Lund University

Econometric Methods and Monte Carlo Simulations for Financial Risk Management

by

A.D.T. Baltaev and I. Chavdarov

A thesis submitted in partial fulfillment for the
degree of Master of Science in Finance

at the

Department of Economics

Lund University

June 2013

Abstract

Value-at-Risk (VaR) forecasting in the context of Monte Carlo simulations is evaluated. A range of parametric models is considered, namely the traditional Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model, the exponential GARCH and the GJR-GARCH, which are put in the context of the Gaussian and Student-t distributions. The returns of the S&P 500 provide the basis for the study. Monte Carlo simulations are then applied in the estimation and forecasting of index returns. Two forecasting periods are employed with respect to the Global Financial Crisis (GFC). The forecasting accuracy of the various models will be evaluated in order to determine the applicability of these VaR estimation techniques in different market conditions. Results reveal that: (i) no model has consistent performance in both volatile and stable market conditions; (ii) asymmetric volatility models offer better performance in the post crisis forecasting period; (iii) all models underestimate risk in highly unstable market conditions.

KEY WORDS: Value-at-Risk, GARCH, Monte Carlo, Global Financial Crisis.

Acknowledgements

We are deeply appreciative towards everyone who took their time to assist us. We would especially thank our supervisors Karl Larsson and Birger Nilsson at the Department of Economics at Lund University, for their knowledge, assistance and time during the completion of this paper.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The importance of financial risk management has been increasing continuously during the last decades in light of a number of severe stock market crashes, starting with "Black Monday" in 1987 and finishing with the Global Financial Crisis (GFC) in 2008/09. In order to prevent the recurrence of such events and to mitigate the losses associated with them, regulators have prompted financial institutions to put greater emphasis on risk management. The makings of GFC became apparent in 2008 when a number of financial institutions in the U.S., such as Fannie Mae and Freddie Mac were nationalized, and Lehman Brothers filed for bankruptcy after being denied governmental support. The downward spiral that followed these and other events shone a light on the pivotal role of forecasting risk in extreme market conditions.

Value-at-Risk (VaR), pioneered in 1993 as part of the RiskMetrics model of J.P. Morgan, presents a viable example of a model useful in the context of managing financial risk. Its inclusion in the 1996 Basel II framework has since made it an integral part of the operations of financial institutions globally. According to the most recent update of the Basel III framework (2010), banks are allowed to internally estimate a VaR forecast, based on which the amount of regulatory required capital is determined. This implies, that VaR can be viewed as both a regulatory mechanism to prevent systemic disturbances due to mismanagement of financial risk, and as internal means for financial institutions to estimate their risk exposure to a particular asset or a portfolio of assets (Gerlach et al., 2011). The two main time horizons in VaR calculations are 10-days and 1-day VaR, with the former being the regulatory requirement and the latter representing the industry standard. A major problem related to this divergence is lack of a theoretically sound mechanism that can be applied in order to convert 1-day VaR estimates into 10-day VaR. The other alternative, the estimation of a value relevant for the entire period of interest, cannot be deemed viable unless restrictions that may be viewed as unreasonable in certain practical contexts are imposed. In light of this, the comparison between the results obtained when using a number of different models to estimate 1- and 10-day VaR can still be of great relevance, particularly when the much more stringent banking supervision nowadays is accounted for.

Considering that VaR is an estimate and not an uniquely defined value, it is important to emphasize that there are a number of methodologies used in order to calculate it, with some of the most important being: Monte Carlo simulations; non-parametric methods which include historical simulations; semi-parametric approaches, such as extreme value theory (EVT) and the dynamic quantile regression CAViaR model (Engle and Manganelli, 2004; Gerlach et al., 2011); and parametric approaches that fully specify model dynamics and distributional assumptions, such as RiskMetrics and generalized autoregressive conditional heteroskedasticity (GARCH) models (Engle, 1982 and Bollerslev, 1986; Gerlach et al., 2011). When these computation alternatives are evaluated, it is vital to point out that even though GARCH models require time-consuming sample parameter estimation, Gerlach et al. (2011) highlight the fact that GARCH models outperform stochastic volatility models in almost all cases related to financial risk management. Since Engle (1982) and shortly thereafter Bollerslev (1986) laid the foundations of the ARCH and GARCH theoretical frameworks, the models in the GARCH family have become some of the most influential tools available to econometricians when analysing financial data. Due to their ability to account for the so-called "ARCH effects", related to a correlation between the residual over time, these models offer superior forecasting accuracy with respect to asset return computations. When Monte Carlo (MC) simulations are taken into consideration, Berry (2013) points to the fact that they exhibit robustness to the occurrence of extreme events, for which MC simulations even provide in-depth details. Furthermore, MC simulations can account for parameter uncertainty not only in point but also in probabilistic forecasting and guarantee efficient and flexible usage of complex models and non-standard parameters. This, coupled with the fact that any statistical distribution can be used to simulate the returns, have made such simulations the industry standard when it comes to measuring risk. Incorporating the fact that GARCH models have wide theoretical application and MC simulations are commonly used in practice, a combination of these two VaR estimation methods would certainly be of interest.

The aim of this paper is to calculate 1- and 10-day VaR forecasts based on the parameter estimations of three different econometric models in two different subsamples. The models used are the traditional GARCH of Bollerslev (1986), the GJR-GARCH of Glosten et al. (1993) and the exponential GARCH (EGARCH) pioneered by Nelson (1991). A regular, normally distributed GARCH model provides the means in order to calculate parameters, which are then incorporated into MC simulations. This enables the construction of probability density functions (pdfs). The relevant 1% and 5% quantiles are then extracted, in order to compute VaR estimates for the different time horizons of interest. The same procedure is repeated when the more sophisticated GJR-GARCH and EGARCH models are applied to the relevant dataset. Since the normal distribution only accounts for the mean and standard deviation of the data, a Student t-distribution, which also incorporates excess kurtosis, will be applied to the GARCH models listed above. This would enable the comparison of a wider range of alternatives that can be used in the VaR estimation procedure. A number of different evaluation criteria, namely back-testing, market risk charge, violation rate and absolute deviation given a violation, will be employed in an attempt to better assess the models' performance, when compared to the actual observed market losses. Finally, the analysis of the VaR estimates obtained from the two forecasting periods will enable the comparison of VaR computation models under both highly volatile and relatively stable market conditions.

Using 15 years data of daily returns from the Standard & Poor's (S&P) 500 index, divided into two subsamples, this paper will argue that the performance of the various VaR estimation techniques employed fluctuates markedly, when put in the context of diverse market conditions. No model showed satisfactory forecasting accuracy in the extremely volatile setting of the GFC, with all estimations being much more precise during the post crisis period. Furthermore, incorporating asymmetric responses to volatility shocks proved to be of higher significance than accounting for excess kurtosis in the relevant computation process.

The rest of this paper is structured in the following fashion: the next section gives the theoretical framework on which the empirical study was build. It includes an overview of the various GARCH models used; details about the VaR framework; explanation of the Monte Carlo simulations technique; and a brief description of density forecasting. Section three is associated with the details related to the dataset being used. Section four deals with VaR estimation procedure, as well as the various criteria applied to evaluate the performance of the underlying forecasting models. Section five portrays the results of the empirical work and offers the associated analysis. The last section contains the concluding remarks.

Chapter 2

Background Theory

2.1 Generalized AutoRegressive Conditional Heteroscedasticity (GARCH)

The Autoregressive Conditionally Heteroscedastic (ARCH) model was pioneered by Engle (1982) and has since become one of the most influential statistical tools for analyzing financial data due to its ability to account for heteroscedasticity. The GARCH model (see Bollerslev, 1986) has further improved Engles framework by allowing the conditional variance to be dependent not only on the immediately previous value of the squared error term, as is the case in a regular ARCH model, but also on its own previous lags. The implication of the added lagged variance is that a first order GARCH model, one that uses only one lag of both the error term and the variance, can incorporate all the relevant information in order to compute the conditional variance and is sufficient to capture the persistence in volatility (Bollerslev et al., 1992).

There has been extensive research regarding the further development of the original GARCH model over the years, which has yielded a wide variety of models. Some of these models have been able to account for factors the original framework somewhat neglected. Dynamic volatility and volatility persistence, fatter tails of the distribution, mean reversion and asymmetric responses to volatility are among the characteristics of financial data, which the original GARCH model was unable to account for (Gerlach et al. 2011). In this paper popular extensions of the GARCH framework are considered, namely the GJR-GARCH and the exponential GARCH (EGARCH). Furthermore, all the models being discussed are not only put in the context of the Gaussian distribution, accounting only for the mean and standard deviation of the data, but also in the Student t-distribution in order to incorporate the potential existence of excess kurtosis in the relevant dataset.

The GJR-GARCH model, also known as asymmetric GARCH, can account for the asymmetric responses of the volatility to positive and negative shocks of the return innovations (Amado et al. 2011). Incorporating such effects could prove beneficial in terms of forecasting accuracy. EGARCH, much like GJR-GARCH, is able to incorporate the fact that positive return shocks have a different impact on volatility that negative shocks (Engle and Ng 1993). Furthermore, the model guarantees that the forecasted conditional variance will be non-negative since it uses the log value of the lagged volatility. Considering these characteristics of both models, their application in the process of analyzing financial data should provide a feasible alternative to the traditional GARCH model, and present an opportunity to evaluate the forecasting accuracy of a number of models with somewhat different characteristics.

2.2 Models

In this paper, VaR forecasting is based on three different models from the GARCH family used in both the Gaussian and Student t-distributions. The distributions' specifications are presented in the Appendix. Each of the models has the following mean equation,

$$r_t = \mu + \varepsilon_t \tag{2.1}$$

where r_t represents the time varying log returns from $r_t = \log(P_t - P_{t-1})$. ε_t is the normalized error, defined accordingly,

$$\varepsilon_t = \sqrt{h_t} z_t \tag{2.2}$$

and z_t is normalized by the conditional standard deviation, σ_t. For this study, z_t follows the Normal- and the Student t-distribution. Furthermore, z_t is also expected to be independent and identically distributed (i.i.d), with mean zero and unit variance given the information up to $t-1$. Finally, h_t expresses the conditional variance of r_t. The variation in the models is represented by the different specification of the conditional variance equation, starting with the traditional GARCH model,

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \tag{2.3}$$

with the imposed restriction that $\alpha + \beta$ <1 in order for the model to be stationary. A final point worth mentioning is that neither the α nor the β coefficients can gain negative values, a restriction known as the non-negativity constraint (Brooks, 2008). Since both the GJR-GARCH and the EGARCH incorporate asymmetric effects into the respective conditional variance estimates, the volatility dynamics specifications for these models differ, when compared to the traditional GARCH model.

The conditional variance for the GJR-GARCH at time t is calculated in the following fashion,

$$h_t = \omega + \alpha\varepsilon_{t-1}^2 + \gamma\varepsilon_{t-1}^2 I(\varepsilon_{t-1} < 0) + \beta h_{t-1} \tag{2.4}$$

where,

$$I = \begin{cases} 1 & \text{if } \varepsilon_{t-1} < 0, \\ 0 & \text{if } \varepsilon_{t-1} > 0. \end{cases}$$

The model allows for positive and negative return innovations to have different impact on the conditional variance, with the effects of positive innovations solely modelled by α, whereas the negative returns are incorporated by $\omega + \alpha + \gamma$.

The variance equation for the EGARCH is expressed as follows,

$$ln\left(h_t\right) = \omega + \beta \ln\left(h_{t-1}\right) + \gamma\frac{\varepsilon_{t-1}}{h_{t-1}} + \alpha\left[\frac{\varepsilon_{t-1}}{h_{t-1}} - \sqrt{\frac{2}{\pi}}\right] \tag{2.5}$$

where γ measures the asymmetric responses to volatility shocks. Thus, if $\gamma = 0$ the model is symmetric. If $\gamma < 0$, then positive shocks have a smaller effect than negative shocks. When $\gamma > 0$ then positive shocks increase the volatility more than negative once.

2.3 Value-at-Risk (VaR)

According to Jorion (1997), VaR is a measure of the highest expected loss, over a given time period, under normal market conditions, and a given confidence level. VaR thus represents a conditional quantile of the asset return loss distribution. Mathematically VaR can be defined as,

$$VaR_\alpha\left(L\right) = \min\left\{l : Pr\left(L > l\right) \leq 1 - \alpha\right\} \tag{2.6}$$

where l is the smallest loss, such that the probability of a portfolio loss L that is larger that l is smaller than or equal to 1-α (Nilsson, 2013). There are two arbitrarily chosen parameters related to this expression: the portfolio holding period and the confidence level. When the confidence level is taken into consideration, it is naturally related to the accuracy of the obtained results, in other words, the higher the confidence level, the less likely it is that a loss greater than the estimated VaR will occur under normal market conditions (Berry, 2013). The holding period on the other hand, is related to the horizon that serves as basis for risk analysis. The two most commonly chosen holding periods are one and ten days. If the 1-day VaR is taken into consideration, one is interested in the worst possible loss that can occur given a predetermined confidence level under normal market conditions (Nilsson, 2013).

The estimation of the 10-day VaR provides the foundations on which the Basel III framework has been build and deviates from the 1-day time frame, which is most commonly employed in practice. This is due to the fact that there is no universally acceptable methodology applicable in the process of 10-day VaR computations. Finger (2009) emphasizes the tensions that arise between regulators and financial institutions because of the divergence in the regulatory and banks' time-horizon for risk management. Hallerbach (1999) presents one solution to this problem, under which VaR is calculated over the entire time frame of interest. A severe drawback of this approach is that if the composition of the portfolio changes rapidly throughout the holding period, the relevance of such calculations can be deemed inadequate. An alternative comes in the form of the Square Root of Time Rule (SRR), deemed as the preferred practical solution. It involves scaling 1-day VaR estimates by taking the square root of the holding period. Even though such an approach is relatively easy to implement, when it comes to its theoretical soundness the method is rather controversial. Kaufman (2004), McNeil, Frey and Embrechts (2005) and Danielsson and Zigrand (2006) all point to the fact that the SRR is most relevant if the risk factors are normally distributed with zero mean and also independently and identically distributed over time. Such assumptions can rarely be justified when it comes to financial risk factors, especially such that have high frequencies. Moreover, time variation in the loss distribution is also ignored, further solidifying the notion that the approach has little theoretical foundations when implemented in conditional VaR estimates. Finally, Diebold et al., (1998) argue that the SRR overestimates VaR for short horizons and underestimates it for longer time frames. Although of small importance for conservative portfolios with short horizons, this characteristic makes the SRR estimates susceptible to sizable mistakes, when the nature of the portfolio is risky and the holding period is relatively long. The end result is the encouragement of risk taking for risky portfolios, implying that the employment of the SRR undermines one of the fundamental principles behind the regulatory methodology. With all this in mind, and with regards to the absence of theoretical soundness, the SRR method will not be further considered in this paper.

Like any theoretical model, there are a number of disadvantages related to VaR and its practical application. Artzner et al. (1999) emphasize that when using this estimation technique it is impossible to predict the amount of a loss in the case of a tail event. Since VaR is calculated based on the assumptions that market conditions will remain normal under the holding period, it is not necessarily feasible to estimate risk in extreme circumstances. Moreover, because of its widespread usage, this method can have a systemically destabilizing effect, in the sense that all financial institutions might become simultaneously inclined to terminate their portfolio positions because they use the same method to manage macroeconomic risk. Additionally, VaR does not always promote diversification, meaning that it is not a coherent risk measure (Nilsson, 2013).

According to Danielson et al. (2005) this is not a pertinent worry considering that sub-additivity violations are not of high relevance for assets that meet the stylized facts of returns.

Although VaR has its drawbacks, its widespread practical application points to the fact that the benefits related to this model far outweigh its pitfalls. If the advantages of this risk management tool are taken into consideration, it needs to be stated that VaR estimates are easy to interpret in monetary terms. More to the point, the method focuses on bad outcomes and is a risk measure applicable to all asset classes. It offers a probabilistic view on losses, meaning that it presents information about the probability of a loss that is bigger than the amount calculated based on the confidence interval in question. Lastly, VaR takes a holistic approach on losses by accounting for all the underlying risk factors via focusing on the distribution of portfolio losses (Nilsson, 2013).

2.4 Monte Carlo Simulations

When the MC simulations approach is taken into account, it was originally a method developed by physicists to use random number generation to compute integrals (Walsh, 2004). This simulation method has however quickly gained traction in other fields, with finance presenting a prominent example. Even though, the practical application of this estimation methodology is rather cumbersome, from a theoretical perspective it offers a number of important benefits, which has enabled MC simulations to gain enormous traction in the financial industry.

Berry (2013) defines MC simulations as an algorithm generating random numbers used to compute a formula that does not have a closed (analytical) form. That implies the simulation process requires some trial and error until an assessment about the results produced by the formula can be made. Since multiple random numbers are drawn over a large number of times the method offers a reasonable indication of what the output in the formula should be. More to the point, when forecasting and inference are taken into consideration, MC simulations provide valid and efficient estimations under parametric conditions. This forecasting technique can produce exact inference for infinite samples and can model instruments with non-linear and path-dependent payoff functions. Regarding the aforementioned complexity of financial instruments and the necessity for ever more sophisticated models required in risk management, MC simulations appear to be a viable choice in the process of estimating VaR.

Due to the uncertainty of forcasting, it is pivotal for to run a large number of simulations when striving to obtain a range, on which further analysis can be based. When the number of iterations in the simulation process is taken into account, there is a variety of alternatives presented in the literature. Berry (2013) states that the financial industry standard is 10,000, even if 1,000 iterations are enough to produce efficient estimates for terminal prices of most assets. 10,000 simulations provide a wide basis from which a better prediction for the movements of the future index prices can be extracted. Gerlach et al. (2011) however, have chosen to use an even higher number of iterations: 20,000, with Moller et al. (2004) using 100,000 iterations in their work. Since the research done by Gerlach et al. (2011) serves as the basis for this work, 20,000 MC iterations can be viewed as a plausible alternative in this particular context.

2.5 Density Forecasting

Tay (2004) defines density forecasting of the realization of a stochastic variable as estimation, related to the probability distribution of the possible future values of that variable. It gives a complete description of the uncertainty related to a forecast, something that can be seen as a palpable advantage when compared to point forecasting, which encompasses no description of the uncertainty in itself. Density forecasts for a time series, rather than the calculation of conditional mean and variance, has gained more and more recognition in the decision-theoretical context. Since financial returns are difficult to predict, it is hardly surprising that this is an area where such an estimation technique is readily applied, especially when it comes to high frequency trading (Timmermann, 2000). Diks et al. (2011) further emphasize the notion, that density forecasts provide the most comprehensive explanation for the uncertainty related to the movement of a random variable. The academic literature points to the fact that there are two main fields in which density forecasts are most commonly applied, namely in Central banks' inflation forecasting and risk management for financial institutions. With respect to the previously discussed strength of density forecasting when dealing with uncertainty and the importance of VaR in risk management, the implementation of the former estimation methodology in the computation of the latter can be viewed as adequate from a theoretical perspective. Finally, it needs to be stated that there is a variety of techniques presented by the literature when it comes to evaluating density forecasts, such as the weighted likelihood ratio test (Amisano et al., 2005), the probability integral transform (Diebold et al., 1998) and the model confidence set (Hansen et al., 2011). The lack of consensus about the optimal way such an assessment should be executed, coupled with the complexity of the process itself, render the evaluation of densities presented below beyond the scope of this paper. With that said, it would still be beneficial to compare the results of the density forecasts developed, based on the different underlying models with the true unobservable densities. Such a comparison will simplify the assessment of the forecasting accuracy of the various GARCH models in the context of financial risk management.

Chapter 3

Data

The data in this study comprises of daily continuously compounded returns of the S&P 500 index obtained from Thompson Financial Datastream. Descriptive statistics regarding the data are presented in Table 3.1. The estimations carried out are based on the daily returns of the S&P 500 index in the period from 04/01/1998 to 30/12/2012. Considering that the U.S. stock market is both the biggest in terms of trading volume and the most empirically studied, it can be viewed as an obvious choice for any study, which is focusing on the management of financial risk. Furthermore, the aforementioned timeframe incorporates two periods of major financial turmoil, the IT crash of 2000 and the GFC, as well as periods of relative systemic stability. In order to examine the accuracy of the above described VaR estimation methods under varying market conditions, two distinct forecasting periods will be examined. The first dataset taken under consideration consists of in-sample period, which begins 04/01/1998 and ends 30/12/2007. Based on the parameter estimates obtained from the different models, a forecasting period starting 04/01/2008 and ending 30/12/2009 will be examined. This timeframe incorporates the effects of the GFC and will determine the accuracy of VaR estimation methods under conditions of high volatility. The second dataset contains an in-sample period starting 04/01/2000 and ending 30/12/2010, with a forecasting period beginning 04/01/2011 and finishing 30/12/2012. This would enable the assessment of the post-crisis performance of the most widely applied financial risk measure and determine the suitability of VaR in the context of a somewhat less turbulent market environment.

TABLE 3.1: Descripitive Statistics

Measure	1998–2007	2000–2010	2008–2009	2011–2012
No. of observations	2607	2869	522	520
Mean (%)	0.0159	0.0051	0.0527	0.0220
Max (%)	5.5732	10.9572	10.9572	4.6317
Min (%)	7.0438	9.4695	9.4695	6.8958
Skewness	0.0383	0.1151	0.1178	0.5256
Excess Kurtosis	2.8565	7.8993	4.5451	4.6723
Yearly SD (%)	17.6936	21.4907	34.3005	18.478
JB	<0.0001	<0.0001	<0.0001	<0.0001
Autocorr. lag 1	0.0326	0.0871	0.1418**	0.0853
Autocorr. lag 2	0.0242	0.0671*	0.1049**	0.1111**
Autocorr. lag 3	0.0212	0.0307*	0.0791**	0.1694**

Note: This table shows the descriptive statistics for the daily Standard & Poor's 500 returns for different relevant periods. Standard errors are presented in parenthsis. JB is the p-value from the Jarque and Bera (1987) test with the null hypothesis of normally distributed returns. *Denotes significance at the 5% level. **Denotes significance at the 1% level.

3.1 Simulated Returns and VaR Quantiles

As the focus of this paper lies in the forecasting process of 1- and 10-day VaR estimates, the MC method is implemented to simulate conditional variances and returns. Initially, parameters are estimated, for the relevant mean and variance equations. The α and the β coefficients are the only estimated parameters in the context of the traditional GARCH framework, with the addition of the leverage coefficient γ for the EGARCH and the GJR-GARCH models. For the Student t-distribution, the degrees of freedom parameter υ is also estimated. The parameter estimates, together with the empirical residuals from the in-sample period are used as inputs for the MC simulations and 20,000 alternative sample paths are generated. The simulations will result in a matrix, denoted M_t, with i number of rows and k number of columns. The first step in order to construct 10-day VaR estimations, involves creating a submatrix, consisting of 10-day cumulative returns at any given day t. (3.1) below is considered for this process,

$$M_{t}* = \sum_{j=1}^{10} M_{t-j+1} \qquad \forall k \tag{3.1}$$

for any day t in the out-of-sample period, the equation sums j number of days for all k colums, from the simulated matrix, M_t. The output is a matrix M_t*, with 10-day cumulative over-lapping returns.

In order to extract 10-day VaR estimates from M_t*, the following equation was implented,

$$VaR_\alpha (10) = -G_\alpha^{-1} (M_t*) \tag{3.2}$$

where (3.2) is the holding period, $-G_\alpha^{-1}$ is the inverse of the return probability density function and α is the confidence level of interest. Finally, by applying (3.1) on (3.2), a vector of 10-day VaR estimates is obtained.

3.2 Testing and comparing VaR models

Since the density forecasts obtained from the MC simulations will be compared with the true unobservable densities, and because of the fact that such an approach may not give sufficient information as to which of the models is optimal in this particular situation, criteria to compare the competing VaR forecasting models need to be implemented (Wilhelmsson, 2011). Gerlach et al. (2011) point to a number of viable alternatives, namely observed violation rates (VRate), market risk charges (MRC), absolute deviation (AD) given a violation and two standard back-testing criteria.

When starting with the back-testing, the most commonly applied tests are the Kupiec unconditional overage (UC) test and the Christoffersen conditional coverage (CC) test. The UC test is aiming to determine whether the reported frequency of VaR violations is significantly more (or less) than the predetermined frequency. Thus, if the number of observed violations is statistically different from the number of expected violations, the null hypothesis of accurate VaR model is rejected (Berkowitz et al., 2011). In the context of this study, the test will be carried out in its two-sided version, where a confidence interval is constructed, and if the number of forecasted violations falls outside this interval the underlying model is rejected. The CC test further improves the stringency of the back-testing methodology by requiring, not only correct unconditional coverage, but also independence of the violations. Thus, the test investigates not only if the evaluated model presents the correct frequency of VaR violation, but also if the violations are independent of each other (Nilsson, 2013). Since the test is essentially composed of two parts, unconditional coverage and independence test, there are two likelihood ratios (LR) comprising the CC test, with the details concerning their specifications presented in the Appendix of this paper.

Regarding the fact that 10-day VaR estimations will be computed based on rolling and over-lapping 10-day intervals, neither of these tests are relevant in this context. Violations cannot be expected to be independent in such a setting, nor should the i.i.d assumption related to the UC or CC tests hold (Gerlach et al., 2011). Both tests will however be applied to the 1-day VaR forecasts.

The VRate is among the simplest criterion used to compare different VaR forecast, and is estimated with the following formula,

$$VRate = \sum_{t=n+1}^{m} I\left(\frac{r_t < -VaR_t}{m}\right) \tag{3.3}$$

where,

$$I = \begin{cases} 1 & \text{if} \quad \mathrm{r}_t < -VaR_t, \\ 0 & \text{if} \quad \mathrm{r}_t > -VaR_t. \end{cases}$$

Thus the VRate is simply the number of VaR violations divided by the forecasted sample size, where n represents the number of in-sample observations and m stands for the forecast sample size. A straightforward implication of this method is the fact that a VRate that is close to the α confidence level is desirable, with models overestimating risk more favourable in the context of the Basel legislative framework (Gerlach et al., 2011).

The MRC is based on a number of quantitative VaR parameters: a 10-trading day horizon; at least a year of in-sample data and a 99% confidence interval (Jorion, 2002). For each day t in the out-of-sample period, the MRC is thereafter calculated as follows,

$$MRC_t = \sup\left\{\left(VaR_{t-1}, \overline{VaR_{60t}} \times 3 + k\right)\right\} \tag{3.4}$$

where MRC_t is equal to the supremum of VaR_{t-1}, which is the VaR estimate for the previous day, and $\overline{VaR_{60t}}$, representing the 60 day VaR moving average, and k is the penalty factor. The k is related to the process of penalizing market risk projections that are not sufficiently conservative, with its respective values being presented in the Table 3.2. A model with lower MRC is deemed to be more adequate in terms of financial risk management (Gerlach et al., 2011). Finally, the mean value of the MRC will be used as an evaluation criterion of the forecasting accuracy of each of the models.

Considering that the size of the expected loss given a violation needs to be perceived as an important characteristic of a risk management tool, the magnitude of the violations and not solely their frequency should be taken into account. McAleer et al. (2008) emphasize the importance of this analytical approach and propose the following estimation method of the size of the loss given a violation,

$$AD_t = |r_t - (-VaR)_t| \tag{3.5}$$

where r_t is only considered when there is a violation. The mean and the maximum for this criterion will be computed, and models with lower AD values will be preferable.

Table 3.2: Modified Basel Accord Penalty Zones.

Zone	Violations	Cumulative Probability	Plus k-factor
Green	0	0.0054	0.0000
	1	0.0336	0.0000
	2	0.1076	0.0000
	3	0.2366	0.0000
	4	0.4051	0.0000
	5	0.5807	0.0000
	6	0.7333	0.0000
	7	0.8459	0.0000
	8	0.9191	0.0000
Yellow	9	0.9611	0.4010
	10	0.9828	0.4982
	11	0.9933	0.6480
	12	0.9973	0.7512
	13	0.9991	0.8505
Red	14	0.9999	1.0000

Note: The Modified Basel Accord Penalty Zones are evaluated by the Basel Committee on Banking Supervision (1996) based on 520 and 522 trading days; true coverage is 99%. The k-factor is obtained by linear interpolation, based on the number of trading days and the binominal cumulative distribution.

Chapter 4

Results

The empirical part of this work begins with the presentation and analysis of the parameter estimates obtained from the different GARCH models. The 1– and 10–day VaR estimates for both subsamples are then presented and compared in an attempt to determine which of models produced the most accurate forecasting results.

4.1 Parameter Estimation

Tables 4.1 and 4.2 display the relevant coefficients obtained from the parameter estimation based on the different models and sample periods. The μ coefficient is stationary, and is related to the mean equation of the models. It is relatively close to zero in all the cases that are being analysed. This should not be seen as surprising, when one takes into account the characteristics of financial data. When turning to the parameters obtained based on the variance equation, it is important to point out that the α and β coefficients in the GARCH model are very close to unity in both distributions and subsamples, relating to the implication that shocks in the conditional variance are highly persistent. The non-negativity constraint is also not breached in none of the cases, proving the theoretical soundness of the model when applied to this particular dataset. Turning the focus to the GJR-GARCH parameters, it is beneficial to point out the fact, that the mean of the Normal distribution is negative in both subsamples, albeit it being very close to zero. In all the cases discussed, the γ coefficients have value of zero, implying that the lagged squared error term only has significance in the estimation of the conditional variance when negative. Furthermore, in both subsamples the γ coefficients are significant at the 1% confidence level, implying that the well-documented leverage effects is present in this dataset. Another intriguing aspect of the estimates is the negative value of the γ coefficient in the context of EGARCH framework. This leads to the implication that positive return shocks generate less volatility. The final point worth exploring in the context of parameter estimates analysis is related to the fact that the γ coefficients for EGARCH and GJR-GARCH have opposite signs. This is the case because both models use different means in order to account for asymmetric volatility responses.

Table 4.1: Parameter estimates from Standard & Poor's index 1998–2007.

Model	μ	ω	α	γ	β	v	LL
1 GARCH-N	0.0376*	0.0111**	0.7532**		0.9179**		-4259.54
	(0.0171)	(0.0017)	(0.0071)		(0.0073)		
2 EGARCH-N	0.0039	0.3878**	0.0932**	-0.1275**	0.9844**		-4194.90
	(0.0158)	(0.0012)	(0.0102)	(0.0078)	(0.0013)		
3 GJR-N	-0.0064	0.0122**	0	0.1280**	0.9260**		-4193.03
	(0.0167)	(0.0016)		(0.0124)	(0.0078)		
4 GARCH-t	0.0482**	0.0076**	0.0753**		0.9179**	7.1835**	-4215.13
	(0.0159)	(0.0028)	(0.0098)		(0.0092)	(0.9833)	
5 EGARCH-t	0.0206	0.0008**	0.8996**	-0.1328**	0.9872**	8.2813**	-4157.74
	(0.0154)	(0.0015)	(0.0136)	(0.0111)	(0.0019)	(1.1721)	
6 GJR-t	0.0165**	0.0093**	0	0.1327**	0.9276**	8.8847**	-4161.82
	(0.0158)	(0.0022)		(0.0164)	(0.0094)	(1.3531)	

Table 4.2: Parameter estimates from Standard & Poors index 2000–2010.

Model	μ	ω	α	γ	β	v	LL
1 GARCH-N	0.0388*	0.0097**	0.0615**		0.9312**		-3674.84
	(0.0177)	(0.0016)	(0.0061)		(0.0066)		
2 EGARCH-N	0.0076	0.0025**	0.0792**	-0.1149**	0.9838 **		-3616.31
	(0.0168)	(0.0012)	(0.0112)	(0.0071)	(0.0016)		
3 GJR-N	-0.0004	0.0118**	0	0.1187**	0.9313**		-3616.34
	(0.0167)	(0.0017)		(0.0116)	(0.0079)		
4 GARCH-t	0.0473**	0.0063**	0.0621**		0.9348**	7.8461**	-3632.93
	(0.0163)	(0.0026)	(0.0090)		(0.0090)	(1.0517)	
5 EGARCH-t	0.0202	0.0005	0.0814**	-0.1187**	0.9866**	9.4984**	-3580.42
	(0.0160)	(0.0015)	(0.0015)	(0.0103)	(0.00123)	(1.3644)	
6 GJR-t	0.0194	0.0089**	0	0.1239**	0.9327**	9.6880**	-3586.43
	(0.0162)	(0.0022)		(0.0161)	(0.0096)	(1.4266)	

Note: These tables report the in-sample parameter estimates for the daily Standard & Poor's 500 returns for the two in-sample periods. Standard errors are given in parentheses. LL is the log likelihood function value. *Dentoes significance at the 5% level. **Denotes significance on the 1% level. All are obtained with the Econometrics Toolbox in MATLAB.

4.2 Subsample I

4.2.1 10-day VaR

The analysis of the VaR estimates begins with the period containing the effects of the GFC. Table 4.3 presents the relevant criteria in the assessment of each of the models for both the 1% and the 5% risk level for 10-day VaR estimates. In an attempt to measure the performance of each of the estimation methods, they are ranked in accordance with the specifications of each of the evaluation criteria. It must be emphasized that, since MRC is incorporated in the Basel II framework, it only focuses on the 1% risk level, thus the computation of this criteria for $\alpha = 5\%$, is deemed redundant.

TABLE 4.3: Summary statistics for 10-day VaR forecast 2008–2009

Model	VRate (%)	MRC mean	AD max	AD min	Penalty	Violations
$\alpha = 1\%$						
1 GARCH-N	6.43	35.9991	20.7020	5.5821	1.0	33
2 EGARCH-N	4.68	44.7815	18.5121	5.228	1.0	24
3 GJR-N	4.87	41.2491	19.3265	6.0471	1.0	25
4 GARCH-t	4.68	42.3760	19.2265	5.8311	1.0	24
5 EGARCH-t	4.68	44.8816	18.5467	5.3600	1.0	24
6 GJR-t	4.68	45.3738	17.9925	5.0219	1.0	24
$\alpha = 5\%$						
1 GARCH-N	12.48	–	24.5895	5.5358	–	64
2 EGARCH-N	9.75	–	23.3536	5.7335	–	50
3 GJR-N	11.50	–	24.1336	5.5635	–	59
4 GARCH-t	12.28	–	24.3383	5.4460	–	63
5 EGARCH-t	9.94	–	23.6295	5.9354	–	51
6 GJR-t	11.31	–	24.1234	5.7872	–	58

Based on the results presented in the table it is apparent that all the models underestimate risk substantially. When starting with the number of violations for $\alpha = 1\%$, they range from 24 to 33, with GARCH-N performing the worst of all models considered. In terms of VRate the two EGARCH models, GJR-t and GARCH-t all yield the same violation rate. This leads to the conclusion that in this particular case the incorporation of asymmetric volatility shocks does not increase the forecasting accuracy of the models at the 1% level. $\hat{\alpha}$, which represents the actual VRate of the models'VaR estimates, is approximately equal to 4.7%, further ascertaining the inadequacy of the forecasting results in the context of such turbulent macroeconomic environment as the one observed during the GFC. This result is however to be expected, since as previously stated, one of the major drawbacks of the VaR framework is the fact that it cannot efficiently account for losses under extreme market conditions. The implication of this pitfall is the somewhat paradoxical fact that the best performing model in terms of MRC is the GARCH-N, which offers the highest number of VaR violations in this dataset. This is partially due to the fact that all the models have a penalty factor of one in accordance with the Basel II framework. Thus, the model that underestimates risk the most and has a $\hat{\alpha} \approx 6.3\%$, or more than six times the desired value, yields the best result in terms of MRC. When the AD criterion is being analysed, GJR-t is the model offering the lowest maximum and mean. GARCH-N is performing the worst in terms of maximum loss, further solidifying the notion that it is the least conservative model in this context.

EGARCH-N performs best at the 5% risk level having 50 violations, with GARCH-N offering the poorest precision. With a VRate of $\hat{\alpha} = 9.75\%$ however, it can hardly be claimed that the EGARCH-N estimates present a reliable source of reference in the context of financial risk management. The difference in the estimation accuracy can be related to the incorporation of asymmetric responses to volatility shocks, with EGARCH in both distributions outperforming the other two frameworks. EGARCH-N is also the estimation method scoring best in relation to the maximum AD criterion, rendering it the most accurate forecasting technique at this confidence level A final point worth exploring is the fact that accounting for excess kurtosis in this dataset does not seem to grant a palpable difference in the VaR estimation precision, with the integration of asymmetric responses to volatility shocks being much more relevant.

4.2.2 1-day VaR

Table A.1 (Appendix) shows the p-values for the UC and CC test for the 1-day VaR estimation models with respect to both the 99% and 95% confidence level. All of the models fail both tests at both confidence levels. With that said, the other available evaluation criteria will still be applied to the forecasting results in an effort to obtain a more detailed overview of the models effectiveness.

The overall performance of the different GARCH models' estimates was even poorer when the 1-day VaR is taken into consideration. It needs to be stated that all models fell in the red penalty zone of the Basel II framework and thus have a penalty factor of one when the MRC is being computed. This streams from the fact that the model having the lowest VRate is GJR-t with $\hat{\alpha} = 4.79\%$. Even though the forecasting accuracy offered by the models is not adequate, accounting for excess kurtosis has proven beneficial in the context of this criterion, with the models using the Student t-distribution outperforming the rest. In terms of MRC, GARCH-N offers the lowest value, due to the fact that the model underestimates risk the most and gives the lowest VaR estimates. GARCH-t has the smallest AD maximum, with GARCH-N having the lowest mean value in this criterion. Accounting for asymmetric responses to volatility shocks and excess kurtosis does not seem to have any significant impact in relation to neither the MRC, nor the AD criteria.

Table 4.4: Summary statistics for 1-day VaR forecast 2008–2009

Model	VRate (%)	MRC mean	AD max	AD min	Penalty	Violations
$\alpha = 1\%$						
1 GARCH-N	7.85	11.5067	6.6045	1.7398	1.0	41
2 EGARCH-N	5.94	12.7156	6.3050	1.9543	1.0	31
3 GJR-N	7.09	12.0010	6.4237	1.8149	1.0	37
4 GARCH-t	5.17	14.0192	5.9180	1.8621	1.0	27
5 EGARCH-t	5.36	13.2332	6.1771	2.0226	1.0	28
6 GJR-t	4.79	13.6477	6.2442	2.1068	1.0	25
$\alpha = 5\%$						
1 GARCH-N	15.52	–	7.6710	1.6392	–	81
2 EGARCH-N	13.98	–	7.4904	1.6404	–	73
3 GJR-N	15.53	–	7.7156	1.6872	–	80
4 GARCH-t	15.13	–	7.6393	1.6342	–	79
5 EGARCH-t	14.18	–	7.5803	1.6961	–	74
6 GJR-t	15.33	–	7.7927	1.6977	–	80

EGARCH-N has the lowest VRate at the 5% risk level, with EGARCH-t being the second best model. However, both estimation methods offer relatively low forecasting precision, with their $\hat{\alpha}$ values being 13.98% and 14.18% respectively. EGARCH-N has the lowest maximum AD value, whereas the GARCH-t has the lowest AD mean. With that said, all the models offer low forecasting accuracy, and it is difficult to view them as adequate estimators of financial risk when the market conditions are as volatile as during the GFC.

4.3 Subsample II

4.3.1 10-day VaR

The empirical results for the 10-day VaR forecasts related to the post crisis forecasting period are presented in Table 4.5. It is of interest to point out the relative conservativeness of all of the models being employed. When it comes to the number of observed VaR violations, all estimation techniques fall in the green zone of the Basel II legislation. This varifies that they are all acceptable from a regulatory point of view and do not encourage risk-taking in this particular context. The model that performs best in terms of VRate is EGARCH-t with a $\hat{\alpha}$ = 0.97%. Although EGARCH-N produces even conservative estimates, its $\hat{\alpha}$ = 0.78% is further away from the desired value of 1%, making it the second best model according to this criterion. Overall, the EGARCH models in the context of both the Gaussian and Student t-distribution outperform the other forecasting methods being used. A further proof of this can be found when the maximum of the AD criterion is examined. This implies that the means this framework uses, in the process of incorporating asymmetric responses to volatility shocks are the most suitable for this dataset. The relevance of the different distributions is however somewhat muted, especially in the case of the GJR-GARCH models, where both distributions produce the same number of VaR violations and the GJR-N yields better results in terms of AD mean. GARCH-N is the model ranked best according to MRC, followed by GJR-N and GJR-t. The models that are most conservative in their estimates are the ones that score the worst in the context of this evaluation criterion. This is due to the fact that their forecasted VaR values are higher and the same penalty factor is applied to all models.

Table 4.5: Summary statistics for 10-day VaR forecast 2011–2012.

Model	VRate (%)	MRC mean	AD max	AD min	Penalty	Violations
$\alpha = 1\%$						
1 GARCH-N	1.56	26.9931	9.5165	3.6692	1.0	8
2 EGARCH-N	0.78	35.1579	6.4950	2.7411	1.0	4
3 GJR-N	1.36	28.0694	8.6561	2.4896	1.0	7
4 GARCH-t	1.36	30.9006	8.5223	3.2144	1.0	7
5 EGARCH-t	0.97	33.1660	6.9722	2.8629	1.0	5
6 GJR-t	1.36	28.6856	8.7487	3.6117	1.0	7
$\alpha = 5\%$						
1 GARCH-N	4.68	–	12.8750	3.0458	–	24
2 EGARCH-N	2.34	–	11.3117	3.9516	–	12
3 GJR-N	4.29	–	12.5071	3.0759	–	22
4 GARCH-t	5.46	–	13.0286	2.7731	–	28
5 EGARCH-t	3.31	–	12.1715	3.4831	–	17
6 GJR-t	6.24	–	13.0854	2.5558	–	32

The EGARCH models are overly conservative when the 5% risk level is analysed. They overestimate risk in the context of both distributions, and naturally offer lower VRate than any of the other models. The best performing model in the context of this criterion is GARCH-N with $\hat{\alpha}$ = 4.68 followed by GARCH-t with $\hat{\alpha}$ = 5.46%, proving that the most conservative estimation method is not necessarily the most optimal in terms of forecasting VaR. The EGARCH-N is however the model performing best in the context of maximum AD, with the GJR-t having the lowest mean AD. Due to the discrepancy in the forecasting accuracy of the models, it is difficult to assess which aspects of the forecasting techniques are most significant in this setting.

4.3.2 1-day VaR

Table A.2 in Appendix shows the p-values for the UC and CC test for the 1-day VaR estimation models, with respect to both the 99% and 95% confidence level for the post crisis subsample. In this context, all models pass the Kupiec test and only the EGARCH-N failed the CC test at $\alpha = 5\%$. The performance of the models is thus much better in light of this evaluation criterion if compared to the results obtained from subsample one.

TABLE 4.6: Summary statistics for 1-day VaR forecast 2011–2012.

Model	VRate (%)	MRC mean	AD max	AD min	Penalty	Violations
$\alpha = 1\%$						
1 GARCH-N	2.12	8.6014	4.1307	1.0975	0.6480	11
2 EGARCH-N	1.35	9.8604	3.5841	1.1798	0.0	7
3 GJR-N	2.12	8.0554	4.2737	1.1288	0.6480	11
4 GARCH-t	1.35	10.1324	3.8107	1.3549	0.0	7
5 EGARCH-t	1.54	9.6801	3.7917	1.1514	0.0	8
6 GJR-t	1.35	10.4593	3.5841	1.1798	0.0	7
$\alpha = 5\%$						
1 GARCH-N	6.35	–	12.8750	3.0458	–	24
2 EGARCH-N	5.38	–	11.3117	3.9516	–	12
3 GJR-N	6.73	–	12.5071	3.0759	–	22
4 GARCH-t	6.35	–	13.0286	2.7731	–	28
5 EGARCH-t	6.35	–	12.1715	3.4831	–	17
6 GJR-t	5.38	–	13.0854	2.5558	–	32

In terms of VRate, the models presenting the most accurate 1-day VaR estimates for the post-crisis forecasting period are EGARCH-N, GARCH-t and GJR-t. All of these models, including the EGRACH-t, are in the green penalty zone of the Basel framework. The GJR-N and the GARCH-N however fall in the yellow zone and are therefore assigned a penalty factor in accordance with the regulatory requirements. This is done by a linear interpolation in order to approximate the k penalty values that are specified in relation to the number of violations and binominal cumulative probabilities in Table 3.2. Although still acceptable in the context of risk management, both models can be regarded as somewhat less conservative than desirable, especially considering that the market conditions during this forecasting period are relatively stable. With that in mind, these models are related to the lowest mean MRC values at this confidence level. This might be deemed, as an additional argument supporting the notion that in some situations, the fact that a model underestimates risk can have a more significant impact on the value than the k penalty factor, particularly when the number of violations is close to the green zone. In terms of maximum AD, EGARCH-N and GJR-t are the top ranked models, while GARCH-N has the lowest mean according to this criterion.

At the 5% level, GJR-t and EGARCH-N provide the most precise VRate forecasts, with the same two models having the lowest mean AD values. EGARCH-N offers the lowest maximum AD as well, making it the best performing model in this context. Accounting for asymmetric responses to volatility shocks can thus be seen as beneficial, while simultaneously it could be claimed that incorporating excess kurtosis is not of great significance in this setting.

4.4 Summary of VaR forecasting

The final part of this analysis is associated with the comparison of the VaR estimates for the two different subsample periods. A general overview of the overall performance of the different models highlights that the accuracy of the forecasted 1-day and 10-day VaR values fluctuates greatly, when the underlying market conditions experience sizable deviations. When the crisis period of 2008 and 2009 is considered, all the models underestimate the inherent systemic risk. Contrary to this, during the period associated with relatively low levels of volatility, the models being employed tend to overestimate risk and produce 10-day VaR forecasts that are somewhat more conservative than required. The 1-day VaR estimates tend to be less restrictive and consequently more fitting in the context of the post-crisis market conditions. The lack of conservativeness in a turbulent systemic environment is however far from desirable. Since financial risk management is most important in highly volatile market conditions, the employment of more liberal estimation methods seems to undermine the most pivotal aspect of the Basel Accord.

The EGARCH models are certainly the most conservative when it comes to the 10-day VaR computations in the period after the GFC. It can be stated that in both distributions, these models emphasize risk more noticeably at the 5% level. With that said, both frameworks are associated with the most accurate VRate results when $\alpha =$ 1%, where GARCH-N is ranked highest according to the MRC criterion simply because in these circumstances the model is presenting the least conservative VaR estimates. At the 1% risk level for the 1-day VaR, no model offers superior violation frequency, with none of the models being overly conservative. On the contrary, the GJR-N and GARCH-N fall in the yellow penalty zone of the Basel framework. With that said, both model are related to the lowest MRC values, even though all other forecasting techniques are assigned a penalty factor of zero. Finally, it should be noted that the EGARCH-N exhibits the highest overall 1-day VaR forecasting accuracy at $\hat{\alpha} = 5\%$, whereas no model can be viewed as superior in the setting of 10-day VaR estimates for the same confidence level.

When the estimates related to the GFC period are taken into consideration, it is very difficult to point out which of the various characteristics of the estimation techniques being used are of highest importance in the process of forecasting VaR. The performance of all of the models at the 1% level is relatively similar, with the GARCH-N yielding worst results when the number of 1- and 10-day VaR violations is evaluated. However, the same model offers the lowest MRC score in both settings, further emphasizing the fairly poor overall performance of the forecasting methods during the financial crisis. With that in mind, the EGARCH models have proven to be slightly more conservative than any of the other estimation techniques in this study at the 5% risk level, for both computation horizons. The precision of these models is nevertheless inadequate, further solidifying the premise that the models being analyzed struggled and substantially underestimated the risk levels during the GFC.

Chapter 5

Conclusion

This paper evaluated the possibility of implementing Monte Carlo simulations in the process of estimating 1- and 10-day VaR forecasts across a variety of parametric heteroscedastic models. Regular GARCH, EGARCH and GJR-GARCH were all considered in both the Gaussian and Student t-distributions. The models were put in the context of diverse market conditions in an attempt to test their forecasting accuracy, in both extreme and relatively stable systemic environment. No model performed consistently well across the various estimation horizons, confidence levels and market settings. For the 10-day VaR estimates during the GFC, none of the models outperformed the others, with the GARCH-N offering the worst forecasting precision at both quantile levels. In terms of 1-day computations, the models'ability to adapt to extreme market conditions was further diminished, with their performance being inferior for this estimation horizon. The forecasting precision of the models was substantially improved in the post crisis subsample, where all 10-day estimates satisfied the requirements of the Basel Accord and fall in the green penalty zone at the 1% risk level. As far as 1-day VaR results for this subsample are concerned, EGARCH-N was the model contributing the most exact forecasts, indicating that volatility asymmetry is an important feature in this setting. An interesting empirical finding was the fact that conservative models tend to present better VaR violation rates, but this conservativeness impairs their performance in terms of mean market risk charge. Thus, a combination of sophisticated models can be seen as a viable managerial strategy in the context of financial risk management and VaR forecasting. The performance of additional parametric models, placed in the context of distributions allowing for additional skewness presents an appealing opportunity for future research.

Appendix A

Appendix

A.1 Equations

Christoffersen conditional coverage test is defined accordingly,

$$LR_{uc} = -2\left[\ln\left(p^{x}\left(1-p\right)^{N-x}\right) - \ln\left(\pi^{x}\left(1-\pi\right)^{N-x}\right)\right] \sim \chi^{2}\left(1\right) \tag{A.1}$$

where N is the number of observations in the testing period, x being the number of observed violations, p representing the anticipated frequency of violations and denoting the actual violation frequency. The test statistic is asymptotically chi-squared distributed, with one degree of freedom.

In the independence part of the test, violations are assigned a value of one and non-violations a value of zero. The test statistics for the independence part of the Conditional Coverage test is presented below,

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

where n_{00} represents the number of transitions from state zero to state zero; n_{01} is the number of transitions from state zero to state one; n_{10} denotes the number of transitions from state one to state zero; and n_{11} signifies the number of transitions from state one to state one. Thus, the conditional frequency of transitions between the states can be calculated accordingly,

$$\pi_{00} = \frac{n_{00}}{n_{00}+n_{01}} \tag{A.3a}$$

$$\pi_{01} = \frac{n_{01}}{n_{00}+n_{01}} \tag{A.3b}$$

$$\pi_{10} = \frac{n_{10}}{n_{10}+n_{11}} \tag{A.3c}$$

$$\pi_{11} = \frac{n_{11}}{n_{10}+n_{11}} \tag{A.3d}$$

n_0 captures the number of non-violations and is equal to $n_0 = n_{00} + n_{10}$. n_1 stands for the number of violations obtained by $n_1 = n_{01} + n_{11}$. n_0 is then equal to $\pi_0 = n_0/N$ and $\pi_1 = n_1/N$ with π_0 and π_1 being the unconditional frequency of violations and non-violations respectively. Finally the results of the two likelihood ratios (LR) are combined,

$$LR_{CC} = LR_{UC} + LR_{IND} \sim \chi^2 (2) \tag{A.4}$$

where the test statistic is asymptotically chi-squared distributed, with two degrees of freedom. All the specifications regarding the CC test stream from the work of Nilsson (2013).

A.2 Tables & Figures

Table A.1: *p*-values for the CC, UC and IND tests 2008-2009.

Model	LR CC (%)	LR UC	LR IND
$\alpha = 99\%$			
1 GARCH-N	<0.0001	<0.0001	0.9416*
2 EGARCH-N	<0.0001	<0.0001	0.9600**
3 GJR-N	<0.0001	<0.0001	0.9554*
4 GARCH-t	<0.0001	<0.0001	0.9064*
5 EGARCH-t	<0.0001	<0.0001	0.113***
6 GJR-t	<0.0001	<0.0001	0.9592**
$\alpha = 95\%$			
1 GARCH-N	<0.0001	<0.0001	0.8697
2 EGARCH-N	<0.0001	<0.0001	0.8456
3 GJR-N	<0.0001	<0.0001	0.8611
4 GARCH-t	<0.0001	<0.0001	0.8611
5 EGARCH-t	<0.0001	<0.0001	0.8456
6 GJR-t	<0.0001	<0.0001	0.8611

Table A.2: *p*-values of unconditional and conditional coverage tests for each model

Model	LR CC (%)	LR UC	LR IND
$\alpha = 99\%$			
1 GARCH-N	0.9853**	0.9741**	0.9384*
2 EGARCH-N	0.7482	0.5491	0.8613
3 GJR-N	0.9851**	0.9743**	0.9384*
4 GARCH-t	0.7484	0.7474	0.8612
5 EGARCH-t	0.8313	0.5492	0.8663
6 GJR-t	0.7482	0.5491	0.8611
$\alpha = 95\%$			
1 GARCH-N	0.6793	0.3162	0.8531
2 EGARCH-N	0.9981***	0.9994***	0.8992
3 GJR-N	0.7411	0.5911	0.8443
4 GARCH-t	0.6783	0.3093	0.8531
5 EGARCH-t	0.9462*	0.9461*	0.8532
6 GJR-t	0.8684	0.7572	0.8994

Note: CC represents the conditional coverage test, UC denotes the unconditional coverage test and IND is the independence test. These tables present the *p*-values for each model. Equations for the tests are presented in this chapter. *Denotes significance at the 10% level. **Denotes significance at the 5% level. ***Denotes significance at the 1% level.

FIGURE A.1: 10-day VaR-forecast for 2008–09 period: (t) $\alpha = 1\%$ (b) $\alpha = 5\%$.

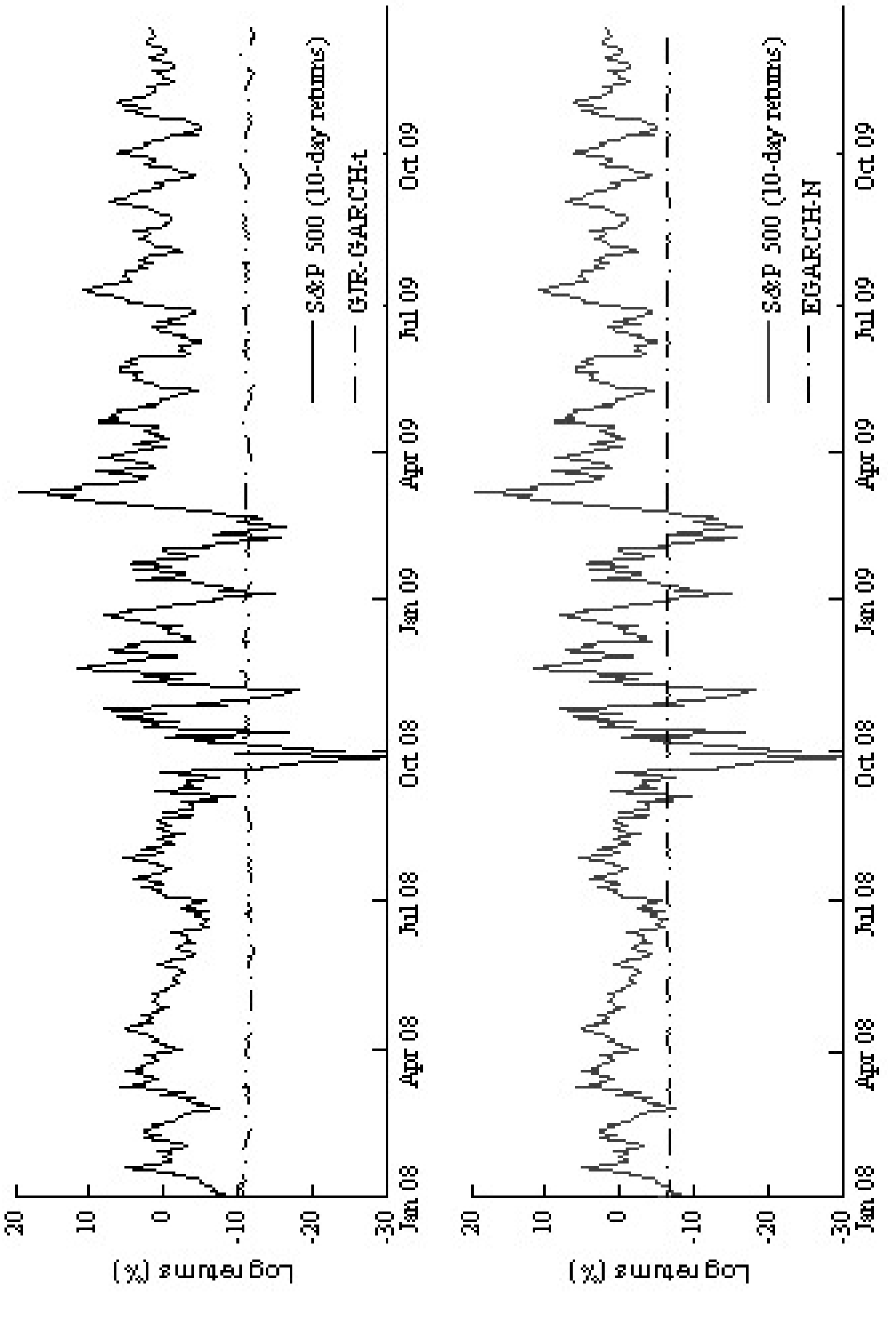


FIGURE A.2: 1-day VaR-forecast for the Global Financial Crisis period: (t) $\alpha = 1\%$ (b) $\alpha = 5\%$.

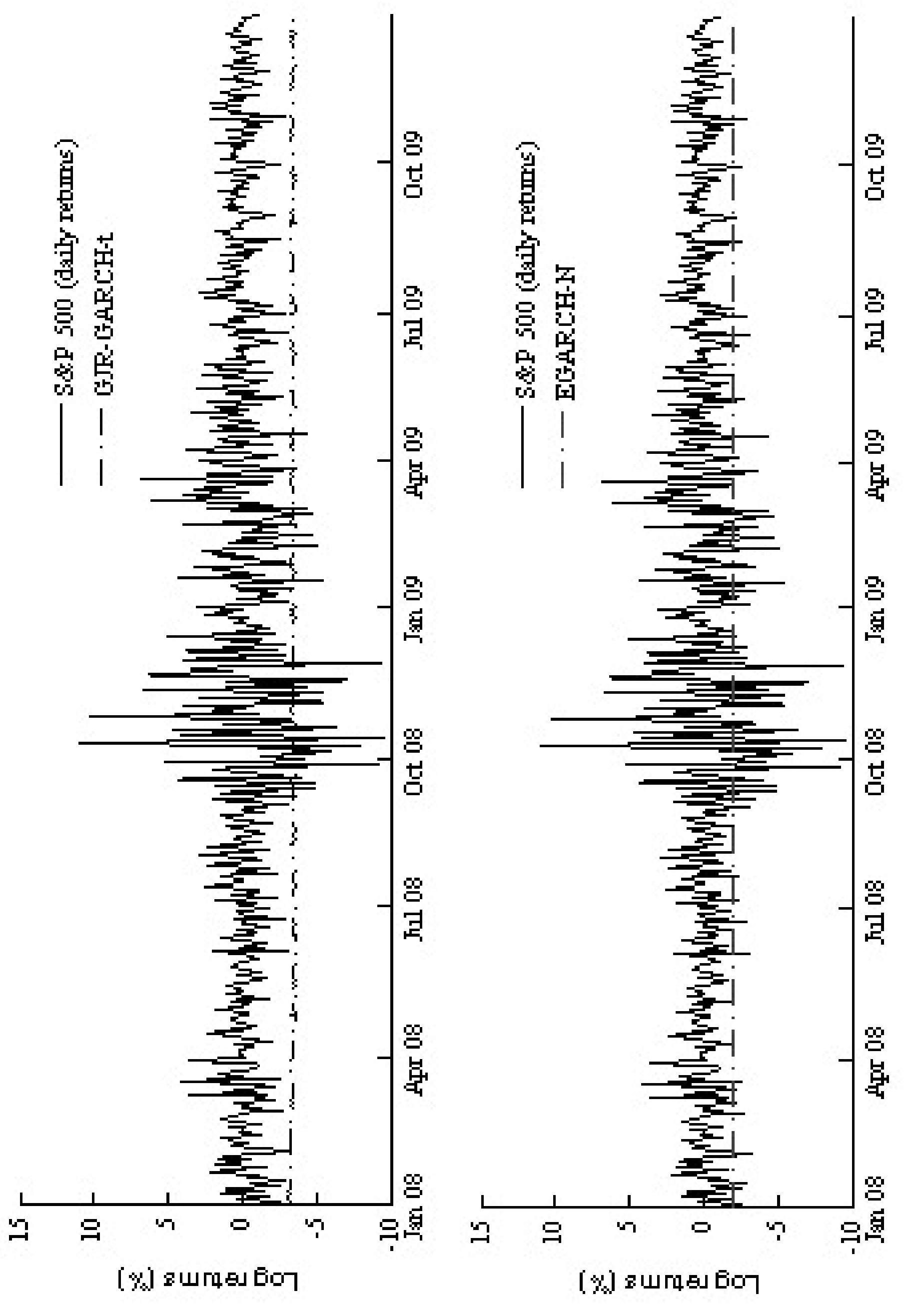


Figure A.3: 10-day VaR-forecast for the period after the Global Financial Crisis period: (t) $\alpha = 1\%$ (b) $\alpha = 5\%$.

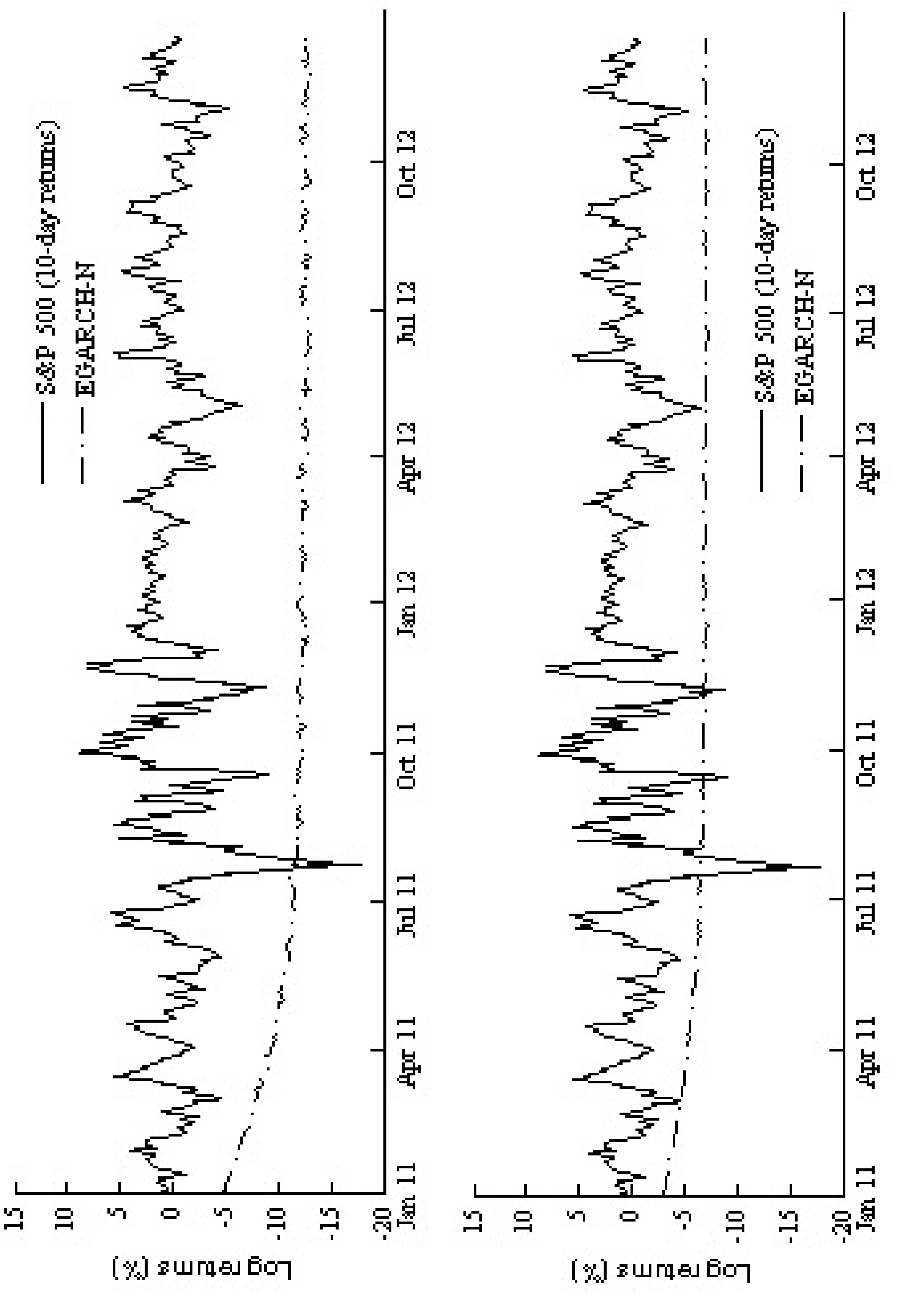
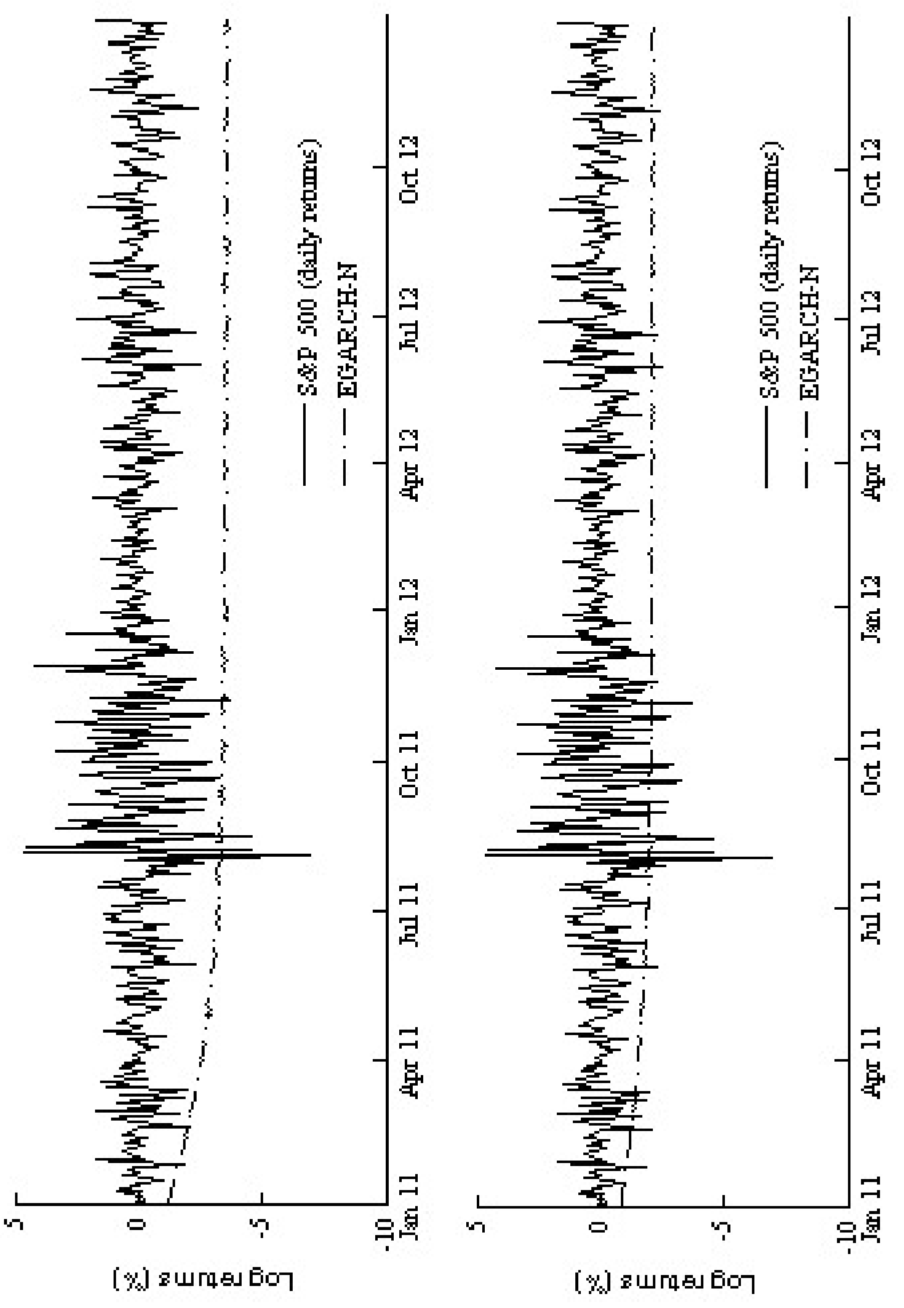


FIGURE A.4: 1-day VaR-forecast for the period after the Global Financial Crisis period: (t) $\alpha = 1\%$ (b) $\alpha = 5\%$.



Bibliography

[1] Cristina Amado and Timo Tersvirta. Conditional Correlation Models of Autoregressive Conditional Heteroskedasticity with Nonstationary GARCH Equations. Creates research papers, School of Economics and Management, University of Aarhus, May 2011.

[2] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

[3] Jeremy Berkowitz, Peter Christoffersen, and Denis Pelletier. Evaluating value-at-risk models with desk-level data. Working paper series, North Carolina State University, Department of Economics, October 2005.

[4] Romain P. Berry. An overview of value-at-risk: Iii – Monte Carlo Simulations Value at risk. Technical report, J.P. Morgan Investments Analytics and Consulting, 2013.

[5] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, April 1986.

[6] Chris Brooks. *Introductory econometrics for finance*. Cambridge University Press, Cambridge, Second edition, 2008.

[7] Cees Diks, Valentyn Panchenko, and Dick van Dijk. Likelihood-based scoring rules for comparing density forecasts in tails. *Journal of Econometrics*, 163(2):215 – 230, 2011.

[8] Robert Engle and Simone Manganelli. Caviar: Conditional autoregressive value at risk by regression quantiles. Econometric society world congress 2000 contributed papers, Econometric Society, August 2000.

[9] Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007, 1982.

[10] Robert F. Engle and Victor K. Ng. Measuring and testing the impact of news on volatility. *The Journal of Finance*, 48(5):1749–1778, 1993.

[11] Christopher C. Finger. Lessons in risk modeling: What works and whats missing. Research monthly, Riscmetrics Group, 2009.

[12] Richard H. Gerlach, Cathy W. S. Chen, and Nancy Y. C. Chan. Bayesian time-varying quantile forecasting for value-at-risk in financial markets. *Journal of Business & Economic Statistics*, 29(4):481–492, 2011.

[13] Lawrence R. Glosten, Ravi Jagannathan, and David E. Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. Staff Report 157, Federal Reserve Bank of Minneapolis, 1993.

[14] Winfried G. Hallerbach. Decomposing portfolio value-at-risk: A general analysis. Tinbergen institute discussion papers, Tinbergen Institute, May 1999.

[15] Philippe Jorion. *Value at Risk: The New Benchmark for Controlling Derivative Risk*. McGraw-Hill, New York, Third edition, 1997.

[16] Philippe Jorion. Fallacies about the effects of market risk management systems. *Journal of Risk*, 5:75–96, 2002.

[17] Roger Kaufman. *Long-term risk management*. PhD thesis, ETH Zurich, 2004.

[18] Michael McAleer and Bernardo da Veiga. Single-index and portfolio models for forecasting value-at-risk thresholds. *Journal of Forecasting*, 27(3):217–235, 2008.

[19] A. J. Mcneil, R. Frey, and P. Embrechts. *Quantitative risk management*. Princeton Series in Finance. Princeton University Press, Princeton, NJ, 2005.

[20] Daniel B Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2):347–70, March 1991.

[21] Birger Nilsson. Financial valuation and risk management. Lecture notes NEKN83/TEK180, Lund University, 2013.

[22] Basel Committee on Banking Supervision. Supervisory Framework for the Use of "Backtesting" in Conjunction With the Internal Models Approach to Market Risk Capital Requirements, 1996.

[23] Anthony S. Tay and Kenneth F. Wallis. *Density Forecasting: A Survey*, pages 45–68. Blackwell Publishing Ltd, 2007.

[24] Allan Timmermann. Density forecasting in economics and finance. *Journal of Forecasting*, 19(4):231–234, 2000.

[25] Bruce Walsh. Markov Chain Monte Carlo and Gibbs Sampling. Lecture Notes EEB 581, University of Arizona, 2004.

[26] Anders Wilhelmsson. Density forecasting with timevarying higher moments: A model confidence set approach. *Journal of Forecasting*, 32(1):19–31, 01 2013.

[27] Jean-Pierre Zigrand and Jon Danielsson. On time-scaling of risk and the square-root-of-time rule. Fmg discussion papers, Financial Markets Group, March 2003.