



LUND UNIVERSITY
DEPARTMENT OF ECONOMICS

BACHELOR THESIS

Practical estimation of Value at Risk and Expected Shortfall: Are complex methods really necessary?

Johannes SOLHEIM KARLSSON
Henning ZAKRISSON

supervised by
Birger NILSSON

May 30, 2016

Abstract

This paper tests the parametric estimation method for Value at Risk and Expected Shortfall estimation together with the historical simulation method to find out if the historical simulation could yield accurate enough estimations in stormy and calm periods. Given that the parametric estimation proved superior, the thesis examines which volatility forecasting models, using which distribution assumptions, would yield the best estimations. To test this, six different GARCH, two different EWMA and two different historical simulation models were examined. Together, all tests were conducted on 10 of the worlds largest stock indices for their relevance in index investing, on two different periods of varying financial stability. The results showed that the EWMA, especially the Gaussian EWMA, consistently gave satisfying results in both the crisis and post crisis period, while the one year-HS also yielded acceptable results in the post crisis period. Other models yielded disappointing results compared to the simpler EWMA model. To answer the initial question: parametric estimation, with the EWMA model, is clearly superior to historical simulation in both stormy and calm periods, though historical simulation yielded acceptable results in calmer periods. Value at Risk and Expected Shortfall estimation should thus be conducted with parametric estimation using the Gaussian distribution EWMA model for all periods. Although, if simplicity is highly regarded by the estimating individual, historical simulation can be used in periods of high financial stability.

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 1 |
| 2 | Theory | 4 |
| 2.1 | Financial asset returns | 4 |
| 2.1.1 | Log-returns | 4 |
| 2.1.2 | Leptokurtic distribution | 4 |
| 2.1.3 | Volatility clustering | 5 |
| 2.1.4 | The Leverage effect | 6 |
| 2.2 | Financial risk measures | 6 |
| 2.2.1 | Value at risk | 6 |
| 2.2.2 | Expected shortfall | 7 |
| 2.3 | Historical simulation | 8 |
| 2.4 | Parametric estimation | 9 |
| 2.5 | Volatility forecasting models | 10 |
| 2.5.1 | GARCH | 10 |
| 2.5.2 | GJR-GARCH | 11 |
| 2.5.3 | EGARCH | 11 |
| 2.5.4 | EWMA/Riskmetrics | 12 |
| 2.6 | Backtesting | 12 |
| 2.6.1 | Backtesting Value at Risk | 12 |
| 2.6.2 | Backtesting Expected shortfall | 13 |
| 3 | Data & Method | 14 |
| 3.1 | Data | 14 |
| 3.2 | Programming | 15 |
| 3.2.1 | GARCH-estimation | 15 |
| 3.2.2 | Choice of models | 16 |
| 3.2.3 | Value at Risk estimation | 16 |
| 3.2.4 | Expected Shortfall estimation | 16 |
| 3.3 | Ranking of models | 17 |
| 4 | Results | 18 |
| 4.1 | Estimated coefficients | 18 |
| 4.2 | Value at Risk | 19 |
| 4.2.1 | Crisis | 19 |
| 4.2.2 | Post crisis | 20 |
| 4.3 | Expected shortfall | 21 |
| 4.3.1 | Crisis | 21 |
| 4.3.2 | Post crisis | 22 |
| 4.4 | Discussion | 23 |
| 5 | Conclusion | 25 |
| | Appendices | 28 |
| A | MATLAB-code | 28 |
| A.1 | Sample statistics | 28 |
| A.2 | One index prediction | 29 |
| A.3 | Multiple index test | 38 |

1 Introduction

Measuring and controlling financial risk has always been important, for both firms and individuals. For firms, proper risk management can, among other things, increase the value of the firm by decreasing the risk of bankruptcy, minimize tax payments by making revenue streams more constant and lower the cost of capital by making debt servicing more secure (Christoffersen, 2011, p. 4-5). A common method of illustrating financial risk is the variance or volatility of the asset. This is a non intuitive measure though, as it only provides an arbitrary number and does not differentiate between positive and negative movements in stock prices. More relevant information for investors is the risk associated with a fall in the stock price. For this purpose there are two favored measures - *Value at Risk* and *Expected Shortfall*, both of which will be evaluated in this thesis.

Value at Risk, henceforth referred to as *VaR*, has been, and continues to be, one of the most popular measures for financial risk due to its understandability (Hull, 2006, p. 472) and its attribute of highlighting the risk of loss and not the chance of gain but has its shortcomings when it comes to its usage as a risk management tool. For this reason, the Basel Committee has decided to phase out the *VaR* measure in favor of the Expected Shortfall, from now on referred to as *ES*, when measuring market risk (Basel Committee on Banking Supervision, 2013, p. 3). The measure gives information about the *tail risk*, meaning the risk in cases where the loss is far from the mean. *ES* gives us a value of expected loss in extreme cases whereas *VaR* only gives information about the threshold value. However, the most common way of estimating these important measures, *historical simulation* (from now on referred to as HS), often yields results which are largely outperformed by more complex methods (Mentel, 2003), (Christoffersen, 2011, p. 32). Better measures for the estimation of *VaR* and *ES* might be granted by so called *parametric estimations* using forecasted volatility or through Monte Carlo-simulation.

There is plenty of earlier research in the field of *VaR* and *ES* estimation. As a result, several estimation methods have been proposed. The major ones are HS, parametric estimation and Monte Carlo-simulation (Bohdalová, 2007). Due to the reported inferiority of the HS method compared to more complex methods (Mentel, 2003), (Christoffersen, 2011, p. 32), an interesting topic is whether investors should stick with the simpler HS approach or invest the time to make more complex estimations. The focus of this thesis will therefore be on comparing the performance of HS to the more complex parametric estimations for *VaR* and *ES* estimation. For the parametric estimations of *VaR* and *ES* we will use the out of sample conditional volatility and the expected loss as parameters. Since we will consistently use an exponentially weighted mean as our expected return, a value that is close to zero in this application, the accuracy of the volatility forecast is what will make the difference in estimation accuracy of *VaR* and *ES*. Another question that therefore arises is which volatility forecasting model should be used and with what specifications.

Like *VaR* and *ES* estimation, volatility forecasting has been heavily studied. As for volatility

forecasting models, Engle (1982) was the first to consider using the observed heteroskedasticity of asset volatility in a model when he proposed the ARCH model. Since then, more models with varying complexity has been proposed such as the univariate and multivariate GARCH-type models, stochastic volatility models as well as simpler EWMA models (Brooks, 2014). Of the proposed models, the univariate GARCH-types, including the EWMA, are among the most commonly applied in practice (Brooks, 2014, p. 428). Even though some of the most commonly used GARCH specifications were proposed 1986, 1991 and 1993, they still hold relevance in volatility forecasting today (Bollerslev, 1986), (Glosten et al., 1993), (Nelson, 1991). For example, Hansen and Lunde (2001) found that no conditional volatility forecasting model, out of 330, produced significantly better results than the GARCH(1,1) when testing on DM/\$ exchange rate data and IBM stock prices. However, when updating their research method by basing their IBM return data on a data set of realised volatility, they found that GARCH(1,1) was significantly outperformed by its asymmetric cousins (Hansen and Lunde, 2005). Yet, Hansen and Lunde drew no conclusions of which model was, overall, the most suitable in either one of their papers, suggesting instead that different models do not fit different data sets equally well (Hansen and Lunde, 2001). Sethapramote et al. (2014) came to a similar conclusion when researching which model best forecast *VaR* in the Thailand stock market index, SET50. Sethapramote found that asymmetric GARCH models such as the EGARCH outperformed even more complex models.

If we can then assume that different models are suitable for different data sets, another reasonable assumption is that the different models are suitable for different time periods. We will therefore look at two different periods, the first being a stormy period with high mean volatility represented by the financial crisis (December 2007 - June 2009) and the second being a calmer period represented by the post crisis period (July 2009 - December 2015). The dates were chosen as the National Bureau of Economic Research (2010) defines December 2007 to June 2009 to be the peak to trough for the financial crisis. Furthermore, there seems to be conflicting evidence regarding whether a Gaussian normal distribution (which will be referred to as Gaussian in this thesis) or a Student's t-distribution should be assumed. Bollerslev (1987) proposed to use the Student's t-distribution based on the theoretical observation that financial asset returns tend to have fatter tails than the Gaussian distribution. On the contrary, Brownlees et al. (2011) found no evidence that Student's t-distributions performed better estimations than did Gaussian distributions. Following our previous assumption that different financial data sets have different characteristics, we will test both of these distributions to leave open the possibility that one works better on average than the other.

All tests will be conducted on 10 world stock indices, where 5 will be taken from developed markets and 5 from developing or emerging markets including China in order to capture financial data set characteristics from a large part of the world.

Due to the width of the field as well as our aim to keep the methods reasonably easy to grasp

and applicable to both individuals and firms, a few delimitations had to be made. First, even though there are many important risks to consider in financial markets, we will only focus on market risk in this thesis. This is mainly due to the scope of the thesis as well as the applicability and relevance of market risk for index investing. Second, even though there are other popular forms of investing for firms and individuals, the models will only be tested on stock indices. This is because of its direct relevance to a lot of people as a contrast to individual stocks, where the model accuracy is relevant to fewer individuals. Third, we will only consider a limited number of GARCH models, as well as the two simpler HS alternatives, despite the abundance of alternatives, since all the presented methods are widely used and heavily studied. An expansion of this study with left-behind models, methods and model specifications is left for future studies. Fourth, we will not consider the VaR and ES that is resulted from so called short sales. In other words, when an individual shorts an asset, the risk of loss is not associated with the asset losing value, as this would increase the value of the position. The risk of loss is associated with the asset gaining in value. The tail with interest for estimation is therefore the right tail. It is very possible that the same models that work on left tail risk will not work as well on right tail risk. However, as long positions are more common among individuals than are short positions, as well as being beyond the scope of this thesis, it will not be considered further. Finally, a promising way to estimate VaR and ES is through Monte Carlo-simulation. However, this will also be left out due to it being beyond the scope of the thesis.

The remainder of the thesis is structured as follows: Chapter two regards the underlying theory for the modelling and estimation. It covers all theoretical aspects that are needed to understand our hypotheses and reasons for conducting the research with the current method. The third chapter regards the data and method used in the testing. It covers what data we are going to use and the practical approach we chose to conduct the research. The fourth chapter presents the results from the estimations and tests as well as a discussion about. The fifth chapter concludes the thesis and presents what could be conceived by the estimations and tests, and also our recommendations to investors.

2 Theory

2.1 Financial asset returns

Financial asset returns have several specific characteristics that will shape our assumptions and choices regarding the models and specifications we are going to use.

2.1.1 Log-returns

When analysing the data of a financial time series, log returns are the preferred method of analysis for their property of being time additive (Brooks, 2014, p. 8). The mathematical definition for log return r at time t can be seen in Equation 1, where P_t is the close price at time t , adjusted for dividends and splits, t refers to a specific market day and $t - 1$ to the previous market day. For the remainder of the thesis we will use r as the measure of the logarithmic returns, and simply refer to it as *returns*.

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (1)$$

2.1.2 Leptokurtic distribution

An observed phenomenon regarding financial asset returns is that they tend to follow distributions that are more *leptokurtic* than the standard Gaussian distribution. This means that the return distributions are more concentrated around their means, μ , and have fatter tails than the standard Gaussian distribution (Brooks, 2014, 416). To combat this problem when analysing financial time series data, Bollerslev (1987) suggested the use of the Student's t-distribution. However, as can be seen in Figure 1, where a Gaussian distribution and a Student's t-distribution with 3 degrees of freedom can be observed, it is apparent that the Student's t-distribution has fatter tails, but is less concentrated around its mean. Neither of the distributions therefore perfectly reflect the theoretically observed leptokurtosis in financial asset returns.

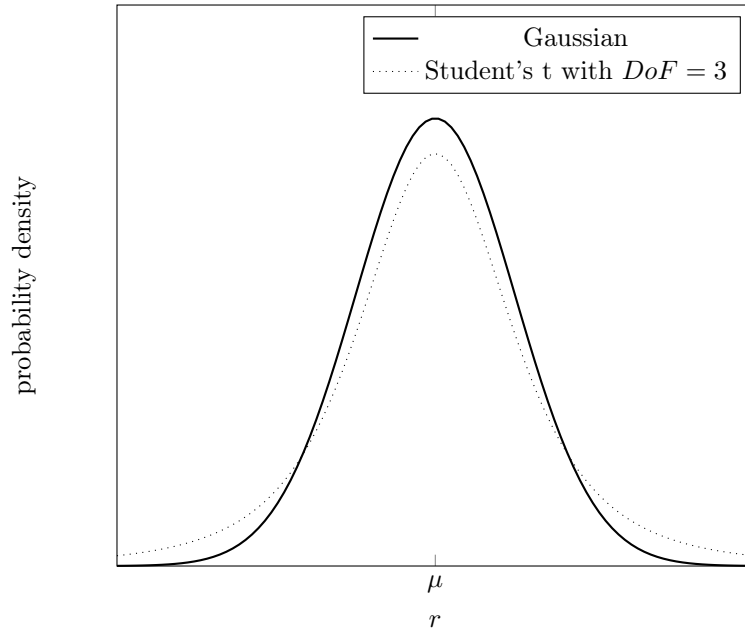


Figure 1: Gaussian- and Student's t-probability distributions

2.1.3 Volatility clustering

Another observation is that of *volatility clustering*, which means that following a period of high volatility, subsequent period's volatility tend to be high as well (Brooks, 2014, p. 416). An example of this can be seen in Figure 2, where the heteroskedasticity of the data can clearly be observed. This has implications for our model selection. For example linear models, such as the historical volatility model for volatility forecasting and the HS approach to *VaR* and *ES* estimation, implicitly assume that the conditional volatility depends as much on yesterdays volatility as that of the volatility n periods back. Because of volatility clustering, more accurate forecasts and estimations could be given by putting more weight on more recent values of volatility rather than putting equal weight on all periods.

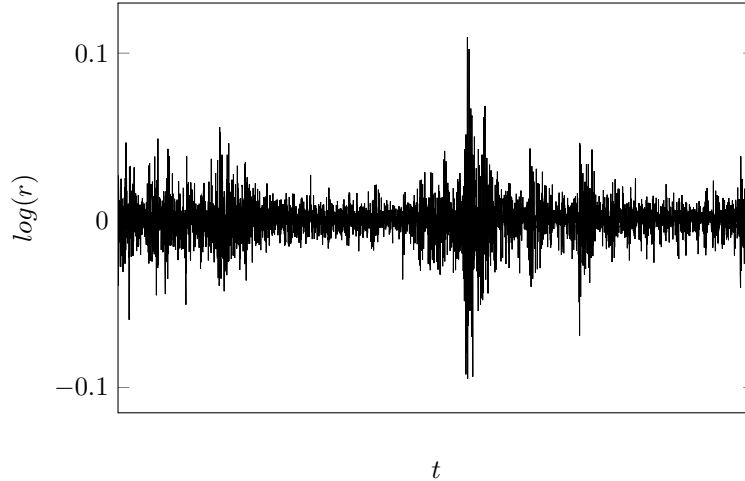


Figure 2: Logged daily returns for the S&P 500-index, 2000-2015

2.1.4 The Leverage effect

An observed effect in financial time series is that volatility tends to be higher when previous periods returns have been negative. This asymmetric effect, often called the leverage effect, can be explained in equities by the fact that when the share price of a company falls, the company's equity drops while debt stays the same. This will raise the debt to equity ratio, thus raising a central risk metric of the company and making it more risky. A negative return should thus increase volatility in the equity or asset price more than a positive return should (Brooks, 2014, p. 440). Another possible explanation for the leverage effect, as described by Bollerslev et al. (2006), is that an expected rise in volatility increases the anticipated risk of the company. The risk-return ratio should therefore increase and the share price needs to fall to accommodate for this. In other words, it says that the returns are dependent on volatility and not the other way around, as discussed above. As the leverage effect is observed, which explanation is dominant is not relevant. This type of asymmetry in the volatility of financial assets is not caught by all volatility forecasting models.

2.2 Financial risk measures

2.2.1 Value at risk

One of the most commonly used methods for estimating financial risk is the Value at risk, VaR measure. The measure is defined as the maximum loss that will occur with a probability level of $1 - q$ (Hull, 2006, p. 471). The often used measure owes much of its popularity to its simplicity as it can be explained to anyone without previous knowledge in financial economics (Hull, 2006, p. 472). Mathematically, the one period VaR , VaR_{t+1}^q , also called the $q\%$ - VaR , is defined as the solution to Equation 2 (Hull, 2006, p. 471), where q is the probability to lose *more* than VaR_{t+1}^q ,

the coming time period. The VaR for period $t + 1$ is estimated using information available in period t . An illustration of the $q\%$ - VaR in a Gaussian distribution can be seen in Figure 3.

$$Pr(r_{t+1} < -VaR_{t+1}^q) = q \quad (2)$$

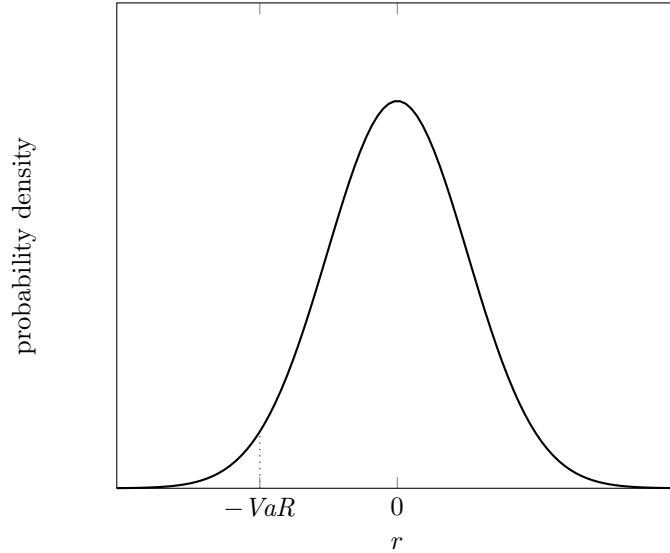


Figure 3: 5%- VaR in a Gaussian distribution.

2.2.2 Expected shortfall

Another measure for financial risk is the Expected shortfall, ES . The measure expands on the theory of VaR and explains what can happen in the cases where the loss does exceed the VaR for that day. It could therefore be claimed that ES contains more information than VaR , as it takes the expected loss into account, not just the threshold loss given by VaR (Acerbi and Tasche, 2002a, p. 4). Its formal definition can be described as the expected loss given that the loss estimated by VaR is exceeded (Christoffersen, 2011, 33), which is defined using VaR in Equation 3 (Hull, 2006, p. 473). The ES of the coming period $t + 1$, is calculated using given information from the current period t . The ES in a Gaussian distribution can be pictured in Figure 4.

$$ES_{t+1}^q = -E(r_{t+1} | r_{t+1} < -VaR_{t+1}^q) \quad (3)$$

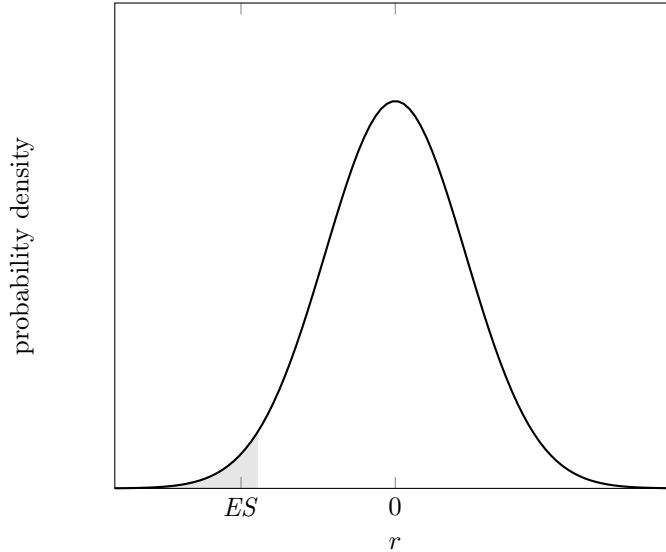


Figure 4: 5%-*ES* in a Gaussian distribution.

2.3 Historical simulation

The HS method is a non-parametric approach to estimate *VaR* and *ES*. The method is widely used among practitioners in the financial industry (Pérignon and Smith, 2010) for its simplicity (Finger, 2006). Besides its simplicity it also has the distinct advantage over more complicated models in that it captures the fat tailed characteristics of financial asset returns.

The method for estimating the *VaR* and *ES* using HS, as done by Acerbi and Tasche (2002b), can be described by sorting the n days' previous returns and choosing the return corresponding to the q :th percentile lowest return to represent the *VaR*. The *ES* is then estimated as the average of the returns below the *VaR*. As for the number of days used in the simulation, there are different opinions. According to the (Basel Committee on Banking Supervision, 2009), 250 days is the minimum threshold when using HS in risk management. Furthermore, according to Mehta et al. (2012), one year, or ca 250 market days, is the most common time period to look at, followed by two and four years. There is a trade-off in how many days are chosen as the rolling window. For short windows, the advantage lies in that it quickly reacts to changes in volatility. A disadvantage of shorter windows is that fewer observations are presumably less representative of the true return distribution than are larger windows. To contrast this, we will in this thesis use two different versions of the HS method, one with a four year time horizon and one with a one year time horizon. Since we will look at 5%-*VaR*, we chose 260 and 1040 market days respectively, to make the percentile account for an actual value of a loss. The estimated *VaR* using HS can be calculated using Equation 4, where r_τ is the q :th percentile lowest return in the sorted list of return over the past n days.

$$\widehat{VaR}_{t+1}^q = r_\tau \quad (4)$$

Similarly, the estimated ES using HS can be calculated using Equation 5, where r_i is the individual return at placement i in the sorting from lowest to highest and τ is the amount of days the loss exceeded the VaR .

$$\widehat{ES}_{t+1}^q = \frac{1}{\tau} \sum_{i=1}^{\tau} r_i \quad (5)$$

Accompanying the advantages of the HS method are also a few disadvantages. First, it assumes that the distribution of returns is the same over time (Finger, 2006). It can clearly be seen in Figure 2 that this is controversial due to volatility clustering. Second, the length of the time period chosen to look at for the simulation can yield differing results, which begs the question of how long the period to look at should be.

2.4 Parametric estimation

To estimate VaR and ES using parametric estimation, there are two different formulas depending on the assumed distribution of the innovation.

If the logarithmic returns, r , are assumed to belong to a Gaussian distribution, an estimation of the next period VaR , \widehat{VaR}_{t+1}^q , can be estimated as according to Equation 6 (McNeil et al., 2005, p. 40) where $-\hat{\mu}_{t+1}$ is the expected *loss* in the coming period, $\hat{\sigma}_{t+1}$ the returns' estimated standard deviation of the coming period and Φ^{-1} the inverse cumulative distribution function of the standard Gaussian probability distribution $N(0, 1)$. ES can then be estimated by Equation 7 (McNeil et al., 2005, p. 45), ϕ is the probability density function of the standard Gaussian probability distribution $N(0, 1)$

$$\widehat{VaR}_{t+1}^q = -\hat{\mu}_{t+1} + \hat{\sigma}_{t+1} \Phi^{-1}(1 - q) \quad (6)$$

$$\widehat{ES}_{t+1}^q = -\hat{\mu}_{t+1} + \hat{\sigma}_{t+1} \frac{\phi(\Phi^{-1}(1 - q))}{q} \quad (7)$$

If the returns are assumed to belong to a *Student's t-distribution*, the VaR can instead be estimated by Equation 8 (McNeil et al., 2005, p. 41) where $-\hat{\mu}_{t+1}$ and $\hat{\sigma}_{t+1}$ are once again estimations of the expected loss and the returns' standard deviation in the coming period, ν are the degrees of freedom and t_{ν}^{-1} is the inverse cumulative distribution function of the Student's t -distribution. The distribution is here scaled by a factor $\sqrt{\frac{\nu-2}{\nu}}$ to give it unit volatility. ES on the other hand be estimated using Equation 9 (McNeil et al., 2005, p. 45), where g_{ν} is the probability density function of the Student's t -distribution with ν degrees of freedom.

$$\widehat{VaR}_{t+1}^q = -\hat{\mu}_{t+1} + \hat{\sigma}_{t+1} \sqrt{\frac{\nu-2}{\nu}} t_{\nu}^{-1}(1 - q) \quad (8)$$

$$\widehat{ES}_{t+1}^q = -\hat{\mu}_{t+1} + \hat{\sigma}_{t+1} \sqrt{\frac{\nu-2}{\nu}} \cdot \frac{g_{\nu}(t_{\nu}^{-1}(1-q))}{q} \cdot \frac{\nu + (t_{\nu}(1-q))^2}{\nu-1} \quad (9)$$

2.5 Volatility forecasting models

2.5.1 GARCH

The ARCH model, proposed by Engle (1982), says that the conditional volatility of a time series can be estimated using the historical volatility of the series as well as the squared previous returns. The model, for which he was later awarded the Nobel price in economics (Nobel Media, 2003), is not frequently employed in practice anymore. This is partly due to its limitations regarding how many lags are included of previous returns, meaning how many previous returns are expected to affect today's volatility (Brooks, 2014, p. 428). A model that solves this problem is the GARCH model, proposed by Bollerslev (1986). The model appends the previous periods conditional volatility times a coefficient to the ARCH model, thus capturing every previous periods returns in descending importance as yesterday's conditional volatility depends on the day before yesterday's conditional volatility and returns, and so on. With this in mind and because the GARCH model is more widely used in practice than the ARCH model (Brooks, 2014, p. 428), only the GARCH model and not the ARCH model, will be evaluated in this thesis. The mathematical definition for the GARCH model can be seen in Equation 10 with \hat{r}_{t+1} and z_{t+1} defined in Equations 11 and 12 respectively. Note that $\hat{\sigma}_{t+1}^2$ is the *estimated* volatility of period $t+1$.

$$\hat{\sigma}_{t+1}^2 = \gamma + \sum_{i=1}^p \alpha_i r_t^2 + \sum_{i=1}^p \beta_i \hat{\sigma}_t^2 \quad (10)$$

$$\hat{r}_{t+1} = z_{t+1} \hat{\sigma}_{t+1} \quad (11)$$

$$z_{t+1} \sim i.i.d.D(0,1) \quad (12)$$

The return r_t , in Equation 10, is assumed to be the squared conditional volatility times an i.i.d (identical and independently distributed) variable following a distribution with 0 mean and a standard deviation of 1. As we do not know what the true distributions of index price time series are, both Gaussian and Student's t-distributions will be tested as $D(0,1)$. The conditional volatility in the future period, $t+1$, is represented by $\hat{\sigma}_{t+1}^2$, the conditional volatility in the current period, t , is represented by $\hat{\sigma}_t^2$ and the return in the current period is represented by r_t^2 . The model uses coefficients to assign weights to each of the variables in the equation based on its presumed importance in the coming periods volatility. These coefficients are β and α as well as a constant γ . A constraint imposed on the coefficients is that $\alpha \geq 0$ and $\beta \geq 0$, as to avoid a negative conditional volatility (Brooks, 2014, p. 425). The summation signs in Equation 10 allows for

the inclusion of several previous periods returns and conditional volatilities rather than just last period's. p and ρ thus refer to how many days should be included. However, according to Hansen and Lunde (2001), setting p and ρ to anything other than 1 will not yield a significant difference in the forecast result, and they will therefore be set to 1 as to satisfy our aim of keeping the method reasonably understandable. The mathematical definition of the GARCH(1,1), the GARCH model with the $p = \rho = 1$ simplification, can be seen in Equation 13. For the remainder of the thesis, the GARCH(1,1) model will simply be referred to as the GARCH.

$$\hat{\sigma}_{t+1}^2 = \gamma + \alpha r_t^2 + \beta \hat{\sigma}_t^2 \quad (13)$$

2.5.2 GJR-GARCH

A shortcoming of the GARCH model is that it does not account for the alleged leverage effect present in the returns of assets, equities and equity indices. A model with characteristics to combat this is the GJR-GARCH model developed by Glosten et al. (1993), which appends a leverage factor to the GARCH model. The mathematical definition for the GJR-GARCH(1,1) can be seen in Equation 14, where I_t , is defined in Equation 15. The coefficient θ , which is expected to be positive, accounts for the leverage effect – that a negative return will have a larger impact on the conditional volatility than a correspondingly large positive return. For the remainder of this thesis, the GJR-GARCH(1,1) model will simply be referred to as the GJR.

$$\hat{\sigma}_{t+1}^2 = \gamma + \alpha r_t^2 + \beta \hat{\sigma}_t^2 + \theta r_t^2 I_t \quad (14)$$

$$I_t = \begin{cases} 1, & \text{if } r_{t-1} < 0 \\ 0, & \text{if } r_{t-1} \geq 0 \end{cases} \quad (15)$$

2.5.3 EGARCH

Another model that accounts for the leverage effect is the Exponential GARCH, or EGARCH, model proposed by Nelson (1991). The model uses the logarithm of the conditional volatility, thus making the non-negativity constraint put on the coefficients in the GARCH model obsolete as a log-function cannot be negative. The mathematical definition of the EGARCH model can be seen in Equation 16, where the δ coefficient accounts for the leverage effect. A negative δ means that the volatility is higher in periods where the previous periods return have been negative, since the expression $\delta \frac{r_t}{\sqrt{\hat{\sigma}_t^2}}$ then will be positive (Brooks, 2014, p. 441). For the remainder of the thesis, the EGARCH(1,1) model will simply be referred to as the EGARCH.

$$\ln(\hat{\sigma}_{t+1}^2) = \gamma + \beta \ln(\hat{\sigma}_t^2) + \delta \frac{r_t}{\sqrt{\hat{\sigma}_t^2}} + \alpha \left[\frac{|r_t|}{\sqrt{\hat{\sigma}_t^2}} - \sqrt{\frac{2}{\pi}} \right] \quad (16)$$

2.5.4 EWMA/Riskmetrics

A simpler approach to volatility forecasting was proposed by RiskMetrics (1996). The model, called either EWMA or Riskmetrics after its authors, is a simplified version of the GARCH that eliminates the need for coefficient estimations. The model authors found that the "decay factor" λ , the β in GARCH, which indicates how much weight should be placed on last period's conditional volatility, should be set to 0.94 and the α coefficient to $1 - \lambda$, or 0.06. The model further sets the γ constant from the GARCH model to zero for simplification. Mathematically, the model can be seen in Equation 17. λ is subsequently set to 0.94 in its original formulation by RiskMetrics (1996), and that is the value used in this thesis.

$$\hat{\sigma}_{t+1}^2 = (1 - \lambda)r_t^2 + \lambda\hat{\sigma}_t^2 \quad (17)$$

2.6 Backtesting

2.6.1 Backtesting Value at Risk

To test the accuracy of our *VaR* estimates, the Kupiec test (proportion of failures- or POF-test) is employed. First introduced by Kupiec (1995), the test measures whether the amount of exceptions is in line with the probability level. In other words, if a $1 - q$ probability level is used when calculating the *VaR*, the loss will exceed the estimated *VaR* in q percent of the periods. If the loss almost never exceeds our estimated *VaR* the estimation is systematically overestimating the risk in the market, meaning it is not an accurate estimate. The same goes if the loss exceeds the *VaR* more than q percent of the days, as this would indicate a systematic underestimation of the risk. Therefore, under the null hypothesis, the proportion of times the actual loss exceeds our *VaR* estimate will be equal to q . Mathematically, the null and alternative hypothesis can be seen in Equation 18 and 19 respectively, where x is the amount of days the loss exceeded estimated *VaR*, \widehat{VaR}_t^q and T is the amount of days *VaR* was estimated.

$$H_0 : q = \frac{x}{T} \quad (18)$$

$$H_1 : q \neq \frac{x}{T} \quad (19)$$

To test the null hypothesis a likelihood ratio test, proposed by Kupiec (1995) is used. The test measures if it is likely that the *VaR* estimates are reliable, given a confidence interval. The test statistic belongs to a χ_1^2 -distribution with 1 noting the degrees of freedom. The test statistic *LR* is defined in Equation 20 where q , T and x are as defined above.

$$LR = -2\ln \left(\frac{(1 - q)^{T-x} q^x}{\left(1 - \left(\frac{x}{T}\right)\right)^{T-x} \left(\frac{x}{T}\right)^x} \right) \quad (20)$$

2.6.2 Backtesting Expected shortfall

As ES has historically been a less popular measure than VaR , the range of backtesting methods is limited in comparison. In addition to this it is more difficult to backtest than VaR , as the distribution of the stochastic loss variable (the actual shortfall value) is needed to test whether the ES estimations are derived from the same distribution. However, the recently proposed test method by Acerbi and Szekely (2014), "Test 2: testing ES directly", finds that the significance threshold, or critical value, for tests with different assumptions about the distribution of the stochastic loss variable is generally the same. This makes practical estimation easier than previously proposed methods. The null hypothesis for the test can be seen in Equation 21, and the alternative hypothesis in Equation 22 where F_t is the unobservable true distribution and P_t is the model distribution. Given that the null hypothesis is not rejected, the ES is assumed to belong to the same, true, distribution as the stochastic loss variable. If the null hypothesis is rejected the ES is assumed not to belong to the same distribution as the stochastic loss variable and instead to a distribution that systematically underestimates the risk.

$$H_0 : F_t^q = P_t^q \quad (21)$$

$$H_1 : ES_{q,F}^t > ES_{q,P}^t \quad (22)$$

The test-statistic is given by Equation 23 where T is the number of observations, $1 - q$ is the probability level used in the VaR^q -measure, L_t is the actual loss at time t , \widehat{ES}_t^q is the estimated ES at time t and J_t is defined in Equation 24. With a correctly posed model the expected value for Z is $E(Z) = 0$, meaning values close to zero indicate that the ES estimations are close to the actual stochastic loss variables. Large negative test statistics indicate an underestimation of the market risk and large positive values indicate an overestimation of the market risk.

$$Z = -\frac{1}{Tq} \sum_{t=1}^T \frac{L_t J_t}{T(1-q) \widehat{ES}_t^q} + 1 \quad (23)$$

$$J_t = \begin{cases} 1, & \text{if } -r_t > \widehat{VaR}_t \\ 0, & \text{if } -r_t \leq \widehat{VaR}_t \end{cases} \quad (24)$$

3 Data & Method

3.1 Data

The data needed for the forecasts were the daily closing prices of the chosen indices for every market day in the periods observed, which is readily available on either Google Finance or Yahoo! Finance. We chose to use data from Yahoo! Finance due to its accessibility. The indices used are presented in Table 1.

Table 1: List of indices to be used in the forecast

| Index | Abbreviation | Country |
|---|--------------|----------------|
| Standard & Poor 500 | S&P 500 | USA |
| Financial Times Stock Exchange 100 Index | FTSE 100 | United Kingdom |
| Nihon Keizai Shimbun 225 | Nikkei 225 | Japan |
| Hang Seng Index | HSI | Hong Kong |
| All Ordinaries Index | AOI | Australia |
| Shanghai Stock Exchange Composite Index | SSE | China |
| Shenzhen Stock Exchange Composite Index | SZSE | China |
| Standard & Poor Bombay Stock Exchange Sensitive Index | S&P BSE | India |
| Russia Trading System Index | RTSI | Russia |
| Bolsa de Valores do Estado de Sao Paulo Index | IBOVESPA | Brazil |

Some statistical data describing the returns of the indices are presented in 2, namely the standard deviation σ , the mean μ , the sample kurtosis κ , the sample skewness S_k , the minimum value min and the maximum value max . The kurtosis κ means how leptokurtic the distribution is and the skewness, S_k , means that a distribution has a tail that is longer than the other, where a positive S_k means the right tail is longer and vice versa for a negative S_k . The sample statistics were all calculated in MATLAB and the code for this can be found in Appendix A.

Table 2: Sample statistics for the indices

| Period | Index | μ | σ | κ | S_k | min | max |
|-------------|------------|-------|----------|----------|-------|-------|-------|
| Crisis | S&P 500 | 0.00 | 0.02 | 6.35 | -0.04 | -0.09 | 0.11 |
| | FTSE 100 | 0.00 | 0.02 | 6.71 | 0.07 | -0.09 | 0.09 |
| | Nikkei 225 | 0.00 | 0.03 | 6.99 | -0.26 | -0.12 | 0.13 |
| | HSI | 0.00 | 0.03 | 6.26 | 0.24 | -0.14 | 0.13 |
| | AOI | 0.00 | 0.02 | 4.94 | -0.35 | -0.09 | 0.05 |
| | SSE | 0.00 | 0.02 | 4.37 | 0.04 | -0.08 | 0.09 |
| | SZSE | 0.00 | 0.03 | 3.77 | -0.13 | -0.09 | 0.09 |
| | S&P BSE | 0.00 | 0.03 | 6.21 | 0.33 | -0.12 | 0.16 |
| | RTSI | 0.00 | 0.04 | 8.61 | -0.14 | -0.21 | 0.20 |
| IBOVESPA | 0.00 | 0.03 | 6.12 | 0.15 | -0.12 | 0.14 | |
| Post crisis | S&P 500 | 0.00 | 0.01 | 6.91 | -0.43 | -0.07 | 0.05 |
| | FTSE 100 | 0.00 | 0.01 | 5.19 | -0.21 | -0.05 | 0.05 |
| | Nikkei 225 | 0.00 | 0.01 | 7.62 | -0.58 | -0.11 | 0.07 |
| | HSI | 0.00 | 0.01 | 5.22 | -0.25 | -0.06 | 0.06 |
| | AOI | 0.00 | 0.01 | 4.29 | -0.24 | -0.04 | 0.03 |
| | SSE | 0.00 | 0.02 | 7.86 | -0.75 | -0.09 | 0.07 |
| | SZSE | 0.00 | 0.02 | 6.06 | -0.50 | -0.09 | 0.09 |
| | S&P BSE | 0.00 | 0.01 | 4.80 | -0.21 | -0.06 | 0.04 |
| | RTSI | 0.00 | 0.02 | 8.35 | -0.19 | -0.13 | 0.13 |
| IBOVESPA | 0.00 | 0.01 | 4.45 | -0.07 | -0.08 | 0.06 | |

3.2 Programming

All estimations and forecasting were done in MATLAB because of the rigorous environment, the applicability of available tools such as the Econometrics Toolbox as well as the ease at which tests can be conducted in the program. All MATLAB-code written by us can be found in Appendix A, while the functions used from the Econometric Toolbox are found in MATLAB (2014).

3.2.1 GARCH-estimation

The coefficient estimation for the models in the GARCH-family was done with the `estimate()`-function on individually specified `garch()`-, `egarch()`- and `gjrr()`-objects. Using the logged returns, MATLAB estimated the coefficients using *maximum likelihood estimation*, or MLE. The method works by estimating the most likely coefficients to fit the data and is used as it works on non-linear models, like the GARCH family of models (Brooks, 2014, p. 431). Every market day, new coefficient estimations were done for the GARCH, EGARCH and GJR model with MLE, using return data for the 2520 previous market days. This particular amount of days was chosen since it equals to roughly 10 years, a satisfying amount of time to yield statistically significant coefficient estimations and to include different periods of financial volatility into the estimations. The models with their estimated coefficients were then used to estimate the following day's volatility σ_{t+1} . For the EWMA the model, no coefficient estimation was done and σ_{t+1} was instead estimated using the fixed value $\lambda = 0.94$. For the models using the Student's t-distribution, the Degrees of Freedom ν were also estimated using MLE.

3.2.2 Choice of models

The non-linear GARCH family models was included for their ability to capture the observed volatility clustering effect discussed in chapter 2.1.3. The asymmetric models GJR and EGARCH was included for their ability to capture the observed leverage effect discussed in chapter 2.1.4. Finally, both the Gaussian and Student's t distribution was included for the sake of discussion as the Student's t distribution should capture the observed fat tails in financial asset returns, as discussed in chapter 2.1.2

3.2.3 Value at Risk estimation

As previously stated, only the parametric estimation and HS methods was evaluated in this thesis. The parametric approach to *VaR*-estimation assumes a probability distribution for future returns, and then finds the value of the *VaR* given the specified probability q . If the distribution was assumed to be Gaussian, Equation 6 was used, and if a Student's t-distribution was assumed, Equation 8 was used. To evaluate these equations, estimated μ and σ were needed. As our different GARCH-models estimated the σ^2 of the coming period, these values were used as the standard deviation of the return distribution as $\hat{\sigma} = \sqrt{\hat{\sigma}^2}$, where $\hat{\sigma}^2$ is the model estimation for the next days volatility. For the expected return $\hat{\mu}$, we used an exponentially weighted moving average of the last 2520 days. The μ - and σ -estimations therefore used the same data. The expected return $\hat{\mu}$ was then calculated using Equation 25 where w_i are weights given by Equation 26.

$$\hat{\mu}_{t+1} = \frac{\sum_{i=t-N+1}^t w_i r_i}{\sum_{i=t-N}^t w_i} \quad (25)$$

$$w_i = \frac{1}{(1 + t - i)^2} \quad (26)$$

For estimations of *VaR* with Gaussian- and t-distributions, GARCH-models with corresponding distribution assumptions were used for the σ^2 -estimations. For the two HS methods, Equation 4 was used with 260 and 1040 days respectively to estimate *VaR*. To test the validity of the estimations made, a Kupiec test was employed yielding test statistics that was compared to significance levels derived from a χ^2 distribution with 1 degree of freedom, which for the 95% confidence interval is 3.84.

3.2.4 Expected Shortfall estimation

The *ES* was also calculated using the parametric and HS methods. By using the same estimations for μ and σ as in the *VaR*-estimations, the *ES* was calculated using Equations 7 and 9 respectively. For the two HS methods, Equation 5 was used with 260 and 1040 days respectively to estimate

ES. To test the validity of the *ES* estimations, test 2 proposed by Acerbi and Szekely (2014) was employed comparing with the critical value of -0.7 .

3.3 Ranking of models

As a means to rank and evaluate the models after the results had been achieved, two comparison methods were used. First, the frequency of best estimation on an index, n_w , was used in the same manner for both *VaR* and *ES* estimations. The measure aims to evaluate what model yielded the best estimation on the most amount of indices. This method of ranking however neglect models that yield consistently good estimations but rarely achieve the best estimation for an index. To solve this problem, a second measure was also used - the average performance. For *VaR*, this was calculated as the sum of all *LRs* for the model in the period, $\sum LR$, while for *ES* it was calculated as the sum of the absolute values of Z , $\sum |Z|$, since negative and positive values summed up could yield misleading results. By using both these measures for both *VaR* and *ES*, the ranking should be able to capture what model performed best in individual indices as well as on average. To test which distributions were most useful, all models belonging to a specific distribution also had their statistics summed up for each period as $\sum \sum LR$ and $\sum \sum |Z|$. To shed light on the overall reliability of the models, models where all estimations proved acceptable in being statistically significant on the 5% confidence level were highlighted as being more reliable than those who did not. For the *LR*-statistic the critical value on the 5% confidence level is 3.84 and for $\sum LR$, it is 18.31. 18.31 comes from the χ_{10}^2 -distribution, since $LR_i \sim \chi_1^2 \Rightarrow \sum_{i=1}^{10} LR_i \sim \chi_{10}^2$. The $\sum \sum LR$ was not be tested for significance but used for internal comparison only. The Z -statistics were compared with the 5% confidence level critical value -0.7 as proposed by Acerbi and Szekely (2014), while $\sum |Z|$ and $\sum \sum |Z|$ do not have any given critical value and only serve as tools for internal comparison of the model accuracy.

4 Results

The results for the statistical tests are presented in six tables. The first two tables, in each subsection, Table 4 and Table 7, show a summary for the models in terms of their summed test statistics, $\sum LR$ and $\sum \sum LR$ for the *VaR* and $\sum |Z|$ and $\sum \sum |Z|$ for the *ES*, as well the number of indices where they are the top performers n_w . Tables 5 and 6 show tests for the *VaR* during and post crisis, while Tables 8 and 9, show the same for the *ES*. The best *VaR* and *ES* estimations for each index are marked by a box. In case of a tie, all the best performers are marked. The test statistics where the null hypotheses was rejected on the 5% confidence level are highlighted in grey whereas the rest, not highlighted nor boxed, are test statistics where the null hypothesis was not rejected but the model did not yield the best result. In the summary tables, the best performers for the entire period are also marked with boxes.

4.1 Estimated coefficients

The minimum, average and maximum values for the coefficient estimations for the two periods are shown in table 3. Here, the column δ / θ refers to the δ coefficient in the EGARCH models and the θ in the GJR models. In total, around 500,000 coefficients were estimated. Here, the coefficients for all 6 estimated models on all of the 10 indices in the relevant periods are used for the average, minimum and maximum values, to show what parameters were used for the models in the two periods. As can be seen, the GARCH and GJR models all seem to neglect the γ -term, estimating it to zero or very close to zero at all times. EGARCH does not do this, and instead has negative γ -values in all cases but for some Gaussian distribution estimations during the crisis. Noteworthy is that some GJR estimations have a negative θ , which would indicate an opposite Leverage effect - meaning that a negative return in the previous period causes a lower volatility in the current period compared to a positive return. This is only in a few cases however, as the mean θ is positive for all periods and both distributions. The same thing seems to happen in some EGARCH estimations, where the δ -value has a positive maximum value for both distributions in the post crisis period.

Table 3: Estimated model coefficients

| | | γ | | | α | | | β | | | δ / θ | | | |
|-------------|-------------|----------|-------|-------|----------|------|------|---------|------|------|-------------------|-------|-------|-------|
| | | min | avg | max | min | avg | max | min | avg | max | min | avg | max | |
| Crisis | Gaussian | GARCH | 0.00 | 0.00 | 0.00 | 0.05 | 0.09 | 0.30 | 0.70 | 0.89 | 0.95 | - | - | - |
| | | EGARCH | 0.75 | -0.23 | -0.07 | 0.08 | 0.15 | 0.27 | 0.91 | 0.97 | 0.99 | -0.14 | -0.08 | -0.02 |
| | | GJR | 0.00 | 0.00 | 0.00 | 0.00 | 0.03 | 0.11 | 0.78 | 0.89 | 0.94 | -0.01 | 0.11 | 0.33 |
| | Student's t | GARCH | 0.00 | 0.00 | 0.00 | 0.04 | 0.10 | 0.19 | 0.80 | 0.89 | 0.96 | - | - | - |
| | | EGARCH | -0.63 | -0.22 | -0.07 | 0.08 | 0.17 | 0.30 | 0.92 | 0.97 | 0.99 | -0.15 | -0.09 | -0.02 |
| | | GJR | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.13 | 0.80 | 0.89 | 0.95 | 0.04 | 0.11 | 0.34 |
| Post crisis | Gaussian | GARCH | 0.00 | 0.00 | 0.00 | 0.04 | 0.09 | 0.18 | 0.81 | 0.90 | 0.96 | - | - | - |
| | | EGARCH | -0.39 | -0.19 | -0.05 | 0.08 | 0.15 | 0.27 | 0.95 | 0.98 | 0.99 | -0.18 | -0.08 | 0.01 |
| | | GJR | 0.00 | 0.00 | 0.00 | 0.00 | 0.03 | 0.15 | 0.74 | 0.90 | 0.95 | -0.02 | 0.11 | 0.30 |
| | Student's t | GARCH | 0.00 | 0.00 | 0.00 | 0.03 | 0.09 | 0.28 | 0.68 | 0.90 | 0.97 | - | - | - |
| | | EGARCH | -0.40 | -0.17 | -0.03 | 0.08 | 0.15 | 0.26 | 0.95 | 0.98 | 1.00 | -0.20 | -0.08 | 0.01 |
| | | GJR | 0.00 | 0.00 | 0.00 | 0.00 | 0.03 | 0.13 | 0.71 | 0.90 | 0.96 | -0.04 | 0.11 | 0.39 |

4.2 Value at Risk

The Kupiec test statistic results for the crisis period can be seen in Table 5 and for the post crisis period in Table 6. A summary of the tests can be seen in Table 4.

Table 4: Summary of *VaR* results

| | | Crisis | | | | Post crisis | | | |
|-------------|---------|-----------|----------------|-------|------------|-------------|----------------|-------|------------|
| | | $\sum LR$ | $\sum \sum LR$ | n_w | $\sum n_w$ | $\sum LR$ | $\sum \sum LR$ | n_w | $\sum n_w$ |
| Gaussian | GARCH | 5.86 | | 1 | | 77.40 | | 0 | |
| | EGARCH | 16.28 | 58.07 | 2 | 8 | 240.80 | 584.96 | 0 | 2 |
| | GJR | 31.40 | | 0 | | 60.49 | | 0 | |
| | EWMA | 4.53 | | 5 | | 6.26 | | 2 | |
| Student's t | GARCH | 8.07 | | 1 | | 44.63 | | 2 | |
| | EGARCH | 15.62 | 55.28 | 1 | 10 | 203.22 | 499.77 | 0 | 5 |
| | GJR | 23.84 | | 3 | | 243.47 | | 0 | |
| | EWMA | 7.75 | | 5 | | 8.44 | | 3 | |
| HS | 1 year | 41.39 | | 1 | | 6.90 | | 4 | |
| | 4 years | 597.40 | | 0 | | 173.60 | | 0 | |

4.2.1 Crisis

The test statistics show that the two EWMA models have the highest frequency of best estimations with 5 for each distribution. Two models did not yield a best estimation on any index this period, namely the Gaussian distributed GJR and the four year-HS. When ranking the models based on average performance, the EWMA with Gaussian distributed innovation yielded the lowest test statistic, followed by the Gaussian distributed GARCH, and the Student's t specification of the EWMA and GARCH models. Worst performers was the four year-HS model where the null hypothesis for all indices was rejected.

As for the distributions, the Student's t distribution seems to perform slightly better than the Gaussian though they usually tie. We arrive at the same result with the average performance for the Student's t and Gaussian distributions, with the Student's t being best on average.

For the reliability of the model estimations, both GARCH and EWMA specifications show significance in all indices with both distributions.

Table 5: Kupiec test for VaR estimates during the financial crisis period (1 December 2007 - 30 June 2009)

| | Gaussian | | | | Student's t | | | | HS | |
|------------|----------|--------|-------|------|-------------|--------|-------|------|--------|---------|
| | GARCH | EGARCH | GJR | EWMA | GARCH | EGARCH | GJR | EWMA | 1 year | 4 years |
| S&P 500 | 1.30 | 0.00 | 0.19 | 1.30 | 1.84 | 0.00 | 0.04 | 1.84 | 6.66 | 106.65 |
| FTSE 100 | 0.28 | 1.38 | 0.10 | 1.38 | 0.93 | 1.38 | 0.02 | 1.91 | 3.22 | 92.97 |
| Nikkei 225 | 1.17 | 0.17 | 2.37 | 0.03 | 0.75 | 0.17 | 0.00 | 0.03 | 3.74 | 54.92 |
| HSI | 1.28 | 2.71 | 2.71 | 1.28 | 1.28 | 2.71 | 1.92 | 0.78 | 1.39 | 67.68 |
| AOI | 0.00 | 1.25 | 3.87 | 0.00 | 0.82 | 1.25 | 1.41 | 0.00 | 3.81 | 66.43 |
| SSE | 0.47 | 0.21 | 0.21 | 0.06 | 1.77 | 0.21 | 0.88 | 1.77 | 0.48 | 36.25 |
| SZSE | 0.55 | 0.09 | 0.27 | 0.09 | 0.55 | 0.27 | 0.09 | 1.39 | 0.09 | 35.06 |
| S&P BSE | 0.54 | 8.55 | 15.26 | 0.00 | 0.00 | 8.55 | 15.26 | 0.00 | 2.41 | 37.32 |
| RTSI | 0.13 | 1.21 | 1.83 | 0.37 | 0.02 | 0.37 | 0.72 | 0.01 | 17.52 | 55.78 |
| IBOVESPA | 0.12 | 0.70 | 4.60 | 0.01 | 0.12 | 0.70 | 3.50 | 0.01 | 2.07 | 44.32 |

4.2.2 Post crisis

The results for the post crisis period paint a slightly different picture than the one derived from the crisis period. Here, the one year-HS and the Gaussian distribution EWMA are the top performers in respect to n_w , with 4 and 2 indices respectively. The more complex model estimations show inferior and, in most indices, insignificant results. The two top performers in the average performance are the same two as in the n_w comparison, but with the Gaussian distribution EWMA now performing better than the one year-HS.

As for the distributions, the Student's t performs the best, producing the best fitting model in more than double the amount of indices as compared to the Gaussian distribution. In the average performance comparison it also gets a lower statistic than the Gaussian distribution.

For the reliability of the model estimations, only the EWMA models and the one year-HS showed significance in all indices, while all others fail at least once. The four year-HS was not significant in any index.

Table 6: Kupiec test for *VaR* estimates for the post crisis period (01-Jul-2009 - 17-Dec-2015)

| | Gaussian | | | | Student's t | | | | HS | |
|------------|----------|--------|-------|------|-------------|--------|-------|------|--------|---------|
| | GARCH | EGARCH | GJR | EWMA | GARCH | EGARCH | GJR | EWMA | 1 year | 4 years |
| S&P 500 | 0.74 | 7.17 | 8.59 | 0.72 | 0.39 | 5.31 | 5.90 | 0.39 | 1.47 | 25.66 |
| FTSE 100 | 1.23 | 14.52 | 23.79 | 3.15 | 1.23 | 13.55 | 21.22 | 3.54 | 0.42 | 19.99 |
| Nikkei 225 | 7.73 | 21.55 | 19.06 | 0.67 | 4.11 | 16.76 | 17.89 | 0.34 | 1.68 | 5.77 |
| HSI | 2.79 | 15.61 | 13.59 | 0.14 | 1.72 | 13.59 | 15.61 | 0.07 | 1.42 | 26.90 |
| AOI | 1.62 | 27.69 | 35.81 | 0.00 | 3.51 | 20.87 | 26.23 | 0.01 | 0.05 | 16.23 |
| SSE | 5.53 | 4.96 | 5.53 | 0.17 | 1.27 | 1.55 | 2.21 | 0.53 | 0.15 | 6.78 |
| SZSE | 3.04 | 5.03 | 4.48 | 0.38 | 0.03 | 1.28 | 1.57 | 1.67 | 0.15 | 5.03 |
| S&P BSE | 22.82 | 56.25 | 66.89 | 0.66 | 14.59 | 58.77 | 89.67 | 0.66 | 0.49 | 42.78 |
| RTSI | 9.52 | 37.60 | 32.34 | 0.16 | 5.41 | 27.59 | 14.99 | 0.91 | 0.00 | 12.08 |
| IBOVESPA | 22.37 | 50.41 | 50.41 | 0.20 | 12.38 | 43.95 | 48.18 | 0.32 | 1.07 | 12.38 |

4.3 Expected shortfall

The Z-test statistic results for the crisis period can be seen in Table 8 and for the post crisis period in Table 9. A summary of the tests can be seen in Table 7.

Table 7: Summary of *ES* results

| | | Crisis | | | | Post crisis | | | |
|-------------|---------|------------|-----------------|-------|------------|-------------|-----------------|-------|------------|
| | | $\sum Z $ | $\sum \sum Z $ | n_w | $\sum n_w$ | $\sum Z $ | $\sum \sum Z $ | n_w | $\sum n_w$ |
| Gaussian | GARCH | 1.31 | | 1 | | 2.72 | | 0 | |
| | EGARCH | 2.01 | | 1 | 8 | 4.57 | | 0 | 5 |
| | GJR | 3.22 | 7.47 | 2 | | 4.99 | 12.98 | 0 | |
| | EWMA | 0.92 | | 4 | | 0.70 | | 5 | |
| Student's t | GARCH | 1.17 | | 1 | | 2.79 | | 0 | |
| | EGARCH | 2.25 | 7.90 | 0 | 3 | 4.60 | 13.91 | 0 | 4 |
| | GJR | 3.44 | | 0 | | 5.08 | | 0 | |
| | EWMA | 1.05 | | 2 | | 0.72 | | 4 | |
| HS | 1 year | 5.31 | | 0 | | 0.66 | | 2 | |
| | 4 years | 24.52 | | 0 | | 3.92 | | 0 | |

4.3.1 Crisis

The Z-test statistic results for the crisis period show that the Gaussian distributed EWMA was the top performer both with regards to number of best estimations, with 4 out of 10, and the average performance, where it had the lowest value. Second contender for most best estimation was the Student's t-distributed EWMA with two. No other model yielded a best estimation on more than one index. Neither the Student's t-distributed EGARCH or GJR nor any HS model

outperformed the others on any index.

As for the distributions, the best performing model used a Gaussian distribution in 8 of the 10 indices with regards to best estimations as compared to the Student's t which was used by 3 best performing models. The average performance comparison also showed the Gaussian distribution as the superior distribution.

For the reliability of model estimations, all specifications showed significance in all indices except for the one and four year-HS.

Table 8: Z test for ES estimates for the financial crisis period (1 December 2007 - 30 June 2009)

| | Gaussian | | | | Student's t | | | | HS | |
|------------|----------|--------|-------|-------|-------------|--------|------|-------|--------|---------|
| | GARCH | EGARCH | GJR | EWMA | GARCH | EGARCH | GJR | EWMA | 1 year | 4 years |
| S&P 500 | -0.25 | 0.00 | 0.18 | -0.24 | -0.19 | 0.06 | 0.18 | -0.21 | -0.82 | -3.79 |
| FTSE 100 | -0.08 | -0.24 | 0.01 | -0.26 | -0.11 | -0.17 | 0.16 | -0.23 | -0.54 | -3.06 |
| Nikkei 225 | -0.17 | -0.01 | 0.39 | -0.05 | -0.10 | 0.02 | 0.17 | 0.00 | -0.68 | -2.63 |
| HSI | 0.22 | 0.35 | 0.37 | 0.21 | 0.26 | 0.40 | 0.39 | 0.23 | -0.38 | -2.78 |
| AOI | 0.03 | -0.17 | 0.47 | -0.04 | -0.12 | -0.13 | 0.35 | 0.02 | -0.51 | -2.73 |
| SSE | -0.12 | -0.11 | 0.16 | -0.05 | 0.00 | 0.16 | 0.42 | -0.12 | 0.15 | -1.54 |
| SZSE | -0.13 | -0.04 | -0.03 | -0.05 | 0.05 | 0.11 | 0.19 | -0.09 | -0.04 | -1.56 |
| S&P BSE | 0.21 | 0.62 | 0.78 | 0.00 | 0.14 | 0.64 | 0.79 | 0.07 | -0.46 | -1.80 |
| RTSI | 0.12 | 0.22 | 0.34 | 0.00 | 0.17 | 0.27 | 0.33 | -0.02 | -1.28 | -2.48 |
| IBOVESPA | 0.00 | 0.25 | 0.50 | 0.00 | 0.02 | 0.27 | 0.46 | 0.05 | -0.46 | -2.14 |

4.3.2 Post crisis

The model that outperformed the others on the most indices in estimating ES in the post crisis period was the Gaussian distributed EWMA with 5 out of 10 indices, followed by the Student's t-distributed EWMA with 4. The top performer in the average performance comparison was the one year-HS, closely followed by the Gaussian distributed EWMA and its Student's t-distributed counterpart.

As for the distributions, the Gaussian distribution outperformed the Student's t distribution slightly, getting the best estimation in 5 indices compared to the Student's t's 4. The average performance comparison also implies that the Gaussian distribution fits the data better.

For the reliability of the model estimations, all models show significance in all estimations, though a slight systematic overestimation of the risk can be seen across the board, as indicated by the mostly positive values of the test statistic.

Table 9: Z test for ES estimates for the post crisis period (01-Jul-2009 - 17-Dec-2015)

| | Gaussian | | | | Student's t | | | | HS | |
|------------|----------|--------|------|-------|-------------|--------|------|-------|--------|---------|
| | GARCH | EGARCH | GJR | EWMA | GARCH | EGARCH | GJR | EWMA | 1 year | 4 years |
| S&P 500 | 0.11 | 0.30 | 0.37 | -0.13 | 0.17 | 0.31 | 0.36 | -0.03 | 0.09 | 0.56 |
| FTSE 100 | 0.17 | 0.41 | 0.53 | -0.19 | 0.25 | 0.44 | 0.54 | -0.11 | 0.03 | 0.48 |
| Nikkei 225 | 0.33 | 0.50 | 0.50 | 0.07 | 0.28 | 0.47 | 0.51 | 0.10 | 0.09 | 0.28 |
| HSI | 0.19 | 0.41 | 0.41 | 0.03 | 0.21 | 0.42 | 0.48 | 0.09 | 0.08 | 0.55 |
| AOI | 0.18 | 0.55 | 0.62 | -0.01 | 0.27 | 0.52 | 0.59 | 0.07 | -0.04 | 0.46 |
| SSE | 0.21 | 0.19 | 0.24 | -0.01 | 0.22 | 0.22 | 0.28 | -0.01 | -0.08 | 0.11 |
| SZSE | 0.18 | 0.20 | 0.24 | -0.08 | 0.15 | 0.21 | 0.26 | -0.04 | -0.09 | 0.10 |
| S&P BSE | 0.51 | 0.73 | 0.78 | 0.09 | 0.47 | 0.76 | 0.87 | 0.16 | 0.05 | 0.61 |
| RTSI | 0.35 | 0.59 | 0.59 | 0.03 | 0.35 | 0.58 | 0.50 | 0.00 | 0.01 | 0.39 |
| IBOVESPA | 0.49 | 0.68 | 0.70 | 0.07 | 0.42 | 0.66 | 0.70 | 0.12 | 0.10 | 0.39 |

4.4 Discussion

The VaR results for the crisis period show clearly superior results for the EWMA model followed by both GARCH specifications using both ranking measurements. The asymmetric GARCH models and one year-HS yielded decent results, though with a few estimations not showing significance. The four year-HS did worse. For the post crisis VaR results, the EWMA performed the best results followed by the one year-HS while the GARCH models and four year-HS did worse. This could be due to the issue regarding the memory of the MLE models and HS, which is a double-edged sword. A long memory increases the significance of the coefficients in the GARCH models, and the simulation in HS, while exposing them to the risk of putting too much weight on the past. Markets can adjust and change characteristics daily and can thus weaken a model or method that focuses too much on the history of an index.

The estimated GARCH models used in this thesis can be described as having two kinds of memory, while the fixed λ -EWMA only has one. The first, and mutual, kind of memory is the lagged σ^2 variable, and the other is the coefficient estimation. The lagged σ^2 will differ in its influence over the estimations and should in a correctly posed model not let the past interfere too much. The coefficient estimation though, which in this thesis was conducted using the maximum likelihood method, face the problems with memory very clearly. Using coefficient estimations from a past period that differs much from the present day can bias the volatility estimations, causing an over- or underestimation of the risk for a long time before stabilizing. Both the crisis and post crisis period examined were preceded by periods with different characteristics. The crisis was preceded by a much calmer period of steady growth, and the post crisis era was preceded by the actual crisis. It is likely that one of the reasons the GARCH family models regularly performed

worse than the EWMA with fixed coefficients, is their problems with memory. However, we cannot say if they are more suitable for longer periods of heterogeneous volatility or if shorter estimation windows would have yielded more accurate estimations. These questions will have to be left for future projects.

As for the four year-HS, it was clearly shown to be the worst performing model of the 10. It does not outperform any of the models or methods in a single index in any of the two periods, and is on average the worst in both. For the *VaR* it is never statistically significant and for the *ES* it is only significant after the crisis, even though it seems to overestimate the risk as indicated by the large positive *Z*-test statistic values. This is likely because it puts too much weight on the past. It will take until 2013 in the second period before data from the crisis will no longer be used, which will result in a big overestimation of the risk after the crisis. By the same logic, it can be argued that it will underestimate the risk during the crisis, since it at no point in that period will base its estimations on market days only from the crisis. Another reason the estimated GARCH, and its family members, yielded worse estimations than did the EWMA in the crisis and post crisis period, could be coefficient estimation error associated with the Maximum Likelihood method. The method is not straightforward and might be conducted in different ways. The one used by MATLAB is not necessarily perfectly posed and might be flawed in this instance. A bad estimation method will result in worse fitting models that can be outperformed by less sophisticated yet more suitable models. Looking at the coefficients estimated in Table 3, it can be seen that some of the estimations of the GARCH model have γ , β and α close or equal to 0, 0.94 and 0.06 respectively – making them similar to the EWMA model. Though, the fact that the mean β is always lower than 0.94, combined with the result that the EWMA yielded better estimations than the GARCH on average, indicates that the MLE produces models that focus too much on previous days returns and too little on previous days estimated volatility. Furthermore, the fact that some GJR models had $\theta < 0$ and some EGARCH models had $\delta > 0$ also goes against the established theory of the Leverage effect, making the coefficient estimations seem less credible. Whatever the reason for the discrepancy, the EWMA performed better than its more complex cousins, implying that the answer to the original question – if the complexity of the other models really is necessary when estimating *VaR* and *ES* on stock indices – is simply no.

For the different distribution assumptions tested in the parametric estimation method, neither the Gaussian- nor the Student's t-distribution was shown to be undoubtedly more accurate. For *VaR*, the Student's t-distribution seemed to perform better while the Gaussian distribution performed the better on the *ES* measure. This implies that the returns do not contain a big tail risk, which would be the case if the Student's t-distribution consistently yielded better estimations for the *ES*. However, no clear answer is given for why the Student's t yields better estimations for the *VaR* and the fact that the *VaR* and *ES* tests do not show the same distribution as superior is conflicting. It can be concluded that no distribution is clearly superior to the other.

5 Conclusion

The question posed in the beginning of the thesis was whether investors, both firms and individuals, should use complex parametric estimations for their *VaR* and *ES* estimations or if using the simpler HS can yield results that are accurate enough. The next question posed, given that HS proved to yield inferior results, was what volatility forecasting models ought to be used – the asymmetric GARCH models or simpler alternatives. The importance of simplicity is of course a factor that matters and is one we leave to the reader to decide for themselves.

In the thesis we have showed the versatility of the EWMA model and issues involved with coefficient estimation in other models. Combining this with the fact that the EWMA is a much simpler model to use in practice, we conclude that investors should not waste time and effort on using more complex models and should instead stick with the EWMA when estimating *VaR* and *ES* on stock indices. As for distribution assumptions, the Gaussian distribution should be used as the difference between the distribution was so slight and the Gaussian proved better for the EWMA model on average. However, as HS yielded acceptable results in the post crisis period, for individuals where simplicity is highly regarded, the HS can be considered as an acceptable method in calm periods. Though, due to the shortcomings of the HS in stormier times, the EWMA and parametric estimation should always be used for more volatile periods. Since we needed to make delimitations in the thesis, Monte Carlo-simulation was not tested, and neither were other specifications and characteristics of the GARCH family models. Conclusions of these methods and their estimation accuracy is therefore left for further studies.

References

- Acerbi, C. and Szekely, B. (2014). Backtesting expected shortfall. *Risk Magazine*.
- Acerbi, C. and Tasche, D. (2002a). Expected Shortfall: A natural coherent alternative to Value at Risk. *Economic Notes*, 31:379–388.
- Acerbi, C. and Tasche, D. (2002b). On the coherence of expected shortfall. *Journal of Banking & Finance*.
- Basel Committee on Banking Supervision (2009). Enhancements to the Basel II framework. *Bank for International Settlements*.
- Basel Committee on Banking Supervision (2013). Fundamental review of the trading book: A revised market risk framework. *Bank for International Settlements*, (2).
- Bohdalová, M. (2007). A comparison of Value-at-Risk methods for measurement of the financial risk. *E-Leader, Prague*.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, (31).
- Bollerslev, T. (1987). A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *Review of Economics and Statistics*, 69:542–547.
- Bollerslev, T., Litvinova, J., and Tauchen, G. (2006). Leverage and Volatility Feedback Effects in High-Frequency Data. *Journal of Financial Econometrics*.
- Brooks, C. (2014). *Introductory Econometrics for Finance*. Cambridge University Press, 3:rd edition.
- Brownlees, C., Engle, R., and Kelly, B. (2011). A practical guide to volatility forecasting through calm and storm. *The Journal of Risk*.
- Christoffersen, P. (2011). *Elements of Financial Risk Management, Second Edition*. Academic Press.
- Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, pages 987–1007.
- Finger, C. C. (2006). How historical simulation made me lazy.
- Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *The Journal of Finance*.
- Hansen, P. R. and Lunde, A. (2001). A Comparison of Volatility Models: Does anything beat a GARCH(1,1) *Working Paper Series*.
- Hansen, P. R. and Lunde, A. (2005). A Forecast Comparison of Volatility Models: Does anything beat a GARCH(1,1) *Journal of Applied Econometrics*.
- Hull, J. C. (2006). *Options, futures, and other derivatives*. Pearson Prentice Hall, 6:th edition.
- Kupiec, P. (1995). Techniques for Verifying the Accuracy of Risk Management Models. *Journal of Derivatives*, (3):73.
- MATLAB (2014). *version 8.3.0.532 (R2014a)*. The MathWorks Inc.
- McNeil, A., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management - Concepts, Techniques and Tools*. Princeton University Press.
- Mehta, A., Neukirchen, M., Pfetsch, S., and Poppensieker, T. (2012). Managing market risk: Today and tomorrow. *McKinsey Working Papers on Risk*, (32).

- Mentel, G. (2003). Parametric or Non-Parametric Estimation of Value-At-Risk. *International Journal of Business and Management*.
- National Bureau of Economic Research (2010). US Business Cycle Expansions and Contractions.
- Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: a New Approach. *Econometrica*.
- Nobel Media (2003). Robert F. Engle III - Facts.
- Pérignon, C. and Smith, D. R. (2010). The level and quality of value-at-risk disclosure by commercial banks. *Journal of Banking Finance*, (2):362.
- RiskMetrics (1996). RiskMetrics Technical Document. (4).
- Sethapramote, Y., Prukumpai, S., and Kanyamee, T. (2014). Evaluation of Value-at-Risk Estimation using Long Memory Volatility Models: Evidence from Stock Exchange of Thailand. *Social Science Research Network*.

Appendices

A MATLAB-code

A.1 Sample statistics

```
1 function [se,k,mu,sk,mi,ma]=sampleStats(filename,startDate,endDate)
2
3 % Get daily close and date data
4 [~,~,rawData]=xlsread(filename,'','','basic');
5 date_cell=rawData(2:length(rawData),1);
6 close_cell=rawData(2:length(rawData),2);
7 date=zeros(1,length(date_cell));
8 close=zeros(1,length(close_cell));
9
10 % Convert the data to string format
11 for i=1:length(date)
12     date(i)=datenum(cell2mat(date_cell(i)));
13     close_i=cell2mat(close_cell(i));
14     if iscellstr(close_i)
15         close(i)=str2num(close_i);
16     else
17         close(i)=close_i;
18     end
19 end
20 date=fliplr(date);
21 close=fliplr(close);
22
23 % Convert the date format to a correct one
24 date=x2mdate(date);
25
26 % First and last days of the forecast will be on index
27 t1=find(date>=datenum(startDate),1);
28 t2=find(date>=datenum(endDate),1);
29
30 % Throw errors if data is missing
31 if isempty(t1) || isempty(t2)
32     msg=cell2mat(strcat('Lacking sufficient data in ',{' '},
33         filename,', data available for period',{' '},datestr(date(1)),{'
34         - '},datestr(date(length(date)))));
35     error(msg)
36 end
37
38 % Calculate the returns and the logarithmized returns
39 r=zeros(length(date),1);
40 for i=t1+1:t2
41     r(i)=close(i)/close(i-1);
42 end
43 logr=log(r);
44 % Cut away non relevant days (those that lack returns and those
45 % after the
46 % period)
```

```

46 date=date(t1:t2);
47 close=close(t1:t2);
48 logr=logr(t1:t2);
49
50 se=std(logr);
51 k= kurtosis(logr);
52 mu=mean(logr);
53 sk=skewness(logr);
54 mi=min(logr);
55 ma=max(logr);
56
57 end

```

A.2 One index prediction

```

1 function [logr ,sigma2 ,VaR ,ES ,date ,LR_POF ,Z ,gamma ,alpha ,beta ,delta]
   = garchPrediction(filename ,startDate ,endDate ,N ,T ,dist ,modelname ,
   q ,lambda ,l)
2
3 % INPUTS
4 % filename - name of the excel workfile, with date in the first
   column and
5 % closes in the second, first row containing row titles and sorted
   with
6 % most recent observation first
7 % startDate - first day of the forecast
8 % endDate - last day of the forecast
9 % N - number of previous observations to use in the coefficient
   estimation
10 % T - period length of coefficient re-estimation
11 % dist - assumed probability distribution of returns ('t' of '
   Gaussian').
12 % dist doesn't matter if model is set to 'all' or HS
13 % model - which model to be used, if 'all', all models and
14 % distributions will be tested
15 % q - probability level of VaR (likely 1-5%)
16 % lambda - EWMA factor
17 % l - number of this run / number of runs (for the progress bar)
18
19 % OUTPUTS
20 % logr - logged daily returns
21 % sigma2 - predicted volatility
22 % VaR - parametric Value at Risk
23 % ES - parametric Expected Shortfall
24 % date - date numbers
25 % LR_POF - VaR test statistic
26 % Z - ES test statistic
27 % gamma - min, mean and max for the gamma coefficient
28 % alpha - min, mean and max for the alpha coefficient
29 % beta - min, mean and max for the beta coefficient
30 % delta - min, mean and max for the delta coefficient
31
32
33
34 % Display that the data fetching has begun
35 msg1=cell2mat(strcat({'Period '},num2str(l(3)),{' of '},num2str(l

```

```

        (4)));
36 msg2=cell2mat(strcat({'Index  '},num2str(1(1)),{' of '},num2str(1
    (2))));
37 disp(msg1)
38 disp(msg2)
39 disp('Fetching data...')
40
41 % Get daily close and date data
42 [~,~,rawData]=xlsread(filename,','',',', 'basic');
43 date_cell=rawData(2:length(rawData),1);
44 close_cell=rawData(2:length(rawData),2);
45 date=zeros(1,length(date_cell));
46 close=zeros(1,length(close_cell));
47
48 % Convert the data to string format
49 for i=1:length(date)
50     date(i)=datenum(cell2mat(date_cell(i)));
51     close_i=cell2mat(close_cell(i));
52     if iscellstr(close_i)
53         close(i)=str2num(close_i);
54     else
55         close(i)=close_i;
56     end
57 end
58 date=fliplr(date);
59 close=fliplr(close);
60
61 % Convert the date format to a correct one
62 date=x2mdate(date);
63
64 % First and last days of the forecast will be on index
65 t1=find(date>=datenum(startDate),1);
66 t2=find(date>=datenum(endDate),1);
67
68 % Throw errors if data is missing
69 if isempty(t1) || isempty(t2)
70     msg=cell2mat(strcat('Lacking sufficient data in ',{' '},
    filename,', data available for period',{' '},datestr(date(1)),{'
    - '},datestr(date(length(date)))));
71     error(msg)
72 elseif (t1-N-1)<0
73     msg=cell2mat(strcat('Lacking sufficient data in ',{' '},
    filename,', data available for period',{' '},datestr(date(1)),{'
    - '},datestr(date(length(date)))));
74     error(msg)
75 end
76
77
78 % Calculate the returns and the logarithmized returns
79 r=zeros(length(date),1);
80 for i=t1-N-1:1:t2
81     r(i)=close(i)/close(i-1);
82 end
83 logr=log(r);
84
85 % Cut away non relevant days (those that lack returns and those

```

```

    after the
86 % period)
87 date=date(t1-N-1:t2);
88 close=close(t1-N-1:t2);
89 logr=logr(t1-N-1:t2);
90
91 % The first day of the forecast will now be on
92 t1=find(date>=datenum(startDate),1);
93
94 % Creating the predicted DoF, volatility, VaR and ES vectors
95 if strcmp(modelname,'all')
96     sigma2=zeros(length(date),8);
97     VaR=zeros(length(date),10);
98     ES=zeros(length(date),10);
99 else
100     sigma2=zeros(length(date),1);
101     VaR=zeros(length(date),1);
102     ES=zeros(length(date),1);
103 end
104
105 % Creating vectors for coefficient estimations
106 if strcmp(modelname,'all')
107     gamma=ones(length(date),6);
108     alpha=ones(length(date),6);
109     beta=ones(length(date),6);
110     delta=ones(length(date),4);
111 elseif strcmp(modelname,'GARCH')
112     gamma=ones(length(date),1);
113     alpha=ones(length(date),1);
114     beta=ones(length(date),1);
115 elseif strcmp(modelname,'EGARCH') || strcmp(modelname,'GJR')
116     gamma=ones(length(date),1);
117     alpha=ones(length(date),1);
118     beta=ones(length(date),1);
119     delta=ones(length(date),1);
120 end
121
122 % Create normalized weight vector for the expected return
    estimation
123 w=zeros(N+1,1);
124 for i=1:N
125     w(N+2-i)=1/i^2;
126 end
127 w=w/sum(w);
128
129
130 % Create models
131 if strcmp(modelname,'all')
132     model_garchN=garch('GARCHLags',1,'ARCHLags',1,'Distribution','Gaussian');
133     model_garchT=garch('GARCHLags',1,'ARCHLags',1,'Distribution','t');
134     model_egarchN=egarch('GARCHLags',1,'ARCHLags',1,'LeverageLags',1,'Distribution','Gaussian');
135     model_egarchT=egarch('GARCHLags',1,'ARCHLags',1,'LeverageLags',1,'Distribution','t');

```

```

136     model_gjrN=gjr('GARCHLags',1,'ARCHLags',1,'LeverageLags',1,'
Distribution','Gaussian');
137     model_gjrT=gjr('GARCHLags',1,'ARCHLags',1,'LeverageLags',1,'
Distribution','t');
138     model_ewmaN=garch('GARCHLags',1,'ARCHLags',1,'Distribution','
Gaussian','Constant',1e-322,'ARCH',(1-lambda)-(1e-14),'GARCH',
lambda-(1e-14));
139     model_ewmaT=garch('GARCHLags',1,'ARCHLags',1,'Distribution','t'
,'Constant',1e-322,'ARCH',(1-lambda)-(1e-14),'GARCH',lambda-(1e
-14));
140 elseif strcmp(modelname,'GARCH')
141     model=garch('GARCHLags',1,'ARCHLags',1,'Distribution',dist);
142 elseif strcmp(modelname,'EGARCH')
143     model=egarch('GARCHLags',1,'ARCHLags',1,'LeverageLags',1,'
Distribution',dist);
144 elseif strcmp(modelname,'GJR')
145     model=gjr('GARCHLags',1,'ARCHLags',1,'LeverageLags',1,'
Distribution',dist);
146 elseif strcmp(modelname,'EWMA')
147     model=garch('GARCHLags',1,'ARCHLags',1,'Distribution',dist,'
Constant',1e-322,'ARCH',(1-lambda)-(1e-14),'GARCH',lambda-(1e
-14));
148 elseif strcmp(modelname,'HS_260') || strcmp(modelname,'HS_1040')
149     model=null(1);
150 else
151     error('Unknown model. Must be GARCH, EGARCH, GJR, EWMA, HS_260,
HS_1040 or all')
152 end
153
154 % Display that the data fetching is done!
155 disp('Done!')
156
157 % Create the progress counter and display that the forecasting has
started
158 prog_1=round(100*(((1(3)-1)*1(2)+1(1))-1)/(1(2)*1(4)));
159 prog_2=round(100*(((1(3)-1)*1(2)+1(1)))/(1(2)*1(4)));
160 prog=prog_1;
161 if prog==0
162     msg2=cell2mat(strcat({'Forecasting: '},num2str(prog),'%'));
163 else
164     msg2=cell2mat(strcat({'Forecasting: '},num2str(prog+1),'%'));
165 end
166 disp(msg2)
167
168 %Estimate coefficients every T market days
169 for t=t1-1:T:length(date);
170     % Vector with relevant returns
171     logr_t=logr(t-N:t);
172     % Estimate model coefficients
173     if strcmp(modelname,'all')
174         estModel_garchN=estimate(model_garchN,logr_t,'print',false)
175         ;
176         estModel_garchT=estimate(model_garchT,logr_t,'print',false)
177         ;
178         estModel_egarchN=estimate(model_egarchN,logr_t,'print',
false);

```

```

177     estModel_egarchT=estimate(model_egarchT,logr_t,'print',
false);
178     estModel_gjrN=estimate(model_gjrN,logr_t,'print',false);
179     estModel_gjrT=estimate(model_gjrT,logr_t,'print',false);
180     estModel_ewmaN=estimate(model_ewmaN,logr_t,'print',false);
181     estModel_ewmaT=estimate(model_ewmaT,logr_t,'print',false);
182     elseif strcmp(modelname,'GARCH') || strcmp(modelname,'EGARCH')
|| strcmp(modelname,'GJR')
183         estModel=estimate(model,logr_t,'print',false);
184     end
185     % Save the estimated coefficients
186     if strcmp(modelname,'all')
187         gamma(t,1)=estModel_garchN.Constant;
188         gamma(t,2)=estModel_garchT.Constant;
189         gamma(t,3)=estModel_egarchN.Constant;
190         gamma(t,4)=estModel_egarchT.Constant;
191         gamma(t,5)=estModel_gjrN.Constant;
192         gamma(t,6)=estModel_gjrT.Constant;
193         alpha(t,1)=cell2mat(estModel_garchN.ARCH);
194         alpha(t,2)=cell2mat(estModel_garchT.ARCH);
195         alpha(t,3)=cell2mat(estModel_egarchN.ARCH);
196         alpha(t,4)=cell2mat(estModel_egarchT.ARCH);
197         if isempty(estModel_gjrN.ARCH)
198             alpha(t,5)=0;
199         else
200             alpha(t,5)=cell2mat(estModel_gjrN.ARCH);
201         end
202         if isempty(estModel_gjrT.ARCH)
203             alpha(t,6)=0;
204         else
205             alpha(t,6)=cell2mat(estModel_gjrT.ARCH);
206         end
207         beta(t,1)=cell2mat(estModel_garchN.GARCH);
208         beta(t,2)=cell2mat(estModel_garchT.GARCH);
209         beta(t,3)=cell2mat(estModel_egarchN.GARCH);
210         beta(t,4)=cell2mat(estModel_egarchT.GARCH);
211         beta(t,5)=cell2mat(estModel_gjrN.GARCH);
212         beta(t,6)=cell2mat(estModel_gjrT.GARCH);
213         delta(t,1)=cell2mat(estModel_egarchN.Leverage);
214         delta(t,2)=cell2mat(estModel_egarchT.Leverage);
215         if isempty(estModel_gjrN.Leverage)
216             delta(t,3)=0;
217         else
218             delta(t,3)=cell2mat(estModel_gjrN.Leverage);
219         end
220         if isempty(estModel_gjrT.Leverage)
221             delta(t,4)=0;
222         else
223             delta(t,4)=cell2mat(estModel_gjrT.Leverage);
224         end
225     elseif strcmp(modelname,'GARCH')
226         gamma(t)=estModel.Constant;
227         alpha(t)=cell2mat(estModel.ARCH);
228         beta(t)=cell2mat(estModel.GARCH);
229     elseif strcmp(modelname,'EGARCH') || strcmp(modelname,'GJR')
230         gamma(t)=estModel.Constant;

```

```

231     if isempty(estModel.ARCH)
232         alpha(t)=0;
233     else
234         alpha(t)=cell2mat(estModel.ARCH);
235     end
236     beta(t)=cell2mat(estModel.GARCH);
237     delta(t)=cell2mat(estModel.Leverage);
238 end
239 % Estimate sigma2 (only for GARCH-family models), VaR and ES
daily
240 for j=t:1:min(t+T-1,length(date))
241     % Vector with N previous returns for the forecast
242     logr_t=logr(j-N:j);
243     % Forecast volatility
244     if strcmp(modelname,'all')
245         sigma2(j,1)=forecast(estModel_garchN,1,'Y0',logr_t);
246         sigma2(j,2)=forecast(estModel_garchT,1,'Y0',logr_t);
247         sigma2(j,3)=forecast(estModel_egarchN,1,'Y0',logr_t);
248         sigma2(j,4)=forecast(estModel_egarchT,1,'Y0',logr_t);
249         sigma2(j,5)=forecast(estModel_gjrN,1,'Y0',logr_t);
250         sigma2(j,6)=forecast(estModel_gjrT,1,'Y0',logr_t);
251         sigma2(j,7)=forecast(estModel_ewmaN,1,'Y0',logr_t);
252         sigma2(j,8)=forecast(estModel_ewmaT,1,'Y0',logr_t);
253     elseif strcmp(modelname,'GARCH') || strcmp(modelname,'
EGARCH') || strcmp(modelname,'GJR')
254         sigma2(j)=forecast(estModel,1,'Y0',logr_t);
255     end
256     % Save estimated Degrees of Freedom for t-distributions
257     if strcmp(modelname,'all')
258         DoF=zeros(8,1);
259         DoF(2)=estModel_garchT.Distribution.DoF;
260         DoF(4)=estModel_egarchT.Distribution.DoF;
261         DoF(6)=estModel_gjrT.Distribution.DoF;
262         DoF(8)=estModel_ewmaT.Distribution.DoF;
263     else
264         if strcmp(dist,'t')
265             DoF=estModel.Distribution.DoF;
266         end
267     end
268
269     % Expected next day loss
270     logr_e=sum(logr(j-N:j).*w);
271     loss_e=-1*logr_e;
272     % Forecasted next day standard deviation using forecasted
273     % volatility
274     if strcmp(modelname,'all')
275         se=zeros(8,1);
276         for i=1:8
277             se(i)=sqrt(sigma2(j,i));
278         end
279     else
280         se=sqrt(sigma2(j));
281     end
282     % Create numbers needed for VaR-estimation of GARCH-family
models
283     if strcmp(modelname,'all')

```

```

284     VaRZ=zeros(8,1);
285     ESZ=zeros(8,1);
286     for i=1:2:7
287         VaRZ(i)=icdf('Normal',1-q,loss_e,1);
288         ESZ(i)=pdf('Normal',VaRZ(i),loss_e,1)/(q);
289     end
290     for i=2:2:8
291         VaRZ(i)=icdf('tLocationScale',1-q,loss_e,1,DoF(i))*
sqrt((DoF(i)-2)/DoF(i));
292         ESZ(i)=sqrt((DoF(i)-2)/DoF(i))*pdf('tLocationScale'
,icdf('tLocationScale',1-q,loss_e,1,DoF(i)),loss_e,1,DoF(i))*
(DoF(i)+(icdf('tLocationScale',1-q,loss_e,1,DoF(i)))^2)/((q)*(DoF
(i)-1));
293     end
294     elseif strcmp(modelname,'GARCH') || strcmp(modelname,'
EGARCH') || strcmp(modelname,'GJR')
295         if strcmp(dist,'t')
296             % Calculate relevant numbers for the VaR and ES
estimations
297             VaRZ=icdf('tLocationScale',1-q,loss_e,1,DoF)*sqrt((
DoF-2)/DoF);
298             ESZ=sqrt((DoF-2)/DoF)*pdf('tLocationScale',icdf('
tLocationScale',1-q,loss_e,1,DoF),loss_e,1,DoF)*(DoF+(icdf('
tLocationScale',1-q,loss_e,1,DoF))^2)/((q)*(DoF-1));
299             else
300                 VaRZ=icdf('Normal',1-q,0,1);
301                 ESZ=pdf('Normal',VaRZ,0,1)/(q);
302             end
303         end
304         % Estimate VaR and ES
305         if strcmp(modelname,'all')
306             for i=1:8
307                 VaR(j,i)=se(i)*VaRZ(i);
308                 ES(j,i)=se(i)*ESZ(i);
309             end
310             % Historical Simulation models
311             logr_HS1=logr(j-260:j);
312             logr_HS2=logr(j-1040:j);
313             logr_HS1=sort(logr_HS1);
314             logr_HS2=sort(logr_HS2);
315             VaR(j,9)=-1*logr_HS1(13);
316             VaR(j,10)=-1*logr_HS2(52);
317             ES(j,9)=-1*mean(logr_HS1(1:13));
318             ES(j,10)=-1*mean(logr_HS2(1:52));
319         elseif strcmp(modelname,'HS_260')
320             logr_HS=logr(j-260:j);
321             logr_HS=sort(logr_HS);
322             VaR(j)=-1*logr_HS(12);
323             ES(j)=-1*mean(logr_HS(1:13));
324         elseif strcmp(modelname,'HS_1040')
325             logr_HS=logr(j-1040:j);
326             logr_HS=sort(logr_HS);
327             VaR(j)=-1*logr_HS(52);
328             ES(j)=-1*mean(logr_HS(1:52));
329         else
330             VaR(j)=se*VaRZ;

```

```

331         ES(j)=se*ESZ;
332     end
333
334     % Step the progress counter and display progress
335     prog_p=prog;
336     prog=prog_1+round((prog_2-prog_1)*(j-t1+1)/(length(date)-t1
+1));
337     if (prog>prog_p && prog>prog_1+1) || (prog>prog_p && prog
==1)
338         msg2=cell2mat(strcat({'          },num2str(prog),'%
'));
339         disp(msg2)
340     end
341 end
342 end
343
344 % Display if the forecast is done
345 if prog==100
346     disp('Forecast complete!')
347 end
348
349
350 % Cut off irrelevant data not belonging to the specified period
351 logr=logr(t1:length(logr));
352 date=date(t1:length(date));
353 sigma2=sigma2(t1:length(sigma2),:);
354 VaR=VaR(t1:length(VaR),:);
355 ES=ES(t1:length(ES),:);
356 if strcmp(modelname,'all') || strcmp(modelname,'EGARCH') || strcmp(
modelname,'GJR')
357     gamma=gamma(t1:length(gamma),:);
358     alpha=alpha(t1:length(alpha),:);
359     beta=beta(t1:length(beta),:);
360     delta=delta(t1:length(delta),:);
361 elseif strcmp(modelname,'GARCH')
362     gamma=gamma(t1:length(gamma),:);
363     alpha=alpha(t1:length(alpha),:);
364     beta=beta(t1:length(beta),:);
365 end
366
367 % Kupiec Test
368 % Vector for days where the negative return exceeded the VaR
369 if strcmp(modelname,'all')
370     I=zeros(length(logr),10);
371     for t=1:1:length(logr)
372         for i=1:10
373             if -1*logr(t)>=VaR(t,i)
374                 I(t,i)=1;
375             end
376         end
377     end
378 else
379     I=zeros(length(logr),1);
380     for t=1:1:length(logr)
381         if -1*logr(t)>=VaR(t)
382             I(t)=1;

```

```

383         end
384     end
385 end
386
387 % number of observations
388 n=length(date);
389 % Number of VaR exceedings
390 x=sum(I);
391 % Kupiec test statistic LR_POF
392 LR_POF=-2*(x.*(log(q)-log(x/n)))+(n-x).*(log(1-q)-log(1-(x/n))));
393
394 % Acerbi & Szekely Z-test
395 % Value of the losses exceeding VaR
396 if strcmp(modelname,'all')
397     L=zeros(n,10);
398     for i=1:10
399         L(:,i)=-1*I(:,i).*logr;
400     end
401 else
402     L=-1*(logr.*I);
403 end
404
405 % Ratio between actual losses and expected shortfall
406 ES_ratio=L./ES;
407 % Test statistic as proposed by Acerbi & Szekely
408 Z=-1*(1/(n*(q)))*sum(ES_ratio)+1;
409
410 % Remove values in coefficient estimations where they were not re-
411 % estimated
412 % and calculate their min, mean and max
413 if strcmp(modelname,'all') || strcmp(modelname,'EGARCH') || strcmp(
414     modelname,'GJR')
415     gamma(all(gamma==1,2),:)=[];
416     alpha(all(alpha==1,2),:)=[];
417     beta(all(beta==1,2),:)=[];
418     delta(all(delta==1,2),:)=[];
419     gamma=[min(gamma);mean(gamma);max(gamma)];
420     alpha=[min(alpha);mean(alpha);max(alpha)];
421     beta=[min(beta);mean(beta);max(beta)];
422     delta=[min(delta);mean(delta);max(delta)];
423 elseif strcmp(modelname,'GARCH')
424     gamma(all(gamma==1,2),:)=[];
425     alpha(all(alpha==1,2),:)=[];
426     beta(all(beta==1,2),:)=[];
427     gamma=[min(gamma);mean(gamma);max(gamma)];
428     alpha=[min(alpha);mean(alpha);max(alpha)];
429     beta=[min(beta);mean(beta);max(beta)];
430     delta=NaN;
431 elseif strcmp(modelname,'EWMA')
432     gamma=0;
433     alpha=1-lambda;
434     beta=lambda;
435     delta=NaN;
436 else
437     gamma=NaN;
438     alpha=NaN;

```

```

437     beta=NaN;
438     delta=NaN;
439 end
440
441
442 end

```

A.3 Multiple index test

```

1 function [LR_POF,Z,gamma,alpha,beta,delta]=multipleGarchPrediction(
2     filenames,startDates,endDates,N,T,q,lambda)
3
4 % INPUTS
5 % filenames - cell with names for the excel workfiles
6 % startDates - cell with the first days of the forecasts
7 % endDates - cell with the last days of the forecasts
8 % N - number of previous observations to use in the parameter
9     estimation
10 % T - period length of parameter re-estimation
11 % q - probability level of VaR (likely 1-5%)
12 % lambda - EWMA factor
13
14 % OUTPUTS
15 % LR_POF - VaR test statistic
16 % Z - ES test statistic
17 % gamma - min, mean and max for the gamma coefficient for all
18     indices
19 % alpha - min, mean and max for the alpha coefficient for all
20     indices
21 % beta - min, mean and max for the beta coefficient for all indices
22 % delta - min, mean and max for the delta coefficient for all
23     indices
24
25 % Look at all models and all distributions
26 garchmodel='all';
27 dist='all';
28
29 % For the progress bar: Index number, total number of indexes,
30     period
31 % number, total number of periods
32 l=[0,length(filenames),0,length(startDates)];
33
34 % Vectors for the test statistics
35 LR_POF=zeros(length(filenames)*length(startDates),10);
36 Z=zeros(length(filenames)*length(startDates),10);
37
38 % Vectors for the coefficients
39 gamma=zeros(3*length(filenames)*length(startDates),6);
40 alpha=zeros(3*length(filenames)*length(startDates),6);
41 beta=zeros(3*length(filenames)*length(startDates),6);
42 delta=zeros(3*length(filenames)*length(startDates),4);
43
44 % Throw error if periods mismatch
45 if length(startDates)~=length(endDates)
46     msg='Need the same amount of startDates as endDates';

```

```

42     error(msg)
43 end
44
45 % Get values from function
46 for j=1:length(startDates)
47     startDate=cell2mat(startDates(j));
48     endDate=cell2mat(endDates(j));
49     l(3)=j;
50     for i=1:length(filenamees)
51         l(1)=i;
52         [~,~,~,~,~,LR_POF((j-1)*length(filenamees)+i,:),Z((j-1)*
length(filenamees)+i,:),gamma((j-1)*length(filenamees)+(i-1)*3+1:(
j-1)*length(filenamees)+(i-1)*3+1+2,:),alpha((j-1)*length(
filenamees)+(i-1)*3+1:(j-1)*length(filenamees)+(i-1)*3+1+2,:),beta
((j-1)*length(filenamees)+(i-1)*3+1:(j-1)*length(filenamees)+(i-1)
*3+1+2,:),delta((j-1)*length(filenamees)+(i-1)*3+1:(j-1)*length(
filenamees)+(i-1)*3+1+2,:)] = garchPrediction(cell2mat(filenamees(
i)),startDate,endDate,N,T,dist,garchmodel,q,lambda,l);
53     end
54 end
55
56 % Calculate the mean coefficients
57
58 gamma_means=zeros(length(filenamees)*length(startDates),6);
59 alpha_means=zeros(length(filenamees)*length(startDates),6);
60 beta_means=zeros(length(filenamees)*length(startDates),6);
61 delta_means=zeros(length(filenamees)*length(startDates),4);
62
63 for i=1:length(filenamees)*length(startDates)
64     gamma_means(i,:)=gamma(3*i-1,:);
65     alpha_means(i,:)=alpha(3*i-1,:);
66     beta_means(i,:)=beta(3*i-1,:);
67     delta_means(i,:)=delta(3*i-1,:);
68 end
69 if length(filenamees)*length(startDates)==1
70     gamma_mean=gamma_means;
71     alpha_mean=alpha_means;
72     beta_mean=beta_means;
73     delta_mean=delta_means;
74 else
75     gamma_mean=mean(gamma_means);
76     alpha_mean=mean(alpha_means);
77     beta_mean=mean(beta_means);
78     delta_mean=mean(delta_means);
79 end
80
81 gamma=[min(gamma);gamma_mean;max(gamma)];
82 alpha=[min(alpha);alpha_mean;max(alpha)];
83 beta=[min(beta);beta_mean;max(beta)];
84 delta=[min(delta);delta_mean;max(delta)];
85
86 % Turn the vectors to make use of them
87 LR_POF=LR_POF.';
88 Z=Z.';
89
90 end

```