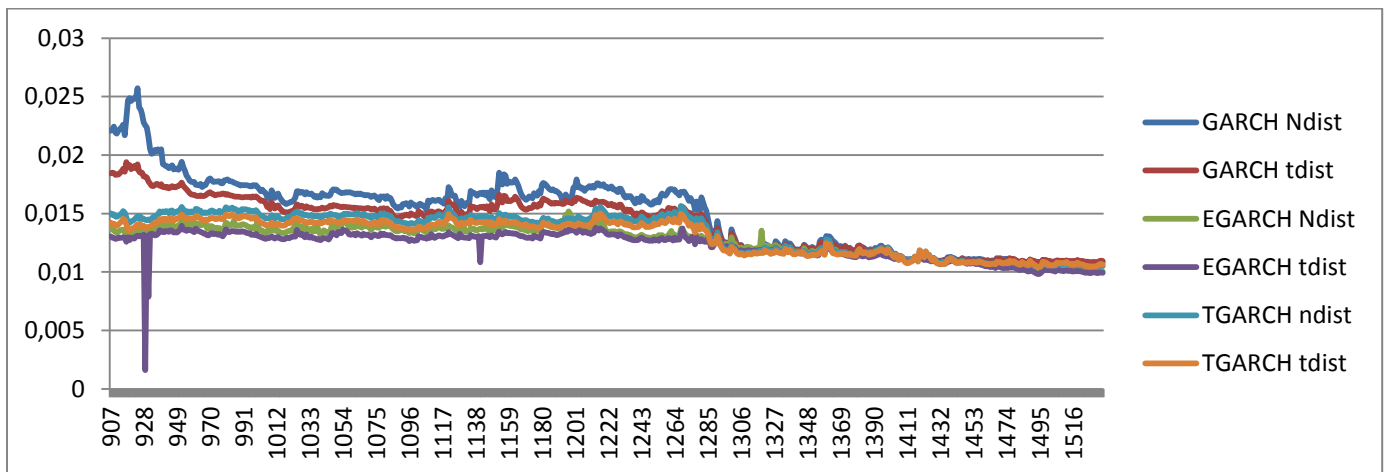

Figure 7: Estimated standard deviations (High-to-Low) for the FTSE 100sample. The standard GARCH model produces higher volatilities than the other models.
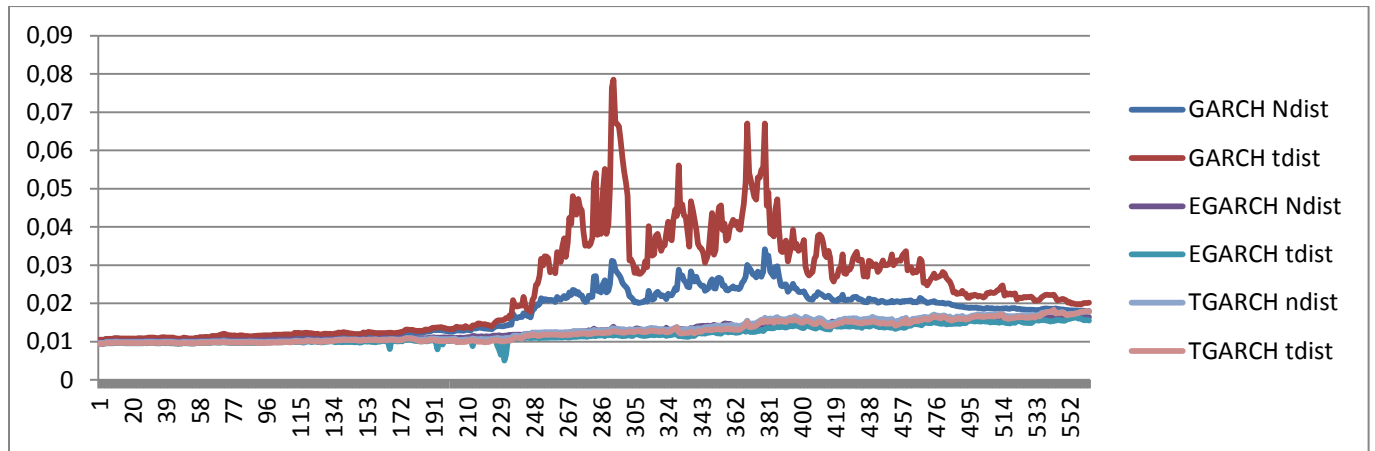
8.1.2 Low-to-High Samples

Figure 8: Estimated standard deviations (Low-to-High) for the OMX 30 sample. The standard GARCH model produces higher volatilities than the other models.
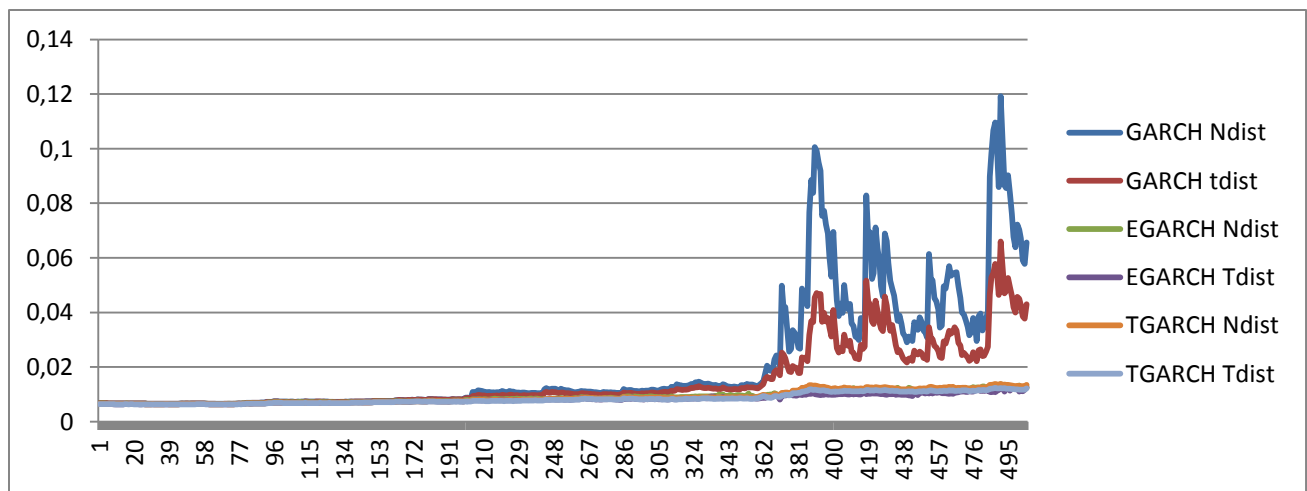
Figure 9: Estimated standard deviations (Low-to-High) for the FTSE 100 sample. The standard GARCH model produces higher volatilities than the other models.

The High-to-Low samples produce more homogenous estimates of standard deviations compared to the Low-to-High samples. For the OMX30 sample, the exceptions in the high-to-low sample, the two EGARCH series, produce lower standard deviations compared to the rest of the models. This is especially true in the early part of the forecasted sample, and it seems to converge with the other model estimates in the tail-end of the sample. This may be due to the fact that the EGARCH model uses a constant lag structure where the lagged asymmetries are

treated the same for each lag, compared to the threshold GARCH model where this is not the case. In general, the Low-to-High sample generates lower estimations of standard deviations, with the exception of the GARCH model with t-distribution and normal distribution. The GARCH model with t-distribution is a notable exception, since the standard deviations are significantly higher than the rest of the model estimates. In general, the standard GARCH models seem to be more explosive. That is; the GARCH models estimate higher volatilities than the two asymmetric models. It is especially visible in the low-to-high samples. This is probably due to the fact that it does not take asymmetries into account. Both positive and negative shocks are treated the same in the standard GARCH models, compared to the EGARCH and TGARCH models where positive and negative shocks do not have the same impact on the estimated volatilities. This results in smoother estimates over time with the asymmetric models compared to the symmetric one (standard GARCH).

8.2 High-to-Low sample

8.2.1 OMX 30

As expected for the High-to-Low sample, the actual number of violations are underestimated. The expected number of violations for the 5% VaR is 28. The corresponding number of expected violations for the 1% level is 6. Since the in-sample period has higher volatility, the VaR estimates are likely to be higher. Since the definition of a violation is that the actual loss exceeds the VaR, the number of violations produced is smaller than the expected number. None of the models passes the Christoffersen test. For some of the models, it is because it fails the unconditional coverage test and for the rest, it is because of a failure to pass the independence test. The reason that so few models pass the independence test is, again, because of the scarce occurrence of violations. For a series to be independent, the probability of violations in succession of each other has to be similar to the probability of successive non-violations. This is not the case for the High-to-Low sample.

In the case of 1% VaR, all six model specifications pass the unconditional (Kupiec) test. The Christoffersen test, however, rejects independence in all the models. This is mainly due to the

fact that there are no consecutive days with violations in the sample, and therefore independence is rejected.

	Normal Distribution			T-Distribution		
	GARCH	EGARCH	TGARCH	GARCH	EGARCH	TGARCH
5% VaR						
Violations	14	19	13	13	23	11
Unconditional Coverage test	9.22	3.58*	10.75	10.75	1.09*	14.30
Independence test	0.88*	5.20	123.77	1.09*	24.89	108.44
Christoffersen test	10.09	8.79	134.52	11.84	25.98	122.74
1% VaR						
Violations	5	9	3	3	9	2
Unconditional Coverage test	0.08*	1.70*	1.51*	1.51*	1.70*	3.17*
Independence test	57.23	92.37	37.41	37.41	92.37	26.57
Christoffersen test	57.31	94.07	38.93	38.93	94.07	29.74

Table 3: Christoffersen test for OMX 30 High-to-Low sample. For the 95% VaR, the two EGARCH models produce significant VaR estimates when it comes to the unconditional coverage test.

8.2.2 FTSE 100

The expected number of violations for the 1% and 5% VaR are 6 and 31 respectively. The number of actual violations is lower than what's expected for all models and significance levels. This is not surprising, since the in-sample has higher volatility than the out-of-sample period. This results in high VaR estimates and therefore few violations. However, the EGARCH model with a student's t-distribution has a significant unconditional coverage test on the 5% VaR level. This is in line with the results for the OMX30 sample. So in two completely different samples, the t-distributed EGARCH model has produced a significant result when it comes to forecasting VaR with high in-sample volatility for both samples. It should be noted that results are significant, but the number of violations is underestimated in this model as well, just to a lower extent than the other models.

	Normal Distribution			T-Distribution		
	GARCH	EGARCH	TGARCH	GARCH	EGARCH	TGARCH
5% VaR						
Violations	12	20	16	15	23	20
Unconditional Coverage test	16.40	5.01	9.67	11.14	2.63*	5.01
Independence test	7.90	12.85	6.65	6.88	15.29	12.85
Christoffersen test	24.31	17.86	16.32	18.02	17.91	17.86
1% VaR						
Violations	2	5	4	2	5	4
Unconditional Coverage test	4.03	0.29*	0.97*	4.03	0.29*	0.97*
Independence test	27.00	58.31	48.44	27.00	58.31	48.44
Christoffersen test	31.02	58.59	49.40	31.02	58.59	49.40

Table 4: Christoffersen test for FTSE 100 High-to-Low sample. For the 95% VaR, the EGARCH model with a t-distribution produces significant VaR estimates when it comes to the unconditional coverage test.

8.3 Low-to-High sample

8.3.1 OMX 30

The table below illustrates the results of the unconditional coverage, independence and Christoffersen tests for the Low-to-High sample. One model specification pass the unconditional coverage test for the 5% VaR, namely the standard GARCH(1,1) with a t-distribution. This model specification is also the only one to pass the Christoffersen test of all the different models. For the 1% VaR, the only model to pass the unconditional coverage test is, again, the GARCH with a t-distribution. None of the models passes the Christoffersen test. However; 7 different models pass the independence test, compared to only two models in the High-to-Low sample. This is in line with the reasoning above, namely that the Low-to-High sample produces lower VaR estimates compared to the High-to-Low sample. This results in a possible underestimation of VaR in this sample. Therefore, the probability of consecutive VaR violations is higher in this sample. Due to the higher share of consecutive VaR violations, the independence test is more likely to be accepted. The reason that the GARCH model with the t-

distribution works so well is the fact that the estimated one day ahead variances are higher than what the other models estimated,which results in fewer VaR violations. Since this is the Low-to-High sample, the violations produced by the GARCH estimations are generally too numerous (the same reason the independence test is accepted more often in this sample). The expected number of violations for the 5% VaR is about 28, and the corresponding number for the 1% VaR is about 6.

	Normal Distribution			T-Distribution		
	GARCH	EGARCH	TGARCH	GARCH	EGARCH	TGARCH
5% VaR						
Violations	49	74	77	37	47	75
Unconditional Coverage test	13.44	55.39	61.87	2.68*	11.16	57.52
Independence test	0.49*	0.01*	0.03*	0.09*	0.28*	0.00*
Christoffersen test	13.94	55.40	61.90	2.77*	11.43	57.52
1% VaR						
Violations	26	46	48	11	48	45
Unconditional Coverage test	39.57	115.51	124.30	4.04*	124.30	111.18
Independence test	210.69	0.02*	0.00*	108.36	0.95*	0.12*
Christoffersen test	250.26	115.52	124.31	112.41	125.25	111.30

Table 5: *Christoffersen test for OMX 30 Low-to-High sample. The GARCH model with a t-distribution produces significant unconditional coverage tests for both 95% and 99% VaR.*

So it can be said that the reason that models pass the independence test is the same reason that they do not pass the unconditional coverage test in the samples analyzed above. If the number of violations exceeds what is expected, then the unconditional coverage test is rejected. But it also means that the probability of consecutive violations is higher. Due to how the independence test is designed, the probability of consecutive violations has to be similar to the probability of consecutive non-violations. With very few violations (leading to an acceptance of the Kupiec test), the probability of consecutive violations is very low, while the probability of consecutive non-violations is very high (and vice versa).

8.3.2 FTSE 100

For the low-to-high sample using data from the FTSE 100, no model on any significance level passes the unconditional coverage test. The expected number of violations for the 1% and 5% level are 5 and 25 respectively. As can be seen in the table below, the violations are far too many to yield a significant result. This is due to the fact that the different GARCH models have underestimated the volatilities. The model closest to being significant is the standard GARCH model with a normal distribution. This is in line with the previous analysis of the estimated volatilities. The standard GARCH seems to be more explosive and yield higher volatility estimates compared to the other models. In this case, where the volatilities are under-estimated, this becomes a positive effect. The same effect is visible in the OMX 30 sample, where the GARCH model with a t-distribution has a significant unconditional coverage test.

	Normal Distribution				T-Distribution		
	GARCH	EGARCH	TGARCH		GARCH	EGARCH	TGARCH
5% VaR							
Violations	47	92	94		56	96	95
Unconditional Coverage test	15.91	114.16	119.98		29.725	125.908	122.931
Independence test	0.58*	0.28*	0.20*		0.31*	0.43*	0.72*
Christoffersen test	16.48	114.44	120.18		30.04	126.342	123.655
1% VaR							
Violations	21	56	55		24	58	57
Unconditional Coverage test	28.47	172.95	167.94		37.64	183.08	177.99
Independence test	174.68	0.61*	0.79*		174.68	0.61*	0.79*
Christoffersen test	203.15	173.56	168.73		203.15	173.56	168.73

Table 6: *Christoffersen test for FTSE 100 Low-to-High sample. No models produce a significant unconditional coverage test.*

The independence tests are mostly significant, due to the same reasons as for the OMX 30 returns. With the actual number of violations being higher the expected number of violations, the probability of consecutive violations increase and the independence test is accepted. This, however, does not indicate much without a significant and unconditional coverage test result.

9. Conclusion

By utilizing the different GARCH models to estimate volatilities and consequently VaR estimates with two different data sets, some patterns have emerged. For the two time series where the in-sample period has high volatility and the out-of-sample period has low volatility, the EGARCH model seems to perform best at estimating Value-at-Risk when looking at the unconditional coverage (Kupiec) test. For the OMX30 sample, it yields significant results both with the normal and the t-distribution. For the FTSE 100, the EGARCH with t-distribution is significant. The reason that EGARCH performs better than the other models is that it produces lower estimates of volatility, whereas the standard GARCH estimates of volatility are more sensitive to the in-sample shocks.

For the two time series where the in-sample period has low volatility and the out-of-sample period has high volatility, the converse relationship seems to be true. That is, the standard GARCH model produces the best results according to the unconditional coverage test. The only significant VaR estimate comes from the standard GARCH model with a t-distribution on the OMX30 sample. The same explanation applies to the low-to-high samples as it does to the high-to-low samples: the standard GARCH model seems to produce higher volatility estimates than the asymmetric counterparts. Since, generally, the VaR estimates using the low-to-high sample underestimates the VaR, this results in more accurate VaR forecasting.

The reason for the lower volatility estimates produces by the EGARCH compared to the standard GARCH model is, as mentioned before, the asymmetric property of the EGARCH model. This model takes into account whether the shocks are positive or negative and that affects the magnitude of the volatility estimate. This is not true for the standard GARCH model, where shocks of both positive and negative signs are treated the same, therefore resulting in higher volatility estimates.

The consequence of these models's behavior is that the standard GARCH model is more fitting in an environment where it is expected that the volatility of the returns will increase in the future. If that is the case, then more accurate VaR estimates will be obtained by using the

standard GARCH which produces higher volatility estimates. If the conditions are the converse, i.e. the volatility in the market is expected to decrease in the future, then the EGARCH with a t-distribution will be a better model to employ, since the estimated volatilities will be lower and more fitting to calculate future VaR's.

10. References

10.1 Primary Data

Nasdaq OMX Nordic price database

http://www.nasdaqomxnordic.com/indexes/historical_prices/?Instrument=SE0000337842

FTSE 100 historical price data

http://uk.finance.yahoo.com/q/hp?s=%5EFTSE&b=1&a=00&c=2004&e=28&d=02&f=2013&g=d

10.2 Secondary Data

Dowd, Kevin (2007), "Measuring Market Risk", Second Edition, Wiley & Sons

Angelidis, Timotheos, Benos, Alexandros & Degiannakis, Stavros (2004), "The use of GARCH models in VaR estimation", Statistical Methodology 1 (2004) 105-128

Basel Committee on banking supervision (2010), "Basel III: A global regulatory framework for more resilient banks and banking systems"

Berkowitz, Jeremy & O'Brien, James (2002), "How Accurate Are Value-at-Risk Models at Commercial Banks?", The Journal of Finance, Vol. 57, No. 3, June 2003

Bollerslev, Tim (1986), "Generalized Autoregressive Conditional Heteroskedasticity", Journal of Econometrics 31, pp. 307-327, North Holland

Brooks, Chris (2008), "Introductory Econometrics for Finance", Second Edition, Cambridge University Press

Brooks, Chris & Persand, Gita (2003), "The Effect of Asymmetries on Stock Index Return Value-at-Risk Estimates", The journal of risk finance, Vol. 4 Iss: 2, pp.29 – 42

Christoffersen, Peter F. (1998), "Evaluating Interval Forecasts", International Economic Review, Vol. 39, No. 4, Symposium on Forecasting and Empirical Methods in Macroeconomics and Finance (Nov. 1998) pp. 841-862

Holton, Glyn A. (2002), "History of Value-at-Risk: 1922-1998" Working paper. Contingency Analysis, Boston, Massachusetts.

Jansky, Ivo & Rippel, Milan (2011), "Value at Risk Forecasting With the ARMA-GARCH Family of Models in Times of Increased Volatility", IES Working Paper 27/2011. IES FSV. Charles University

Nelson, Daniel B. (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach", Econometrica, Vol. 59, No. 2 (Mar. 1991) pp. 347-370

Schwert, G. William (1989), "Why Does Stock Market Volatility Change Over Time?", The journal of finance, Vol. 44, No. 5

Slovik, Patrick & Cournède, Boris (2011), "Macroeconomic Impact of Basel III", OECD Economics Department Working Papers No. 844 OECD Publishing. http://dx.doi.org/10.1787/5kghwnhkkjs8-en

So, Mike K.P. & Yu, Philip L.H. (2006), "Empirical Analysis of GARCH models in value at risk estimation", Journal of International Financial markets, Institutions & Money, Vol 16, pp. 180-197

Zakoian, Jean-Michel (1991), "Threshold Heteroskedastic Models", Journal of Economic Dynamics and Control 18, (1994), pp. 931-955, North Holland