Risk Theory

Examining the methods for measuring Market Risk

By

ANDREAS THEODOULOU

Department of Mathematics
UNIVERSITY OF BRISTOL

BSc Mathematics with Statistics Project (20cp)

MAY 2018

Acknowledgement of Sources

For all ideas taken from other sources (books, articles, internet), the source of the ideas is mentioned in the main text and fully referenced at the end of the report. All material which is quoted essentially word-for-word from other sources is given in quotation marks and referenced. Pictures and diagrams copied from the internet or other sources are labelled with a reference to the web page,book, article etc.

Signed: Andreas Theodoulou Dated:01/05/2018

SIGNED: ANDREAS THEODOULOU....... .... DATE: 01/05/2018 ....

Table of Contents

Page

List of Tables iii

List of Figures iv

1 Introduction 1

1.1 Background . . . 1

1.2 Project Overview . . . 2

2 Theory 5

2.1 Financial Asset Returns . . . 5

2.1.1 Log-returns . . . 5

2.1.2 Portfolio Returns . . . 5

2.1.3 Volatility clustering . . . 6

2.1.4 Leptokurtic distributions . . . 6

2.1.5 Asymmetry and leverage effect . . . 7

2.1.6 Uncorrelated Daily returns . . . 7

2.2 Risk Measures . . . 7

2.2.1 Value at Risk . . . 7

2.2.2 Expected Shortfall . . . 8

2.2.3 Evaluation of VaR and ES . . . 9

2.2.4 VaR and ES Implementation . . . 10

2.3 Historical Simulation . . . 11

2.4 Parametric Method . . . 12

2.4.1 Volatility Forecasting . . . 14

2.4.2 Exponential Weighted Moving Average . . . 14

2.4.3 GARCH model . . . 15

2.4.4 GJR-GARCH . . . 16

2.5 Backtesting . . . 17

2.5.1 Backtesting VaR - Unconditional Coverage testing . . . 17

2.5.2 Backtesting ES - Unconditional test . . . 18

3 Data & Methodology **21**
3.1 Data . . . 21
3.2 Programming . . . 22
3.3 Methodology . . . 23
3.3.1 General Methodology and Hypotehses . . . 23
3.3.2 VaR and ES Methods and Backtesting . . . 25
3.4 Methods Ranking . . . 25

4 Results **27**
4.1 Value at Risk . . . 27
4.1.1 95% Value at Risk . . . 27
4.1.2 99% Value at Risk . . . 28
4.2 Expected Shortfall . . . 29
4.2.1 95% Expected Shortfall . . . 31
4.2.2 99% Expected Shortfall . . . 31
4.3 Discussion . . . 33

5 Conclusion **37**
5.1 Recommendations for further research . . . 38

A Derivation of parametric method equations **39**
A.1 Value at Risk . . . 39
A.1.1 Normal Distribution . . . 39
A.1.2 Student's t Distribution . . . 40
A.1.3 Expected Shortfall . . . 40
A.1.4 Normal Distribution . . . 40

B Matlab Code **43**
B.1 Matlab Code . . . 43

Bibliography **55**

List of Tables

TABLE **Page**

3.1 List of the Stock Indices, representing the portfolios to be tested . . . 22

3.2 Sample statistics of the returns of our "portfolios" 24

4.1 Ranking statistics of 95%-VaR estimates 28
4.2 Unconditional Coverage test LR results for 95%-VaR estimates during the Crisis period 29
4.3 Unconditional Coverage test LR results for 95%-VaR estimates during the Post-Crisis period 29
4.4 Ranking statistics of 99%-VaR estimates 30
4.5 Unconditional Coverage test LR results for 99%-VaR estimates during the Crisis period 30
4.6 Unconditional Coverage test LR results for 99%-VaR estimates during the Post-Crisis period 30
4.7 Ranking statistics of 95%-ES estimates 31
4.8 Unconditional test, Z-statistic for 95%-ES estimates during the Crisis period 32
4.9 Unconditional test Z-statistic results for 95%-ES estimates during the Post-Crisis period 32
4.10 Ranking statistics of 99%-ES estimates 32
4.11 Unconditional test Z-statistic for 99%-ES estimates during the Crisis period 33
4.12 Unconditional test Z-statistic results for 99%-ES estimates during the Post-Crisis period 34

LIST OF FIGURES

FIGURE **Page**

2.1 Autocorrelation of squared daily S&P 500 returns January 1, 2010-December 31, 2010 [4] 6
2.2 Histogram of daily S&P 500 returns and the normal distribution January 1, 2010-December 31, 2010 [4] 7
2.3 Autocorrelation of daily S&P 500 returns January 1, 2010-December 31, 2010 [4] . . 8
2.4 Value at Risk from the normal distribution return probability distribution [4] 9
2.5 Expected Shortfall in terms of the Probability density function of returns [12] 10

3.1 S&P500 graph of daily prices. Source: Google 24

CHAPTER 1

INTRODUCTION

1.1 Background

As pointed out by Dowd, "Everything changes and changes can be good or bad for those affected by them. Change therefore leads to risk, the prospect of gain or loss." [7, p.3] Consequently, this holds true for any financial institution and in order to survive in the long-term they need to deal with them. They must assess the risks they currently bear, decide on the risks they want to bear, and change their exposures as the landscape changes so that they bear the risks they want to. The risks a firm faces can be generally divided into the following: Business risks, Market risks, credit risks, liquidity risks, operational risks, legal risks.[7, p.3-4] This project will focus on Market risk[1] and the measures currently used in the industry to interpret it.

The Value at risk (VaR) measure is probably the most important measure for quantifying financial risk[2] in the past 20 years. In the meantime, we have seen the rise in popularity for Expected Shortfall (ES). Below, I aim to give a brief overview to the problem of measuring market risk, its history and its importance.

Work on internal models by a number of major financial institutions had already started in the late 1970s to help understand the bank's exposure across the entire trading portfolio. In particular, the first VaR report originated when the chairman of J.P. Morgan asked his staff to give him, after the close of trading each day, a one-page report indicating risk and potential losses over the next 24 hours across the bank's entire trading portfolio.[21, p.307] This problem of course had by no means an easy solution as often a financial institution's portfolio depends on hundreds, or even thousands, of market variables . In about 1990, his subordinates came up

[1] Market Risk is the risk of losses owing to movements in the level or volatility of market prices

[2] Financial risk is the risk which relate to possible losses owing to financial market activities such as loss occurred as a result of interest-rate movements or defaults of financial obligations

with the following answer: "We are X percent certain that we will not lose more than V dollars in time T", where the value V followed from the Value at Risk measure that was invented. The measure was groundbreaking at the time as it essentially aggregated all the risks of a portfolio into a single measure. [1, p183]

In fact, shortly after the fall of the world famous Barings bank at the time, the Basel committee in 1998 gave the option to financial organisations to report their capital requirements using the Value at Risk measure, which since then has been the popular approach for many financial organisations. [1, p.266] [13] Even though, VaR has been heavily studied, and several models were used in the financial world in the meantime, almost all of them seems that failed to predict the severe financial crisis of the 2007. This seems to be one of the reasons the Basel III accord to be introduced in 2019, has decided to shift its requirements from the Value at Risk measure to the Expected Shortfall.[9] [5] Therefore, through this project I aim to examine the best methods to estimate the VaR and ES measures. We will also evaluate the future shift from VaR to the ES measure as the solution to capturing market risk and discuss whether VaR could still remain a useful measure in the industry.

1.2 Project Overview

We will begin with the Chapter 2 of the project by introducing the theory around the material that is going to applied later on through the project. We start by introducing the main characteristics of the financial asset returns and defining our portfolio returns. We then introduce the two risk measures to be studied, the Value at Risk and the Expected shortfall, mathematically and qualitatively. From there on we discuss the challenge of estimating them and introduce the methods to be studied, namely the Historical Method, the Parametric method, and the Volatility forecasting methods used to aid the parametric method. Lastly, we review one Backtesting method for each of the measures, namely the Unconditional Coverage test and the Unconditional test for VaR and ES respectively.

In Chapter 3, we introduce the data to be used in applying the methods, explain the specifics our methodology and our reasoning behind it. We first introduce the 10 portfolios which we will try to forecast. Then we explain the programming used to aid the application of the methods in our VaR and ES forecasts. Moreover, we explain our general methodology around these simulations as well as specifically how each method is going to be applied, express the hypotheses we make and what we expect to review from these simulations. Lastly, we introduce the method that will be used for ranking the VaR and ES methods.

In Chapter 4, we introduce and discuss our results. These consist of the backtesting test statistics of the 95% and 99% VaR and ES estimates and our ranking statistics used to evaluate our models. We perform forecasts on two periods, one during the financial crisis of 2007-2008 and the other from 2013 onwards, which we refer to as the Post Crisis period. We then discuss our results for each situation. Concluding, we give an overall view on our results and conclude on the best model given these results.

CHAPTER 2

THEORY

2.1 Financial Asset Returns

Market risk is caused by movements in financial asset prices or equivalently asset returns. Therefore, we begin by defining returns and then give an overview of the characteristics of typical financial asset returns. These will provide part of the intuition based on which we will form our models.

2.1.1 Log-returns

In this project we will follow log-returns, instead of simple returns, as this follows the usual method in the literature. The main reasons for this is that they preserve better the lower bound of zero on the prices and allow to easily calculate the compound return at the K-day horizon by simply summing the daily returns.[4] From now on we will refer to log returns as just returns.

$$R_{t+1} = ln(S_{t+1}) - ln(S_t)$$

where R_{t+1} is the daily log return at time t+1, S_t is the final price at which a security is traded at time t, and t refers to a specific market day

2.1.2 Portfolio Returns

Consider a portfolio of n assets. Then the value of a portfolio at time t, $V_{PF,t}$ is the weighted average of the asset prices, Rt, using the current holdings of each asset, N_i as weights is $V_{PF,t} = \sum_{i=1}^{N} N_i S_{i,t}$. The (log) return of the portfolio between day t+1 and day t is then defined as

$$R_{PF,t+1} = ln(V_{PF,t+1}) - ln(V_{PF,t}) \tag{2.1}$$

Note that we assume that the portfolio value on each day includes the cash from accrued dividends and other asset distributions.

2.1.3 Volatility clustering

As first noted by Mandelbrot, volatility clustering is that "large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes." The effect can be seen more clearly in the Figure 2.1. This results in much stronger autocorrelations of the squared returns than the returns, for shorter time lags. [21] As squared returns are used in the calculations of volatility[1], the changing dependence of volatility over time can be captured by statistical models, which will be introduced later.

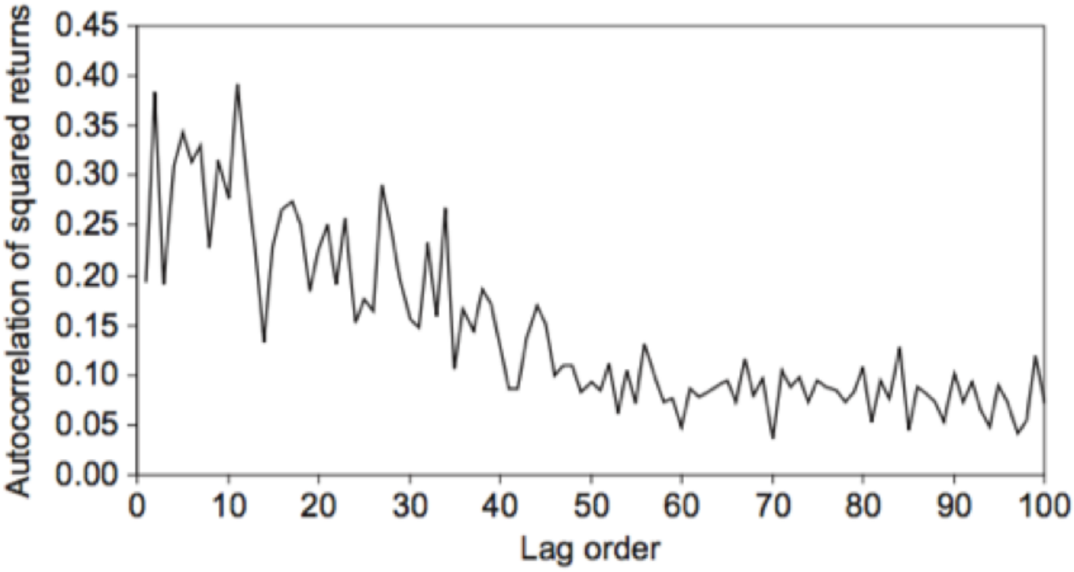

Figure 2.1: Autocorrelation of squared daily S&P 500 returns January 1, 2010-December 31, 2010 [4]

2.1.4 Leptokurtic distributions

The distribution of daily returns tends to follow distributions that are more leptokurtic than the standard Gaussian distribution. This means that the distributions tend to be more peaked around zero and have more small positive and fewer small negative returns, also known as fatter tails. Fatter tails mean a higher probability of large losses (and gains) than the normal distribution would suggest. These characteristics can be shown at Figure 2.2 where a histogram of the daily S&P 500 return data with the normal distribution is imposed. [4]

[1]Please note that the volatility term is simply the term used in finance for standard deviation.

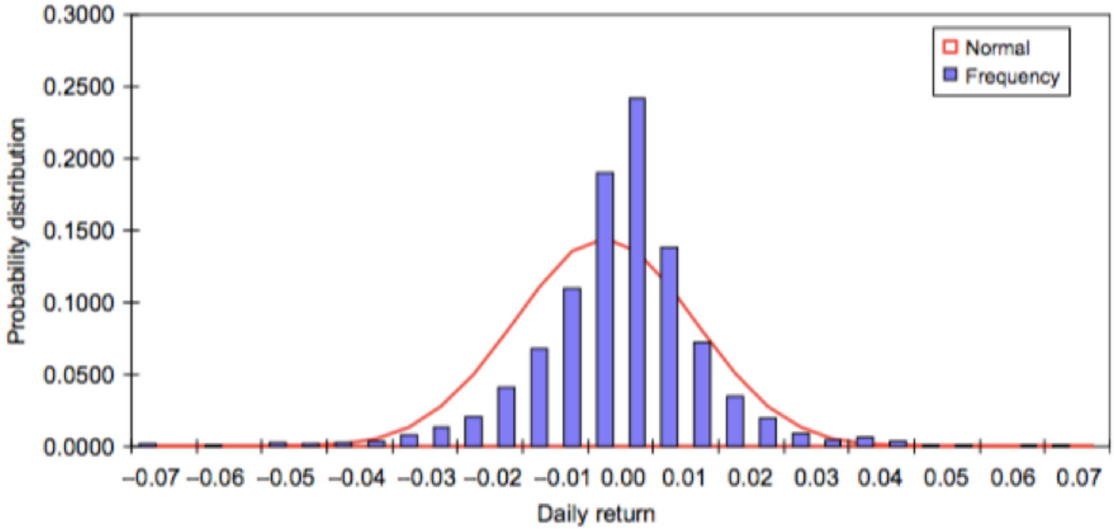

Figure 2.2: Histogram of daily S&P 500 returns and the normal distribution January 1, 2010-December 31, 2010 [4]

2.1.5 Asymmetry and leverage effect

Asymmetry of magnitudes in upward and downward movements of asset returns is typical in equity markets. In particular, the volatility response to a large positive return is considerably smaller than that to a negative return of the same magnitude. This asymmetry is sometimes referred to as a leverage effect.[21] The reasoning behind this characteristic is based on economic principles, which are beyond the scope of this project. We will model the leverage effect in some of our models.

2.1.6 Uncorrelated Daily returns

The autocorrelation of daily returns is generally around 0. In other words, returns are almost impossible to predict using past data. Also, for short time daily conditional mean of returns is very close to 0. Therefore, focusing just on volatility modelling of the returns distribution for this project and assuming conditional mean is constant, seems to be a reasonable simplification. Figure 2.3 shows the correlation of daily S&P 500 returns with returns lagged from 1 to 100 days. [4]

2.2 Risk Measures

2.2.1 Value at Risk

VaR summarizes the worst loss over a target horizon that will not be exceeded with a given level of confidence. More formally, VaR describes the quantile of the projected distribution of gains

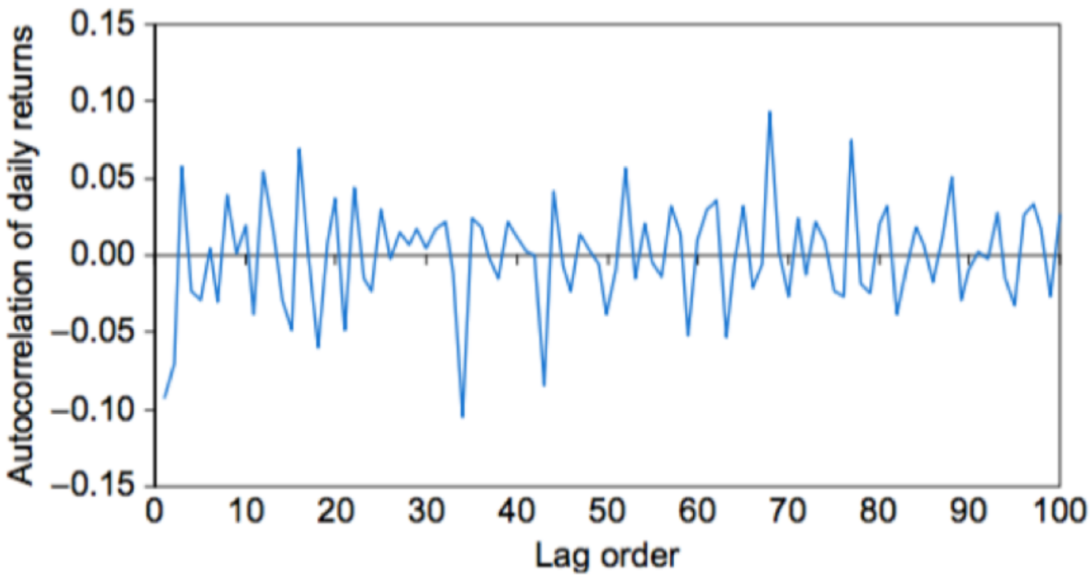

Figure 2.3: Autocorrelation of daily S&P 500 returns January 1, 2010-December 31, 2010 [4]

and losses over the target horizon.[11] Mathematically, the q-%VaR for the 1-day ahead return, satisfies the following equation:

$$P(R_{PF,t+1} < -VaR^{q}_{t+1}) = q \tag{2.2}$$

Note that VaR depends on the following 3 parameters, which from now on will be indicated as follows:

- q, quantile: The probability to lose more than VaR^{q}_{t+1} in the specified period.
- t, time period: t is assumed to be the present day, so t+1 indicates the one-period VaR
- $R_{PF,t+1}$: The return distribution. This is to be calculated using information available in period t. The specific methods will follow in the later chapters.

Figure 2.4 illustrates the $VaR^{0.01}_{t+1}$ from a normal distribution of returns. Multiplying the $VaR^{0.01}_{t+1}$ by the total value of the portfolio will give you the worst loss over a day with a probability of 99%.

2.2.2 Expected Shortfall

The Expected shortfall, also know as conditional Value at Risk (CVaR), is a risk measure that accounts for the magnitude of large losses as well as their probability of occurring.[4, p.33] More specifically it tells us the expected value of tomorrow loss, conditional on it being worse than the VaR. The Expected Shortfall measure aggregates this information into a single number by

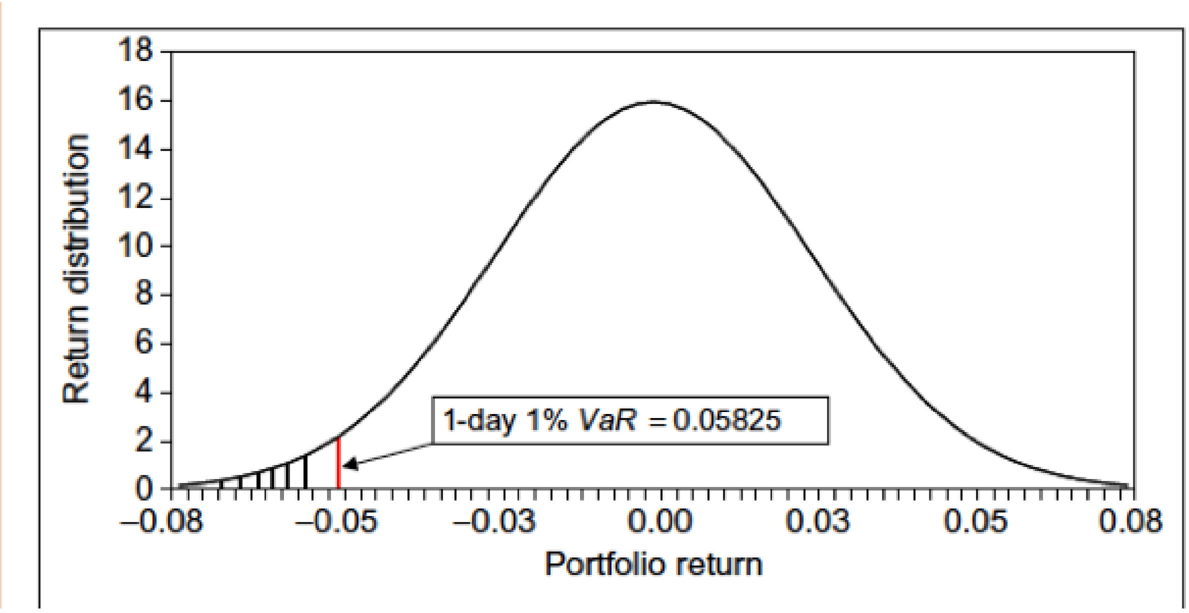

Figure 2.4: Value at Risk from the normal distribution return probability distribution [4]

computing the average of the tail outcomes weighted by their probabilities. Mathematically ES is defined as follows:

$$ES^{q}_{t+1} = -E_t[R_{PF,t+1} \mid R_{PF,t+1} < -VaR^{q}_{t+1}] \tag{2.3}$$

where the negative signs in front of the expectation and the VaR are needed because the ES and the VaR are defined as positive numbers.

In the figure 2.5 below, the shaded area represents the losses that exceed the VaR. Cumulative probability in the shaded area is equal to 1-q. ES or CVaR, can be represented by the average of the shaded area.

2.2.3 Evaluation of VaR and ES

Value at risk captures an important aspect of risk, namely how bad things can get with a certain probability, q.[4] Furthermore, it is easily communicated and easily understood.VaR does, however, have some drawbacks. Most important, extreme losses are ignored. The VaR number only tells us that q% of the time we will get a return below the reported VaR number, but it says nothing about what will happen in those q% worst cases. A second major criticism of VaR is that the measure is not subadditive. Subadditivity holds that adding the risk of Asset A and the risk of Asset B will not result in an overall risk that is greater than the sum of the two risks together. This can undermine the principle of diversification, often important when managing assets. [12] Lastly, the Value at Risk measure has developed a wide variety of applications through the years,

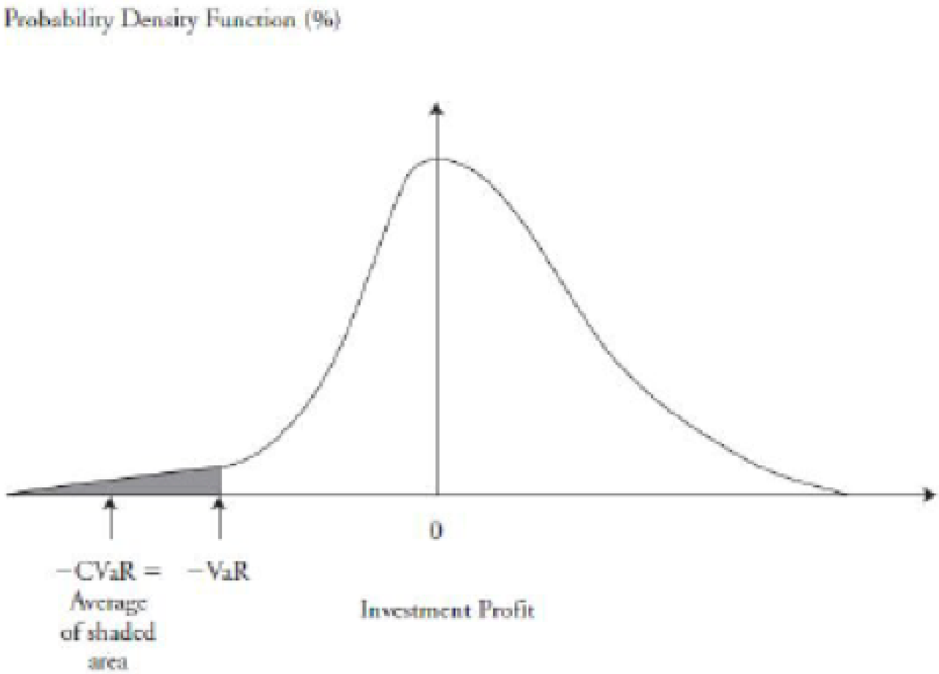

Figure 2.5: Expected Shortfall in terms of the Probability density function of returns [12]

including portfolio optimization. For the purpose of this project, we will only focus on its capability of capturing the market risk.

Now Expected shortfall, satisfies both of the above VaR disadvantages, however it is not perfect either. ES often requires a large number of observations to generate a reliable estimate, and it is more sensitive to estimation errors than VaR [12]. Moreover, the reliability of ES depends substantially on the accuracy of the tail model used. It should also be noted that ES is based on an average loss beyond VaR, and thus it is not a measure of the most extreme potential loss.

Overall, it is important to remember that both measures are only an estimate. This means that different samples will lead to different results. Therefore, the processes of calculating the risk measures are as important as the number itself. [11] We will now review the processes of deriving these risk measures, and more specifically the returns distribution needed to calculate them.

2.2.4 VaR and ES Implementation

While VaR and ES are easy and intuitive concepts, their measurement are a challenging statistical problem. The existing models for calculating VaR employ different methodologies, however all follow a common general structure; 1) Mark-to-market the portfolio, 2) Estimate the distribution of portfolio returns, 3) Compute the VaR of the portfolio. [14] The main challenge is due to the difficulty of modelling the evolution of a portfolio containing hundreds of assets[19] [10].

The existing models can be classified into 3 broad categories [18]

- Parametric models (e.g. RiskMetrics and GARCH)
- Nonparametric models (e.g. Historical Simulation and the Hybrid model)
- Semiparametric models (e.g. Extreme Value Theory and Conditional Autoregressive Value at Risk)

In this project we will focus on parametric and nonparametric methods as one can extend to the semi-parametric once the former are understood. We refer the interested reader to some of the most popular semiparametric methods to [14]. Moreover, the number and types of approaches to VaR estimation within parametric and nonparametric methods is growing exponentially [14], therefore we will try to focus on the most popular ones. These will be the Historical Simulation and the Risk Metrics, Garch and GJR models for the parametric method. Also, note that the approaches considered only cover the univariate approaches, which consider portfolio-level risk, whereas multivariate approaches, which consider individual asset level risk and aggregate them into a portfolio-level risk, are out of the scope of this paper. The reader can refer to [4] Chapters 7 to 9 for the use of multivariate approaches. Lastly, we will only focus on estimating the next day VaR. The next k-day VaR estimate can follow from there, and the methods are introduced in [4] Chapter 8. In the following parts, estimated quantities will be denoted by a hat (^).

2.3 Historical Simulation

Historical Simulation (HS) is one of the most used methods for VaR and ES estimation. This approach drastically simplifies the procedure for computing the VaR and ES, since it doesn't make any distributional assumption about portfolio returns. [14]

We will define the historical simulation following Christoffersen at [4, p,22]. Let today be day t, consider a portfolio of n assets and consider the value of this portfolio and its corresponding returns. Using today's portfolio holdings, one can consider its historical asset prices and compute the history of the past m daily "pseudo" portfolio returns, call it $\{R_{PF,t+1-\tau}\}_{m=1}^{\tau}$,that would be true if today's portfolio allocation had been the same in the past. The HS technique simply assumes that the distribution of tomorrow's portfolio returns, $\{R_{PF,t+1-\tau}\}_{m=1}^{\tau}$, is well approximated by the empirical distribution of the past m observations, or simply the histogram of, $\{R_{PF,t+1-\tau}\}_{m=1}^{\tau}$. The q%-VaR is then calculated as $100q$th percentile of the sequence of past portfolio returns. Mathematically, we write

$$\widehat{VaR_{t+1}^{q}} = -Percentile(\{R_{PF,t+1-\tau}\}_{\tau=1}^{m}, 100q)$$

In practice, we simply sort the returns R_{PF} in ascending order and choose the VaR to be the number at the qth percentile of this list. As the VaR typically falls in between two observations, linear interpolation can be used to calculate the exact number.

Similarly, the estimate of ES using HS can be calculated by taking the average of the returns that were less that VaR in the specified time window.

$$\widehat{ES^q_{t+1}} = \frac{1}{\tau}\sum_{i=1}^{\tau} r_i$$

Where τ is the number of times the returns were below VaR and r_i the value of these returns.

The reasons for the popularity of the HS method are generally its convenience to implement and it's model free nature. [4] [6] The former implies that the method makes no distributional assumption about the returns, and the benefit of this is that model building approaches can be misleading if the model is poor. However, from the model-free advantage follows one of its biggest drawbacks, which is that it assumes that the distribution of portfolio returns doesn't change within the window. One of the implications of this is that its performance may be affected by the choice of sample length, m, mainly due to the volatility clustering periods which are difficult to identify. [14] Also, this assumption means the HS is likely to result in VaR estimates with high variance due to the discreteness of extreme returns. To see why, assume a rolling window of 180 days and that today's return is a large negative number. It is easy to predict that the VaR estimate will jump upward, because of today's observation. [14]

2.4 Parametric Method

The parametric approach is put simply a model building approach. It aims to parameterize and engulf in a single model, the characteristics of financial asset returns by using appropriate statistical distributions and techniques. Once this is done, the calculations of VaR and ES can follow from Equations 2.2 and 2.3, given the distribution of the model has an analytic solution. As the complexity of a parametric model can always be furthered and in this particular topic could cover a research paper in itself, for the scope of this project we will keep to the simpler approaches and try to refer the reader where appropriate.

We start by considering the a generic model for individual asset returns, R_t, following [4, p11-12]. Based on the asset return characteristics reviewed in Chapter 2.1, the following generic form is employed

$$R_{t+1} = \mu_{t+1} + \sigma_{t+1} z_{t+1}, \qquad \text{with } z_{t+1} \sim \text{i.i.d } D(0,1)$$

The random variable z_{t+1} is an innovation term, which we assume is identically and independently distributed (i.i.d.) according to the distribution D(0,1), which has mean zero and variance one. The conditional mean of the return, $E_t[R_{t+1}]$, is thus μ_{t+1}, and the conditional variance, $E_t[R_{t+1} - \mu_{t+1}]^2$, is σ^2_{t+1}. As mentioned in Subsection 2.1.6, for simplicity we will assume conditional mean of returns to be 0. For longer horizons, the risk manager may want to estimate a model for the conditional mean. Considering Equation 2.1, allowing for a dynamic variance model, to be introduced later, and applying the assumption on conditional mean, we can now define a generic model for the portfolio returns as follows,

$$R_{PF,t} = \sigma_{PF,t} z_t \qquad \text{with } z_{t+1} \sim \text{i.i.d } D(0,1) \tag{2.4}$$

We will first consider the appropriate models for the innovation term and then review the volatility models used to estimate the variance of the portfolio. The innovation term models will be the Normal distribution and the Student's t-distribution.

Although the Normal distribution does not satisfy the financial data characteristics as explained earlier, it remains a popular method because of its simplicity and the quasi-maximum likelihood GARCH result. [14] For more information on the result the reader can refer to [14, p.17] and [4, p.75]. The use of the Student's t-distribution is one of the methods used capture the non-normality of the financial data. [21, p.144] However although it captures the fatter tails, one should note that it still has limitations as it is less concentrated around the mean. For more methods used to capture the non-normality of the financial asset returns the reader can refer to [4] Chapter 6.

Under a Normal distribution of the innovation terms, $N(0,1)$, the VaR and ES estimations can be calculated through the following equations.

$$\widehat{VaR^{q}_{t+1}} = -\widehat{\sigma_{t+1}}\Phi^{-1}(q) \tag{2.5}$$

$$\widehat{ES^{q}_{t+1}} = -\widehat{\sigma_{t+1}}\frac{\phi(\Phi^{-1}(q)}{1-q} \tag{2.6}$$

where ϕ and Φ are the probability density function and cumulative density functions of $N(0,1)$ respectively.

If the innovation terms belong to a Student's t-distribution, with ν degrees of freedom, then the VaR and ES are estimated as follows,

$$\widehat{VaR^{q}_{t+1}} = -\widehat{\sigma_{t+1}}t_{\nu}^{-1}(q)\sqrt{\frac{\nu-2}{\nu}} \tag{2.7}$$

$$\widehat{ES^{q}_{t+1}} = -\widehat{\sigma_{t+1}}\frac{g_{\nu}t_{\nu}^{-1}(q)}{1-q}\frac{\nu+(t_{\nu}^{-1}(q))^2}{\nu-1}\sqrt{\frac{\nu-2}{\nu}} \tag{2.8}$$

where t_{ν} and g_{ν} is the cumulative density function and probability density function of the Student's t distribution respectively.

Please note that the distribution is scaled by the $\frac{\nu-2}{\nu}$ factor, or simply divided by its variance, in the equations above in order to satisfy the unit variance assumption of the innovation's distribution. We introduce a full derivation of the above Equations 2.5, 2.6, 2.7 and 2.8 in the Appendix A. Moreover, one of the methods of estimating the degrees of freedom for the Student's t-distribution with variance 1 and mean 0, is using the maximum likelihood method. The method chooses the number of degrees of freedom that maximizes the log-likelihood function of the our t-distribution over our chosen time frame. [4, p.129-130]

As part of the parametric method, we will now review volatility forecasting models.

2.4.1 Volatility Forecasting

The focus of Volatility Forecasting is to introduce the models to be used for forecasting tomorrow 's variance. Again, these will aim to capture some of the characteristics of the financial asset returns and more specifically in this case the squared returns. For example, as mentioned in Subsection 2.1.3, the varying autocorrelation of squared returns over time is a characteristic to be considered. The models to be reviewed will be the Risk Metrics, the GARCH model and the GJR model, through which we will follow [4] at Chapter 4. Please note that the models defined below can be used for calculating both the variance of an individual asset and the portfolio as a whole. For now, we will define their individual asset form. Their portfolio wide form follows by simply replacing the individual asset variables with their equivalent portfolio wide ones. Moreover, these models can be used to approach the volatility forecasting problem in a multivariate approach. This essentially means that each asset can be modelled individually and are then aggregated into a portfolio wide forecast, by considering their interdependence with methods like correlations and copulas modelling. As for the purpose of risk measurement at a portfolio level the univariate approaches are considered to be sufficient [4, p.68], we will only consider the univariate ones and refer the interested reader to [4] Chapters 7 to 9 for the multivariate approaches.

2.4.2 Exponential Weighted Moving Average

The Exponential Weighted Moving Average (EWMA) model, also known as RiskMetrics model, was first introduced by J.P Morgan's for market risk management in 1996. It is a simpler approach to volatility foracasting, where the weights on past squared returns decline exponentially as time moves backwards. More specifically,

$$\sigma^2_{t+1} = (1-\lambda)\sum_{\tau=1}^{\infty} \lambda^{\tau-1} R^2_{t+1-\tau} \qquad \text{for } 0 < \lambda < 1$$

In this project we are going to use the following version of the equation, which can be found by separating the first term of the sum in the above equation,

$$\sigma^2_{t+1} = \lambda\sigma^2_t + (1-\lambda)R^2_t \tag{2.9}$$

RiskMetrics advantages consist of the following. First, it is broadly consistent in capturing the volatility clustering. Second, the model is quite simple to implement as it contains only one unknown parameters, namely λ. Moreover, for a large number of assets, the parameter estimates have been found to be quite similar across assets, and it is therefore suggested to simply set $\lambda = 0.94$ for every asset for daily variance forecasting. Lastly, little data is needed to be stored as $\lambda^{\tau-1}$ decays quickly to 0. On the other hand though, it does have drawbacks. For example, it provides unrealistic longer-horizon forecasts and does not allow for the leverage effect characteristic of returns. Its shortcomings led us to consider more sophisticated but still simple models from the GARCH family.

2.4.3 GARCH model

The family of Generalized autoregressive conditional heteroskedasticity (GARCH) models were introduced Bollerslev (1986), as an extension to the nobel awarded ARCH models suggested by Engle. The model has been successfully applied to financial data since then. [14]. The GARCH(p,q) model can defined as follows,

$$\sigma_{t+1}^2 = \omega + \sum_{i=1}^{p} \alpha_i R_{t+1-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t+1-j}^2$$

where ω, α_i and β_i are weights to be estimated.

The difference between the generalized ARCH and the ARCH model lies in the addition of the weighted summation over squared volatility term. This essentially means that the GARCH model also allows squared volatility to depend on previous squared volatility, σ_t^2, on top of the squared return. [2, p.145] For short-term volatility forecasting,the simplest model of the family, namely GARCH(1,1), is often found to be sufficient.[4, p.78] Therefore for this project we consider reasonable to only use the GARCH(1,1) model and from now on we will simply refer to it as the GARCH model. Mathematically, it is defined as follows

$$\sigma_{t+1}^2 = \omega + \alpha R_t^2 + \beta \sigma_t^2 \qquad \text{with } \alpha + \beta < 1 \tag{2.10}$$

Through this model periods of high volatility tend to be persistent. This is because $|\sigma_{t+1}|$ has a chance of being large if either $|\sigma_{t+1}|$ is large or σ_t is large. The same effect can be achieved by ARCH models of higher order but lower-order GARCH models tend to capture it much more persistently.[2, p.145] Moreover, notice here that the RiskMetrics model is a special version of the GARCH models, if we set $\alpha = 1 - \lambda$, $\beta = \lambda$ so that $\alpha + \beta = 1$ and $\omega = 0$. However the models still have a major difference. That is the RiskMetrics model does not account for the fact that long-run average variance tends to be relatively stable over time. [4, p.71] This can be more explicitly seen be defining the (long-run average) variance, σ to be,

$$\begin{aligned}\sigma^2 &= E[\sigma_{t+1}^2] = \omega + \alpha E[R_t^2] + \beta E[\sigma_t^2] \\ &= \omega + \alpha\sigma^2 + \beta\sigma^2 \quad \text{so that,} \\ \sigma^2 &= \omega/(1 - \alpha - \beta)\end{aligned}$$

One can now see that if $\alpha + \beta = 1$, as it is true in the RiskMetrics model, then the long-run variance is not well defined in that model. Moreover, the GARCH model does take into account the long-run variance, as it implicitly relies on it. [4, p.71] By solving for ω in the above equation and substituting into the the GARCH model then,

$$\sigma_{t+1}^2 = (1-\alpha-\beta)\sigma^2 + \alpha R_t^2 + \beta\sigma^2 = \sigma^2 + \alpha(R_t^2 - \sigma^2) + \beta(\sigma_t^2 - \sigma^2)$$

Therefore, one can see that under the GARCH model also depends on the long-run average variance, on top of the today's squared return and today's variance. Although this difference might not be as important for the one-day horizon forecast, it does grows in importance for longer-horizon forecasting. [4, p.71]

Lastly, the unknown coefficients of the model can be estimated through the maximum likelihood estimation method. We are now going to review our last volatility forecasting model which is an extension of the GARCH model.

2.4.4 GJR-GARCH

As we have seen in Subsection 2.1, the leverage effect is a characteristic we have not yet tried to capture through our parametric models. The GJR-GARCH model introduced by Glosten et al. in 1993 as an extension of the GARCH model, aims to do this. One can notice that with a θ larger than 0, the model defined as follows, will capture the leverage effect, [4, p.77]

$$\sigma_{t+1}^2 = \omega + \alpha R_t^2 + \alpha\theta \mathbb{1}_t R_t^2 + \beta\sigma_t^2 \tag{2.11}$$

$$\text{for } \mathbb{1}_t = \begin{cases} 1, & \text{if } R_t < 0. \\ 0, & \text{otherwise.} \end{cases}$$

and where, ω, α, β and θ are weights to be estimated.

Now to compute the variance forecast in each of the GARCH models, we first need to estimate the unknown parameters of our model. For this we will choose the Maximum likelihood Estimation method, which finds the parameters, α, β, ω and θ, in each case accordingly, by choosing their value that maximizes the log-likelihood function of the portfolio returns, as defined in Equation 2.4, over our chosen time frame. [4, p.73]

Overall, the family of GARCH models are widely used because of their ability to capture important characteristics of the returns data and are flexible enough to accommodate specific characteristics of individual assets. Various extensions to the GARCH model introduced exist, which aim to capture even more characteristics of the returns data, as indicated by the GJR-GARCH model. For further extensions the reader can review Chapter 4 Section 5 of [4]. On the downside, the GARCH models requires a non-linear parameter estimation. Moreover, there are different methods used to individually evaluate each volatility forecasting models. As one of the scopes of the project is to evaluate the models as a whole for the measurement of VaR and ES, we will not review these and the interested reader can refer to [4] Chapter 4 Section 6.

2.5 Backtesting

Although some of the methods introduced for VaR and ES implementation may make sense in theory, a considerate risk manager would want to evaluate them in practice. Therefore, we will use backtesting on the methods introduced for this reason but also as a means of evaluating the most suitable ones. For the Value at Risk measure we will use the Unconditional Coverage testing,or Proportion of failuers test, as introduced in [4, p.301-304], whereas for the Expected Shortfall measure we will use the "Test 2: testing ES after VaR" method as recently introduced in [3].

2.5.1 Backtesting VaR - Unconditional Coverage testing

Consider the the time series of the past portfolio returns, $\{R_{PF,t+1-\tau}\}_{\tau=1}^{T}$, and of the past VaR foreacsts, $\{\widehat{VaR^{q}_{t+1-\tau}}\}_{\tau=1}^{T}$. Then we can construct an indicator function, name it the "hit sequence", which returns 1 if the loss on the t+1 day was greater than the VaR forecast, or simply if VaR was violated, and 0 if not.

$$I_t + 1 = \begin{cases} 1, & \text{if } R_{PF,t+1} < -VaR^{q}_{t+1}. \\ 0, & \text{otherwise.} \end{cases}$$

One can see that the hit sequence, $\{I_{t+1-\tau}\}_{\tau=1}^{T}$, allows us to see when the past VaR violations occurred. Moreover, recall that given the perfect model, the VaR measure promises that the actual return will only be worse that the VaR forecast q100% of the time. In other words, the hit sequence of violations should show 1 with a probability of q every next day, should be completely unpredictable, and therefore distributed independently over time. It essentially provides us with the general framework under which we can test our models, or an initial null hypothesis, which can be put formally as follows,

$$H_0 : I_{t+1} \sim \text{i.i.d. Bernoulli(p)}$$

This is idea is usually the starting for most VaR backtesting methods. A more intuitive point of interest at this point could be whether q is close to the fraction of violations obtained by the VaR model, call it π. This is what the Unconditional Coverage test aims to examine, which can be described as follows,

$$\begin{aligned} H_0 &: \pi = q \\ H_1 &: \pi \neq q \end{aligned} \tag{2.12}$$

Let $L(\cdot)$ denote the likelihood function for an i.i.d. Bernoulli random variable, T_0 and T_1 be the number of 0s and 1s in the sample of the hit sequence, and T the total days of the sequence. Also, note that we can estimate π from $\hat{\pi} = T_1/T$. Then for the sample hit sequence $L(\hat{\pi}) = (1-\hat{\pi})^{T_0}\hat{\pi}^{T_1}$, and $L(q) = (1-q)^{T_0}q^{T_1}$. We can now check the unconditional coverage hypothesis using a likelihood ratio test,

$$LR_{UC} = -2ln[L(q)/L(\hat{\pi})] \sim \chi_1^2 \tag{2.13}$$

Using the Likelihood ratio test, if our likelihood ratio value LR_{UC} is larger than the critical value of the significance level, from the χ_1^2 distribution, then we reject our null hypothesis, otherwise we do not reject the null hypothesis. Alternatively through the p-value approach, we reject the null hypothesis if the p-value is less than the significance level value. The p-value is calclulated by $p - value = 1 - F_{\chi_1^2}(LR_{UC})$.

2.5.2 Backtesting ES - Unconditional test

Expected shortfall has traditionally been more limited in terms of the available backtesting methods. This is mainly because of the following challenge. A VaR backtesting has two possibilities for every day, either there is a VaR violation or not. On the other side, for ES the possibilities every day are infinite, it may exceed VaR by 1%, 50% or 500%. Moreover, if the VaR was exceeded by 30% on average and it was estimated to have been 20%, this error in estimation will be much more significant for a thin-tailed distribution than for a fatter-tailed one. This means that information about the distribution of returns for each day is needed. However, this problem has been overcome recently when [3] introduced the "unconditional" test or "Test 2" where you can get approximate test results without providing the distribution information.[16] This was a very important result as it made the backtesting of ES much simpler. This paper was also very important as a discovery in 2011 that ES is not elicitable, raised a debate over the risk management community at the time, on whether ES can be actually backtested. This paper comes to prove that it can be actually backtested.

Recall that the ES risk measure promises that whenever we violate the VaR , the expected value of the violation will be equal to ES_{t+1}^q. Also, let F_{t+1} denote the real (unknowable) return distribution and $P_t + 1$ the model predictive distribution, used to estimate VaR_{t+1}^q. Moreover, let $ES_{t+1,q}^P$ denote the value of ES when returns are distributed to P_{t+1}, with q and t as defined earlier, and $ES_{t+1,q}^F$ the value of ES when returns are distributed to F_{t+1}. Then given the above, we can formulate the following hypothesis,

$$\begin{aligned} &H_0 : P_{t+1}^{[q]} = F_{t+1}^{[q]}, \text{for all t} \\ &H_1 : ES_{t+1,q}^P \geq ES_{t+1,q}^F \end{aligned} \tag{2.14}$$

Now assuming H_0, and with I_t as defined earlier, we can define ES as follows,

$$ES_{t+1}^q = -E\left[\frac{R_{PF,t+1}I_{t+1}}{q}\right]$$

By expanding the expectation and rearranging the equation we can consider the following test statistic,

$$Z_2(R_{PF,t+1}) = \frac{1}{qT}\sum_{t=0}^{T}\frac{R_{PF,t+1}I_t}{ES_{t+1}^q} + 1 \tag{2.15}$$

In other words our H_0 and H_1 could have been simply $E[Z_2] = 0$ and $E[Z_2] < 0$ respectively. Moreover, the test statistic allows us to formulate our framework for testing our hypothesis through the p-value approach as follows,

$$\begin{aligned}
&\text{simulate independent} \quad && R^i_{PF,t+1} \sim P_t && \text{for all t and for all } i = 1,\dots,M \\
&\text{compute} && Z^i = Z\left(R^i_{PF,t+1}\right) && \\
&\text{estimate} && p = \sum_{i=1}^{M}\left(Z^i < Z\left(r^i_{PF,t+1}\right)\right)/M &&
\end{aligned} \tag{2.16}$$

where M is a suitably large number of scenarios. Given a preassigned significance level ϕ, the test is rejected if p < ϕ.

CHAPTER 3

DATA & METHODOLOGY

3.1 Data

In this section we will review the data used for the simulations, the details of our methodology in general as well as for the specific methods, the programming used to run them, and introduce our way for ranking the performance of the models. The data chosen are the stock indices listed in Table 3.1, and each essentially represents a single portfolio. We have extracted our data from Yahoo! Finance because of it is widely used and it's easily accessible. Specifically, the choice of our data are 10 stock indices of a pool of countries ranging from developed to developing economies. This was firstly because of the simplicity that using a single index as a portfolio allows, but at the same time a relatively representative sample of the movements of a country's financial assets. As indicated below each index consists of a relatively high number of stocks, and more importantly these are the top stocks by market capitalisation[1] in the country. For example, even for the Hang Seng Index, which consists of 50 stocks, they represent about 58% of the market capitalisation of Hong Kong Stock exchange. [Wikipedia] Secondly, we have deliberately chosen stock indices only because of the linearity of the stock returns. For non-linear assets, like bonds and derivatives, our methods do no apply and one has to resort to other ones like the delta-normal or delta-gamma approach. For the scope of this project we have chosen to exclude these types of portfolios as linear portfolios tend to be more common and their risk measures can be calculated by simply separating the portfolio into its linear and the non-linear part and apply the appropriate methods. Two sources suggested on calculating VaR and ES for non-linear portfolios are Chapter 11 and 12 [4] and [15].

[1] Market Capitalisation is the value of a company that is traded on the stock market, calculated by multiplying the total number of shares by the present share price.

Table 3.1: List of the Stock Indices, representing the portfolios to be tested

Stock Index	No. of Stocks	Region
Standard & Poor 500 (S&P 500)	500	USA
Financial Times Stock Exchange 100 (FTSE 100)	100	UK
Nihon Keizai Shimbun 225 (Nikkei 225)	225	Japan
EuroNext 100	100	Europe
DAX Performance Index (DAX)	30	Germany
Hang Seng Index (HSI)	50	Hong Kong
ASX 200	200	Australia
Standard & Poor Bombay Stock Exchange Sensitive Index (S&P BSE)	30	India
Bolsa de Valores do Estado de Sao Paulo Index (IBOVESPA)	60	Brazil
Indice de Precios y Cotizaciones (IPC)	35	Mexico

3.2 Programming

For our simulations, and any related numeric results posted in this Chapter we have chose MATLAB because of its easy of use given the multiple available toolboxes and its widely used environment. We have specifically used functions from the Risk Management Toolbox and the Econometric Toolbox, as well as a few of their ideas in the relevant code constructions. Our codes are listed in Appendix B.

We will now give an insight to the reader to the functions used from the above toolboxes. For volatility foreacsting, we have first used the garch() and gjr() models, which were used to create the EWMA, GARCH and GJR models. They allow to input the p and q parameters of the models, the distribution of our innovation term, and specify the parameters if desirable. For the EWMA we have therefore used the garch() function with specified parameters. Moreover, we have used the estimate() function on these models, inserting the log returns. Given these inputs the function estimates the coefficients of our GARCH family models using the maximum likelihood estimation (mle) method. It is the method generally used for estimating the coefficients of GARCH models and finds the value of coefficients that are most probable to fit the data. [4, p.73-74] If the innovation term distribution is a t-distribution the function also estimates the degrees of freedom of the distribution using the mle, given its stated mean and variance. Lastly, the forecast() function was used to provide the forecast of our variance, with equations as described in 2.1.3.

For the backtesting methods, we have used the varbacktest(), esbacktestbysim() and esbacktest() functions. The varbacktest() functions provides a report of the results various backtesting methods, including the unconditional coverage test, or proportion of failures test. The esbacktestbysim() and esbacktest() produce again a report of various backtesting methods, including our chosen one, with their difference being that the esbacktestbysim() allows us to choose the parameters of the model which will be used for simulations as explained in Equation 2.16.

3.3 Methodology

3.3.1 General Methodology and Hypotehses

We have split the periods of our estimates into two periods, the "Crisis Period" ranging from 1st of June 2006 to 1st of August 2009, and the "Post-Crisis Period" ranging from 1st of January 2013 to 1st of January 2018. We provide the sample statistics of the returns of each of our "portfolios" at the Table 3.2, to give a broader view of how they differ between the two periods. The results in this table and any table that follows are provided to two decimal places.

As expected, firstly we can notice a significance difference in the volatility, or standard deviation values, of our returns between the Crisis and the Post-Crisis period. Also the mean of around 0 is not surprising as discussed earlier in the characteristics of asset returns. The kurtosis statistic essentially describes how leptokurtic or not a distribution is. The value of the kurtosis of a standard Normal distribution is 3 and the greater it is the more it tends towards a t-distribution. As we can see in our returns data, our kurtosis is generally significantly above 3. Moreover, the a positive skewnness implies that a distribution has a right tail that is longer, or simply more data are concentrated towards the negative values of a distribution, and vice versa. From there we could potentially argue that the generally positive skewnness of the Crisis period relative to the Post Crisis, means more large negative returns and leads to the higher standard deviation. This is roughly what the leverage effect argues, which recalling to section 2.1, it suggests that the volatility response to a large positive return is considerably smaller than that to a negative return of the same mangitude.Therefore, these are reasons that lead us to test the following models in the parametric method. The t-distribution to test whether it captures the leptokurtic aspects of our returns data, the EWMA and GARCH models to check whether they adapt to changes in volatility in the recent time period, or in technical terms volatility clustering, and lastly the GJR which on top of that aims to also capture the leverage effect.

On top of the data for the forecasts time periods, our models need to be fed with data before this in time. For example, the GARCH family of models need an amount of past data in order to compute their coefficients. As suggested by [4, p.71], for the GARCH models one should use at least 1,000 daily past observations. Given the year of trading days consists of roughly 250 days this equates to about past 4 years of data. We have chosen to stick to roughly 1,000 daily past observations for our estimation periods for the following reason. During the Crisis period it allows to feed the GARCH models with a period of relatively "calm" past data, and then ask it to forecast "stormy" estimates, while for the Post-Crisis period we feed it with "calm" past data and ask for "calm" forecasts. We illustrate this more explicitly in the Figure 3.1 for the S&P500 index, which is a generally fair representative of our 10 portfolios. The reasoning behind this is that we would like to test how much the GARCH models can stretch their capabilities in the first case, and at the same time be an appropriate family of models under "normal" conditions.

Lastly, we have chose to test our risk measures at 95% and 99% quantiles. This allows us to

Table 3.2: Sample statistics of the returns of our "portfolios"

Period	Index	Mean	Standard Deviation	kurtosis	skewness	min	max
Crisis	S&P500	0	0.02	7.33	-0.10	-0.09	0.11
	FTSE100	0	0.02	7.07	0.00	-0.09	0.09
	Nikkei225	0	0.02	8.46	-0.32	-0.12	0.13
	Euronext100	0	0.02	8.01	0.15	-0.09	0.10
	DAX	0	0.02	8.41	0.32	-0.07	0.11
	HSI	0	0.03	6.41	0.15	-0.14	0.13
	ASX200	0	0.02	5.26	-0.32	-0.09	0.06
	S&P BSE	0	0.03	6.73	0.20	-0.12	0.16
	IBOVESPA	0	0.03	6.84	0.07	-0.12	0.14
	IPC	0	0.02	7.33	0.17	-0.07	0.10
Post-Crisis	S&P500	0	0.01	5.88	-0.43	-0.04	0.04
	FTSE100	0	0.01	5.66	-0.20	-0.05	0.04
	Nikkei225	0	0.01	7.62	-0.038	-0.08	0.07
	Euronext100	0	0.01	6.66	-0.40	-0.07	0.04
	DAX	0	0.01	5.51	-0.37	-0.07	0.05
	HSI	0	0.01	5.47	-0.31	-0.06	0.04
	ASX200	0	0.01	4.57	-0.30	-0.04	0.03
	S&P BSE	0	0.01	5.95	-0.39	-0.06	0.04
	IBOVESPA	0	0.01	4.96	-0.09	-0.09	0.06
	IPC	0	0.01	4.91	-0.19	-0.05	0.04

test more thorouhgly the behaviour of the tail of the distribution and thus determine whether a t-distribution does prove a better one in these cases. This might also help us to determine the best distribution for the 97.5% required by the Fundamental Review of the Trading Book (FRTB) regulation.[8]

Figure 3.1: S&P500 graph of daily prices. Source: Google

3.3.2 VaR and ES Methods and Backtesting

As stated earlier, the parametric methods and HS method introduced will be used to estimate the one day VaR and ES. For the parametric method we will test both a Normal Distribution and a Student's t for the innovation term, and for the Volatility forecasting part of the method we will test the EWMA, GARCH and GJR models. For the quantiles of VaR and ES we will test the 95% and 99% quantiles. For the Historical Simulation method we will chose a time window of 250 days as past data, which is the most popular time window as mentioned in [17]. For the Volatility Forecasting models, we will choose a time window of 1000 days, for the estimation of the model coefficients. These will be estimated by the mle method, and the estimation will be updated on every day of the forecast. The volatility estimate will be calculated with the equations introduced in Chapter 2. For the mean estimate in the parametric method, as mentioned earlier we will simply assume it to be 0. Lastly, given these and the equations defined in Chapter 2 for parametric and HS estimates, we will calculate our VaR and ES estimates. We will simulate roughly 1800 forecasts for our 10 "portfolios", introduced in Table 3.1, in the time periods of "Crisis" and "Post-Crisis" as defined earlier.

To test these methods we will use one backtesting method for each risk measure. For the VaR we will use the Unconditional Coverage test with a 95% confidence interval and report the Likelihood Ratio of our methods. We will then compare these to the critical values of a χ_1^2 distribution, which for a 95% confidence interval is 3.841. For the ES we will use the Unconditional test. We will simulate scenarios from a t-distribution with 15 degrees of freedom, for our t-distribution parametric methods, as this was roughly the average of the degrees of freedom calculated by the mle method using the estimate() function in MATLAB. These degrees of freedom also fall in line with the values successfully tested by [3]. For the normal distribution parametric method and for HS we will simply simulate the values from a N(0,1) distribution. The Z-statistics of our backtest will be then reported which will be compared to the critical value of Z for a 95% confidence level, which is -0.7.

3.4 Methods Ranking

To compare and rank the performance of our methods we will use two methods. Both of these methods are based on the test statistics produced by our backtests. The first method will be a measure of the frequency of best estimation, n_ω over all our portfolios. In simple terms this measure will count the times that a method performed best over the rest. Our indicator for best performance will be the test statistics produced in each case relative to our critical values. In the VaR case the smallest LR of a method will indicate the method performed best. Therefore through this measure, we aim to capture which method performed best the most times. However, the limitation that arises with this is that we do not have any information on how much "well" did it perform, and we might also miss methods that on average performed well but not the best in a single portfolio. Therefore, we will also consider a second measure, name it the average performance measure. In the VaR case this measure will calculate the sum of the Likelihood Ratios, $\sum LR$, of a single model over all the portfolios. For the ES, this will be the sum of the absolute value of the Z-statistic, $\sum |Z|$, of a model over all portfolios. We have chosen the absolute value of Z since the optimal Z lies in the value closest possible to 0, and therefore without the absolute value our measure would not be representative of performance. Therefore, will allow us to determine which method performed best on average.

As mentioned previously, apart from the best individual model, we would like to determine which distribution over the Normal and t-distribution, represents best the asset returns distribution in the parametric method. For this we will make use of the two model ranking measures introduced above and adjust them to this case. Therefore for the VaR we will sum the sum of the LR ratios of a single normal parametric method over the portfolios, over the 3 normal distribution methods, $\sum\sum LR$ and equivalently for the t-distribution methods. For the ES case we will do the same with the Z statistics and thus our measure will be, $\sum\sum |Z|$.

Lastly, we will also consider which methods satisfied our hypothesis testing, and consequently were below the critical value over all. For the VaR case, as a stated earlier the critical value of a 95% confidence interval test, for a single LR-statistic is 3.841, for the $\sum LR$ is 18.307 and for $\sum\sum LR$ is 124.342. As summing a χ^2_1 distributions n times is equal to a χ^2_n distribution, the $\sum LR$ and $\sum\sum LR$ are distributed to a χ^2_{10} and and χ^2_{100} respectively, and therefore the critical values are derived from there. For ES, we will use the critical value of Z-statistic is -0.7 for a 95% confidence interval, where as for the for the $\sum |Z|$ and $\sum\sum |Z|$, we do not have any reasonable critical value to use and will therefore just compare their values as explained above.

CHAPTER 4

RESULTS

In this Chapter we will report our results from the simulations and discuss them, with the aim of evaluating our methods. Our results are reported in the 12 tables below. We will first discuss the results of VaR in each of the two time periods, Crisis and Post-Crisis, move on to the ES shortfall, and lastly discuss the general picture. In each of the VaR and ES sections we firstly introduce our Ranking statistics for the 95% and 99% quantiles and then for each period we report our individual portfolios results for the 95% and 99% quantiles. The ranking statistics table include our ranking measures as introduced in Section 3.4, where as the individual portfolios results report the test statistics for each corresponding backtesting method as explained in Subsection 3.3.2. In the ranking statistics tables, we highlight with a grey area the methods that performed best for each ranking measure, and write with red coloured numbers the methods that are rejected by the null hypothesis.

4.1 Value at Risk

In this section we post the tables below with the ranking statistics results, in Tables 4.1 and 4.4 and the Unconditional Coverage results for the 95% and 99% VaR results and for the Crisis and the Post Crisis period, Tables 4.2, 4.3, 4.5 and 4.6

4.1.1 95% Value at Risk

Reviewing our ranking statistics for the Crisis period, we can first determine that EWMA the best model contender. It outperforms the other models in both the average performance measure and the frequency of best estimation. The only other fair competitors are the GARCH and the GJR t-distributed models which have performed relatively well by the frequency of best estimation

Table 4.1: Ranking statistics of 95%-VaR estimates

		Crisis				Post Crisis			
		$\sum LR$	$\sum\sum LR$	n_ω	$\sum n_\omega$	$\sum LR$	$\sum\sum LR$	n_ω	$\sum n_\omega$
Normal	HS	93.93		1		2.3		6	
	EWMA	14.99		3		3.28		4	
	GARCH	72.38	176.45	0	4	44.11	294.29	0	4
	GJR	89.08		1		246.91		0	
Student's t	EWMA	12.76		1		9.09		0	
	GARCH	26.08	139.51	2	5	33.11	375.93	0	0
	GJR	100.67		2		333.73		0	

measure. The rest of the models seem to be quite far away in terms of the ranking statistics, with the GARCH normal model in particular achieving no best estimation over our 10 portfolios. For the Post Crisis period the picture changes slightly. HS becomes the best method according to both of our ranking statistics. The EWMA remains a strong competitor, especially for the normal distribution. The rest of the models remain far behind with the worst performing one being the GJR, especially for the t-distribution.

For the distributions, we have a slightly better performance of the normal one over the t during the Crisis period, for both of the ranking statistics. During the Post-Crisis period, the normal distribution seems to be the clear winner, with the t-distribution achieving no best estimation.

Lastly, in terms of our hypothesis tests, during the Crisis period all of our models but the EWMA, are rejected at the 95% confidence interval test. More specifically, the EWMA is the only model that is not rejected for every single of our portfolio, for both the normal and the t-distributions. On the other hand no single method has been rejected for all of the portfolios. For the Post Crisis period, the only significant changes is that the HS is not rejected in any of the portfolios, why the GJR is rejected for all, for both of the distributions. Lastly, under our critical value both the normal distribution and the t-distribution are rejected for both periods.

4.1.2 99% Value at Risk

Regarding the 99% Value at Risk estimation, our test statistics again rank the EWMA method first for both the Crisis and the Post-Crisis period. For both periods, the normal EWMA model performs best in the frequency of best estimation statistic, while the t-distributed one best for the average performance. Its only fair competitor is the GARCH model, especially for the normal distribution for both periods. For the Post-Crisis period HS becomes another good competitor, performing relatively well in both of our ranking statistics. The worst performer over both periods is the GJR model, which performed relatively worse in both of the test statistics for both of the periods.

For the distributions, the change form the 95% VaR estmation is that the Normal becomes

Table 4.2: Unconditional Coverage test LR results for 95%-VaR estimates during the Crisis period

Portfolio		Parametric					
	HS	EWMA-N	GARCH-N	GJR-N	EWMA-t	GARCH-t	GJR-t
S&P 500	18.07	3.75	7.78	3.10	3.75	9.75	1.10
FTSE 100	11.51	2.77	0.31	0.40	1.69	0.55	0.18
Nikkei 225	6.81	1.90	47.99	42.03	1.44	0.20	56.73
Euronext 100	20.14	4.15	6.50	3.12	3.48	6.50	3.12
DAX	15.22	1.44	1.44	3.79	1.44	2.41	3.79
HSI	4.24	0.31	0.13	2.54	0.31	0.05	3.28
ASX 200	7.61	0.08	0.08	12.80	0.11	0.22	14.70
S&P BSE	0.13	0.31	2.27	7.79	0.48	0.68	4.89
IBOVESPA	5.79	0.05	5.79	2.54	0.05	4.99	1.91
IPC	4.41	0.23	0.09	10.97	0.01	0.73	10.97

Table 4.3: Unconditional Coverage test LR results for 95%-VaR estimates during the Post-Crisis period

Portfolio		Parametric					
	HS	EWMA-N	GARCH-N	GJR-N	EWMA-t	GARCH-t	GJR-t
S&P 500	0.84	0.84	12.06	88.36	4.06	11.03	111.16
FTSE 100	0.05	0.75	7.519	41.91	0.66	4.78	61.36
Nikkei 225	0.11	0.04	2.89	9.94	0.22	0.54	15.39
Euronext 100	0.00	0.41	6.53	42.91	1.07	6.53	59.41
DAX	0.01	0.04	3.20	19.85	0.20	2.29	44.27
HSI	0.23	0.12	1.04	8.21	0.12	0.39	7.39
ASX 200	0.72	0.48	2.69	6.90	1.51	2.26	8.52
S&P BSE	0.03	0.24	3.46	12.08	0.39	2.95	15.53
IBOVESPA	0.13	0.31	2.27	7.79	0.48	0.68	4.89
IPC	0.18	0.05	2.45	8.96	0.38	1.66	5.81

the best distribution in terms of both our test statics for the Crisis period, while it remains the better one for the Post-Crisis. Especially, in terms of best frequency estimation it outperforms the t-distribution by at least 3 times in both periods.

In terms of our hypothesis testing, the EWMA is again proven to be a reliable model, as it is not rejected for any portfolio for both of the periods. On the other hand, for the Crisis period we have no model rejected for all portfolios, while for the Post-Crisis the GJR for both of the distributions is rejected over all the portfolios.

4.2 Expected Shortfall

For the Expected shortfall estimates, our relevant results are presented in the tables below. These include the ranking statistics, in Tables 4.7 and 4.10 and the Unconditional test Z-statistics for the 95% and 99% estimations for the Crisis and Post Crisis period, in Tables 4.8, 4.9, 4.11, and

Table 4.4: Ranking statistics of 99%-VaR estimates

		Crisis				Post-Crisis			
		$\sum LR$	$\sum\sum LR$	n_ω	$\sum n_\omega$	$\sum LR$	$\sum\sum LR$	n_ω	$\sum n_\omega$
Normal	HS	64.66		1		18.22		2	
	EWMA	7.96		5		9.35		4	
	GARCH	40.55	138.9	3	8	58.55	246.66	2	6
	GJR	90.39		0		178.76		0	
Student's t	EWMA	7.13		0		6.6		0	
	GARCH	49.06	152.46	0	1	123.42	306.25	2	2
	GJR	96.27		1		176.23		0	

Table 4.5: Unconditional Coverage test LR results for 99%-VaR estimates during the Crisis period

Porftolio		Parametric					
	HS	EWMA-N	GARCH-N	GJR-N	EWMA-t	GARCH-t	GJR-t
S&P 500	17.16	0.85	5.38	0.29	0.29	0.29	0.09
FTSE 100	4.45	1.01	0.42	10.92	1.01	2.89	10.92
Nikkei 225	4.42	0.36	11.12	11.12	0.36	11.12	11.12
Euronext 100	3.82	0.60	3.30	11.56	0.60	6.04	11.56
DAX	2.66	1.57	1.57	11.48	1.57	5.97	11.48
HSI	15.36	0.95	0.06	3.05	0.95	1.43	11.18
ASX 200	2.66	0.27	0.01	11.48	0.01	0.58	11.48
S&P BSE	1.51	1.84	5.82	7.87	1.83	7.87	5.82
IBOVESPA	5.66	0.49	1.43	11.18	0.49	1.43	11.18
IPC	6.96	0.02	11.44	11.44	0.02	11.44	11.44

Table 4.6: Unconditional Coverage test LR results for 99%-VaR estimates during the Post-Crisis period

Portfolio		Parametric					
	HS	EWMA-N	GARCH-N	GJR-N	EWMA-t	GARCH-t	GJR-t
S&P 500	0.03	2.85	25.31	25.31	1.41	25.31	25.31
FTSE 100	1.37	0.83	14.00	25.41	0.42	18.31	25.41
Nikkei 225	4.08	0.22	2.74	24.74	0.15	7.68	24.74
Euronext 100	1.28	0.76	1.26	25.69	0.76	25.69	25.69
DAX	2.00	0.01	4.42	18.39	0.04	14.06	25.49
HSI	3.13	2.31	0.04	5.68	0.55	5.68	5.68
ASX 200	0.82	0.14	3.05	18.35	1.19	8.16	14.03
S&P BSE	3.11	0.22	0.01	13.49	0.04	7.72	13.49
IBOVESPA	1.51	1.84	5.82	7.87	1.84	7.87	5.82
IPC	0.89	0.17	1.90	13.83	0.20	2.94	10.57

Table 4.7: Ranking statistics of 95%-ES estimates

		Crisis				Post-Crisis			
		$\sum\lvert Z\rvert$	$\sum\sum\lvert Z\rvert$	n_ω	$\sum n_\omega$	$\sum\lvert Z\rvert$	$\sum\sum\lvert Z\rvert$	n_ω	$\sum n_\omega$
Normal	HS	9.57	9.67	0	3	1.26	8.27	3	7
	EWMA	1.77		2		0.45		7	
	GARCH	2.85		1		2.56		0	
	GJR	5.05		0		5.26		0	
Student's t	EWMA	1.25	8.18	4	7	1.35	9.98	0	0
	GARCH	1.76		2		2.65		0	
	GJR	5.17		1		5.98		0	

4.12.

4.2.1 95% Expected Shortfall

Looking at our test statistics, one can determine the EWMA is again the winner for both of the periods. For the Crisis periods, it's t-distribution model performs best and its competitors involve mainly the GARCH model with the t model performing better. Regarding the Post-Crisis period, the EWMA Normal model has the most frequencies of best estimation by far. The Historical Simulation again proves to perform better during the Post-Crisis period, becoming the main competitor to the EWMA normal model. The worst model for the Crisis period is the HS, with the GJR normal model as a slight second. Both of the models achieve no best estimations. Regarding the post-crisis period, the GJR model seems to be the overall worst model according to both of the ranking statistics, while the GARCH model achieved was close next, achieving no best estimations for both of its distributions.

For our distributions, the t-distribution seems to be the better one for the Crisis period. It has a lower average performance value and more than double the number of best estimations. During the post-crisis period, the normal distribution proves to be the better one, achieving 7 best estimations to the 0 of the normal one.

Regarding our hypothesis tests the only rejected method is the HS for the crisis period. Specifically, it is also rejected for 7 out of the 10 portfolios, while the other methods are not rejected for any of the portfolios. However, the GJR model in particular seems to have a large number of positive values for both of the distributions, which indicates an overestimation of the Expected Shortfall. As for the Post Crisis period, the only difference is that the HS method performs much better and is not rejected, or otherwise proves reliable, for all of our portfolios.

4.2.2 99% Expected Shortfall

The 99% estimates of the Expected Shortfall estimates, generally have a similar picture with the 95% ones. According to our ranking statistics, the EWMA outperforms any other model in the Crisis period, being the only model with best estimates. The worst performing model is

Table 4.8: Unconditional test, Z-statistic for 95%-ES estimates during the Crisis period

Portfolio		Parametric					
	HS	EWMA-N	GARCH-N	GJR-N	EWMA-t	GARCH-t	GJR-t
S&P 500	-1.46	-0.37	-0.60	-0.29	-0.29	-0.52	0.05
FTSE 100	-1.12	-0.31	-0.07	0.22	-0.20	-0.04	0.24
Nikkei 225	-1.25	-0.26	0.96	0.93	-0.11	0.16	1
Euronext 100	-1.30	-0.33	-0.37	0.42	-0.23	-0.31	0.44
DAX	-1.10	-0.21	-0.16	0.41	-0.14	-0.17	0.45
HSI	-0.67	0.10	0.08	0.33	0.16	0.06	0.39
ASX 200	-0.83	-0.08	-0.04	0.64	0.02	-0.03	0.67
S&P BSE	–0.45	-0.03	0.16	0.85	0.03	0.15	0.92
IBOVESPA	-0.77	-0.01	-0.39	0.35	0.03	-0.30	0.38
IPC	-0.62	-0.07	0.02	0.61	0.04	-0.02	0.63

Table 4.9: Unconditional test Z-statistic results for 95%-ES estimates during the Post-Crisis period

Portfolio		Parametric					
	HS	EWMA-N	GARCH-N	GJR-N	EWMA-t	GARCH-t	GJR-t
S&P 500	-0.10	0.05	0.45	0.92	0.24	0.48	0.98
FTSE 100	-0.11	-0.11	0.37	0.73	0.13	0.35	0.83
Nikkei 225	-0.11	-0.00	0.24	0.42	0.08	0.17	0.55
Euronext 100	-0.04	0.05	0.31	0.72	0.16	0.37	0.83
DAX	-0.05	-0.02	0.23	0.55	0.11	0.28	0.75
HSI	-0.01	0.00	0.11	0.35	0.08	0.14	0.39
ASX 200	-0.20	0.06	0.21	0.36	0.17	0.23	0.41
S&P BSE	-0.07	0.05	0.21	0.45	0.13	0.25	0.53
IBOVESPA	-0.04	0.10	0.22	0.37	0.14	0.17	0.34
IPC	-0.15	-0.01	0.21	0.39	0.11	0.21	0.37

Table 4.10: Ranking statistics of 99%-ES estimates

		Crisis				Post Crisis			
		$\sum\|Z\|$	$\sum\sum\|Z\|$	n_ω	$\sum n_\omega$	$\sum\|Z\|$	$\sum\sum\|Z\|$	n_ω	$\sum n_\omega$
	HS	17.06		0		4.24		2	
	EWMA	2.67		5		2.24		4	
Normal	GARCH	6.21	17.66	0	5	4.71	15.78	2	4
	GJR	8.78		0		8.83		0	
	EWMA	2.07		5		10.84		2	
Student's t	GARCH	6.84	18.27	0	5	7.75	27.48	0	2
	GJR	9.36		0		8.89		0	

Table 4.11: Unconditional test Z-statistic for 99%-ES estimates during the Crisis period

Portfolio		Parametric					
	HS	EWMA-N	GARCH-N	GJR-N	EWMA-t	GARCH-t	GJR-t
S&P 500	-2.89	-0.33	-1.05	-0.12	0.01	-0.03	0.36
FTSE 100	-1.74	-0.42	0.31	1	-0.26	0.68	1
Nikkei 225	-2.29.	-0.29	1	1	0.13	1	1
Euronext 100	-1.71	0.28	0.67	1	0.38	0.85	1
DAX	-1.48	0.42	0.50	1	0.50	0.82	1
HSI	-2.16	-0.32	0.15	0.66	-0.07	0.58	1
ASX 200	-1.63	-0.23	-0.07	1	0.04	0.36	1
S&P BSE	-1.39	0.06	1	1	0.20	1	1
IBOVESPA	-1.43	0.29	0.46	1	0.36	0.52	1
IPC	-1.28	-0.03	1	1	0.12	1	1

the HS one, as it has the highest average performance value. As for the Post Crisis period, the EWMA model is still the better model, with its best performing being the Normal one, while its competitors seems to be the HS and the GARCH normal model. The rest of the models have no best estimation. The worst performing model seems to be again the GJR, underperfoming in both of the ranking statistics.

For the distributions, in the Crisis period the Normal distribution slightly outperforms the t one, with its only edge being the slightly less average performance value. In the Post-Crisis period, it seems to be more clearly the better distribution, outperforming the t distribution significantly in both of the ranking statistics.

Regarding our hypothesis tests, for the Crisis period the HS proves again to be an unreliable model, as it's rejected for all of the portfolios. All other models except the GARCH Normal, are not rejected by the critical value, however the GJR model seems to again overestimate the ES. For the Post Crisis period the main difference concerns the improved performance of the HS, which has only two rejects out of the ten, and no other model is being rejected over all the portfolios.

4.3 Discussion

Given the above results and methods of comparison, the EWMA proves to be the best method quite clearly for both "calm" periods and "stormy" periods. For the Crisis period this was generally followed by the GARCH model where as for the Post-Crisis by the HS model. The overall worst performing model was the GJR model, which performed particularly bad in the Post-Crisis period. This was followed by the HS method for the Crisis Period and with the GARCH model for the Post Crisis.

Firstly, the outperformance of the EWMA might sensibly suprise. This is because compared to the parametric methods it takes the least factors into account. Theoretically, it should not adapt as quickly to volatility clustering compared to the GARCH model, as its coefficient is fixed,

Table 4.12: Unconditional test Z-statistic results for 99%-ES estimates during the Post-Crisis period

Portfolio		Parametric					
	HS	EWMA-N	GARCH-N	GJR-N	EWMA-t	GARCH-t	GJR-t
S&P 500	-0.10	-0.56	1	1	-0.11	1	1
FTSE 100	-0.56	-0.25	0.86	1	0.03	0.94	1
Nikkei 225	-0.72	-0.17	0.47	1	0.18	0.71	1
Euronext 100	-0.49	-0.26	0.34	1	-0.03	1	1
DAX	0.04	-0.02	0.54	0.93	0.22	0.87	1
HSI	-0.63	-0.41	-0.02	0.62	0.02	0.65	0.69
ASX 200	-0.41	-0.06	0.48	0.93	0.38	0.72	0.87
S&P BSE	-0.74	-0.11	0.03	0.84	0.21	0.67	0.87
IBOVESPA	-0.35	0.35	0.58	0.66	0.42	0.68	0.65
IPC	-0.49	-0.05	0.39	0.85	9.24	0.51	0.81

and it does not account for the leverage effect that the GJR model does. However, on a second thought this raises concern over the estimation of the coefficients of the GARCH family of models. Firstly, the mle method has different calculation variations and it is regarded by no means as an optimal method for coefficient estimation. Secondly, another concern around the mle might be the amount data it receives. Given that it needs more than 1,000 data to perform well, this means that after all our coefficients might be slow to adapt. Therefore, it a large coefficient estimation error using the mle might be a factor for the EWMA proving to be a better method. Moreover, another limitation over the mle that we have is whether using the minimum suggested amount to achieve our testing of "calm" and "stormy" periods, has proved costly.

On the positive side for the GARCH model, the fact that it has second best and better than the HS method in the Crisis period, indicates that volatility clustering might indeed be worth capturing. The HS method did not seem to be able to cope with the higher volatility of the Crisis period and this might be mainly the reason for performing significantly worse during this period. This agrees with the theory, as the method relies on a relatively large amount of past data and therefore, it is slow to adapt to changing circumstances. Even using the least amount of past data suggested by the Basel committee, the HS proved to perform worse to any other models tested. Therefore, we would suggest that the Historical Simulation method is avoided during "stormy" periods. On the bright side, under normal conditions Historical method seems to be a strong one. This again agrees with the theory and it could have been expected by looking at the figure 3.1.

Regarding the distributions of the parametric methods, we have seen the normal distribution to overall outperform the t-distribution, with the only case of the t-distribution performing better being the 95% quantile for the Crisis period for both VaR and ES. This suggests that our data did not have much fatter tails than a standard normal, and probably much less than a student's t. Moreover, our belief for the worse performance of the Student's t in the Post Crisis period, is that our returns tend to be more peaked around 0, as suggested by the min and max values in the

Table 3.2. This is one of the limitations of using a Student's t as described in the theory part of the project, since it is less peaked around 0 than a normal distribution. Another reason, for the worse performance of the Student's t distribution might again lie around the mle, as it is used to calculate the degrees of freedom of the distribution. Therefore, these led us to suggest the use of a normal distribution for the innovation term in the parametric method, given our results and the easier nature of use of the Normal distribution. However, at the same time we acknowledge that the Normal distribution has been long empirically rejected as a model for the returns of financial assets, and therefore suggest this topic, as a field for further research.

Concerning, the use of the ES over the VaR estimates, we have seen that the ES models have performed much better given the backtests used. This might therefore make the use of ES over VaR in the future a more reliable method of capturing market risk. On the other hand, we will hold our limitations on this point as this has been based on a single backtesting method, which has not been in the industry for too long.

Lastly, for these results, we have seen no significant difference in the performance of our models over the 95% and 99%, and therefore we suggest that the same model is used for these quantiles, as well as potentially any quantile in between them.

Chapter 5

Conclusion

The aim of the project as introduced in the beginning was to determine the best methods for computing VaR and ES. We also aimed on discussing how the two measures differ, and therefore what the change in regulations from VaR to ES means.

Overall, regarding our models we have seen that going for more complex parameterisations, like in the GJR model, has not paid out well. On the other hand the relatively simple EWMA has proved to be consistent over almost all of our portfolios, for both of the time periods, and for both of our risk measures. Therefore, given the simplicity of the model we suggest that this model is used for stock portfolios. Lastly, the Student's t distribution did not seem to prove competent enough to capture the fatter tails of the asset returns characteristics, and therefore the more easy to use Normal Distribution is suggested to be used.

Regarding the introduction of the ES over VaR for regulation purposes is expected to have some effects on market risk calculations as explained. Firstly, as determined in Chapter 4 much more methods for ES were proved to be reliable under hypothesis testing compared to VaR, and therefore we expect the ES to capture market risk generally more accurately over different institutions. Moreover, theoretically as discussed in Chapter 2, the ES covers some of the disadvantage of VaR like the subadditivity and the consideration of extreme losses. Therefore, we expect it to be a more complete measure. However as explained, it still faces some challenges itself. Also, the shift to it might pose a challenge to financial institutions in general, due to it being less widely used for now. Concerning the future of Value at Risk, we expect it to remain relevant and be used for internal purposes, due to its advantages, its wide use and understanding, and its various applications developed, as discussed in Chapter 2.

5.1 Recommendations for further research

As discussed in Chapter 4, we recommend further research on the best distribution fitting the financial asset returns, as the Normal and Student's t have proved inadequate in this project.

Moreover, we would suggest further research in the methods for backtesting ES. During our research for this project we have noticed that backtesting methods on ES have been relatively scarce and not as solid as for VaR. The method used in this project is currently seen as one of the best in the literature, however given the short period it has been used, we would encourage that it is tested further.

Lastly, since the volatility forecasting models with adaptable coefficients have not proved to be adequate in this project, we encourage further research in volatility forecasting. More specifically, through our research amongst others, we have came across the [20] article which illustrates a integrated statistical model with deep neural networks that outperforms volatility forecasting models such as GARCH. Although this is a field traditionally more relevant to Computer Science, we encourage the keen risk manager, or statistician, to further research how fields like deep neural networks and machine learning could be used to improve volatility forecasting, in a period where they are becoming ever more important.

Appendix A

Derivation of parametric method equations

A.1 Value at Risk

A.1.1 Normal Distribution

Let Φ denote the cumulative distribution function of the Standard normal distribution.

Recall the definition of the Cumulative distribution function, F, for a random variable X, over a value x.

$$F_X(x) = P(X < x)$$

We start with using the Equation 2.2, used to define Value at Risk.

$$P(R_{PF,t+1} < -VaR^q_{t+1}) = q$$

By substituting our assumed model for the Returns of the portfolio distribution as defined in equation 2.4 we have the following

$$P(\sigma_{PF,t+1} z_{t+1} < -VaR^q_{t+1}) = q$$

Suppose σ is estimated using a volatility forecasting model as introduced in Section 2.4.1, then since a standard deviation is positive we have

$$P\left(z_{t+1} < \frac{-VaR^q_{t+1}}{\widehat{\sigma_{PF,t+1}}}\right) = q$$

Assuming z_{t+1} is distributed to a N(0,1) distribution, we can now take the inverse cumulative function on both sides of the equation and get our VaR equation 2.5 as follows,

$$\frac{-VaR^q_{t+1}}{\widehat{\sigma_{PF,t+1}}} = \Phi^{-1}(q)$$

$\iff$

$$\widehat{VaR^q_{t+1}} = \widehat{\sigma_{PF,t+1}}\Phi^{-1}(q)$$

A.1.2 Student's t Distribution

For the Student's VaR equation we begin from the following step of the Normal distribution derivation.

$$P\left(z_{t+1} < \frac{-VaR^q_{t+1}}{\widehat{\sigma_{PF,t+1}}}\right) = q$$

Now recall that the variance of a Student's t distribution with ν degrees of freedom is $\frac{\nu}{\nu-2}$. Assuming that z_{t+1} is distributed to a Student's t distribution with variance 1 and mean 0 as introduced in Equation 2.4, then it must be that $z_t \sim t_\nu\sqrt{\frac{\nu-2}{\nu}}$ or $z_t\sqrt{\frac{\nu}{\nu-2}} \sim t_\nu$. Multypling the above equation by $\sqrt{\frac{\nu}{\nu-2}}$ gives,

$$P\left(z_{t+1}\sqrt{\frac{\nu}{\nu-2}} < \frac{-VaR^q_{t+1}}{\widehat{\sigma}_{PF,t+1}}\right)) = q\sqrt{\frac{\nu}{\nu-2}}$$

Taking the inverse of the Student's t distribution and rearranging the equation gives the desired VaR equation 2.7 as follows,

$$\widehat{VaR^q_{t+1}} = -\widehat{\sigma_{PF,t+1}}t^{-1}_\nu(q)\sqrt{\frac{\nu}{\nu-2}}$$

A.1.3 Expected Shortfall

For the derivation of the Expected Shortfall parametric equations we will follow [2, p.45-46]

A.1.4 Normal Distribution

Assume that z_{t+1} is distributed to a N(0,1).

We start from the ES equation as defined in Equation 2.3

$$ES^q_{t+1} = -E_t[R_{PF,t+1} \mid R_{PF,t+1} < -VaR^q_{t+1}]$$

Substituting equation 2.4 gives

$$ES^q_{t+1} = -E_t[\sigma_{PF,t+1}z_{t+1} \mid \sigma_{PF,t+1}z_{t+1} < -VaR^q_{t+1}]$$

By taking the standard deviation out and using that $VaR^q_{t+1} = \Phi^{-1q}$, we have

$$ES^q_{t+1} = -\sigma_{PF,t+1} E_t[z_{t+1} \mid z_{t+1} < -\Phi^{-1}(q)]$$

Using the equation of the conditional expectation we have,

$$ES^q_{t+1} = -\sigma_{PF,t+1} \frac{1}{1-q} \int_{\Phi^{-1}}^{\infty} z\phi(z)$$

$\Longleftrightarrow$

$$ES^q_{t+1} = -\sigma_{PF,t+1} \frac{1}{1-q} [-\phi(z)]^{\infty}_{\Phi^{-1}(q)}$$

$\Longleftrightarrow$

$$ES^q_{t+1} = -\sigma_{PF,t+1} \frac{1}{1-q} \phi(\Phi^{-1}(q))$$

As derivation of the student's distribution includes similar steps except for the integration of a student's t distribution with mean 0 and standard deviation 1, we will stop our derivations here.

APPENDIX B

MATLAB CODE

B.1 Matlab Code

```
%% Value−at−Risk and Expected Shortfall Estimation and Backtesting
%
% This code estimates the value−at−risk (VaR) and Expected Shortfall(ES)
%using the following methods, and performs a VaR and ES backtesting
%analysis. The methods are:

% # Historical simulation
% # Exponential weighted moving average (EWMA) − Normal Distribution
% # GARCH(1,1) − Normal Distribution
% # GJR − Normal Distribution
% # Exponential weighted moving average (EWMA) − t−Distribution
% # GARCH(1,1) − t−Distribution
% # GJR − t−Distribution
%
% The estimation methods used, estimate the VaR at 95% and 99%
% confidence levels.

%% Start Function
function [] = VaRestimates(filename, estimationyear, estimationmonth,PortfolioID)

%filename − Data to be imported from excel, 1 column of dates and 1 column
```

```
%of closing prices.
% estimation year – year on which the VaR and ES estimations begin.
%estimation month – month on which the VaR and ES estimations begin.

%% Load the Data and Define the Test Window

% Get daily close and date data
[~,~ , rawData ]= xlsread (filename,'','' ,'basic');
date_cell = rawData (2:length(rawData),1);
close_cell = rawData (2:length(rawData),2);
date = zeros (1,length(date_cell));
close = zeros (1,length(close_cell));

% Convert the data to string format
for i =1: length (date)
date(i) = datenum(cell2mat(date_cell(i)));
close_i = cell2mat(close_cell(i));
if iscellstr (close_i)
close (i)= str2num (close_i);
else
close (i)= close_i;
end
end
date = fliplr(date);
% Convert the date format to a correct one
DateReturns = datestr(x2mdate(date));
DateReturns = flipud(DateReturns);

Returns = zeros(length(DateReturns),1);
for i= 2:length(DateReturns)
Returns(i)= log(close(i)/ close(i-1));
end

%Define samplesize
SampleSize = length(Returns);

%%
% Define the estimation window as 250 trading days for HS and 1000 for
```

```
% parametric methods. For parametric methods the estimation window is used
% to calculate the coefficients  of the GARCH family models.

%%
TestWindowStart     = find(year(DateReturns)==estimationyear & month(DateReturns)==
    estimationmonth,1);
TestWindow          = TestWindowStart : SampleSize;
EstimationWindowSizeHS = 250;
EstimationWindowSizeGarch = 1000;

%%
% For a VaR confidence level of 95% and 99%, set the complement of the VaR level.
q = [0.05  0.01];

%% Compute the VaR Using the Historical Simulation Method

%Create the vector for the Historical Simulation VaR
VaR_Historical95 = zeros(length(TestWindow),1);
VaR_Historical99 = zeros(length(TestWindow),1);
ES_Historical95 = zeros(length(TestWindow),1);
ES_Historical99 = zeros(length(TestWindow),1);

%run iterations and save each of the length(TestWindow) Historical Simulation VaR
for t  = TestWindow
    i  = t  - TestWindowStart + 1;
    EstimationWindow = t-EstimationWindowSizeHS:t;
    X = Returns(EstimationWindow);
    %VaR estimates
    VaR_Historical95(i) = -quantile(X,q(1));
    VaR_Historical99(i) = -quantile(X,q(2));
    %ES estimates method
    N = length(X);
    k1 = ceil(N*(1-q(1)));
    k2 = ceil(N*(1-q(2)));
    z  = sort(X);

     if  k1 < N
       ES95 = ((k1 - N*(1-q(1)))*z(k1) + sum(z(k1 +1:N)))/(N*q(1));
```

```
    else
       ES95 = z(k1);
     end

    if k2 < N
       ES99 = ((k2 − N*(1−q(2)))*z(k2) + sum(z(k2+1:N)))/(N*q(2));
    else
       ES99 = z(k2);
     end
    %ES estimate
    ES_Historical95(i) = ES95;
    ES_Historical99(i) = ES99;

end

%% Compute VaR using the Exponential Weighted Moving Average(EWMA) model.

%Set lambda = 0.94
lambda = 0.94;
Zscore    = norminv(q);

%Create EwmaN and GwmaT models
EwmaN = garch ('GARCHLags',1,'ARCHLags',1,'Distribution','Gaussian','Constant',1e−322 ,'
    ARCH',(1−lambda)−(1e−14) ,'GARCH',lambda−(1e−14));
EwmaT = garch ('GARCHLags',1,'ARCHLags',1,'Distribution','t','Constant',1e−322,'ARCH',(1−
    lambda)−(1e−14),'GARCH',lambda −(1e−14));

%  Create vectors for GarchN and GarchT VaR estimates, sigma2, Degrees of
%  Freedom.
VaR_EwmaN95 = zeros(length(TestWindow),1);
VaR_EwmaN99 = zeros(length(TestWindow),1);
VaR_EwmaT95 = zeros(length(TestWindow),1);
VaR_EwmaT99 = zeros(length(TestWindow),1);

ES_EwmaN95 = zeros(length(TestWindow),1);
ES_EwmaN99 = zeros(length(TestWindow),1);
ES_EwmaT95 = zeros(length(TestWindow),1);
ES_EwmaT99 = zeros(length(TestWindow),1);
```

```
Sigma2EwmaN = zeros(length(TestWindow),1);
Sigma2EwmaT = zeros(length(TestWindow),1);
DoFs1 = zeros(length(TestWindow),1);

% Estimate coefficients every day and compute GARCHN VaR estimates
for t= TestWindow
    %create a variable that starts from 1
    i = t - TestWindowStart + 1;

% Vector with relevant returns
EstimationWindow = t-EstimationWindowSizeGarch:t;
R_t = Returns(EstimationWindow);
% Estimate model coefficients
estEwmaN = estimate(EwmaN,R_t,'Display','off');
estEwmaT = estimate(EwmaT,R_t, 'Display','off');

% Forecast volatility
Sigma2EwmaN(i)= forecast(estEwmaN,1,'Y0',R_t);
SigmaEwmaN = sqrt(Sigma2EwmaN(i));

Sigma2EwmaT(i) = forecast(estEwmaT,1,'Y0',R_t);
SigmaEwmaT = sqrt(Sigma2EwmaT(i));

%degrees of freedom
DoF = estEwmaT.Distribution.DoF;
DoFs1(i) = DoF;
%VaR estimates
    VaR_EwmaN95(i) = -Zscore(1)*SigmaEwmaN;
    VaR_EwmaN99(i) = -Zscore(2)*SigmaEwmaN;
    VaR_EwmaT95(i) = -Zscore(1)*SigmaEwmaT;
    VaR_EwmaT99(i) = -Zscore(2)*SigmaEwmaT;

%ES estimates
   ES_EwmaN95(i) = SigmaEwmaN*(normpdf(Zscore(1))/q(1));
   ES_EwmaN99(i) = SigmaEwmaN*(normpdf(Zscore(2))/q(2));
   ES_EwmaT95(i) = SigmaEwmaT*(tpdf(tinv(1-q(1),DoF),DoF)/(q(1))*((DoF + tinv(1-q(1),
       DoF)^2)/(DoF-1))*sqrt((DoF-2)/DoF));
```

```
    ES_EwmaT99(i) = SigmaEwmaT*(tpdf(tinv(1-q(2),DoF),DoF)/(q(2))*((DoF + tinv(1-q(2),
        DoF)^2)/(DoF-1))*sqrt((DoF-2)/DoF));

end

%% Compute VaR using the GARCH(1,1).

%Create GarchN and GarchT model
GarchN = garch('GARCHLags', 1, 'ARCHLags', 1, 'Distribution', 'Gaussian' );
GarchT = garch('GARCHLags', 1, 'ARCHLags', 1, 'Distribution', 't' );

%  Create vectors for GarchN and GarchT VaR estimates, sigma2, Degrees of
%  Freedom
VaR_GarchN95 = zeros(length(TestWindow),1);
VaR_GarchN99 = zeros(length(TestWindow),1);
VaR_GarchT95 = zeros(length(TestWindow),1);
VaR_GarchT99 = zeros(length(TestWindow),1);

ES_GarchN95 = zeros(length(TestWindow),1);
ES_GarchN99 = zeros(length(TestWindow),1);
ES_GarchT95 = zeros(length(TestWindow),1);
ES_GarchT99 = zeros(length(TestWindow),1);

Sigma2GarchN = zeros(length(TestWindow),1);
Sigma2GarchT = zeros(length(TestWindow),1);
DoFs2 = zeros(length(TestWindow),1);

% Estimate coefficients every day and compute GARCHN VaR estimates
for t= TestWindow
    %create a variable that starts from 1
    i  = t  - TestWindowStart +1;
% Vector with relevant returns
EstimationWindow = t-EstimationWindowSizeGarch:t;
R_t = Returns(EstimationWindow);
% Estimate model coefficients
estGarchN = estimate(GarchN,R_t,'Display','off');
estGarchT = estimate(GarchT,R_t, 'Display','off');
```

```
% Forecast volatility  and compute standard deviation
Sigma2GarchN(i)= forecast(estGarchN,1,'Y0',R_t);
SigmaGarchN = sqrt(Sigma2GarchN(i));

Sigma2GarchT(i) = forecast(estGarchT,1,'Y0',R_t);
SigmaGarchT = sqrt(Sigma2GarchT(i));

%store degrees of freedom
DoF = estGarchT.Distribution.DoF;
DoFs2(i) = DoF;

%VaR estimates
    VaR_GarchN95(i) = -SigmaGarchN*Zscore(1);
    VaR_GarchN99(i) = -SigmaGarchN*Zscore(2);
    VaR_GarchT95(i) = -SigmaGarchT*(tinv(q(1),DoF)*sqrt((DoF-2)/DoF));
    VaR_GarchT99(i) = -SigmaGarchT*(tinv(q(2),DoF)*sqrt((DoF-2)/DoF));

%ES estimates
   ES_GarchN95(i) = SigmaGarchN*(normpdf(Zscore(1))/q(1));
   ES_GarchN99(i) = SigmaGarchN*(normpdf(Zscore(2))/q(2));
   ES_GarchT95(i) = SigmaGarchT*(tpdf(tinv(1-q(1),DoF),DoF)/(q(1))*((DoF + tinv(1-q(1),
       DoF)^2)/(DoF-1))*sqrt((DoF-2)/DoF));
   ES_GarchT99(i) = SigmaGarchT*(tpdf(tinv(1-q(2),DoF),DoF)/(q(2))*((DoF + tinv(1-q(2),
       DoF)^2)/(DoF-1))*sqrt((DoF-2)/DoF));

end

%% Compute VaR using the GJR(1,1).

%Create GarchN and GarchT model
GjrN = gjr('GARCHLags',1,'ARCHLags',1,'LeverageLags',1,'Distribution', 'Gaussian' );
GjrT = gjr( 'GARCHLags',1,'ARCHLags',1,'LeverageLags',1,'Distribution','t' );

%  Create vectors for GarchN and GarchT VaR estimates, ES estimates, sigma2, Degrees of
%  Freedom
VaR_GjrN95 = zeros(length(TestWindow),1);
VaR_GjrN99 = zeros(length(TestWindow),1);
VaR_GjrT95 = zeros(length(TestWindow),1);
```

```
VaR_GjrT99 = zeros(length(TestWindow),1);

ES_GjrN95 = zeros(length(TestWindow),1);
ES_GjrN99 = zeros(length(TestWindow),1);
ES_GjrT95 = zeros(length(TestWindow),1);
ES_GjrT99 = zeros(length(TestWindow),1);

Sigma2GjrN = zeros(length(TestWindow),1);
Sigma2GjrT = zeros(length(TestWindow),1);
DoFs3 = zeros(length(TestWindow),1);

% Estimate coefficients every day and compute GARCHN VaR estimates
for t= TestWindow
    %create a variable that starts from index no.1
    i = t - TestWindowStart + 1;
% Vector with relevant returns
EstimationWindow = t-EstimationWindowSizeGarch:t;
R_t = Returns(EstimationWindow);
% Estimate model coefficients
estGjrN = estimate(GjrN,R_t,'Display','off');
estGjrT = estimate(GjrT,R_t, 'Display','off' );
% Forecast volatility
Sigma2GjrN(i)= forecast(estGjrN,1,'Y0',R_t);
SigmaGjrN = sqrt(Sigma2GjrN(i));

Sigma2GjrT(i) = forecast(estGjrT,1,'Y0',R_t);
SigmaGjrT = sqrt(Sigma2GjrT(i));

%degrees of freedom
DoF = estGarchT.Distribution.DoF;
DoFs3(i) = DoF;

%VaR estimates
    VaR_GjrN95(i) = -Zscore(1)*SigmaGjrN;
    VaR_GjrN99(i) = -Zscore(2)*SigmaGjrN;
    VaR_GjrT95(i) = -Zscore(1)*SigmaGjrT;
    VaR_GjrT99(i) = -Zscore(2)*SigmaGjrT;
```

```
%ES estimates
   ES_GjrN95(i) = SigmaGjrN*(normpdf(Zscore(1))/q(1));
   ES_GjrN99(i) = SigmaGjrN*(normpdf(Zscore(2))/q(2));
   ES_GjrT95(i) = SigmaGjrT*(tpdf(tinv(1-q(1),DoF),DoF)/(q(1))*((DoF + tinv(1-q(1),DoF)^2)
       /(DoF-1))*sqrt((DoF-2)/DoF));
   ES_GjrT99(i) = SigmaGjrT*(tpdf(tinv(1-q(2),DoF),DoF)/(q(2))*((DoF + tinv(1-q(2),DoF)^2)
       /(DoF-1))*sqrt((DoF-2)/DoF));

end

%% VaR Backtesting
%Set up the returns and dates vector for our the time window to be
%backtested.
ReturnsTest = Returns(TestWindow);
DatesTest  = DateReturns(TestWindow);

%Name VaR data, models
VaRData = [VaR_Historical95 VaR_EwmaN95 VaR_EwmaT95 VaR_GarchN95 VaR_GarchT95
     VaR_GjrN95 VaR_GjrT95 VaR_Historical99 VaR_EwmaN99 VaR_EwmaT99
     VaR_GarchN99 VaR_GarchT99 VaR_GjrN99 ...
    VaR_GjrT99];
Models = {'HS95','EwmaN95','EwmaT95','GarchN95','GarchT95', 'GjrN95', 'GjrT95' ...
     'HS99','EwmaN99','EwmaT99', 'GarchN99','GarchT99', 'GjrN99', 'GjrT99'};
VaRLevel = [0.95 0.95 0.95 0.95 0.95 0.95 0.95 0.99 0.99 0.99 0.99 0.99 0.99 0.99];

% Run the Unconditional Coverage backtest on our methods.

%Set up the backtesting results
vbt = varbacktest(ReturnsTest,VaRData,'PortfolioID',PortfolioID,'VaRID',Models,'VaRLevel',
     VaRLevel);
%Report the Unconditional coverage backtesting result
pof(vbt)

%% ES Backtesting

% Run the Unconditional backtest on our methods.

rng('default'); % for reproducibility; the esbacktestbysim constructor runs a simulation
```

```
%Name VaR data, ES data, and degrees of freedom.

ESData = [ES_Historical95 ES_EwmaN95 ES_EwmaT95 ES_GarchN95 ES_GarchT95
    ES_GjrN95 ES_GjrT95 ES_Historical99 ES_EwmaN99 ES_EwmaT99 ES_GarchN99
    ES_GarchT99 ES_GjrN99 ...
   ES_GjrT99];

DoFs = [round(mean(DoFs1)), round(mean(DoFs2)), round(mean(DoFs3)),round(mean(
    DoFs1)), round(mean(DoFs2)), round(mean(DoFs3))];


%set up the ES backtesting results
VaRDataT = [VaR_EwmaT95 VaR_GarchT95 VaR_GjrT95 VaR_EwmaT99 VaR_GarchT99
    VaR_GjrT99];
ESDataT = [ES_EwmaT95 ES_GarchT95 ES_GjrT95 ES_EwmaT99 ES_GarchT99 ES_GjrT99
    ];
ModelsT = {'EwmaT95', 'GarchT95', 'GjrT95', 'EwmaT99', 'GarchT99', 'GjrT99'};
VaRLevelT = [0.95 0.95 0.95 0.99 0.99 0.99];

VaRDataN = [VaR_EwmaN95 VaR_GarchN95 VaR_GjrN95 VaR_EwmaN99 VaR_GarchN99
    VaR_GjrN99];
ESDataN = [ES_EwmaN95 ES_GarchN95 ES_GjrN95 ES_EwmaN99 ES_GarchN99
    ES_GjrN99];
ModelsN = {'EwmaN95', 'GarchN95', 'GjrN95', 'EwmaN99', 'GarchN99', 'GjrN99'};
VaRLevelN = [0.95 0.95 0.95 0.99 0.99 0.99];

VaRDataHS = [VaR_Historical95 VaR_Historical99];
ESDataHS = [ES_Historical95 ES_Historical99];
ModelsHS = {'HS95','HS99'};
VaRLevelHS = [0.95 0.99];

ebtsT = esbacktestbysim(ReturnsTest,VaRDataT,ESDataT,"t",...
        'DegreesOfFreedom',15,...
        'Location' ,0,...
        'Scale' ,1,...
        'PortfolioID',PortfolioID ,...
        'VaRID',ModelsT,...
```

```
        'VaRLevel',VaRLevelT);

ebtsN = esbacktestbysim(ReturnsTest,VaRDataN,ESDataN,"normal",...
        'Mean',0,...
        'StandardDeviation',1,...
        'PortfolioID',PortfolioID ,...
        'VaRID',ModelsN,...
        'VaRLevel',VaRLevelN);

ebtsHS = esbacktest(ReturnsTest,VaRDataHS,ESDataHS,'PortfolioID',PortfolioID,'VaRID',
     ModelsHS,'VaRLevel',VaRLevelHS);
   %Report the Unconditional test

unconditional(ebtsT)
unconditional(ebtsN)
unconditionalNormal(ebtsHS)

%% Compute Sample Statistics for the Returns
Mean = mean(Returns);
Standard_Deviation = std(Returns);
Kurtosis = kurtosis(Returns);
Skewness = skewness(Returns);
Min = min(Returns);
Max = max(Returns);

SampleStatistics = table(Mean,Standard_Deviation ,Kurtosis,Skewness,Min,Max)
end
```

BIBLIOGRAPHY

[1] *Risk Management and Financial Institutions*, John C.Hull, 2012.

[2] P. E. ALEXANDER J. MCNEIL, RUDIGER FREY, *Quantitative Risk Management*, Princeton University Press, 2005.

[3] B. S. CARLO ACERBI, *Backtesting Expected Shortfall*, MSCI Inc., 2014.

[4] P. F. CHRISTOFFERSEN, *Elements of Financial Risk Management*, 2012.

[5] J. DANIELSSON AND C. G. DE VRIES, *Value-at-risk and extreme returns*, first version from 1997, working paper, Financial Markets Group, London School of Economics, London., (2000).

[6] J. DANIELSSON AND C. G. DE VRIES, *Value-at-risk and extreme returns*, Annales d'Economie et de Statistique, (2000), pp. 239–270.

[7] K. DOWD, *Beyond value at risk: the new science of risk management*, John Wiley Sons, 1998.

[8] B. FOR INTERNALTIONAL SETTLEMENTS, *Minimum capital requirements for market risk*.

[9] ——, *Revised framework for market risk capital requirements issued by the basel committee*.

[10] J. HULL, *Options, futures and other derivatives*, Prentice Hall, 2000.

[11] P. JORION, *Value at Risk: The New Benchmark for Managing Financial Risk*, McGraw-Hill, 2007.

[12] D. KIDD, *Value at risk and condition value at risk: A comparison*, Investment Risk and Performance Articles 2012.1, (2012).

[13] T. J. LINSMEIER AND N. D. PEARSON, *Value at risk*, Financial Analysts Journal vol. 56, no. 2, (2000).

[14] S. MANGANELLI AND R. ENGLE, *Value at risk models in finance*, European Central Bank - Working Papers, (2001).

[15] C. MARSHALL AND M. SIEGEL, *Value-at-risk: Implementing a risk measurement standard*, (1996).

[16] MATHWORKS®, *Overview of expected shortfall*.

[17] M. N. S. P. T. P. MCKINSEYCOMPANY, AMIT MEHTA, *Managing market risk: Today and tomorrow*, McKinsey Working Papers on Risk, Number 32, (2012).

[18] S. MITRA AND T. JI, *Risk measures in quantitative finance*, International Journal of Business Continuity and Risk Management, 1 (2010), pp. 120–135.

[19] M. MUSIELA AND M. RUTKOWSKI., *Martingale Methods in Financial Modelling*, Springer, 2005.

[20] X. X. RUI LUO, WEINAN ZHANG AND J. WANG, *A neural stochastic volatility model*, arXiv:1712.00504v1 30 Nov 2017, (2017).

[21] H. X. TZE LEUNG LAI, *Statistical Method for Financial Markets*, 2007.