The impact of estimation frequency on Value at Risk (VaR) and Expected Shortfall (ES) forecasts:

An empirical study on conditional extreme value models

Thesis presented for the degree of Masters of Mathematical Statistics in the Department of Statistics.

University of Cape Town

2019

Author - Alice Elizabeth Coyne

Supervisor – Allan Clark

Acknowledgements

To my mother, Janella Beatrix Coyne, thank you for teaching me to be a strong independent woman. May you rest in peace – you will live on in my heart forever.

To my father, Rodney Cecil Coyne, thank you for your patience and unconditional love.

To my sister, Myra Jane Coyne, thank you for being a shoulder to lean on through both the good and bad time.

To my aunts, uncles, cousins, and the rest of my family, thank you for all of your support.

To Kelly-Ann De Villiers and the rest of my Chistlehurst family, you have helped shape me into the person I am today.

Lastly, to two of my closest friends, Barbara Jean Still and Zosia Kocznur you are both like family to me, thank you for all of the special memories.

Table of Contents

Acknowledgements **2**

Table of Contents **4**

List of Figures **8**

List of Tables **11**

Abstract **14**

Glossary of Terms **15**

Aims and Rational of Dissertation **16**

Structure of Dissertation **18**

Chapter 1 – Introduction **20**

1.1 The Basel Accord **23**

Chapter 2 – The Data **26**

2.1 Statistical Computing **29**

2.2 Summary of the data **30**

2.2.1 Standards and Poor's Index **32**

2.2.2 Johannesburg Stock Exchange **35**

2.3 Emerging Markets – A South African Perspective **38**

Chapter 3 – Generalised Autoregressive Conditional Heteroskedasticity (GARCH) **41**

Chapter 4 – Extreme Value Theory **44**

4.3 Extreme Value Theory Defined **46**

4.3.1 Fisher-Tippett-Gnedenko Theorem 47

4.4 Generalized Extreme Value Distribution 50

4.4.1 Block Maxima Method 53

4.4.2 Maximum Likelihood Estimation (for GEV): 55

4.5 Generalized Pareto Distribution 58

4.5.1 Peaks Over Threshold (for GPD): 60

4.5.2 Maximum Likelihood Estimation (for GPD): 62

Chapter 5 – Risk Measures 63

5.1 Value at Risk 64

5.2 Expected Shortfall 67

5.2.1 Expected Shortfall (Generalized Extreme Value Distribution) 68

5.2.2 Expected Shortfall (Generalized Pareto Distribution) 69

Chapter 6 – Backtesting 70

6.1 – Value at Risk backtesting 71

6.1.1 Kupiec Test (Violation based test) 71

6.1.2 Christoffersen's Test (Independence based test) 72

6.2 Expected Shortfall Backtesting 74

Chapter 7 – Methodology and Analysis 75

7.1 Generalized Autoregressive Conditional Heteroscedasticity Model 75

7.1.1 GARCH Model Parameter Estimates 80

7.1.2 Ljung-Box Test & ARCH Lagrange multiplier 81

7.1.2 Auto-Correlation Function 84

7.1.2 Jarque-Bera Test 86

7.1.3 Q-Q Plots 88

7.1.4 Weakness of GARCH modelling 89

7.2 Generalized Extreme Value Distribution 90

7.3 Generalized Pareto Distribution 98

7.4 Backtesting Results 105

7.4.1 Generalized Extreme Value Distribution Backtesting 105

7.4.2 Generalized Pareto Distribution Backtesting 108

Chapter 8 – Impact of Estimation Frequency 110

8.1 GARCH Model Results 112

8.2 Generalized Extreme Value Distribution 120

8.2.1 Generalized Extreme Value Distribution Backtesting Results 126

8.3 Generalized Pareto Distribution 130

8.3.1 Generalized Pareto Distribution Backtesting Results 136

Chapter 9 – Discussion and Conclusion 140

References 141

Appendix 1 150

Extreme Value Theorem (Fisher-Tippett-Gnedenko theorem) 150

Proof: 150

Jarque-Bera Test 155

Ljung-Box Test 156

ARCH Langrange Multiplier Test 157

Akaike Information Criterion 158

Bayesian Information Criterion 159

Expected Shortfall (GARCH (1,1)) 160

Properties of Generalized Extreme Value Distribution 161

Properties of Generalized Pareto Distribution **162**

Update frequency colour key **163**

List of Figures

Figure 1: Structure of Dissertation 19

Figure 2: Basel III Framework 24

Figure 3: Closing price for S&P 500 33

Figure 4: Daily returns for S&P 500. 33

Figure 5: Histogram of S&P 500 34

Figure 6: Closing price for JSE All Share Index 36

Figure 7: Daily Returns for JSE All Share Index 36

Figure 8: Histogram for JSE All Share Index 37

Figure 9: Extreme Value Distributions 49

Figure 10: Visual representation of Block Maxima Method 54

Figure 11: Visual representation of Peaks over Threshold 61

Figure 12: Standardized Residuals for S&P 500 76

Figure 13: Plot of daily returns of S&P 500 76

Figure 14: S&P 500 residuals extracted from GARCH model 77

Figure 15: Standardized Residuals for JSE All Share Index 77

Figure 16: Plot of daily returns of JSE All Share Index 78

Figure 17: JSE All Share Index residuals extracted from GARCH model 78

Figure 18: ACF of standardized residuals for S&P 500 84

Figure 19: ACF of standardized residuals for JSE All Share Index 85

Figure 20: ACF of squared standardized residuals for S&P 500 85

Figure 21: ACF of squared standardized residuals for JSE All Share Index 86

Figure 22: Q-Q Plot for S&P 500 88

Figure 23: Q-Q Plot for JSE All Share Index 89

Figure 24: GEV diagnostic plots for S&P 500 data 91

Figure 25: GEV diagnostic plots for JSE All Share Index 91

Figure 26: Value at Risk for S&P 500 (using Block Maxima Method) 95

Figure 27: Value at Risk for JSE All Share Index (using Block Maxima Method) 95

Figure 28: Expected Shortfall for S&P 500 (using Block Maxima Method) 96

Figure 29: Expected Shortfall for JSE All Share Index (using Block Maxima Method) 96

Figure 30: GPD diagnostic plots for S&P 500 data 98

Figure 31: GPD diagnostic plots for JSE All Share Index data 99

Figure 32: Value at Risk for S&P 500 (using Peaks over Threshold Method) 102

Figure 33: Value at Risk for JSE All Share Index (using Peaks over Threshold Method) ... 102

Figure 34: Expected Shortfall for S&P 500 (using Peaks over Threshold Method) 103

Figure 35: Expected Shortfall for JSE All Share Index (using Peaks over Threshold Method) 103

Figure 36: GEV Value at Risk estimations at 99th quantile using S&P 500 data 123

Figure 37: GEV Value at Risk estimations at 99th quantile using JSE All Share Index data 123

Figure 38: GEV Expected Shortfall estimations at 99th quantile using S&P 500 data 124

Figure 39: GEV Expected Shortfall estimations at 99th quantile using JSE All Share Index data 124

Figure 40: GPD Value at Risk estimations at 99th quantile using the S&P 500 data 133

Figure 41: GPD Value at Risk estimations at 99th quantile using the JSE All Share Index data 133

Figure 42: GPD Expected Shortfall estimations at 99th quantile using the S&P 500 data.. 134

Figure 43: GPD Expected Shortfall estimations at 99th quantile using the JSE All Share Index data 134

List of Tables

Table 1: Glossary of terms 15

Table 2: Summary statistics 31

Table 3: Properties of Generalized Extreme Value Distribution 51

Table 4: Properties of Generalized Pareto Distribution 59

Table 5: Error Analysis of GARCH model for S&P 500 80

Table 6: Summary of GARCH model for JSE All Share Index 80

Table 7: Standardised Residuals Tests for S&P 500 82

Table 8: Standardised Residuals Tests for JSE All Share Index 82

Table 9: GARCH Model Information Criterion Statistics 83

Table 10: Jarque-Bera Test 87

Table 11: Parameter estimates for Generalized Extreme Value Distribution using Maximum Likelihood parameter estimation 91

Table 12: Value at Risk and Expected Shortfall summary for GEV ($p = 0.99$) 93

Table 13: Value at Risk and Expected Shortfall summary for GEV ($p = 0.975$) 93

Table 14: Value at Risk and Expected Shortfall summary for GEV ($p = 0.95$) 94

Table 15: Parameter estimates for Generalized Pareto Distribution using Maximum Likelihood parameter estimation 99

Table 16: Value at Risk and Expected Shortfall summary for GPD ($p = 0.99$) 100

Table 17: Value at Risk and Expected Shortfall summary for GPD ($p = 0.975$) 101

Table 18: Value at Risk and Expected Shortfall summary for GPD ($p = 0.95$) 101

Table 19: Value at Risk results for Generalized Extreme Value Distribution using S&P 500 data. 105

Table 20: Value at Risk results for Generalized Extreme Value Distribution using JSE All Share Index data. 106

Table 21: Results for Value at Risk of Generalized Pareto Distribution using S&P 500 data. 108

Table 22: Results for Value at Risk of Generalized Pareto Distribution using JSE All Share Index data. 108

Table 23: Error Analysis of GARCH model for S&P 500 113

Table 24: Error Analysis of GARCH model for the JSE All Share Index 114

Table 25: Standardised Residuals Tests for the S&P 500 115

Table 26: Standardised Residuals Tests for the JSE All Share Index 117

Table 27: GARCH Model Information Criterion Statistics 118

Table 28: Jarque-Bera Test 119

Table 29: Parameter estimates for Generalized Extreme Value Distribution using Maximum Likelihood parameter estimation 120

Table 30: Value at Risk and Expected Shortfall summary for GEV (p = 0.99) 121

Table 31: Value at Risk and Expected Shortfall summary for GEV (p = 0.975) 122

Table 32: Value at Risk and Expected Shortfall summary for GEV (p = 0.95) 122

Table 33: Results for Value at Risk of Generalized Extreme Value Distribution using S&P 500 data. 126

Table 34: Results for Value at Risk of Generalized Extreme Value Distribution using JSE All Share Index data. 128

Table 35: Parameter estimates for Generalized Pareto Distribution using Maximum Likelihood parameter estimation 130

Table 36: Value at Risk and Expected Shortfall summary for GPD (p = 0.99) 131

Table 37: Value at Risk and Expected Shortfall summary for GPD (p = 0.975) 132

Table 38: Value at Risk and Expected Shortfall summary for GPD (p = 0.95) 132

Table 39: Results for Value at Risk of Generalized Pareto Distribution using the S&P 500 data. 136

Table 40: Results for Value at Risk of Generalized Pareto Distribution using JSE All Share Index data. 138

Abstract

This study investigates extreme market events which occur in the tails of a distribution. The extreme events occur with a very low probability, but with significant consequences, which is what makes them of interest. In this study 20 years of data from both the S&P 500 and the JSE All Share index have been used. An extreme value approach has been taken to quantify the risks associated with extreme market events. To achieve this a two phased process is used to calculated the Value at Risk and Expected Shortfall. The first phase involved running the daily returns through the GARCH model, and then extracting the residuals. The second phase involves using the Block Maxima Method, or Peaks over Threshold method to fit the residuals to the Generalized Extreme Value Distribution or the Generalized Pareto Distribution. Finally, the impact of estimation frequency is considered for each of the models. In conclusion, taking an extreme value approach to provide a statistically sound method to calculate risk, even when the parameters of the model are updated less frequently, this is preferable to simpler models where the parameter estimates are updated daily.

Glossary of Terms

Term	Definition
ACF	Auto-Correlation Function
AIC	Akaike Information Criterion
AR	Autoregressive
ARMA	Autoregressive Moving Average
BIC	Bayesian information criterion
BM (M)	Block Maxima (Method)
CI	Confidence Interval
CVaR	Conditional Value at Risk
ES	Expected Shortfall
EVT	Extreme Value Theory
FTSE	Financial Times Stock Exchange
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GEV	Generalized Extreme Value
GPD	Generalized Pareto Distribution
IDE	Integrated Development Environment
IID	Independent and Identically Distributed
JSE	Johannesburg Stock Exchange
LR	Likelihood Ratio
MLE	Maximum Likelihood Estimator
POT	Peaks Over Threshold
Q-Q plot	Quantile-Quantile plot
QMLE	Quasi-Maximum Likelihood Estimation
S&P 500	Standard & Poor's 500 Index
VaR	Value at Risk

Table 1: Glossary of terms

Aims and Rational of Dissertation

The aim of this dissertation is to investigate the effects of estimation frequency on the performance of extreme value GARCH models, by focusing on tail estimation of financial returns. The impact of estimation frequency will be measured by using both Value at Risk and Expected Shortfall to predict potential market losses. Currently most models of financial returns are updated on a daily basis.

This study looks at the effects of updating the GARCH models parameters less frequently. Initially, the model parameters were updated every day, then every 5 days, 10 days, and 20 days. This is used to measure the extent to which the update frequency effects the performance of forecasting Value at Risk and Expected Shortfall.

Previous research by McNeil & Frey (2000) has shown that taking more the advanced approach of combining the pseudo-maximum-likelihood fitting of the GARCH model with Extreme Value Theory to calculate both Value at Risk and Expected Shortfall statistics performs significantly better than simpler models such as the GARCH model on its own. Ardia & Hoogerheide (2014) found that updating the parameter estimates daily of the GARCH equation only marginally improved the performance of the model when comparing it to a model which is updated less frequently.

Previous studies have largely looked at the S&P 500 (Ardia & Hoogerheide, 2014). The aim of this research is to extend this to more volatile markets, such as the JSE All share index. This will allow a comparison between how more mature markets perform compared to less mature and more volatile markets. Secondly, an extreme value methodology will be used, where the residuals are extracted from the GARCH model and fitted to either the Generalized Extreme Value Distribution and the Generalized Pareto Distribution (McNeil & Frey, 2000).

In summary this master's thesis aims to determine if the use a wider data window to estimate the Value at Risk and Expected Shortfall in the current South African markets, provide an accurate estimate of the potential market losses. If it is successful, how frequently does the model need to be updated in order to maintain accurate results? Can taking an Extreme Value approach to calculating the Value at

Risk and Expected Shortfall be used as an accurate risk measure in a South African context, if a wider time frame and data window are used? Following this, which distribution best fits the tails of the financial returns?

Structure of Dissertation

Chapter 1 discusses the background of financial markets, the legislation, and current influences on financial returns. Previous research of the GARCH model and Extreme Value Theory is reviewed, and briefly summarized.

Chapter 2 is an in depth analysis of the data. Particularly focusing on the properties of financial time series data. This helps to motivate for the use of an extreme value approach. There is also a brief section on emerging markets, since the JSE All Share Index is a South African index.

Chapter 3 looks at the theory behind the GARCH model. In this chapter the positives and negatives of the GARCH model are discussed, and the use of the GARCH model is explained.

Chapter 4 is a detailed look at Extreme Value Theory. Some of the early extremist are studied, as well as the theory behind both the Generalized Extreme Value Distribution and the Generalized Pareto Distribution.

Chapter 5 discusses the risk measures used in this thesis. The theory behind both Value at Risk and Expected Shortfall is explained.

Chapter 6 explains the backtesting process which is used to determine how well the models performed.

Chapter 7 contains the results from the GARCH model are explored, and then fed into either the Generalized Extreme Value Distribution or the Generalized Pareto Distribution. From here the Value at Risk and Expected Shortfall estimates are calculated. Finally, the results are backtested to determine which model fits the data best.

Chapter 8 repeats the previous chapters, but changes the update frequency of the parameters of the models. In this chapter the models are updated daily, weekly, fortnightly, and monthly.

Chapter 9 summarizes the previous chapters, and suggests room for further research.

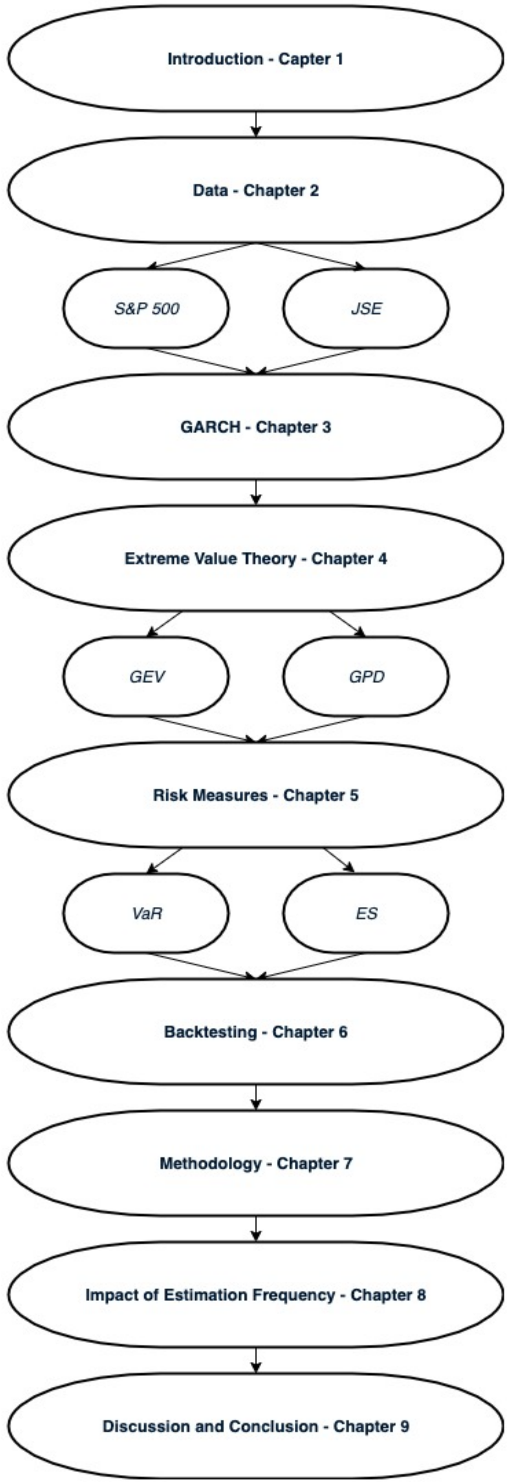

Figure 1: Structure of Dissertation

Chapter 1 – Introduction

Extreme events are events which have a low probability of occurring, but generally have major consequences when they do occur. These extreme events are often difficult to predict, such as the terrorist attack on the World Trade Center in 2001 which killed almost 3 000 people. Extreme events can also be found in nature such as the 2004 Indian Ocean tsunami which killed over 200 000 people and caused an estimated $15 billion in damages. Financial markets are also prone to extreme events such as the Wall Street Crash in 1929.

Extreme events in the financial markets are of particular interest to financial risk managers, who are constantly monitoring changes in the market. They are especially interested in larger losses, since large gains are not seen as a risk, and build models to quantify market risks. Value at Risk and Expected Shortfall are commonly used by financial risk managers to measure risk. Taking an Extreme Value approach to calculating Value at Risk and Expected Shortfall often leads to better predictions (McNeil & Frey, 2000).

2016 was a year of extreme change. Both BREXIT and the US election have had large impacts on two of the world's largest economies (Writer, 2017). These effects have rippled down to smaller economies (Mc Grattan & Waddle, 2017), such as South Africa. Due to South Africa's turbulent past, the economy is already fairly weak and unstable (Mc Grattan & Waddle, 2017).

This means that, when measuring financial risk, advanced measures are required to quantify the risk. The study of volatility plays a crucial part in financial risk management (Ardia, 2008). Because of this, advanced techniques are required to predict the risk associated with a particular financial time series or index.

Generalized Autoregressive Conditional Heteroskedasticity (GARCH) is one of the commonly used techniques to estimate the volatility in financial markets (Bollerslev, 1986). GARCH models are particularly useful for financial models where the volatility changes over time (Chan, Deng, Peng, & Xia, 2007), such as that of the South African markets.

In a study by Ardia & Hoogerheide (2014), where they looked at the performance of the S&P 500 over a period of 12 years, they found that even when the parameters of the GARCH model was updated less frequently had a marginal effect on the Value at Risk and Expected Shortfall forecasts. Updating the GARCH equation daily, only marginally improved the performance of the model when compared to weekly, monthly, or quarterly updates (Ardia & Hoogerheide, 2014).

Value at Risk is used to quantify a financial risk over a specific time frame (Ardia & Hoogerheide, 2014). It helps to determine the amount of potential loss, and the probability associated with the potential loss faced by banks and investment companies (Ardia & Hoogerheide, 2014). This allows one to quantify the risk of an investment. Therefore, Value at Risk represents the risk from market movement as one number: the maximum loss expected on an investment, over a given time period at a specific level of confidence (Ardia & Hoogerheide, 2014).

Therefore, increasing the frequency at which the GARCH models parameters are updated only marginally improves the model's performance (Ardia & Hoogerheide, 2014). Models that require daily updates becomes computationally heavy. Updating the parameters of the model less frequently allows for better predictions of the daily returns, but with significantly fewer computations (Ardia & Hoogerheide, 2014).

According to a study by Van De Venter (2000), when calculating the Value at Risk statistic in a volatile market, such as that in South Africa, risk is often overestimated in periods of high volatility and underestimated during periods of low volatility. This particular study looked at Value at Risk in South Africa from July 1993 until July 1995. Being the end of Apartheid, this would have been a particularly volatile period in the South African economy. In this particular study, a 100-day and 250-day data window, whereas the study by Ardia (2014) used a much wider estimation window of 1000-day.

Using a normal distribution is a convenient and simple to use and considers the data in the entire distribution, but it does not capture the kurtosis seen in most financial returns (Milwidsky & Mate, 2010). Taking an extreme value approach, just focuses to the tails of the distribution. Two different distributions will be considered. These include the Generalized Extreme Value distribution, and Generalized Pareto distribution. For comparison, two different data sets will be used, which are the Standard & Poor 500 (S&P 500), and the Johannesburg Stock Exchange (JSE) All Share index.

All of the above-mentioned distributions are fat tailed distributions which implies that, unlike the standard normal distribution, they cannot be described with only its first two moments.

Financial data is generally fat tailed, which means the normal distribution doesn't adequately capture the probability of extreme events which occur in the tails of the distribution (Schmitt, Chetalova, Schafer, & Guhr, 2013). Generally, the tails are associated with events that have a low frequency, but a high severity (McNeil & Frey, 2000). This is why it is so important to consider the 4^{th} moments, when looking at fat tailed distributions.

Extreme Value Theory focuses on the extreme values which occur in the tails of a distribution (McNeil & Frey, 2000). It is a tool used to predict the probability of extreme or rare events (Chinhamu, Huang, Huang, & Chikobvu, 2015). It explains the behaviour of maxima and minima of random variables (Chinhamu, Huang, Huang, & Chikobvu, 2015). These events occur in the tails of the distribution (Chinhamu, Huang, Huang, & Chikobvu, 2015).

There are two main approaches to Extreme Value Theory. The first approach is known as the Block Maxima Method. In this approach, the variable considered is divided into *n* equal blocks (Coles, 2004). From each block, the maximum value is selected and used to fit a Generalized Extreme Value Distribution to the data (Coles, 2004). Therefore, the local block maxima from each of the n-blocks is used to fit the Generalized Extreme Value Distribution (Coles, 2004).

The second approach is known as a threshold method (Coles, 2004). In this approach, a threshold is selected, and the Generalized Pareto Distribution is fitted using all the values above the threshold (Coles, 2004). Everything above the threshold represents an extreme loss or extreme gain (Coles, 2004). This is known as the Peaks over threshold approach (Coles, 2004). For the purposes of this study, only extreme losses will be studied.

1.1 The Basel Accord

The Basel Committee on Banking Supervision has issued three Basel Accords, Basel I, Basel II, and most recently Basel III (Moody's Analystics, 2013). The Basel Accord currently consists of recommendations for regulation on the banking system. The most recent Basel Accord, Basel III, was developed in response to the shortfalls in financial regulation during the 2007 – 2008 financial crisis. It aimed to strengthen the capital requirements of banks, by decreasing bank leverage and increasing the banks liquidity requirement (Moody's Analystics, 2013).

Basel Accord

The third Basel Accord consists of the following three pillars:

- Pillar 1: Enhanced Minimum Capital & Liquidity Requirements
- Pillar 2: Enhanced Supervisory Review Process for Firm-wide Risk Management and Capital Planning
- Pillar 3: Enhanced Risk Disclosure & Market Discipline

(Moody's Analystics, 2013)

The most noticeable change from Basel II to Basel III occurred in Pillar 1, with the enhanced minimum capital and liquidity requirements (Moody's Analystics, 2013). The Tier 1 Capital Ratio has increased from 4% to 6%, and the Common Equity Requirement has increased from 2% to 4.5% (Moody's Analystics, 2013).

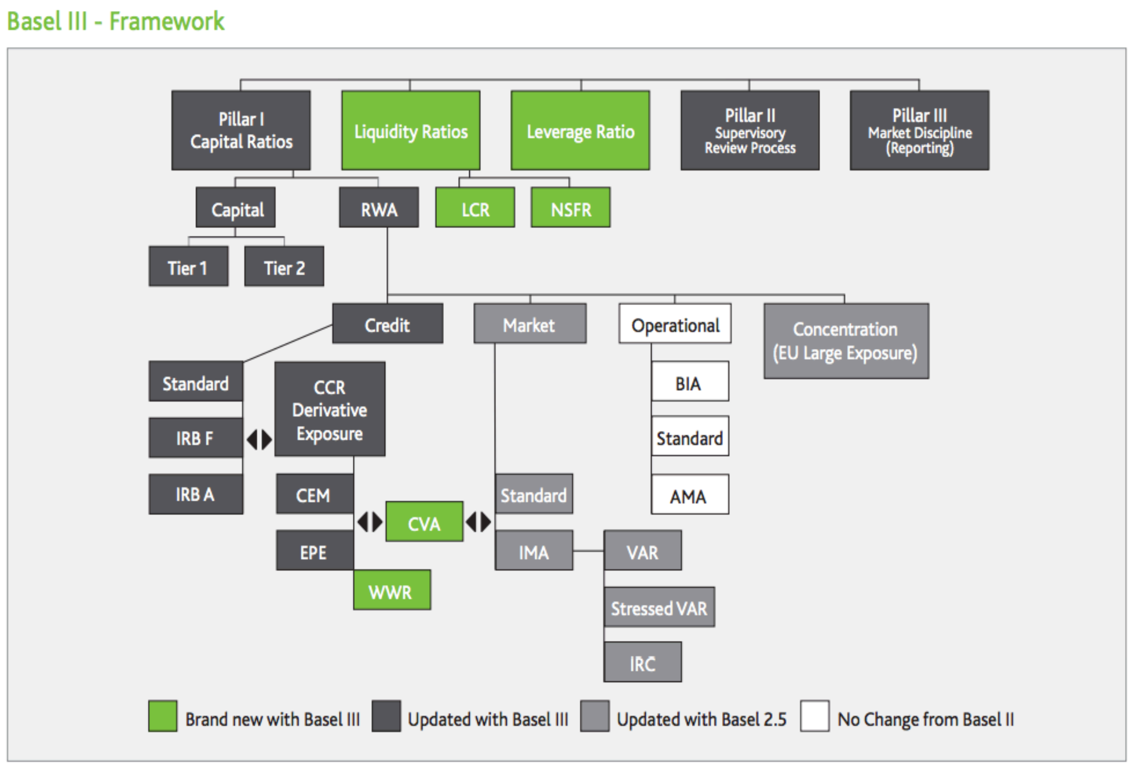

Figure 2: Basel III Framework

(Moody's Analystics, 2013, p. 3)

Financial markets are constantly changing, and these changes can be both positive gains or negative losses. In general, large losses have more of an impact than large gains, which makes them more interesting to study. Risk cannot be avoided, instead it needs to be managed and understood.

The third Basel Accord is a voluntary framework which regulates banking capital requirements and manages risk (Moody's Analystics, 2013). It is scheduled to be implemented on 31 March 2019. It was developed after the 2007 – 2008 financial crash, to respond to the failures of the current financial regulation (Chang, Jimenez-Martin, Maasoumi, Mc Aleer, & Perez-Amaral, 2015).

Market risk is the risk that a financial institution will make a loss due to adverse market conditions (McNeil & Frey, 2000). The new quantile requirements of Basel II and Basel III indicate financial institutions are required to use more advanced methods to understand market risk (Moody's Analystics, 2013). Extreme Value Theory is a useful tool to better understand the tails of financial distributions (Miller, 2012).

Banks are required to hold a certain amount of capital to cushion against market movements (McNeil & Frey, 2000). Value at Risk is the most widely used risk measure in the banking industry (Hu & Kercheval, 2007), and is used to allocate capital to protect against exposer to market risk (Ardia, 2008).

Under the Basel II, regulations recommend Value at Risk to be calculated at a 99% confidence level, with a two week (or 10 working day) holding period (Ardia, 2008). In practice, many consider this to be too stringent, and calculate Value at Risk at 95% only looking one day ahead (Ardia, 2008).

Chapter 2 – The Data

The data sets analysed in this study are the daily returns for both the S&P 500 and the JSE All Share index. Twenty years of data points have been used from 25 March 1997 until 24 March 2017, this giving a total of 5 220 data points. When the daily closing price is converted to the daily return one data point is lost, therefore 5 219 daily returns are used in this study.

The data was obtained from the University of Cape Town's library using the Bloomsberg system. The log of the daily returns was calculated, and then saved in a CSV file, which was then loaded into R. The daily returns are calculated as follows:

Definition

$$r_i = \ln\left(\frac{x_i}{x_{i-1}}\right) , \qquad i = 1, \dots , N$$

Where x_i represents the current day's closing price, and x_{i-1} represents the previous day's closing price.

The S&P 500 data set was selected as a baseline, since the majority of studies referenced in this dissertation look at the application of Extreme Value Theory to the S&P 500. Some of these studies include Bystrom (2004), McNeil & Frey (2000), and Ardia & Hoogerheide (2014). The JSE All Share index was selected as a comparison since there are fewer studies looking at Extreme Value Theory in emerging markets or an African perspective.

Both the S&P 500 or the JSE All share index are examples of time series data. In general, a financial time series of stock prices, is considered to be non-stationary (Schmitt, Chetalova, Schafer, & Guhr, 2013). This means that the statistical proprieties of the random process, such as the mean and variance, change over time (Schmitt, Chetalova, Schafer, & Guhr, 2013).

Properties of a financial time series

Modelling financial time series data is a complex problem, and generally have the following properties:

1. *Non-stationarity of returns.* The returns generally oscillate around zero. These oscillations vary in magnitude, especially in volatile markets.
2. *Absence of auto-correlation for price variations.* Thus, making it close to white noise.
3. *Auto- correlation of squared price returns, or absolute error.*
4. *Volatility clustering.* High volatility periods are normally followed by low volatility periods. This results in clusters of large absolute returns.
5. *Fat tailed distributions.* Financial time series data seldom follow a Gaussian distribution.
6. *Leverage effect.* Negative returns generally have a larger effect on the volatility than positive returns.
7. *Seasonality.* Different holidays, seasons, and day of the week all have significant effects on the returns.

(Francq & Zakoian, 2010)

This is especially common in financial data, which goes through business cycles, changes in inflation, impacts from natural disasters, or advances in technology. Given the non-stationary nature of financial time series data, naïve analysis of the data can produce misleading results.

The stochastic nature of financial returns results in clustered periods of high volatility and extreme values contrasted with periods of calm (Chinhamu, Huang, Huang, & Chikobvu, 2015). The study of volatility plays a crucial part in financial risk management (Ardia, 2008). Extreme Value Theory focuses on the tails of the distribution, where extreme events occur, and has been shown to produce more accurate results than traditional methods (Bystrom, 2004).

2.1 Statistical Computing

The data was analysed using the statistical program, R, and R Studio was used as the integrated development environment (IDE). Both R and R Studio are open source software and can be downloaded on https://cran.r-project.org/bin/windows/base/ and https://www.rstudio.com/products/rstudio/download/ respectively.

It is a widely used statistical programming language and is designed to perform statistical analysis (R Core Team, 2017). The version of R used is 3.4.1 and was released on the 30 June 2017 (R Core Team, 2017). The following packages were extensively used for analysis:

- fGarch (Wuertz, et al., 2017)
- evir (Pfaff & Mc Neil, 2012)
- VaRES (Nadarajah, Chan, & Afuecheta, 2013)
- Dowd (Dowd, 2005)

2.2 Summary of the data

As can be seen from the summary statistics below, both the S&P 500 and the JSE All Share Index have a mean and median very close to zero, and a standard deviation of approximately 0.5%. The S&P 500 has a minimum daily return of -0.047 and a maximum daily return of 0.041. The JSE All Share Index has a minimum daily return of -0.032 and a maximum daily return of 0.055.

Both the S&P 500 and the JSE All Share Index have a positively skewed distribution, although the JSE All share index is slightly more skewed at 0.451, compared to the S&P 500 with a skewness of 0.231. The kurtosis of the distributions are also noticeable larger than that of a normal distribution, which has a kurtosis of three. The S&P 500 has a much sharper point than the JSE All Share Index, each with a kurtosis of 8.079 and 6.101 respectively.

Summary Statistics

Data	**S&P 500**	**FTSE**
Minimum	-0.047	-0.032
Maximum	0.041	0.055
Skewness	0.231	0.451
Kurtosis	8.079	6.101
Mean	< -0.001	< -0.001
Median	< -0.001	< -0.001
Standard Deviation	0.005	0.005
1% Quantile	-0.014	-0.014
2.5% Quantile	-0.011	-0.011
5% Quantile	-0.008	-0.008
95% Quantile	0.007	0.008
97.5% Quantile	0.01	0.01
99% Quantile	0.014	0.013
Sample Size	N = 5219	N = 5219

Table 2: Summary statistics

2.2.1 Standards and Poor's Index

The top 5 companies on the S&P 500 are all well-known brands, and include: Apple Inc., Alphabet, Microsoft Corp., Amazon.com Inc., and Facebook Inc. In this study, 20 years of data was used taking the daily closing price of the index and calculating the daily return. The daily closing price of the S&P 500 is illustrated in Figure 3.

The Figure 4 show the daily returns of the S&P 500, which are used to build the Extreme Value model. One of the features of the daily returns is the volatility clustering, which can clearly be seen in the graphs above. The S&P 500 experiences high volatility around the 2007 – 2008 financial crash. During this period, the daily returns vary rather drastically. This period of extreme volatility can be seen in the red block in the image below.

After the global recession, the S&P 500 appears to recover well, and has continued to increase in value overtime. The histogram in Figure 5 illustrates that the daily returns of the S&P 500 do not follow a normal distribution. It can clearly be seen that the S&P 500 daily returns are leptokurtic. The grey bars represent the daily returns of the S&P 500, and the red curve represents a normal distribution.

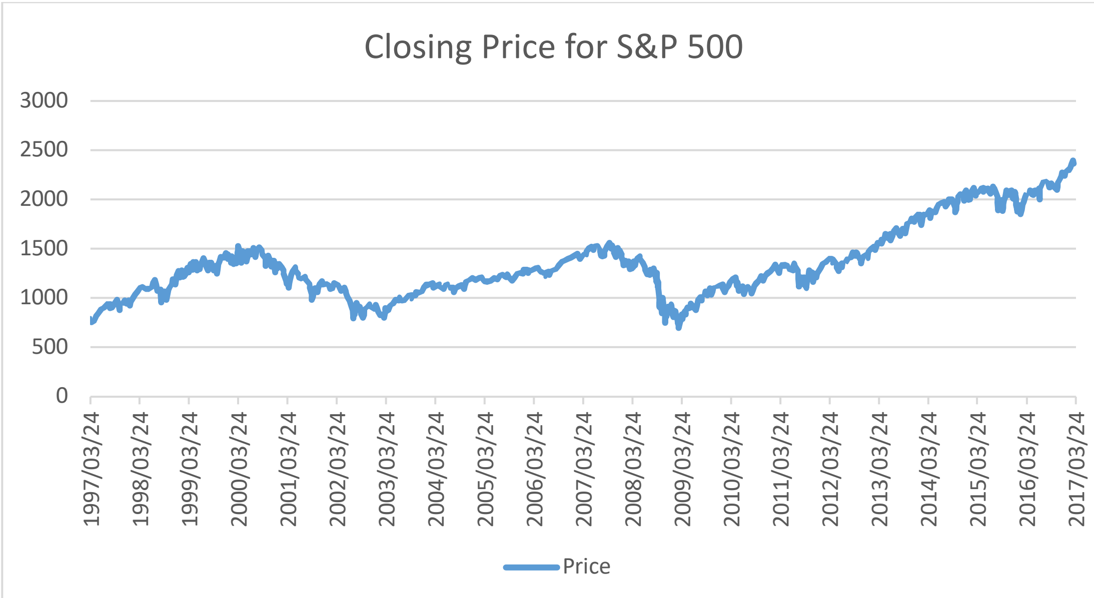

Figure 3: Closing price for S&P 500

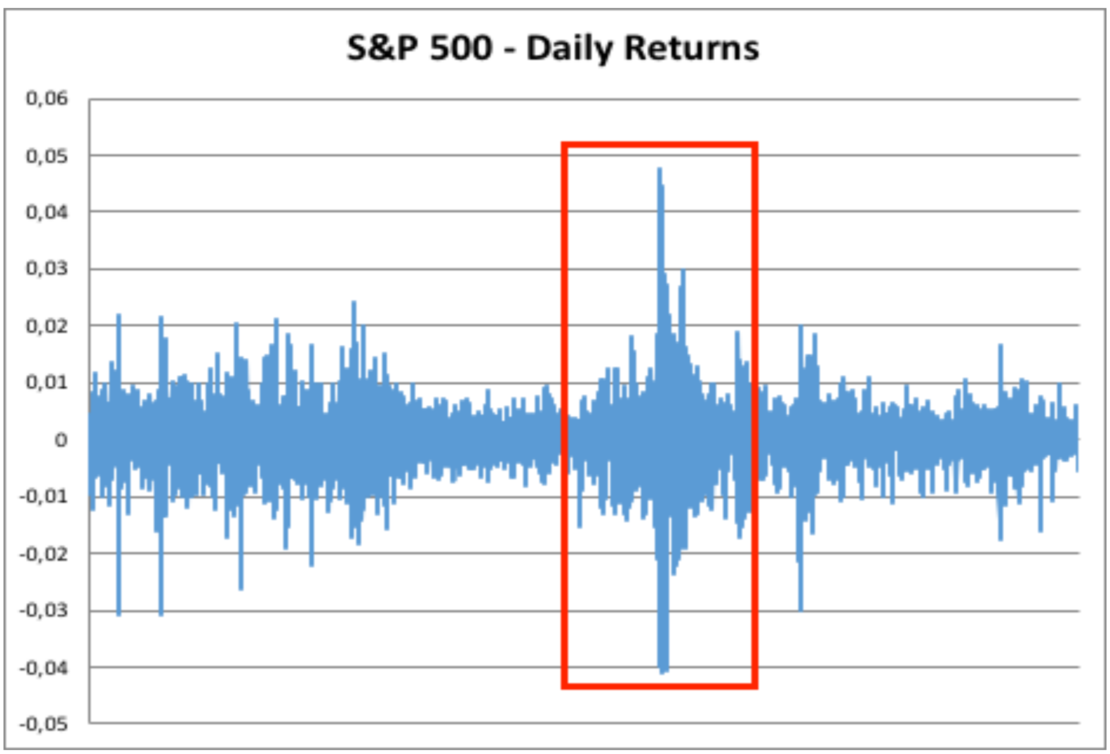

Figure 4: Daily returns for S&P 500.

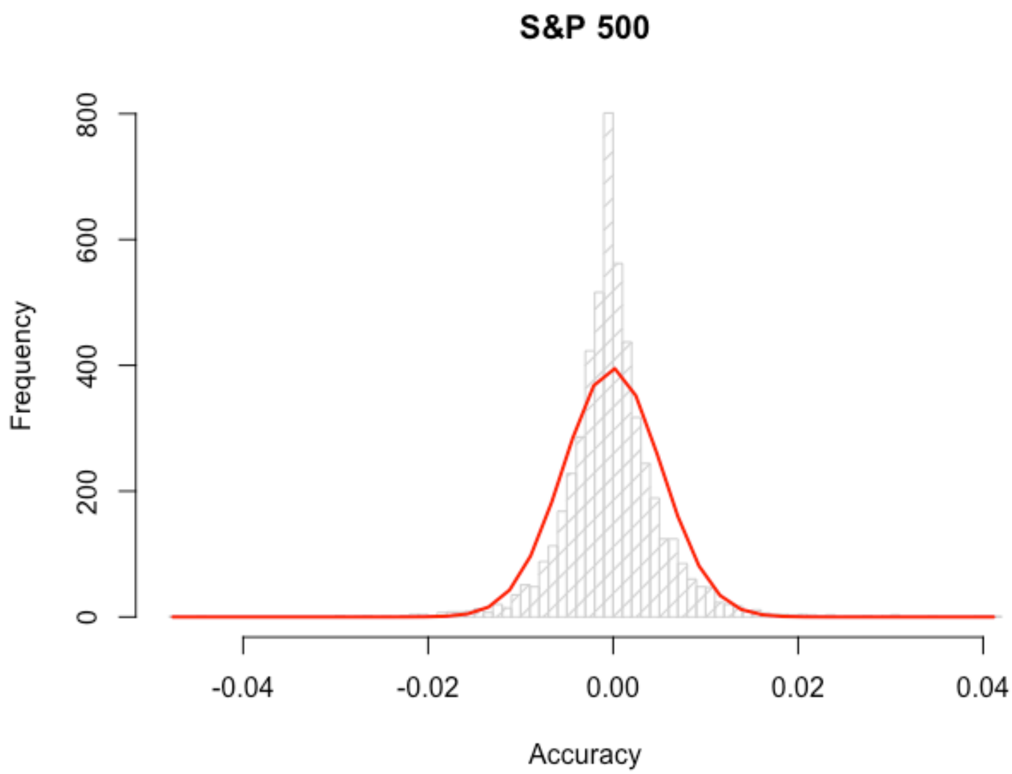

Figure 5: Histogram of S&P 500

2.2.2 Johannesburg Stock Exchange

The Financial Time Stock Exchange (FTSE) /Johannesburg Stock Exchange (JSE) All Share Index is designed to represent the performance of South African companies. It represents 99% of the capital value in South Africa. The index was first launched in June 1995 as the JSE, but in June 2002 it became known as the FTSE.

The Figure 6 and Figure 7 show the daily closing price and the daily returns of the JSE All Share Index respectively. As seen with the S&P 500, the daily returns of the JSE All Share Index has volatility clustering which can clearly be seen in the graphs below.

The JSE All Share index experienced the highest period of volatility in 1997, which was the end of Apartheid. These periods of high volatility clustering are followed by clusters of low volatility. High volatility around the 2007 – 2008 financial crash, can be seen in the red block in the image above. After the global recession, the JSE All Share Index appears to recover well, and has continued to increase in value overtime.

The histogram shown below illustrates that the daily returns of the JSE All Share Index do not follow a normal distribution. It can clearly be seen that the daily returns are leptokurtic. The grey bars represent the daily returns of the JSE All Share Index, and the red curve represents a normal distribution.

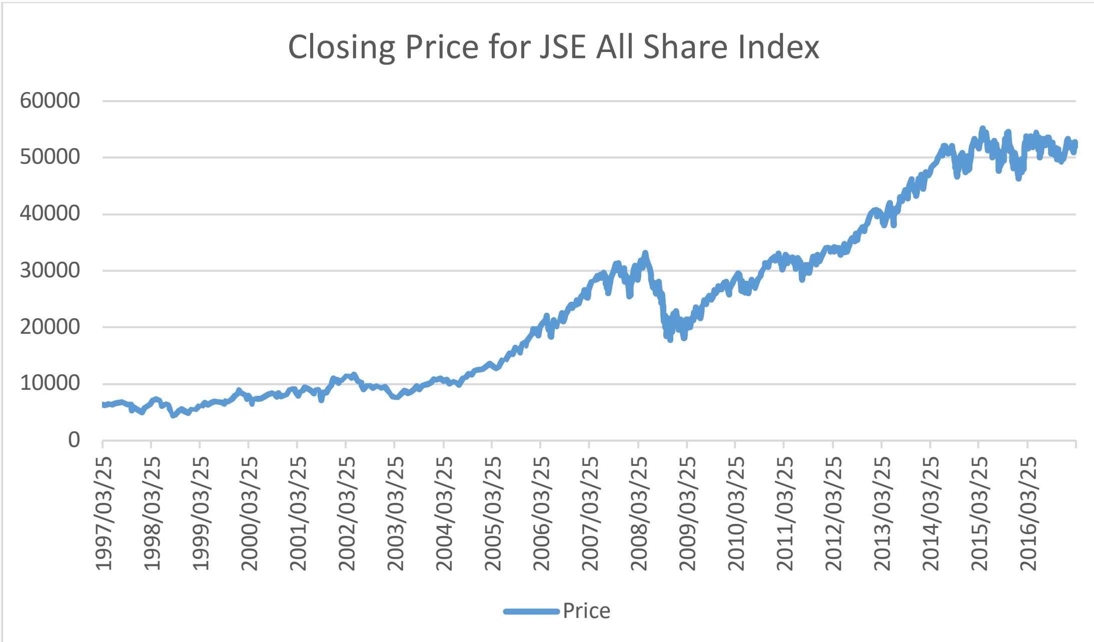

Figure 6: Closing price for JSE All Share Index

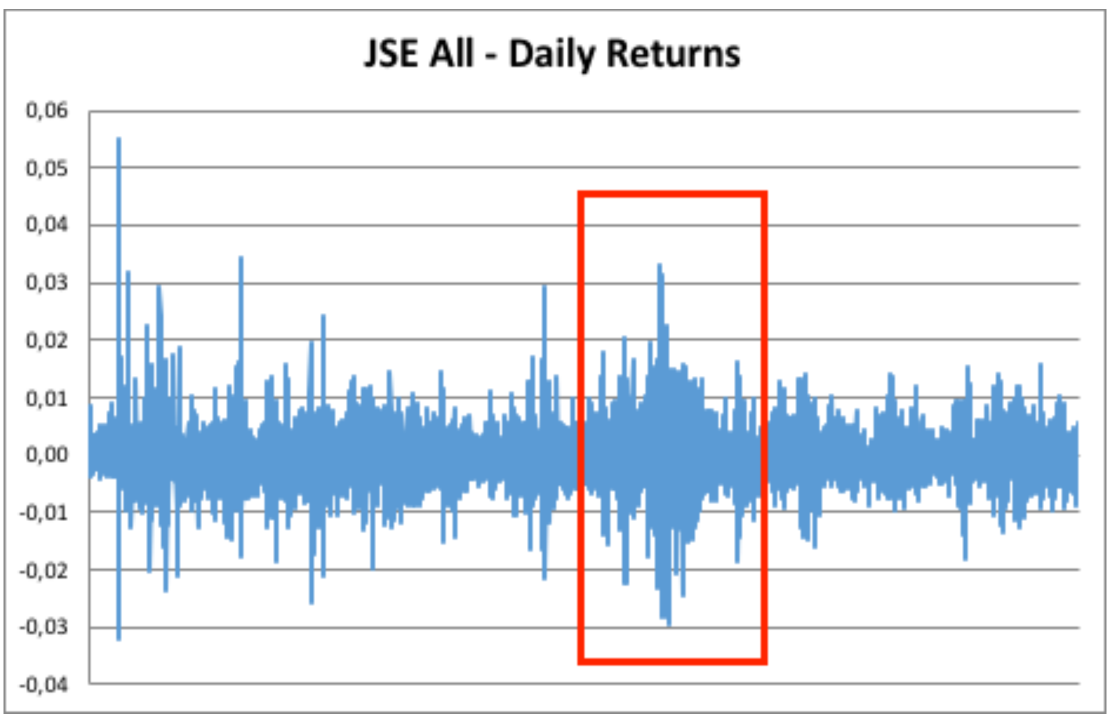

Figure 7: Daily Returns for JSE All Share Index

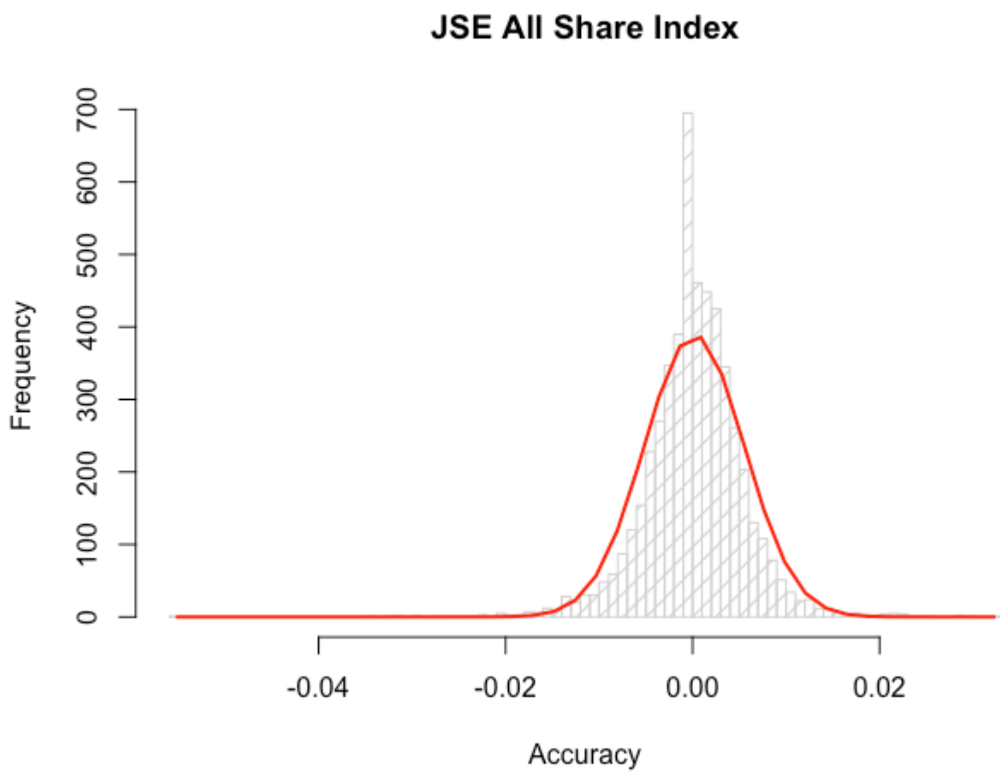

Figure 8: Histogram for JSE All Share Index

2.3 Emerging Markets – A South African Perspective

From 1948 until the early 1990s, South Africa was divided by racial segregation, known as Apartheid (Davenport, 1991). The word ‘Apartheid’ is an Afrikaans word which means ‘separateness’. It was adopted as the slogan for the Afrikaner National Party, commonly known at the National Party, for the 1948 elections (Davenport, 1991).

The National Party was founded in 1914 and promoted the interests of Afrikaans speakers in South Africa, to protect South Africa against the influences of the United Kingdom. In 1924 the National Party, led by James Barry Munnik Hertzog, won the national election and became the governing party in South Africa (Davenport, 1991).

One of the focuses of the National Party was to promote social and economic welfare for working class white South Africans (Feinstein, 2005). They introduced minimum wage and pensions for white unskilled workers in the form of the Wage Act (1925) and The Old Age Pensions Act (1917) respectively. The agricultural sector was also given various forms of assistance to help boost the economy, most notably, white farmers were given preferential interest rates from the Land Bank.

All of this helped improve the living conditions of white South Africans and allowed many unskilled white South Africans to increase their social economic position (Feinstein, 2005).

The Population Registration Act (1950), was used to divide all South Africans into one of four racial groups, either ‘Black’, ‘White’, ‘Coloured’, or ‘Indian’ (Davenport, 1991). This lead to further segregation in one of the largest mass evictions in modern history (Davenport, 1991). From 1960 - 1963 approximately 3,5 million non-white South Africans were forced to leave their homes and move into segregated neighbourhoods (Feinstein, 2005).

Apartheid caused significate opposition both from within South Africa, and internationally (Feinstein, 2005). The United Nations condemned South Africa and implemented economic trade sanctions (Feinstein, 2005). From the early 1960s, became internationally criticised.

On 7 August 1963, the United Nations passed Resolution 181, which requested all states to voluntarily stop selling fire arms to South Africa. On 4 November 1977, the United Nations issued Resolution 418, which was a mandatory arms prohibition (Feinstein, 2005). By the late 1980s, 25 different nations had passed laws containing various trade sanctions on South Africa.

There was a mass movement of disinvestment from South Africa amongst several countries and cities. This severely impacted the South African economy (Feinstein, 2005). This ultimately lead to the demise of Apartheid (Feinstein, 2005).

This lack of trade negatively affected the South African economy (Feinstein, 2005). This led to South Africa experiencing capital flight, which is when a country's assets or money rapidly leave the country. This led to a decline of the rand on the international exchange rate, which had a ripple effect causing inflation to increase by 12 - 15% per year. In 1990, president Frederick Willem de Klerk acknowledged the economic instability in South Africa, and began to reverse Apartheid.

Finally, on 17 June 1991, the Apartheid legislation was repealed (Hamann, Khagram, & Rohan, 2008). This left the South African economy fairly weak and unstable. In the first democratic election in South Africa in 1994, the newly elected African National Congress inherited an economy weakened by long years of internal conflict and external sanctions (Hamann, Khagram, & Rohan, 2008).

Currently, South Africa has the second largest economy in Africa, preceded by Nigeria. According to the World Bank, South Africa is currently classified as an upper-middle-income economy. The economy of South Africa has increased since 1996, when the international trade sanctions were lifted from South Africa, and the Gross Domestic Product increased to a peak of $400 billion in 2011.

In the late 2000s when South Africa was affected by the global recession and struggled to recover significantly more than other emerging markets (Hamann, Khagram, & Rohan, 2008). Currently, private investments and export volumes have not yet fully recovered. From 2000 - 2009 the South African Gross Domestic Product only grew 2.2% per year, compared to the rest of the world which grew by 3.1% per year during the same period.

Adding to all of this, high levels of unemployment have led to an increase in crime, which has slowed investment and growth in South Africa (Hamann, Khagram, & Rohan, 2008). Currently, the South African economy is considered to be so weak, it has been graded as sub-investment standard by Fitch Ratings.

Chapter 3 – Generalised Autoregressive Conditional Heteroskedasticity (GARCH)

Time series analysis aims to build a model that captures the underlying stochastic process of financial data (Francq & Zakoian, 2010). The study of volatility plays a crucial part in financial risk management (Ardia, 2008). Volatility in financial returns is due to the inconsistent variance of the data, where the variance constantly changes with time (Miller, 2012). This implies that risk changes over time, and therefore more advance methods are needed to analyse financial returns (Miller, 2012).

The Autoregressive Conditional Heteroscedasticity (ARCH) first introduced in 1982 by Engle was used to deal with the conditional variance of time series data (Engle R. F., 1982). Later, in 1986 Bollerslev generalized the ARCH model (Bollerslev, 1986).

The GARCH (1,1) is the most commonly used model when it comes to predicting volatility according to (Yang, Chaptpatanasiri, & Sattayatham, 2016). For the purposes of comparability, the GARCH (1,1) model is used to make the data more independent and identically distributed (IID), as done in the studies by McNeil & Frey (2000) and Ardia & Hoogerheide (2013). A sequence of random variables is said to be independent and identically distributed if each random variable has the same probability distribution as the others, and all of the random variables are mutually independent. Mathematically this can be defined as follows:

Definition

a. Let the random variables be defined to have the values in $\mathbb{I} \subseteq \mathbb{R}$. Two random variables X and Y are said to be identically distributed if and only if

$$P[x \geq X] = P[x \geq Y], \qquad \forall\, x \in \mathbb{I}.$$

b. Two random variables X and Y are said to be independent if and only if

$$P[y \geq Y] = P[y \geq Y \mid x \geq X] \; and \; P[x \geq X] = P[x \geq X \mid y \geq Y] \; \forall\, x, y \in \mathbb{I}.$$

Auto Regressive Moving Average (ARMA) models are commonly used when the time series is a second order stationary process (Francq & Zakoian, 2010). The Generalized Autoregressive Conditionally Heteroscedastic (GARCH) model is used to capture the conditional volatility of the daily returns (Engle R. F., 1982). This ensures the time varying nature of the mean and variance are accounted for.

Definition

The GARCH (1,1) model has the general form:

$$R_t = \mu_t + \sigma_t Z_t$$

Where R_t is the daily returns at time t, μ_t is the mean function of the series, ${\sigma_t}^2$ is the conditional variance, and Z_t is the residuals at time t.

GARCH models are particularly useful for analysing and forecasting volatility (Charpentier, 2015). Using a combination of both ARMA + GARCH leads to a more accurate representation of the temporal dependencies (Charpentier, 2015).

By definition, a process (ϵ_t) is called a GARCH (1,1) process if its first two conditional moments exist, and satisfies the following two conditions (Francq & Zakoian, 2010):

$$E(\epsilon_t \,|\epsilon_u\,, u < t) = 0, \qquad t \in \mathbb{Z}.$$

There exist constants ω, α and β, such that:

$$\sigma_t^2 = Var(\epsilon_t \mid \epsilon_u, u < t) = \omega + \alpha\, \epsilon_{t-1}^2 + \beta\, \sigma_{t-1}^2, \quad t \in \mathbb{Z}\,.$$

Where $\epsilon_t = \sigma_t Z_t$, and ω, α, β, are all non-negative.

The mean equation of a univariate time series can be described as follows:

$$x_t = E(x_t \mid \Omega_{t-1}) + \epsilon_t,$$

where $E(x_t \mid \Omega_{t-1})$ represents the conditional expected value given Ω_{t-1}, and ϵ_t represents the residuals of the time series (Charpentier, 2015). ϵ_t explains the uncorrelated disruptions, with a mean of zero, and is associated with the unpredictable part of the financial time series (Charpentier, 2015).

There are several variations of the GARCH model, but the GARCH (1,1) model is most commonly used both in academia and practice (Ardia, 2008). This study only focuses on the GARCH (1,1) model. When comparing the performance of the GARCH (1,1) model, it had better prediction than other models such as the Student-t, even when the latter parameters are updated more frequently (Ardia, 2013). This is especially noticeable when looking at Expected Shortfall (Ardia & Hoogerheide, 2014).

Chapter 4 – Extreme Value Theory

Extreme market risk is defined as the risk due to extreme changes in prices (Ruppert, 2004). These risks have a small probability of occurrence, but with drastic consequences. Financial returns are very unpredictable; this results in clustered periods of high volatility and extreme values (Chinhamu, Huang, Huang, & Chikobvu, 2015). Extreme Value Theory focuses on the tails of the distribution where these extreme events occur, and has been shown to produce more accurate results than traditional methods (Bystrom, 2004).

The returns of financial indexes generally have fatter tailed distributions, in other words, they do not follow a normal distribution (Schmitt, Chetalova, Schafer, & Guhr, 2013). Heavy tailed distributions are those whose tails are not exponentially bounded (Peters & Shevchenko, 2015). These heavy tails make it more difficult to predict extreme events which occur in the tails of the distribution (Bystrom, 2004).

This is of particular importance, since events in the tails result in more extreme losses, especially in the left tail of the distribution, since this results in significant losses, whereas events in the right tail results in significant gains. They therefore rely on models that focus on the tails of the distributions (Bystrom, 2004).

Extreme Value Theory is used to deal with low probability events that occur in the tails of the distribution (Bystrom, 2004). It is therefore very important to use models that capture the behaviour of these extreme events in the tails (Bystrom, 2004). It is of particular interest the events that occur in the left tail, as this is associated with large losses and large losses are of more interest than large gains.

Small probabilities in the far ends of the tails are of particular importance, since as the probability of an event becomes smaller, the consequences of that event become significantly larger (Bystrom, 2004). It is therefore vital to be able to model these extreme events as accurately as possible.

One of the main difficulties in statistics, is coming up with estimate for the tails of the given distribution (McNeil & Frey, 2000). Extreme events have a low probability of occurring, but high consequences when they do occur. The main interest is in the tails of the distribution, since this is where the extreme events occur (McNeil & Frey, 2000). What makes them particularly difficult to predict, is the lack of data since these extreme events don't occur very frequently (Ardia & Hoogerheide, 2014).

Most financial indices are not normally distributed (Schmitt, Chetalova, Schafer, & Guhr, 2013). In general, their tails are not exponentially bounded, and therefore have thicker tails than the normal distribution (Schmitt, Chetalova, Schafer, & Guhr, 2013). Extreme Value Theory based approaches are useful for estimating the tails of a distribution because they are based on sound statistical theory, and they offer a parametric form for the tails of a distribution (McNeil & Frey, 2000).

4.3 Extreme Value Theory Defined

Assume $X_1, \dots, X_n$ is a sequence of independent and identically distributed random variables with a common distribution function F. The maximum value of these random variables over time is denoted by $M_n = \max(X_1, \dots, X_n)$. The distribution of the maximum M_n can be derived as follows:

Definition

$$
\begin{aligned}
P(M_n \leq z) &= P(X_1 \leq z, \dots, X_n \leq z) \\
&= P(X_1 \leq z) \dots P(X_n \leq z) \\
&= \left(F(z)\right)^n
\end{aligned}
$$

Since the distribution F is unknown, it can be approximated by looking at extreme values in the tails of the distribution of X. It is important to capture the behaviour in the tails of the distribution, since extremes can be found in both the upper and lower ends of the distribution (Coles, 2004).

The indicator function is a finite sequence defined as, $I_n = I(M_n > z)$, which is a Bernoulli process with probability of success: $p(z) = 1 - \left(F(z)\right)^n$. This is determined by the impact of z on the extreme events (Coles, 2004). In practice, we might not have the distribution function F but the Fisher–Tippett–Gnedenko theorem provides an asymptotic approximation of the distribution (Coles, 2004).

4.3.1 Fisher-Tippett-Gnedenko Theorem

Maurice René Fréchet was a French mathematician, who lived from 2 September 1878 to 4 June 1973 (Bru & Hertz, 2001). His major field of study was topology, but also contributed significantly towards statistics, probability, and calculus (Bru & Hertz, 2001). During his 60-year career, he wrote over 300 publications, contributing to both pure and applied mathematics (Bru & Hertz, 2001).

He was one of the early pioneers of the study of extreme events, in his work on the asymptotic maximum stable distributions for large values which can be seen in his 1927 publication (Fréchet, 1927). This became the start of the study of extreme values in statistics (Bru & Hertz, 2001). Late in 1933, Fréchet started working with Emil Julius Gumbel, since they both shared an interest in extreme events (Bru & Hertz, 2001).

Emil Julius Gumbel was born in Munich, Germany on 18 July 1891 (Hertz, 2001). In 1934 Gumbel became an assistant at the Institut de Science Financiere et d' Assurances in Lyon, France (Hertz, 2001). When the war broke out, Gumbel was forced into exile in the United States of America, where he remained until his death on 10 September 1966. While living in New York, he summarised all of his contributions in a book called 'Statistics of Extremes' which was published in 1958 (Hertz, 2001).

The last distribution which falls under the Generalized Extreme Value distribution was developed by a Swedish engineer Waloddi Weibull (Broberg, 1997). Weibull was known for having a wide variety of scientific interests, which led him to statistics and probability (Broberg, 1997).

He became interested in probability theory and statistics, and in 1939 published a paper on the Weibull distribution (Broberg, 1997). Unfortunately, this paper did not receive much recognition since World War II broke out. In 1951, after the war, he published another paper where he included his results from the 1939 paper, which has become his most cited paper (Broberg, 1997). He continued to publish papers until his death on 12 October 1979 (Broberg, 1997).

The above three distributions are grouped together in a family of continuous probability distributions known as Generalized Extreme Value distribution. This is proved by the Fisher-Tippett-Gnedenko theorem, named after three notable mathematicians.

Theorem

Let $X_1, \ldots, X_n$ be a sequence of independent and identically distributed random variables, and $M_n = \max\{X_1, \ldots, X_n\}$. If there exist sequences of constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that:

$$P\left\{\frac{(M_n - b_n)}{a_n} \leq z\right\} \rightarrow G(z) \text{ as } n \rightarrow \infty \text{ then } G(z) \propto exp\left[-(1 + \xi z)^{-\frac{1}{\xi}}\right]$$

where $z = \frac{x - \mu}{\sigma}$, and ξ depends on the tail shape of the distribution.

When normalized, G belongs to either the Weibull, Gumbel, or Fréchet distribution.

Weibull distribution:

$$We(z) = \begin{cases} 1 - \left(exp\left(-\left(\frac{z - b}{a}\right)\right)^{\alpha}\right) & , z < b \\ 1 & , z \geq b \end{cases}$$

when the distribution of M_n has a light tail with finite upper bound (Coles, 2004).

Gumbel distribution:

$$Gu(z) = exp\left(-exp\left(-\left(\frac{z - b}{a}\right)\right)\right) \text{ for } z \in \mathbb{R},$$

when the distribution of M_n has a tail, which is bounded exponential (Coles, 2004).

Fréchet distribution:

$$Fr(z) = \begin{cases} 0 & , z \le b \\ exp\left(-\left(\frac{z-b}{a}\right)^{-\alpha}\right) & , z > b \end{cases}$$

when the distribution of M_n has a heavy tail which decay's exponentially (Coles, 2004).

In all cases, $\alpha > 0$.

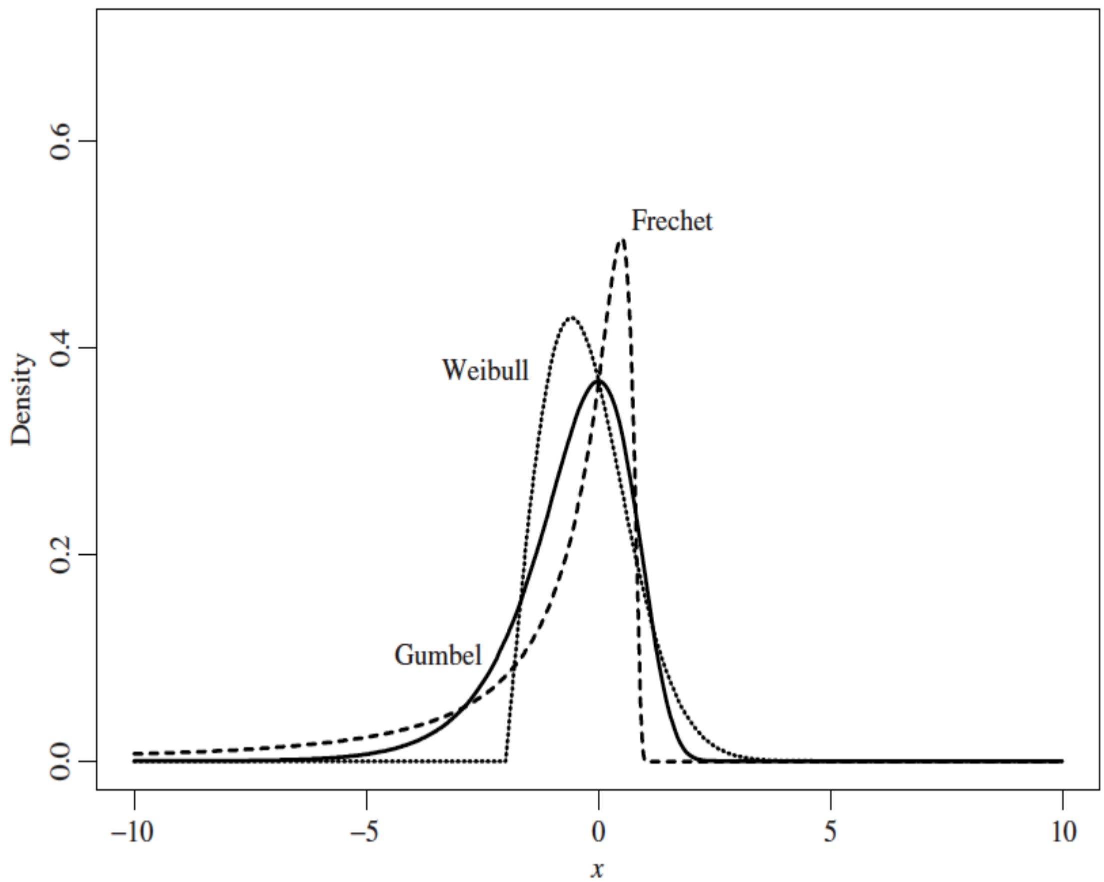

Figure 9: Extreme Value Distributions

(Tsay, Analysis of Financial Time Series, 2002, p. 272)

4.4 Generalized Extreme Value Distribution

This generalization, is known as the generalized extreme value (GEV) distribution. Each of the above distributions give a different representation of Extreme Value Theory (Coles, 2004). They can be combined to form one single family of models, which has the following distribution:

$$F(z,\mu,\sigma,\xi) = \begin{cases} exp\left(-\left(1+\xi\left(\frac{z-\mu}{\sigma}\right)^{-1/\xi}\right)\right), & \xi \neq 0 \\ exp\left(-\,exp\left(-\frac{z-\mu}{\sigma}\right)\right), & \xi = 0 \end{cases}$$

This Generalized Extreme Value (GEV) distribution has the following properties:

Parameters	$\mu \in \mathbb{R}$ - Location parameter $\sigma \in \mathbb{R}$ - Scale parameter $\xi \in \mathbb{R}$ - Shape parameter
Cumulative Distribution Function	$F(z,\mu,\sigma,\xi) = \begin{cases} exp\left(-\left(1+\xi\left(\frac{z-\mu}{\sigma}\right)^{-1/\xi}\right)\right), & \xi \neq 0 \\ exp\left(-exp\left(-\frac{z-\mu}{\sigma}\right)\right), & \xi = 0 \end{cases}$
Probability Density Function	$f(z,\mu,\sigma,\xi) = \frac{1}{\sigma}\begin{cases} \left(1+\xi\left(\frac{z-\mu}{\sigma}\right)\right)^{-1/\xi-1} exp\left(-\left(1+\xi\left(\frac{z-\mu}{\sigma}\right)\right)^{-1/\xi}\right) \\ exp\left(-\frac{z-\mu}{\sigma}\right) exp\left(-exp\left(-\frac{z-\mu}{\sigma}\right)\right) \end{cases}$
Mean	$\begin{cases} \mu - \frac{\sigma}{\xi} + \frac{\sigma\Gamma(1-\xi)}{\xi}, & 0 < \xi < 1 \\ \mu + \sigma\gamma, & \xi = 0 \\ \infty, & \xi \geq 1 \end{cases}$

Median	$\begin{cases}\mu + \sigma \dfrac{\ln(2)^{-\xi} - 1}{\xi}, & \xi \neq 0 \\ \mu - \sigma \ln(\ln(2)), & \xi = 0\end{cases}$
Mode	$\begin{cases}\mu + \sigma \dfrac{(1+\xi)^{-\xi}}{\xi}, & \xi \neq 0 \\ \mu, & \xi = 0\end{cases}$
Variance	$\begin{cases}\dfrac{\sigma^2}{\xi^2}(\Gamma(1-\xi) - \Gamma^2(1-\xi)^2), & \xi \neq 0 \\ \sigma^2 \dfrac{\pi}{6}, & \xi = 0\end{cases}$
Support	$1 + \dfrac{\xi(x-\mu)}{\sigma} > 0, \quad \forall\ \mu, \xi \in \mathbb{R}, \quad \sigma > 0$

Table 3: Properties of Generalized Extreme Value Distribution

The quantiles associated with the upper tails of the GEV distribution can be obtained by inverting the equation:

$F(z,\mu,\sigma,\xi) = \begin{cases} exp\left(-(1+\xi\left(\frac{z-\mu}{\sigma}\right)^{-1/\xi}\right), & \xi \neq 0 \\ exp\left(-\ exp\left(-\ \frac{z-\mu}{\sigma}\right)\right), & \xi = 0\end{cases}$ and solving for $F(z_p) = 1 - p$.

Solving this equation, it can easily be shown that

$$z_p = \begin{cases}\mu - \frac{\sigma}{\xi}\left[1 - \{-\ln(1-p)\}^{-\xi}\right], & for\ \xi \neq 0 \\ \mu - \sigma\ln\{-ln(1-p)\}, & for\ \xi = 0\end{cases}.$$

These quantiles z_p, represent the return level for the corresponding return period $\frac{1}{p}$ and is associated with the upper tail of the distribution (Coles, 2004). For example, when considering daily returns, the return level z_p will be exceeded every $\frac{1}{p}$ days.

4.4.1 Block Maxima Method

The Block Maxima Method is used to divide the daily return series into blocks of equal length n, as represented in Figure 10 (Coles, 2004). This results in a series of consecutive blocks, from which the maximum return is selected from each block and used to fit the Generalized Extreme Value distribution (Ardia & Hoogerheide, 2014). From the random sample of data $\begin{pmatrix} y_{11} & \cdots & y_{1n} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{nN} \end{pmatrix}$, these data points are divided into a series of consecutive blocks of length n. From each block, the maximum value is selected generating the following series:

Definition

$$x_1, x_2, \dots, x_N \text{ where } x_i = \max\left(y_{n(i-1)+1}, y_{n(i-1)+2}, \dots, y_{ni}\right)$$
$$\text{for } i = 1,2, \dots, N.$$

Selecting the correct block size is vitally important (Coles, 2004) since the block size needs to be large enough that the asymptotic result of the Fisher-Tippet theorem still holds (Coles, 2004). If the blocks are too small, this results in estimation bias (Coles, 2004). Alternatively, if the blocks are too large, there are too few blocks which results in a large estimation variance (Coles, 2004).

Once the data has been divided into blocks, there are various methods can be applied to calculate the parameters $(\mu,\ \sigma,\ \xi)$ of the Generalized Extreme Value Distribution (Coles, 2004). One such method is known as the Maximum Likelihood Estimation method (MLE) which is used to estimate the parameters of the statistical model given the observations.

The block maxima method involves a careful balance between bias and variance when selecting block size (Coles, 2004). This is one of the major downfalls of the block maxima method (Coles, 2004).

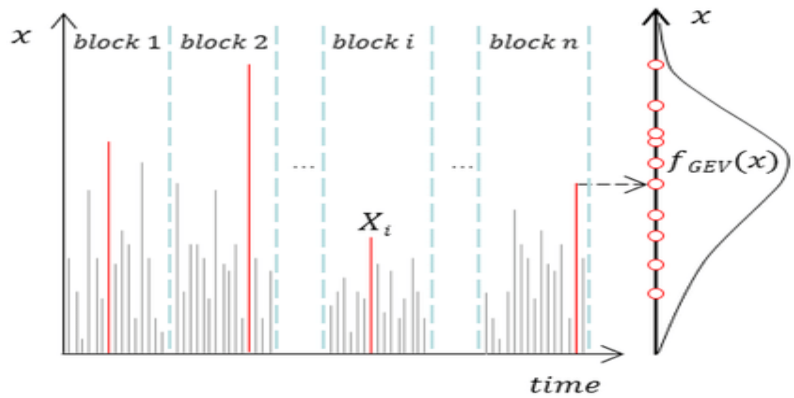

Figure 10: Visual representation of Block Maxima Method

(Hamdi, Bardet, Duluc, & Rebour, 2015)

4.4.2 Maximum Likelihood Estimation (for GEV):

Maximum Likelihood Estimator for the Generalized Extreme Value distribution with probability density function $f(x;\ \mu,\ \sigma,\ \xi)$, can be defined as follows:

Definition

$$L(\mu,\sigma,\xi|x_1,\dots,x_N) \ = \ f(x_1,\dots,x_N;\ \mu,\sigma,\xi) \ = \prod_{i\,=\,1}^{N} f(x_i;\ \mu,\sigma,\xi)$$

The log-likelihood function can be defined as follows:

$$l(\mu,\sigma,\xi|\ x_1,\dots,x_N) \ = \ \ln\left(L(\mu,\sigma,\xi\ |\ x_1,\dots,x_N)\right) \ = \sum_{i\,=\,1}^{N} \ln\big(f(x_i;\ \mu,\sigma,\xi)\big)$$

Where $l(\mu,\ \sigma,\ \xi)$ denotes the log-likelihood function.

Log-likelihood of GEV when $\xi \ \neq \ 0$:

$$l(\mu,\sigma,\xi) \ = \sum_{i\,=\,1}^{N}\Bigg(-\ \ln(\sigma) \ - \ \left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\,1/\xi} - \ \left(\frac{1}{\xi}+1\right)\ \ln\left[1+\xi\left(\frac{x_i\ -\ \mu}{\sigma}\right)\right]\Bigg)$$

provided that $1 + \ \xi\left(\frac{x_i -\ \mu}{\sigma}\right) \ > \ 0,\ \forall\ i \ = \ 1,\dots,N.$

Differentiating this equation with respect to μ, σ, and ξ gives the following:

$$\frac{\partial l(\mu,\sigma,\xi)}{\partial\mu} = \sum_{i=1}^{N}\left(-\frac{1}{\sigma}\left[1+\ \xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\left(1/\xi\,+\,1\right)} + \frac{(1+\xi)}{\sigma}\left[1+\xi\ \left(\frac{x_i-\mu}{\sigma}\right)\right]^{-1}\right) = 0$$

$$\frac{\partial l(\mu,\sigma,\xi)}{\partial\sigma} = \sum_{i=1}^{N}\left(-\frac{1}{\sigma}\left(\frac{x_i-\mu}{\sigma}\right)\left[1+\ \xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\left(1/\xi+1\right)} + \left(\frac{1+\ \xi}{\sigma}\right)\left(\frac{x_i-\mu}{\sigma}\right)\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-1} - \frac{1}{\sigma}\right) = 0$$

$$\frac{\partial l(\mu,\sigma,\xi)}{\partial\xi} = \sum_{i=1}^{N}\left(-\frac{1}{\xi^2}\left[1+\ \xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-1/\xi}\ \ln\left(1+\ \xi\left(\frac{x_i-\mu}{\sigma}\right)\right) + \frac{1}{\xi}\left(\frac{x_i-\mu}{\sigma}\right)\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\left(1/\xi+1\right)} + \frac{1}{\xi^2}\ln\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right] - \left(\frac{1}{\xi}+1\right)\left(\frac{x_i-\mu}{\sigma}\right)\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-1}\right) = 0$$

In the case where $\xi = 0$, the Gumbel log-likelihood function is:

$$l(\mu,\sigma,0) = -N\ln(\sigma) - \sum_{i=1}^{N}\left(\frac{x_i-\mu}{\sigma}\right) - \sum_{i=1}^{N}\exp\left[-\left(\frac{x_i-\mu}{\sigma}\right)\right]$$

And can be differentiated as follows:

$$\frac{\partial l(\mu,\sigma,0)}{\partial \mu} = \sum_{i=1}^{N}\left(\frac{1}{\sigma}\left[1+\ \xi\ \left(\frac{x_i-\ \mu}{\sigma}\right)\right]^{-\left(1/_{\xi\ +\ 1}\right)} + \frac{(1+\xi)}{\sigma}\left[1+\xi\left(\frac{x_i-\ \mu}{\sigma}\right)\right]^{-1}\right)$$
$$=\ 0$$

$$\frac{\partial l(\mu,\sigma,0)}{\partial \sigma} = \frac{1}{\sigma}\sum_{i=1}^{N}\left(-1+\left(\frac{x_i-\ \mu}{\sigma}\right)\left(1-\exp\left(-\left(\frac{x_i-\ \mu}{\sigma}\right)\right)\right)\right) =\ 0$$

Once the maximum likelihood parameters have been calculated, the quantiles of the Generalized Extreme Value Distribution can be obtained by substituting these parameters as follows:

$$\widehat{z_p} = \begin{cases} \mu\ -\ \dfrac{\hat{\sigma}}{\hat{\xi}}\left[\,1-\ y_p{}^{-\hat{\xi}}\right], & for\ \hat{\xi}\ \neq 0 \\ \hat{\mu}\ -\ \hat{\sigma}\log\left(y_p\right)\ , & for\ \hat{\xi} = 0 \end{cases}$$

Where $y_p\ =\ \log\,(1-\ p)$.

4.5 Generalized Pareto Distribution

Let $X_1, X_2, \ldots$ be a sequence of independent random variables, with a common distribution F, and let

$$M_n = \max\{X_1, X_2, \ldots, X_n\}$$

Denote an arbitrary term in the X_i sequence by X, so that for large n,

$$P\{M_n \leq z\} \approx F(z),$$

Where

$$F(z) = exp\left\{-\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$$

For some $\mu, \sigma > 0$ and ξ. Then, for large enough u, the distribution function of $(X - u)$, conditional on $X > u$, is approximately

$$H(y) = 1 - \left(1 + \frac{\xi\, y}{\tilde{\sigma}}\right)^{-1/\xi}$$

Defined on $\left\{y : y > 0 \ and \ \left(1 + \frac{\xi\, y}{\tilde{\sigma}}\right) > 0\right\}$, where $\tilde{\sigma} = \sigma + \xi(u - \mu)$.

This Generalized Pareto Distribution (GPD) has the following properties:

Parameters	$\mu \in \mathbb{R}$ - Location parameter $\sigma \in (0, \infty)$ - Scale parameter $\xi \in \mathbb{R}$ - Shape parameter
Cumulative Distribution Function	$F(z, \mu, \sigma, \xi) = \begin{cases} exp\left(-(1 + \xi\left(\frac{z - \mu}{\sigma}\right)^{-1/\xi}\right), & \xi \neq 0 \\ exp\left(-exp\left(-\frac{z - \mu}{\sigma}\right)\right), & \xi = 0 \end{cases}$
Probability Density Function	$f(z, \mu, \sigma, \xi) = \frac{1}{\sigma}\begin{cases} \left(1 + \xi\left(\frac{z - \mu}{\sigma}\right)\right)^{-1/\xi - 1} exp\left(-\left(1 + \xi\left(\frac{z - \mu}{\sigma}\right)\right)^{-1/\xi}\right), & \xi \neq 0 \\ exp\left(-\frac{z - \mu}{\sigma}\right) exp\left(-exp\left(-\frac{z - \mu}{\sigma}\right)\right), & \xi = 0 \end{cases}$
Mean	$\begin{cases} \mu - \frac{\sigma}{1 - \xi}, \; \xi < 1 \end{cases}$
Median	$\begin{cases} \mu + \frac{\sigma(2^{\xi} - 1)}{\xi} \end{cases}$
Variance	$\begin{cases} \frac{\sigma^2}{(1 - \xi)^2 (1 - 2\xi)}, & \xi < \frac{1}{2} \end{cases}$
Support	$\begin{cases} z \geq \mu, & \xi \geq 0 \\ \mu \leq z \leq \mu - \frac{\sigma}{\xi}, & \xi < 0 \end{cases}$

Table 4: Properties of Generalized Pareto Distribution

4.5.1 Peaks Over Threshold (for GPD):

The Generalized Pareto Distribution is considered to be a threshold model (Coles, 2004). It is often preferred over the Block Maxima Method, since dividing the data into blocks, and only selecting one extreme from each block can be seen as wasteful (Coles, 2004). The Peaks over Threshold method accounts for all extreme values above a set threshold (Coles, 2004).

As with the Block Maxima Method, the data consists of a series of independent and identically distributed events $x_1, \ldots, x_n$ (Coles, 2004). Extreme events are selected using a high threshold, u, for which the exceedances are $\{ x_i : x_i > u\}$. Figure 11 illustrates how the peaks over threshold method works.

Selecting the threshold for the Generalized Pareto Distribution, has the same issues as the Block Maxima Method. In the peaks over threshold method, it is important to carefully select the threshold (Coles, 2004). If the threshold is too low it will violate the asymptotic basis of the model (Coles, 2004). Alternatively, if the threshold is too high, there will be too few exceedances, which results in a high variance (Coles, 2004).

The parameters of the Generalized Pareto Distribution for the given threshold are exclusively determined by the related Generalized Extreme Value distribution with the block maxima method (Coles, 2004). This can be seen by the ξ parameter of the Generalized Pareto Distribution which is equal to the corresponding Generalized Extreme Value distribution (Coles, 2004).

Changing the block size of the Generalized Extreme Value Distribution, would affect the parameters of the Generalized Extreme Value Distribution, but not the corresponding Generalized Pareto Distribution (Coles, 2004). The parameter ξ does not vary with changes to μ and σ which are both self-compensating (Coles, 2004).

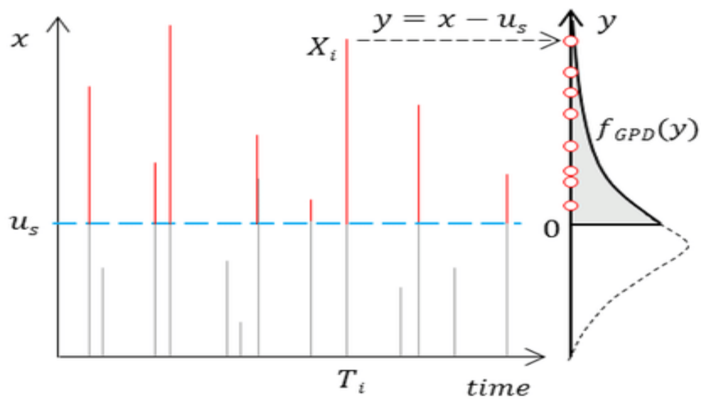

Figure 11: Visual representation of Peaks over Threshold

(Hamdi, Bardet, Duluc, & Rebour, 2015)

4.5.2 Maximum Likelihood Estimation (for GPD):

As with the Generalized Extreme Value distribution, the parameters of the Generalized Pareto Distribution can be predicted using Maximum Likelihood Estimation (Coles, 2004).

Let $y_1, \dots, y_k$ be the k values that exceed the threshold u (Coles, 2004). When $\xi \neq 0$, the log-likelihood is:

$$l(\sigma, \xi) = -k \log(\sigma) - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{k} \log\left(1 + \frac{\xi\, y_i}{\sigma}\right)$$

Provided that $\left(1 + \frac{\xi\, y_i}{\sigma}\right) > 0$ for $i = 1, \dots, k$. Otherwise, $l(\sigma, \xi) = -\infty$ (Coles, 2004).

When $\xi = 0$, the log-likelihood function is defined as follows:

$$l(\sigma) = -k \log(\sigma) - \frac{1}{\sigma} \sum_{i=1}^{k} y_i$$

Maximizing the above function gives parameter estimates for the Generalized Pareto Distribution (Coles, 2004). The quantiles of the Generalized Pareto Distribution can be obtained by substituting the parameters as follows:

$$\widehat{z_p} = \begin{cases} \mu - \frac{\hat{\sigma}}{\hat{\xi}}\left[1 - y_p{}^{-\hat{\xi}}\right], & for\ \hat{\xi} \neq 0 \\ \hat{\mu} - \hat{\sigma} \log(y_p)\ , & for\ \hat{\xi} = 0 \end{cases}$$

Chapter 5 – Risk Measures

Once the residuals of the daily returns have been fitted to either the Generalized Extreme Value Distribution or the Generalized Pareto Distribution, the Value at Risk and Expected Shortfall are calculated in order to understand the risks associated with extreme events. This is used to determine if the actual losses are in line with the forecasted Value at Risk results.

One of the main problems with Extreme Value Theory, is there is very little data in the tails of the distribution (McNeil & Frey, 2000). Value at Risk done using a parametric approach attempts to identify the worst expected loss based on some confidence level (McNeil & Frey, 2000).

Expected Shortfall, or Conditional Value at Risk (CVaR) is a risk measure, similar to Value at Risk, but is more sensitive to the shape of the tails of the loss distribution (McNeil & Frey, 2000). Expected shortfall is “the expected size of a return exceeding Value at Risk” (McNeil & Frey, 2000, p. 271). The expected shortfall measures the expected loss, given that the loss exceeds Value at Risk (McNeil & Frey, 2000).

Therefore, Value at Risk and Expected Shortfall allow one to quantify the level of risk. These risk measures are used to protect against severe financial loss (Chinhamu, Huang, Huang, & Chikobvu, 2015).

5.1 Value at Risk

Governments are increasingly requiring stricter regulations regarding banks capital reserves, and lending practices. Banks are required to hold a required amount of capital to protect against the negative effects of market movements (McNeil & Frey, 2000). The risk capital of a bank is expected to be large enough to cover the losses of the bank's trading portfolio for a 10-day holding period (McNeil & Frey, 2000). This value is referred to as Value at Risk (McNeil & Frey, 2000).

Value at Risk is a widely used risk measure of downside risk (Ardia & Hoogerheide, 2014), and is commonly used all over the world (Miller, 2012). Since is it relatively easy to compute it is one of the most popular risk measures (Ardia, 2008). It has emerged as a prominent risk measurement technique in financial institutions across the world. Value at Risk can be defined as the α quantile of the profit and loss distribution in terms of the generalized inverse of the distribution function (Miller, 2012).

There are 3 main approaches for estimation the Value at Risk:

Non-parametric historical simulations: *This uses the empirical distribution of past increases and decreases to estimate the profit and loss distribution* (McNeil & Frey, 2000).

Parametric method: *which are based on econometric models for volatility* (McNeil & Frey, 2000). *The data is assumed to be conditionally normal, and GARCH type models are used to estimate value at risk* (McNeil & Frey, 2000). *The main drawback of this approach is that most financial data has heavy tails and is therefore not conditionally normal* (McNeil & Frey, 2000).

Extreme Value Theory: *to calculate value at risk is a fairly recent approach* (McNeil & Frey, 2000). *In this case, parametric estimation techniques are used to fit a distribution to the extreme values* (McNeil & Frey, 2000).

Recently Extreme Value Theory has been applied to Value at Risk to study the markets behaviour during periods of extreme stress (Chinhamu, Huang, Huang, & Chikobvu, 2015). This provides a more robust Value at Risk statistic, which makes it an attractive method to calculate risk (Chinhamu, Huang, Huang, & Chikobvu, 2015). This is largely due to the fact that Extreme Value Theory is derived from rigorous statistical theory and provides a parametric approach to understanding the tails of the distribution (Chinhamu, Huang, Huang, & Chikobvu, 2015).

In practical application, the distribution of the financial returns is unknown, so the Value at Risk has to be estimated from the sample data (Ardia, 2008). There is empirical evidence that the distribution of financial returns is typically skewed, with a peak around the mode, and has fatter tails (Bali, 2007). From a mathematical point of view, Value at Risk provides a single quantifiable estimate of potential financial losses (McNeil & Frey, 2000).

Definition

$$VaR(\alpha, T) = \begin{array}{l} \inf\{l \in \mathbb{R} \mid 1 - F_x(l) \leq 1 - \alpha\} \\ \inf\{l \in \mathbb{R} \mid F_x(l) \geq \alpha\} \end{array}$$

Extreme events are more likely to happen in practice, than predicted by a normal thin tailed distribution (Bali, 2007). This means the normality assumption produces Value at Risk estimates which are not a true measure of the risk faced by in reality (Bali, 2007). To account for the drawbacks of the normal distribution a skewed fat-tailed distribution is used (Bali, 2007). This accounts for the non-normality of the financial returns and the effects of relatively infrequent events (Bali, 2007).

Value at Risk estimates can be seen as a confidence interval of the lower tail (Miller, 2012), which is associated with large losses (Bali, 2007). These forecasts are evaluated both conditionally (with reference to information at each point in time) and unconditionally (without reference to available information at each point in time) (Bali, 2007). In this study, Value at Risk statistic is calculated for each estimate window and repeated for each rolling window (Ardia, 2008). This creates a series of Value at Risk estimates.

The advantages and disadvantages of parametric Value at Risk are discussed below (McNeil & Frey, 2000):

1. Not very data intensive, easy to use, and computationally fast
2. Relies on a statistical distribution to characterize potential losses
3. Value at Risk is a function of the parameters of the distribution (eg. GEV or GPD)
4. Value at Risk can be scaled over time
5. Value at Risk gives the worst possible expected loss, given some level of confidence.
6. One of the current shortfalls of parametric Value at Risk, is that it normally relies on a normal distribution and financial data doesn’t follow a normal distribution.

5.2 Expected Shortfall

Value at Risk is often seen as a suitable measure of risk, it does not explain all aspects of risk (Chinhamu, Huang, Huang, & Chikobvu, 2015). One of the main criticisms of Value at Risk, is it does not describe the tail of the distribution (Miller, 2012). Therefore, it does not indicate the potential size of the loss if the confidence level is exceeded (Ardia, 2008). This makes it less sensitive to the shape of the tails of the distribution (Francq & Zakoian, 2010).

Two different distributions could have the same Value at Risk, at a given α level, but the tails of the distribution could be very different beyond that point (Miller, 2012). Expected Shortfall measures the expected loss of a portfolio, given that the loss exceeds the Value at Risk (McNeil & Frey, 2000). It is often referred to as the conditional expected loss (Miller, 2012), and can be mathematically expressed as follows:

Definition

$$E[L|L > VaR_\alpha] = \frac{1}{1-\alpha}\int_{-\infty}^{VaR} x\,f(x)\,dx$$

In short, Expected Shortfall can be seen as a risk measure of the expected loss should the financial returns fall below the Value at Risk level (Ardia, 2008). The latest Basel requirements recommend Expected Shortfall is used as an internal model-based approach to calculate risk (Chinhamu, Huang, Huang, & Chikobvu, 2015). This is due to the fact that Expected Shortfall captures both the magnitude and likelihood of exceedances above the threshold (Chinhamu, Huang, Huang, & Chikobvu, 2015).

5.2.1 Expected Shortfall (Generalized Extreme Value Distribution)

Let r_t denote the financial return at time t, where r_1 denotes the minimum return in each block. According to Extreme Value Theory, the cumulative distribution function of r_1 can be approximated by

$$F(x) = \exp\left\{-1 \left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-1/\xi}\right\}$$

For $\left(1 + \xi \frac{x-\mu}{\sigma}\right) > 0 \, , \mu \in \mathbb{R} \, , \sigma > 0 \, , \xi \in \mathbb{R}$.

The corresponding probability density function is

$$f(x) = \frac{1}{\sigma}\left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-1/\xi - 1} \exp\left\{-1 \left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-1/\xi}\right\}$$

For $\left(1 + \xi \frac{x-\mu}{\sigma}\right) > 0 \, , \mu \in \mathbb{R} \, , \sigma > 0 \, , \xi \in \mathbb{R}$.

The Expected Shortfall of the Generalized Extreme Value Distribution can be computed as follows

$$ES_p(x) = \int_{-\infty}^{u} x\left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-1/\xi - 1} \exp\left\{-1 \left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-1/\xi}\right\} dx$$

Where $u = \mu - \frac{\sigma}{\xi}\left[1 - \{-\ln(p)\}^{-\xi}\right]$

5.2.2 Expected Shortfall (Generalized Pareto Distribution)

Suppose the daily returns $X_1, \dots, X_n$ follow the Generalized Pareto Distribution, with the following cumulative distribution function

$$F(x) = 1 - \left(1 + \xi\,\frac{x-u}{\sigma}\right)^{-1/\xi}$$

Where either $u < x < \infty$ when $\xi \geq 0$, or $u < x < u - \frac{\sigma}{\xi}$ when $\xi < 0$.

The Expected Shortfall of the Generalized Pareto Distribution can be computed as follows

$$ES_p(x) = \frac{1}{1-\xi}\left\{u + \frac{\sigma}{\xi}\left[\left\{\frac{n\,(1-p)}{N_u}\right\}^{-\xi}\right]\right\} + \frac{\beta - \xi u}{1-\xi}$$

Where N_u is the number of returns exceeding the threshold u.

Chapter 6 – Backtesting

Several methods of backtesting have been used in this study. Firstly, a violation based approach has been used to backtest the Value and Risk statistic. This is done using the Kupiec likelihood ratio test for unconditional coverage (Kupiec, 1995). Next, two different independence based tests have been used. The first of these independence based tests is the Christoffersen conditional coverage test (Christoffersen P. F., 1998). Lastly, the Expected Shortfall is backtested using the same procedure proposed by McNeil and Frey (2000), this is implemented both with and without bootstrapping.

6.1 – Value at Risk backtesting

Backtesting can be defined as a set of statistical procedures designed to check if the real losses are in line with Value at Risk forecasts (Jorion, 2007). It involves a historical simulation of an algorithmic investment strategy (Bailey, Borwein, Lopez de Prado, & Jim Zhu, 2014). The Value at Risk forecasts are generated as an internal risk model and are used to produce a sequence of pseudo out-of-sample Value at Risk forecasts for a past time period. Backtesting is used to compare the observed profit and loss to the Value at Risk forecasts (Singh, Allen, & Robert, 2013).

6.1.1 Kupiec Test (Violation based test)

Value at Risk is back tested using the Kupiec test for unconditional convergence (Kupiec, 1995). This is then compared to the corresponding failure rate to α, the confidence level (Chinhamu, Huang, Huang, & Chikobvu, 2015). To account for clustering of the residuals, the block-bootstrap test is used to ensure the residuals are independent and identically distributed.

If we denote $l_t(\alpha)$ as the indicator variable associated to the ex-post observation of an α % Value at Risk violations at time t:

Definition

$$l_t(\alpha) = \begin{cases} 1, & if\ r_t < VaR_{t|t-1}(\alpha) \\ 0, & \end{cases}$$

The Kupiec test for unconditional convergence calculates the number of times the daily returns fall below or above the estimated Vale at Risk at a given α value (Chinhamu, Huang, Huang, & Chikobvu, 2015). The Kupiec statistic is used to test the null hypothesis, that the expected fraction of violations is equal to α, can be calculated as follows:

Definition

$$LR_{UC} = 2\ln\left(\left(\frac{x^{\alpha}}{N}\right)^{x^{\alpha}} \times \left(1 - \frac{x^{\alpha}}{N}\right)^{N-x^{\alpha}}\right) - 2\ln\left(\alpha^{x^{\alpha}} \times (1-\alpha)^{N-x^{\alpha}}\right) \sim \chi_1^2$$

Where x^{α}, is the number of violations below the Value at Risk estimate for a given α. The Kupiec statistic is said to be asymptotically distributed and under the null hypothesis of a correct violation probability follows a Chi-Squared distribution with one degree of freedom (Chinhamu, Huang, Huang, & Chikobvu, 2015).

6.1.2 Christoffersen's Test (Independence based test)

The Christoffersen's test for correct conditional convergence is considered standard practice in financial risk management and is frequently used to back test Value at Risk (Chinhamu, Huang, Huang, & Chikobvu, 2015). It forms an extension to the Kupiec test by testing for independence of the violations, such as clustering of extreme returns (Chinhamu, Huang, Huang, & Chikobvu, 2015).

A sequence of Value at Risk forecasts, with a confidence level of $1 - \alpha$, is said to have correct conditional converge if V_t form a sequence of independent and identically distributed Bernoulli random variables, with parameter α, and $V_t = I\{r_t < VaR_{t\,|\,t-1}\}$ where V_t is a random sequence (Christoffersen P. F., 1998). When testing the null hypothesis (H_0: The exceedances have the correct probability and are independent), it will be rejected if the fraction of Value at Risk violations is much lower or higher than α, or if the Value at Risk violations occur in clusters.

According to Christoffersen (1998) the forecast produced by Value at Risk are valid if and only if the violation process $l_t(\alpha)$ satisfies the following assumptions:

1. The unconditional coverage hypothesis (UC)
2. The independence hypothesis (IND)

This lead to his definition of conditional converge, which can be defined by the following likelihood ratio (Christoffersen P. F., 1998):

Definition

$$LR_{CC} = (LR_{UC} + LR_{IND}) \sim \chi_2^2$$

Where $LR_{IND} = 2\ln\left(\frac{\lfloor(1-\pi_0)^{\varphi_{00}} \times \pi_0^{\varphi_{01}} \times (1-\pi_1)^{\varphi_{10}} \times \pi_1^{\varphi_{11}}\rfloor}{\ln\lfloor(1-\pi_0)^{(\varphi_{00}+\varphi_{10})} \times \pi_1^{(\varphi_{01}+\varphi_{11})}\rfloor}\right)$.

φ_{ij} represents the number of Value at Risk violations in the previous period and the current period (Christoffersen P. F., 1998). When $i = 0$, this is the number of returns that did not violate the Value at Risk estimate in the previous period. When $i = 1$, this is the number of returns that did violate the Value at Risk estimate in the previous period. Similarly, when $j = 0$, this is the number of returns that did not violate the Value at Risk estimate in the current period. When $j = 1$, this is the number of returns that did violate the Value at Risk estimate in the current period. Finally, π_i is the conditional probability of having a Value at Risk violation, given that there either was ($i = 1$) or was not ($i = 0$) a violation on the previous day (Christoffersen P. F., 1998).

6.2 Expected Shortfall Backtesting

The same procedure used in McNeil & Frey (2000) is used to backtest the Expected Shortfall. The Null hypothesis is that the excess Expected Shortfall is independent and identically distributed, and has a mean of zero (Chinhamu, Huang, Huang, & Chikobvu, 2015). It is a one sided t-test against the alternative, that the excess shortfall has mean greater than zero and therefore the Expected Shortfall is systematically underestimated (Chinhamu, Huang, Huang, & Chikobvu, 2015). If the model correctly forecast the Expected Shortfall, then it should have a conditional mean equal to zero.

The test statistic is calculated as:

Definition

$$T = \frac{\bar{r} - \mu_0}{\bar{\sigma}/\sqrt{n}}$$

Where $\bar{r}$ is the mean of the exceedance residuals, and $\bar{\sigma}$ is the standard deviation. The residuals exceedances are defined as $\{r_1, \ldots, r_n\}$ (McNeil & Frey, 2000).

The Expected Shortfall results are tested to determine if the actual Value at Risk violations is independent and identically distributed and has a mean of zero. The test is a one sided t-test against the alternative hypothesis, which tests if the shortfall has a mean greater than zero. This implies that the Expected Shortfall is systematically underestimated. Using bootstrapping to obtain the p-value helps to remove any bias with respect to assumptions about the underlying distribution of the Expected Shortfall.

Chapter 7 – Methodology and Analysis

7.1 Generalized Autoregressive Conditional Heteroscedasticity Model

The Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model is fitted to the data, with a rolling estimation window of 1000 days. The GARCH model is used since the financial returns are heteroskedastic, which means the variance of the returns changes with time. This means the volatility of the financial time series varies with time, and that the time series tends to display volatility clustering.

The GARCH model is used since the daily returns of the S&P 500 and the JSE All Share index are not independent and identically distributed. In particular, the AR (1) + GARCH (1,1) model is used, with a Quasi-Maximum Likelihood Estimation (QMLE) to estimate the parameters. The QMLE assumes an normal distribution.

The log-returns of the AR (1) + GARCH (1,1) can be defined as follows:

Definition

$$
\begin{aligned}
r_t &= \mu + \rho r_{t-1} + u_t, \ \forall\, t = 1, \dots, T \\
u_t &= \sigma_t\, \varepsilon_t, \mid \varepsilon_t \sim iid\ f_\varepsilon \\
\sigma_t^2 &= \alpha_0 + \alpha_1\, \varepsilon_{t-1}^2 + \alpha_2\, \sigma_{t-1}^2, \qquad \mid \alpha_0 > 0\,, \alpha_{1,2} \geq 0\,, \alpha_1 + \alpha_2 < 1
\end{aligned}
$$

The AR (1) + GARCH (1,1) model is used as a consistent and unbiased estimator of the daily returns of the S&P 500 and the JSE All Share Index. As done in McNeil & Frey (2000) and Ardia & Hoogerheide (2013), the model starts with an Autoregressive moving average (ARMA) component to filter out a possible Autoregressive portion of the daily returns. Once the daily returns have been fitted to the GARCH model, the residuals are extracted.

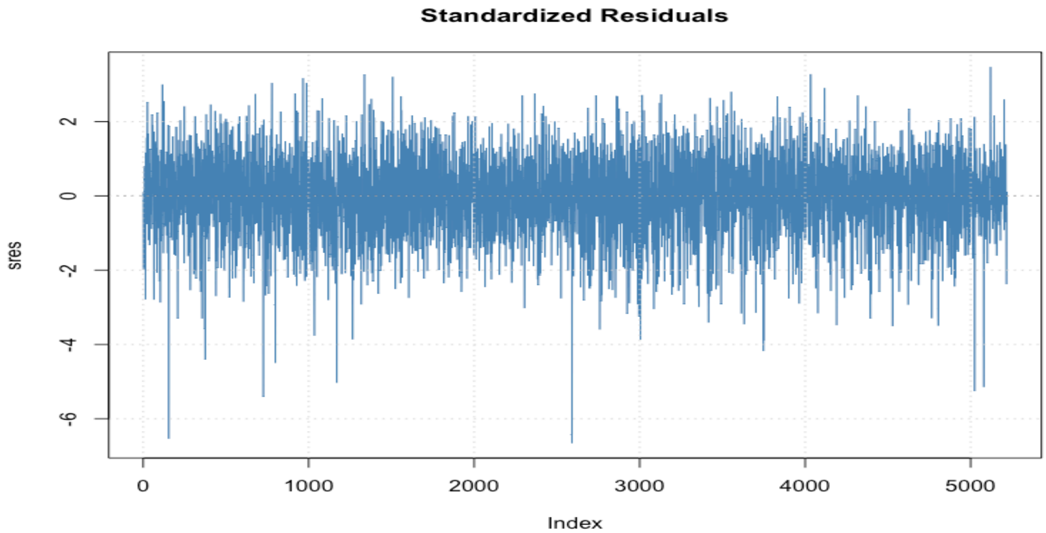

Figure 12: Standardized Residuals for S&P 500

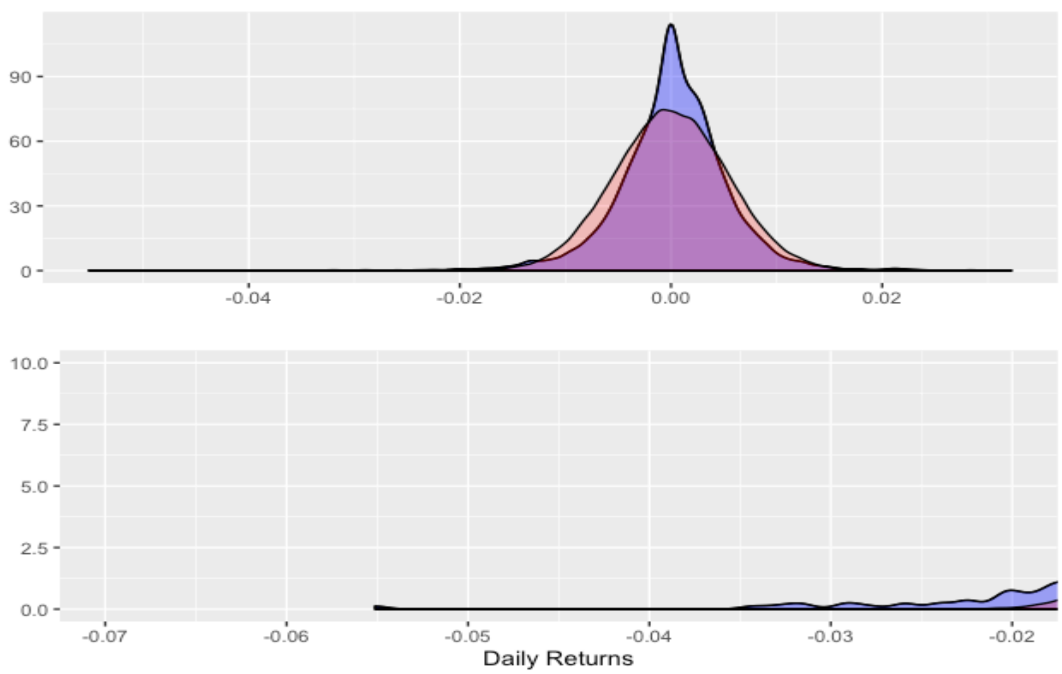

Figure 13: Plot of daily returns of S&P 500

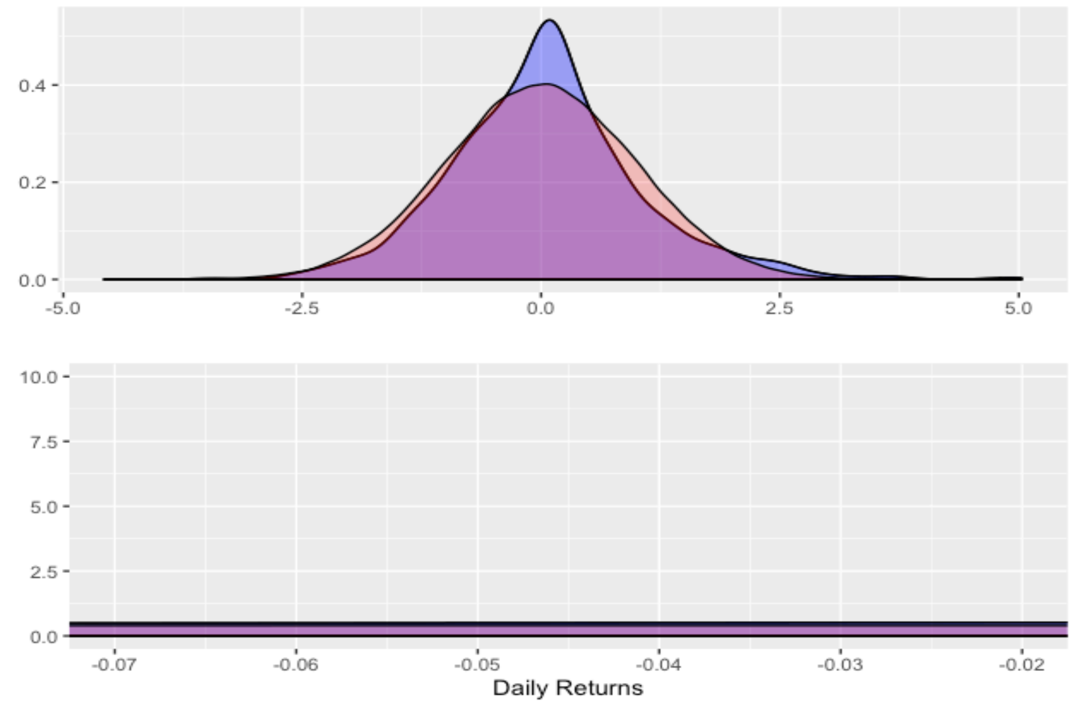

Figure 14: S&P 500 residuals extracted from GARCH model

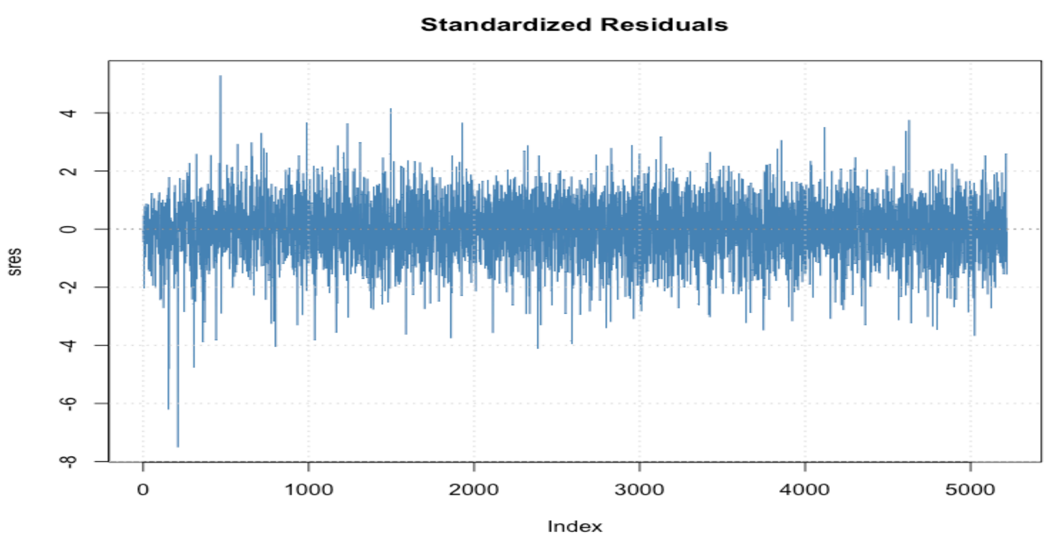

Figure 15: Standardized Residuals for JSE All Share Index

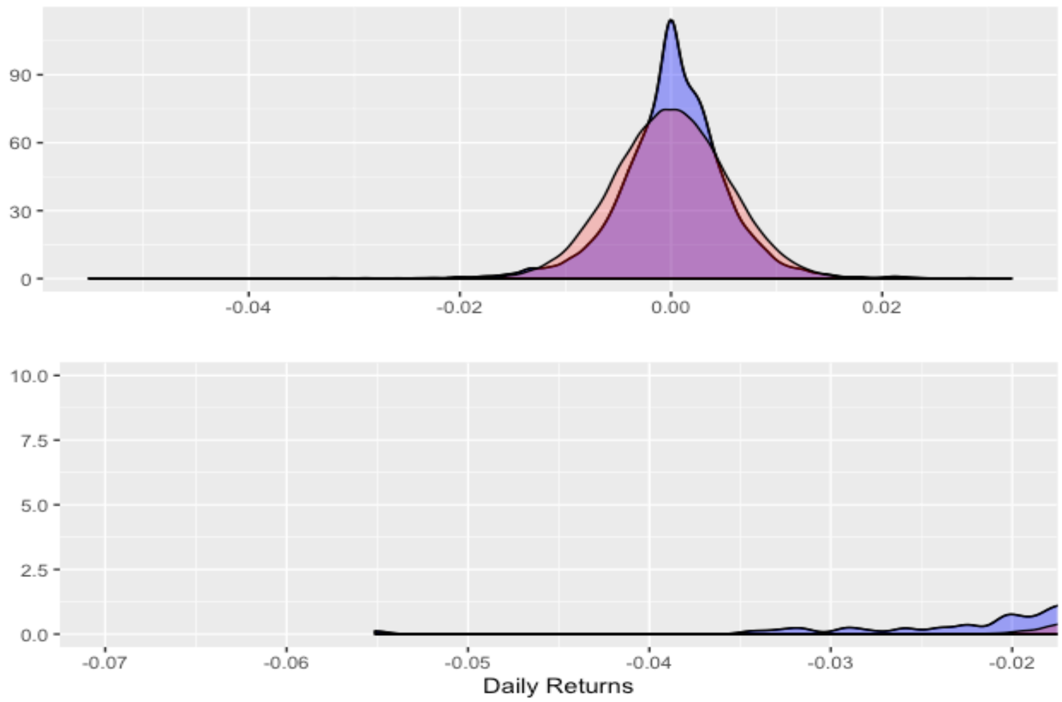

Figure 16: Plot of daily returns of JSE All Share Index

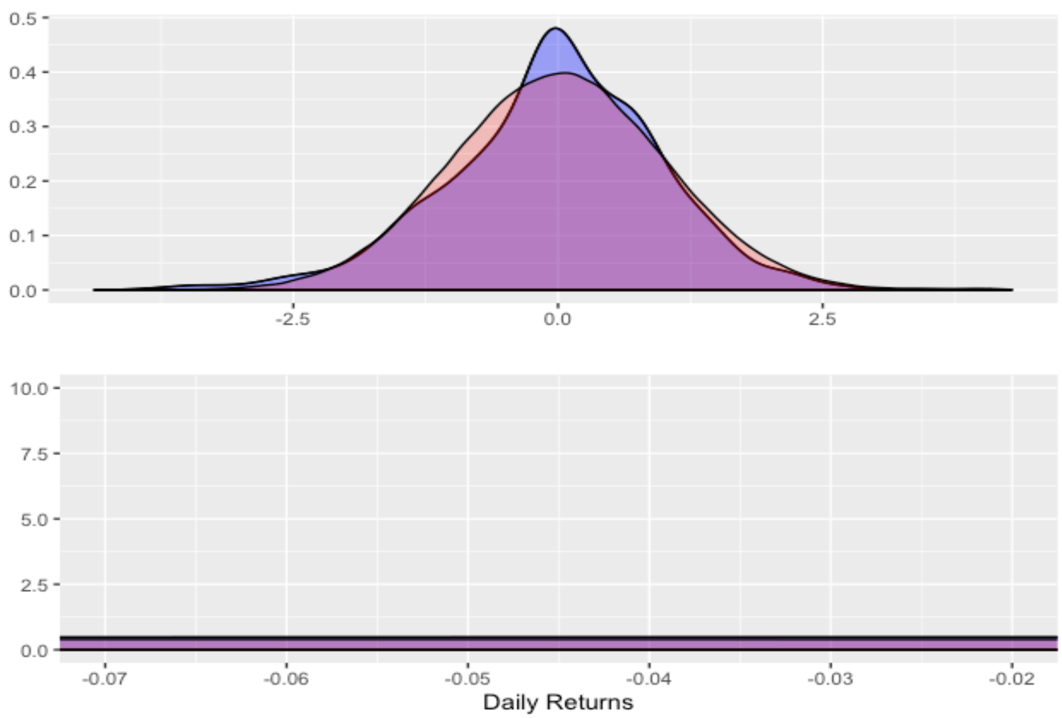

Figure 17: JSE All Share Index residuals extracted from GARCH model

The standardized residuals can be seen as an estimate for the innovations of the GARCH model. Figure 13 and Figure 15 show the daily returns (in red) plotted against the normal distribution (in blue). When examining the tails of the distribution, it is evident that the daily returns do not follow the normal distribution, and is fat tailed.

Figure 14 and Figure 17 show the residuals extracted from the GARCH model plotted against the normal distribution. As seen from the figures above, GARCH model does a good job at making the returns more Independent and Identically distributed since the standardized residuals are not correlated.

7.1.1 GARCH Model Parameter Estimates

Error Analysis of GARCH model for S&P 500

Parameter	Estimate	Std Error	t-value	p-value
Mu	0.00023883535	0.00005255635	4.544	0.000005510052
Ar1	-0.04789787773	0.01427314808	-3.356	0.000791
Omega	0.00000034331	0.00000008723	3.936	0.000083003168
Alpha1	0.09305574347	0.01239432569	7.508	$< 0,1 \times 10\text{-}8$
Beta1	0.89391021770	0.01284023354	69.618	$< 0,1 \times 10\text{-}8$

Table 5: Error Analysis of GARCH model for S&P 500

Error Analysis of GARCH model for JSE All Share Index

Parameter	Estimate	Std Error	t-value	p-value
Mu	0.0002942296	0.0000580959	5.065	0.0000004094
Ar1	0.0605795616	0.0144914800	4.180	0.0000291052
Omega	0.0000004336	0.0000001577	2.749	0.00598
Alpha1	0.0973425294	0.0170586811	5.706	0.0000000115
Beta1	0.8896780815	0.0210729830	42.219	$< 0,1 \times 10\text{-}8$

Table 6: Summary of GARCH model for JSE All Share Index

The above tables give the Maximum Likelihood parameter estimates for the GARCH (1,1) model. All of the parameter estimates are statistically significant at 0.1% for both the S&P 500 and the JSE All Share Index data set. Once the GARCH (1,1) model has been fitted to the data, diagnostic tests are performed to determine if the residuals are stationary.

7.1.2 Ljung-Box Test & ARCH Lagrange multiplier

The Ljung-Box Test was developed by Greta Ljung and George Box (Ljung & Box, 1978). It is used to test for autocorrelations in a time series. This determines if the time series is Independently Distributed, and the model fits the data well.

The test statistic is defined as follows:

Definition

$$\boxed{Q(L) = n(n+2)\sum_{k=1}^{L}\frac{\hat{\rho}_k^2}{n-k}}$$

The ARCH Lagrange multiplier was developed by Robert Engle (1982). It is used to test for autoregressive conditional heteroskedasticity in the standardized residuals. This is done using linear regression as follows:

Definition

$$\boxed{LM = TR^2}$$

Where T represents the sample size and R represents the sample multiple correlation coefficient obtained from:

$$z_t^2 = \alpha_0 + \alpha_1 z_{t-1}^2 + \cdots + \alpha_m z_{t-m}^2 + e_t$$

Where $t = m + 1, \ldots, T$.

Standardised Residuals Tests for S&P 500

			Statistic	**p-value**
Ljung-Box Test	**R**	**Q(10)**	15.53537	0.1137279
Ljung-Box Test	**R**	**Q(15)**	20.89065	0.1403595
Ljung-Box Test	**R**	**Q(20)**	23.10243	0.2837929
Ljung-Box Test	**R^2**	**Q(10)**	17.65603	0.06105056
Ljung-Box Test	**R^2**	**Q(15)**	22.57233	0.09365199
Ljung-Box Test	**R^2**	**Q(20)**	23.40643	0.2692811
LM Arch Test	**R**	**TR^2**	19.43357	0.0785885

Table 7: Standardised Residuals Tests for S&P 500

Standardised Residuals Tests for JSE All Share Index

			Statistic	**p-value**
Ljung-Box Test	**R**	**Q(10)**	6.537721	0.7682465
Ljung-Box Test	**R**	**Q(15)**	11.30035	0.7310271
Ljung-Box Test	**R**	**Q(20)**	17.15641	0.642795
Ljung-Box Test	**R^2**	**Q(10)**	14.39141	0.1558751
Ljung-Box Test	**R^2**	**Q(15)**	16.38764	0.3567672
Ljung-Box Test	**R^2**	**Q(20)**	21.68154	0.358038
LM Arch Test	**R**	**TR^2**	15.9098	0.1954024

Table 8: Standardised Residuals Tests for JSE All Share Index

Table 7 and Table 8 show the Ljung-Box statistic and the ARCH Lagrange multiplier. $Q(L)$ is used to test for auto-correlation, $Q^2(L)$ and TR^2 tests for higher order heteroskedasticity between the residuals. Since the p-value is greater than 0.05 in all cases, the null hypothesis of the data being independently distributed is not rejected. This indicates that the GARCH (1,1) model successfully captures the volatility clustering.

The $Q(L)$ corresponds to what can be seen in the Autocorrelation Function (ACF) plots and $Q^2(L)$ and TR^2 corresponds to the ACF squared plots. The ACF is used as a graphical representation to compare the standardised residuals against the lags. For complete details of the Ljung-Box Test and the ARCH Lagrange multiplier test see Appendix 3.

Model Selection Tests

	AIC	**BIC**	**Log likelihood**
S&P 500	-8.056176	-8.049891	21027.59
JSE	-7.899148	-7.892863	20617.83

Table 9: GARCH Model Information Criterion Statistics

Table 9 illustrates that the GARCH (1, 1) model performs better when fitted to the S&P 500 as opposed to the JSE All Share Index. It should be noted that these statistics one their own do not give much information about the goodness of fit for the GARCH Model. It should be noted that all of the above statistics have been calculated by running the full data set of 20 years through the GARCH model.

7.1.2 Auto-Correlation Function

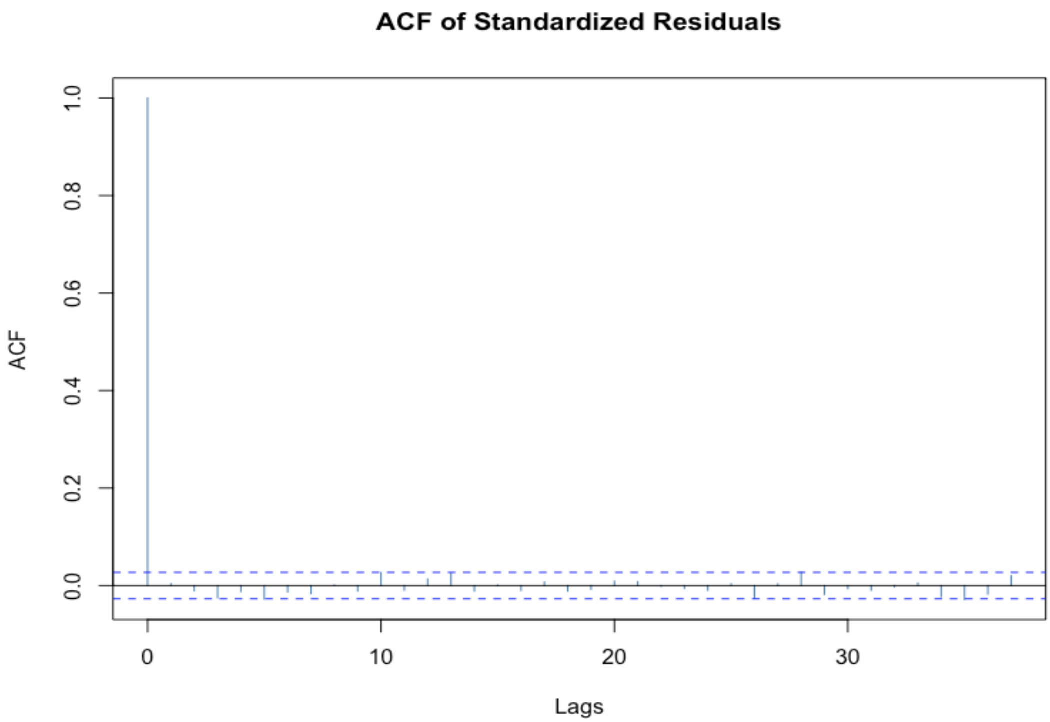

Figure 18: ACF of standardized residuals for S&P 500

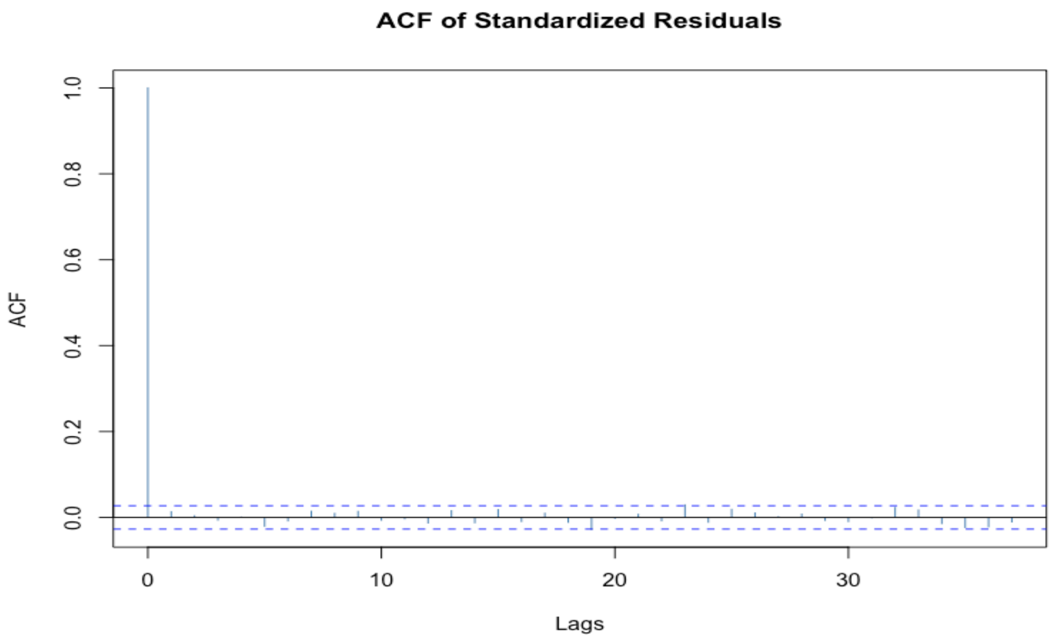

Figure 19: ACF of standardized residuals for JSE All Share Index

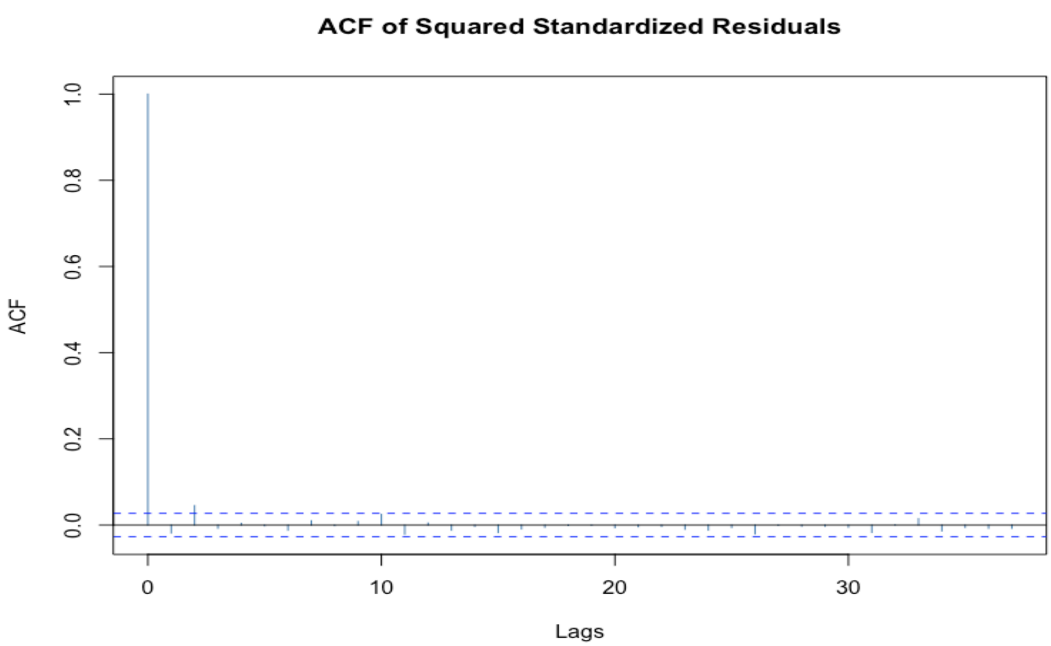

Figure 20: ACF of squared standardized residuals for S&P 500

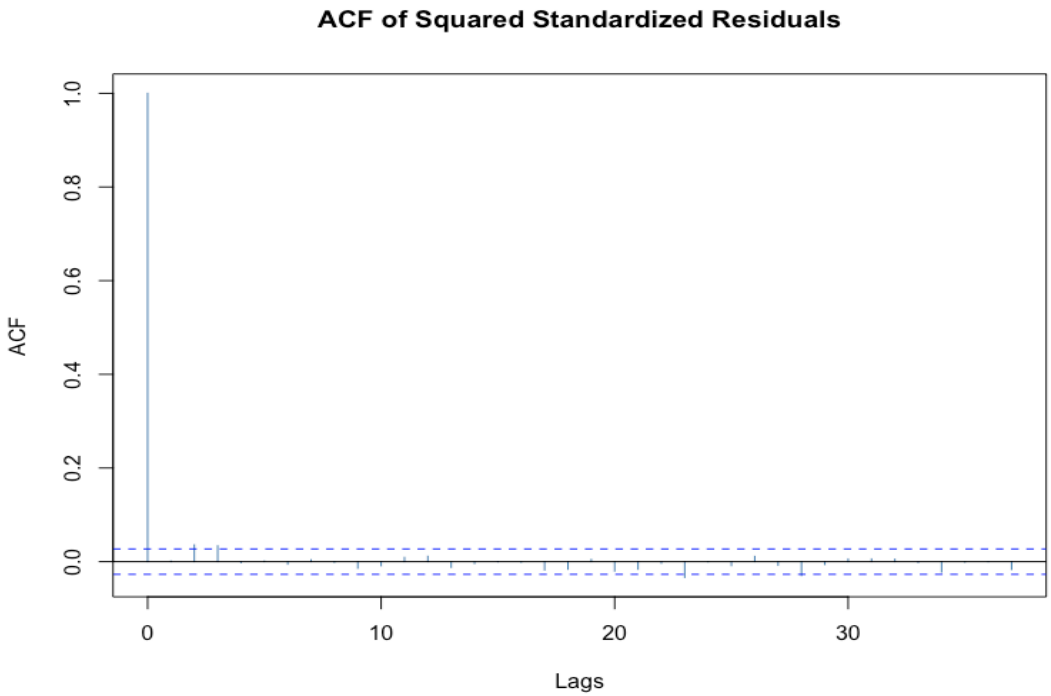

Figure 21: ACF of squared standardized residuals for JSE All Share Index

The Auto-Correlation Function (ACF) represents the correlation between the daily returns and the daily returns with a time lag. It is used to find patterns in the data which might be hidden by noise, such as periodic signally. A good volatility model should be able to capture and explain these stylized facts (Chinhamu, Huang, Huang, & Chikobvu, 2015). As can be seen from the ACF plots of standardized residuals, neither the S&P 500 or the JSE All Share Index show any underlying patterns or correlations once the daily returns have been fitted to the GARCH model. This shows that the GARCH (1, 1) model does a good job of making the residuals on Independent & Identically Distributed.

7.1.2 Jarque-Bera Test

The Jarque-Bera test is a goodness of fit test developed by Carlos Jarque and Anil Bera. It tests whether the sample data have skewness and kurtosis corresponding to a normal distribution. The test statistic is defined as follows:

Definition

$$JB = \frac{n-k+1}{6}\left(S^2 + \frac{1}{4}(C-3)^2\right)$$

Jarque-Bera Test Results

		Statistic	**p-value**
S&P 500	**Chi^2**	939.3655	0
JSE All Share Index	**Chi^2**	833.7646	0

Table 10: Jarque-Bera Test

Table 10 displays the statistic for the Jarque-Bera Test, which tests for goodness of fit against the normal distribution. As can be seen in the above table, both the S&P 500 and the JSE All Share index have a p-value less than 0.05 so that the null hypothesis for normality is rejected at a 5% significance level. This confirms the notion that neither the S&P 500 or the JSE All Share Index follow a normal distribution as can be seen it the Q-Q plots. The residuals are still heavy-tailed, and therefore do not follow a normal distribution. See Appendix 3 for complete details on the Jarque-Bera Test.

7.1.3 Q-Q Plots

The Q-Q plots are graphical representation of how closely the distribution of the data resembles a normal distribution. This is achieved by plotting the quantiles of a theoretical distribution against a normal distribution. As can be seen from the Q-Q plots below, both the Generalized Extreme Value Distribution and the Generalized Pareto Distribution model fit the excesses in the tails relatively well.

Definition

Given an ordered sample of independent observations

$$x_1 \leq x_2 \leq \cdots \leq x_n$$

from a population with estimated distribution function $\hat{F}$, a quantile plot consists of the points

$$\left\{\left(\hat{F}^{-1}\left(\frac{i}{n+1}\right), x_i\right) : i = 1, \ldots, n\right\}$$

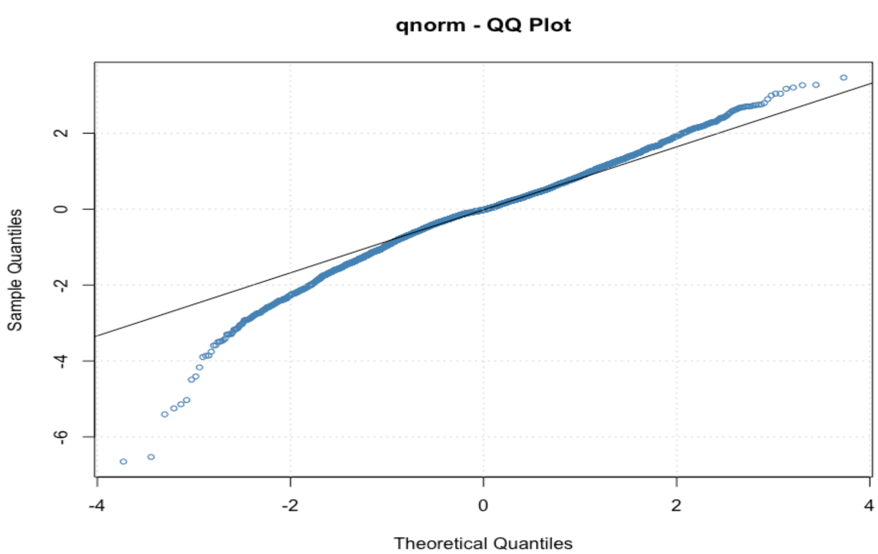

Figure 22: Q-Q Plot for S&P 500

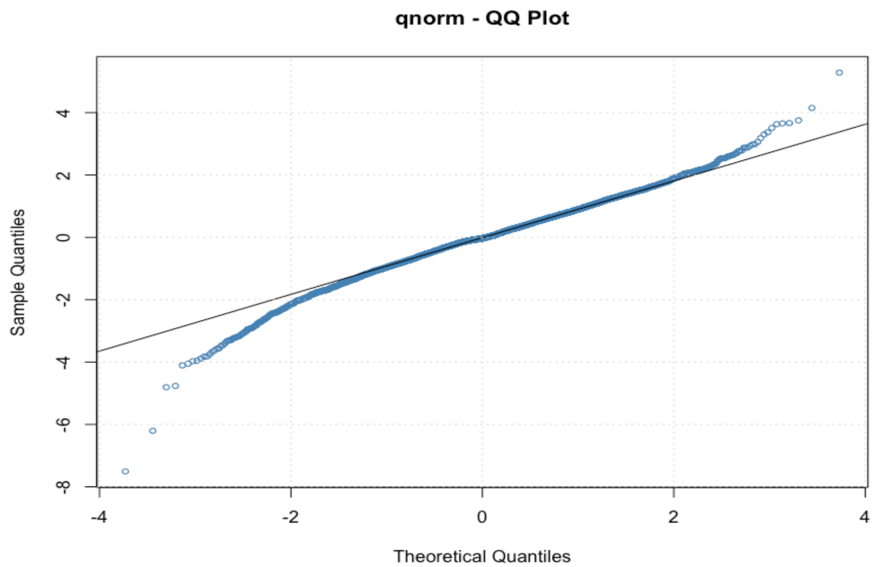

Figure 23: Q-Q Plot for JSE All Share Index

The Q-Q plots above very clearly indicate that neither the S&P 500 or the JSE All Share index follows a normal distribution. Both of the returns are heavy tailed in both the left and right hand tail of the distribution.

7.1.4 Weakness of GARCH modelling

The GARCH (1,1) model is frequently used to calculate Value at Risk and Expected Shortfall, but it does have a few drawbacks. Firstly, the GARCH (1,1) model may violate the non-negativity condition. Secondly, the GARCH (1,1) model cannot account for the leverage effect, or the asymmetric nature of volatility. Lastly, the GARCH (1,1) model does not allow for feedback between the conditional variance and the conditional mean (Brooks, 2008). Using an extreme value approach can help to account for some of these draw backs.

7.2 Generalized Extreme Value Distribution

Once the residuals are extracted, they are fitted to one of two heavy-tailed distributions, either the Generalized Extreme Value distribution, or the Generalized Pareto Distribution. The residuals are used to fit the data to the tails of the distribution. This is to calculate probability estimates associated with extreme or rare events.

When the data is fitted to the Generalized Extreme Value distribution, the Block Maxima Method (BMM) is used to divide the data into equally sized blocks. Let z_i represent the residuals of the daily returns, which are assumed to be independent and identically distributed (Bystrom, 2004). From these blocks, the most extreme value is selected from each block, and used to fit the Generalized Extreme Value distribution. In this study, the data is broken into blocks with 10 data points in each block. The block size of 10 was selected since smaller blocks produce more accurate results, in agreement with the asymptotic property of Maximum Likelihood Estimator (Coles, 2004).

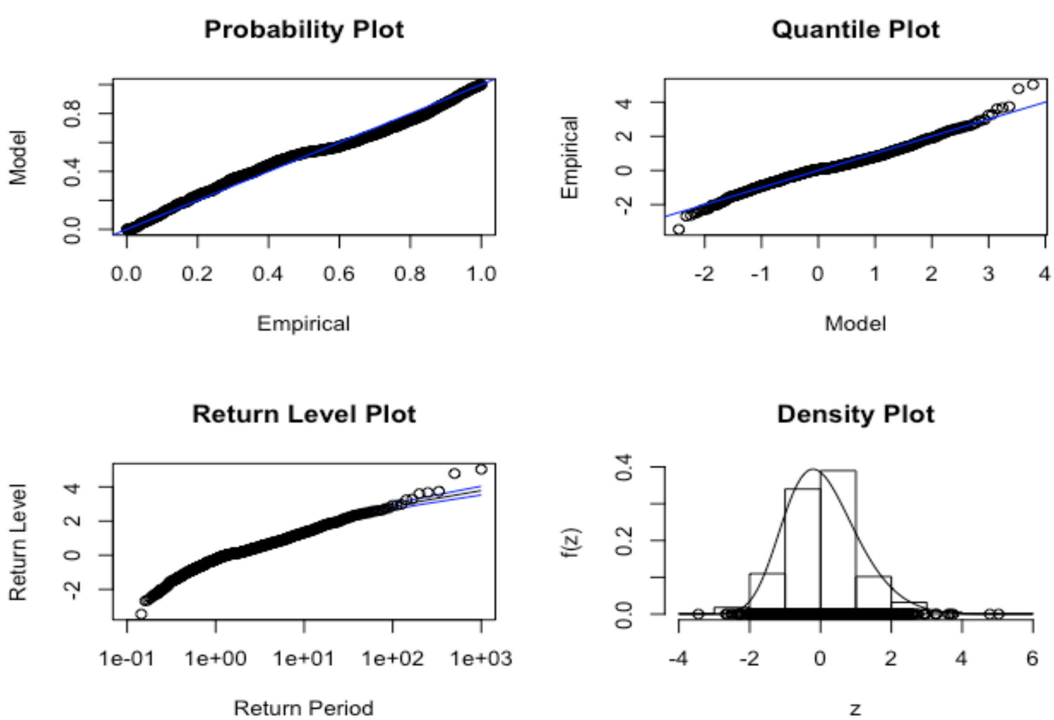

Figure 24: GEV diagnostic plots for S&P 500 data

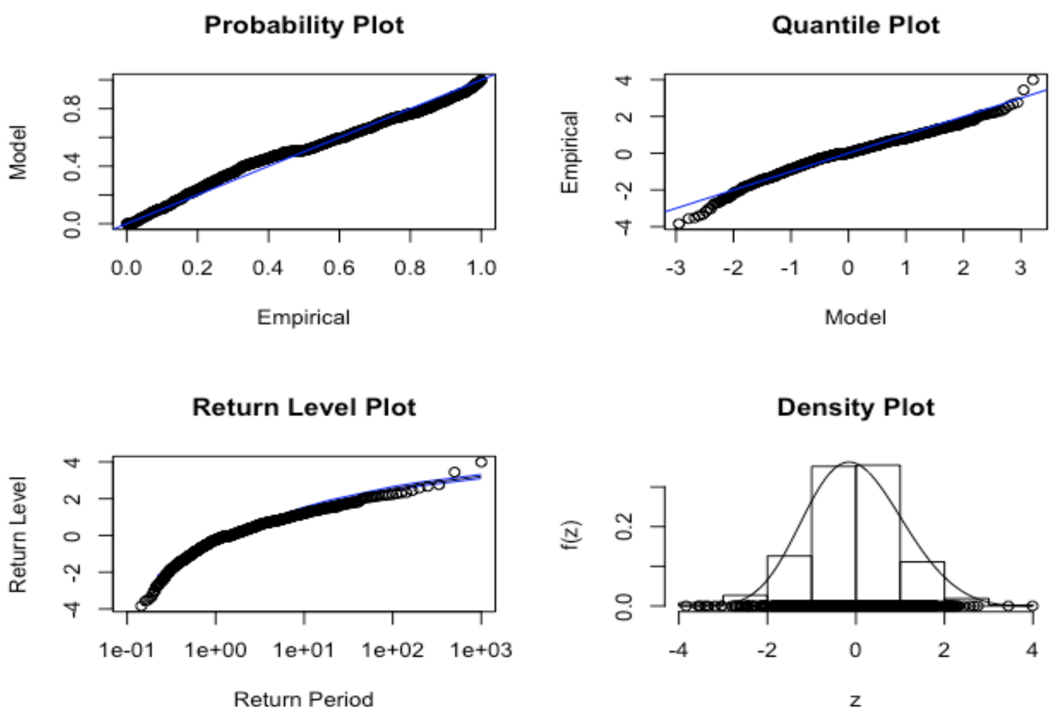

Figure 25: GEV diagnostic plots for JSE All Share Index

Parameter Estimates for Generalized Extreme Value Distribution

	xi	**sigma**	**mu**
S&P 500	-0.09551837	0.49109976	1.19499994
JSE`	-0.07941019	0.52832616	1.23506914

Table 11: Parameter estimates for Generalized Extreme Value Distribution using Maximum Likelihood parameter estimation

As can be seen from Table 11, the shape parameter (xi) is negative. This suggests the standardised residuals for both the S&P 500 and the JSE All Share index follow a Weibull distribution. Once the data has been fitted to the Generalized Extreme Value

Distribution, the Value at Risk and Expected Shortfall are calculated. This done at the 99^{th} , 97.5, and 95^{th} percentile.

Value at Risk and Expected Shortfall summary for GEV (p = 0.99)

	Value at Risk		**Expected Shortfall**	
	S&P 500	**JSE**	**S&P 500**	**JSE**
Mean	-0.012295	-0.011508	-0.00001691	-0.0019745
Median	-0.010428	-0.010573	-0.00010441	-0.0016828
Maximum	-0.067559	-0.037230	-0.01210814	-0.0109319
Minimum	-0.004835	-0.005239	0.01018629	0.0007204
1st Quantile	-0.014077	-0.013175	-0.00086265	-0.0024083
3rd Quantile	-0.007915	-0.008637	0.00093892	-0.0011582

Table 12: Value at Risk and Expected Shortfall summary for GEV (p = 0.99)

Value at Risk and Expected Shortfall summary for GEV (p = 0.975)

	Value at Risk		**Expected Shortfall**	
	S&P 500	**JSE**	**S&P 500**	**JSE**
Mean	-0.009738	-0.009532	-0.0001298	-0.0020605
Median	-0.008364	-0.008711	-0.0001864	-0.0017624
Maximum	-0.053052	-0.031679	-0.0124652	-0.0110996
Minimum	-0.004003	-0.004408	0.0095984	0.0005961
1st Quantile	-0.011031	-0.010804	-0.0009491	-0.0024911
3rd Quantile	-0.006342	-0.007253	0.0008267	-0.0012395

Table 13: Value at Risk and Expected Shortfall summary for GEV (p = 0.975)

Value at Risk and Expected Shortfall summary for GEV (p = 0.95)

	Value at Risk		**Expected Shortfall**	
	S&P 500	**JSE**	**S&P 500**	**JSE**
Mean	-0.007708	-0.007952	-0.0003032	-0.0021930
Median	-0.006636	-0.007234	-0.0003133	-0.0018880
Maximum	-0.042620	-0.027179	-0.0130272	-0.0113615
Minimum	-0.003231	-0.003683	0.0087178	0.0004053
1st Quantile	-0.008722	-0.008948	-0.0010858	-0.0026145
3rd Quantile	-0.005031	-0.006089	0.0006608	-0.0013549

Table 14: Value at Risk and Expected Shortfall summary for GEV (p = 0.95)

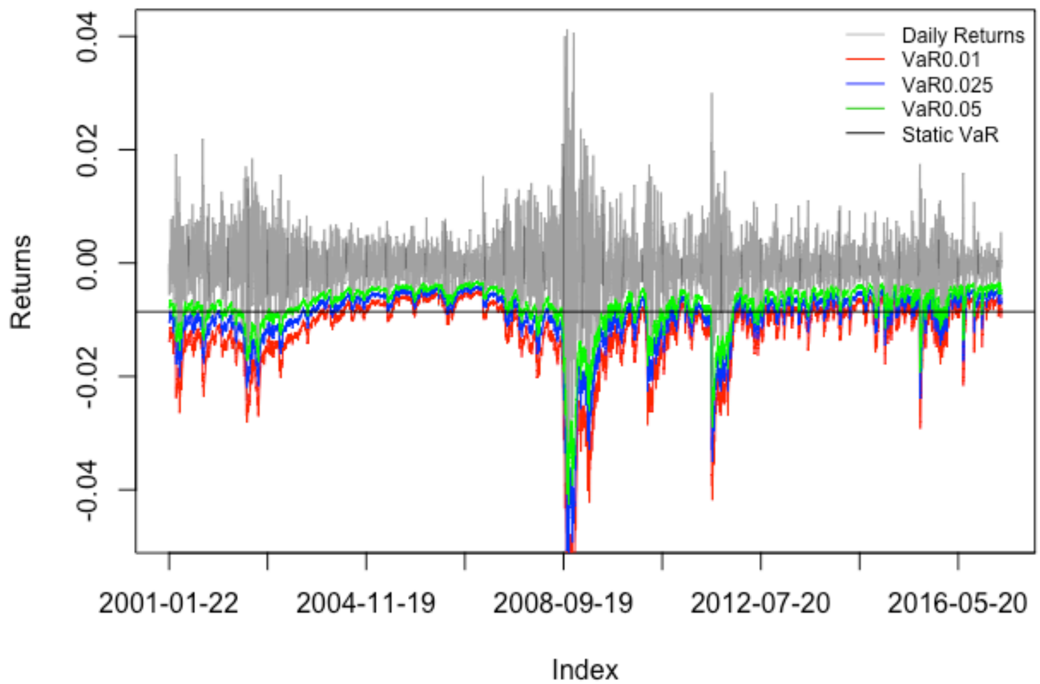

Figure 26: Value at Risk for S&P 500 (using Block Maxima Method)

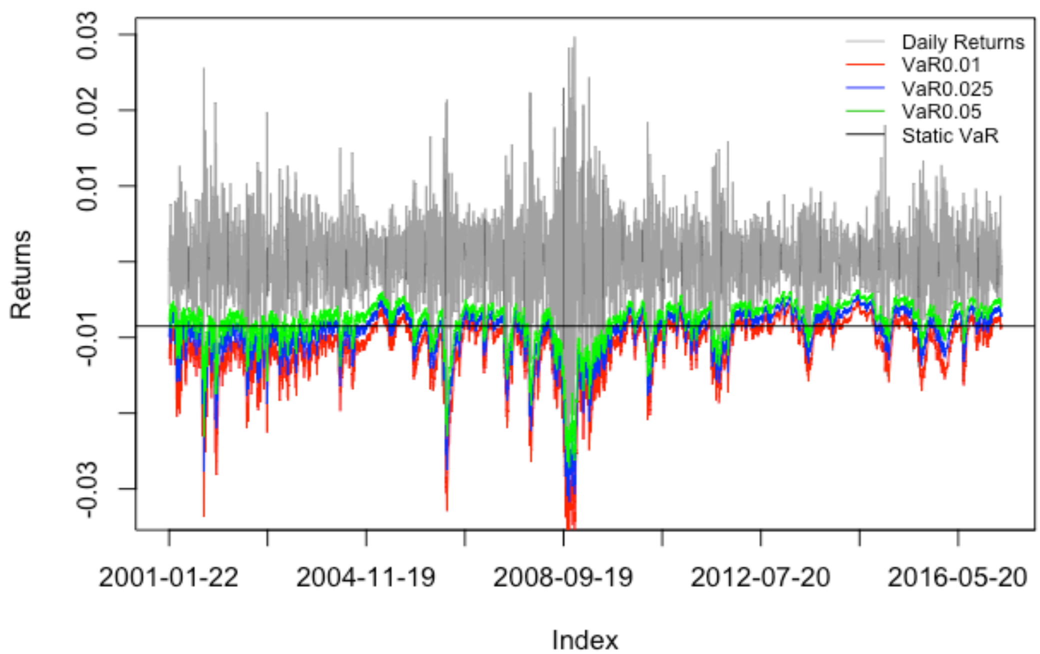

Figure 27: Value at Risk for JSE All Share Index (using Block Maxima Method)

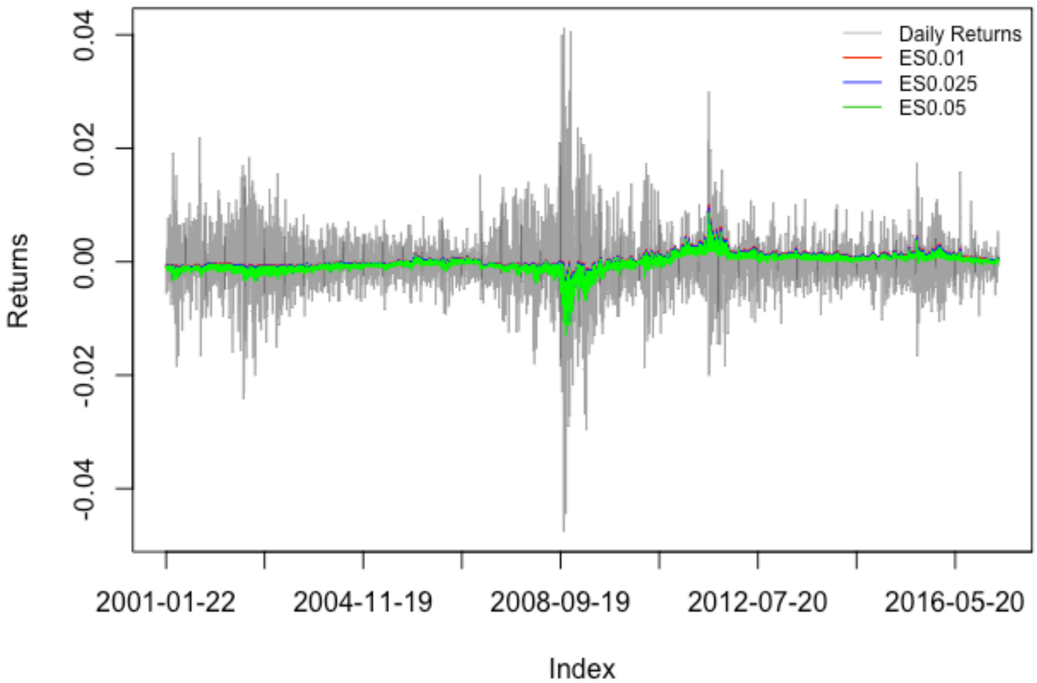

Figure 28: Expected Shortfall for S&P 500 (using Block Maxima Method)

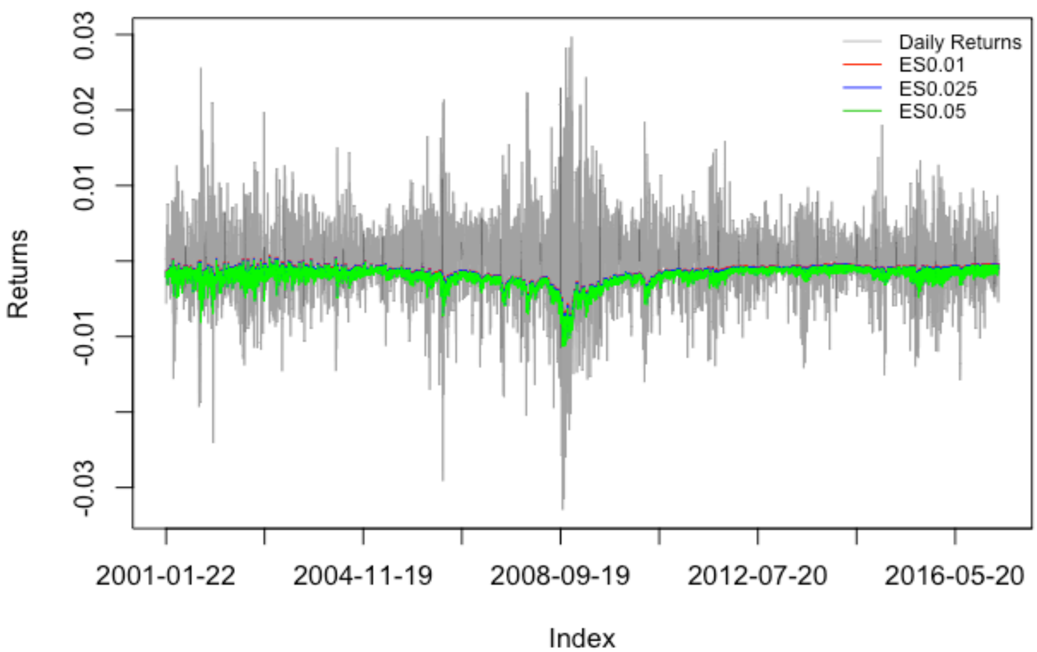

Figure 29: Expected Shortfall for JSE All Share Index (using Block Maxima Method)

The tables above show the Value at Risk and Expected Shortfall estimates for both the S&P 500 and the JSE All Share Index when calculated using the block maxima method.

Table 12 show that at the 99th quantile, the average losses expected to be 0.012% for the S&P 500, and 0.011% for the JSE All Share Index. At the 97.5th quantile the average losses are expected to be 0.009% for both the S&P 500 and the JSE All Share Index, and the maximum expected losses are 0.05% and 0.03% for the S&P 500 and the JSE All Share Index respectively, as seen in Table 13. In Table 14 the expected losses at the 95th quantile are displayed. The average losses for the S&P 500 and JSE All Share index is 0.07%, and the maximum expected losses are 0.042% for the S&P 500 and 0.027% for the JSE All Share Index.

The changes in the Value at Risk and Expected Shortfall over time can be seen in Figure 26, Figure 27, Figure 28, and Figure 29.

7.3 Generalized Pareto Distribution

When the data is fitted to the Generalized Pareto Distribution, the Peaks-Over-Threshold (POT) method is used. In this case, the threshold selected is 90%, and all the residuals above the threshold are used to fit the Generalized Pareto Distribution. This means that only the 10% most extreme residuals will be used to fit the Generalized Pareto Distribution.

The Peaks over Threshold method uses the residuals are extracted using the GARCH (1, 1) model, and a high threshold is selected of 90%. This means only the top 10% most extreme residuals are used to fit the Generalized Pareto Distribution.

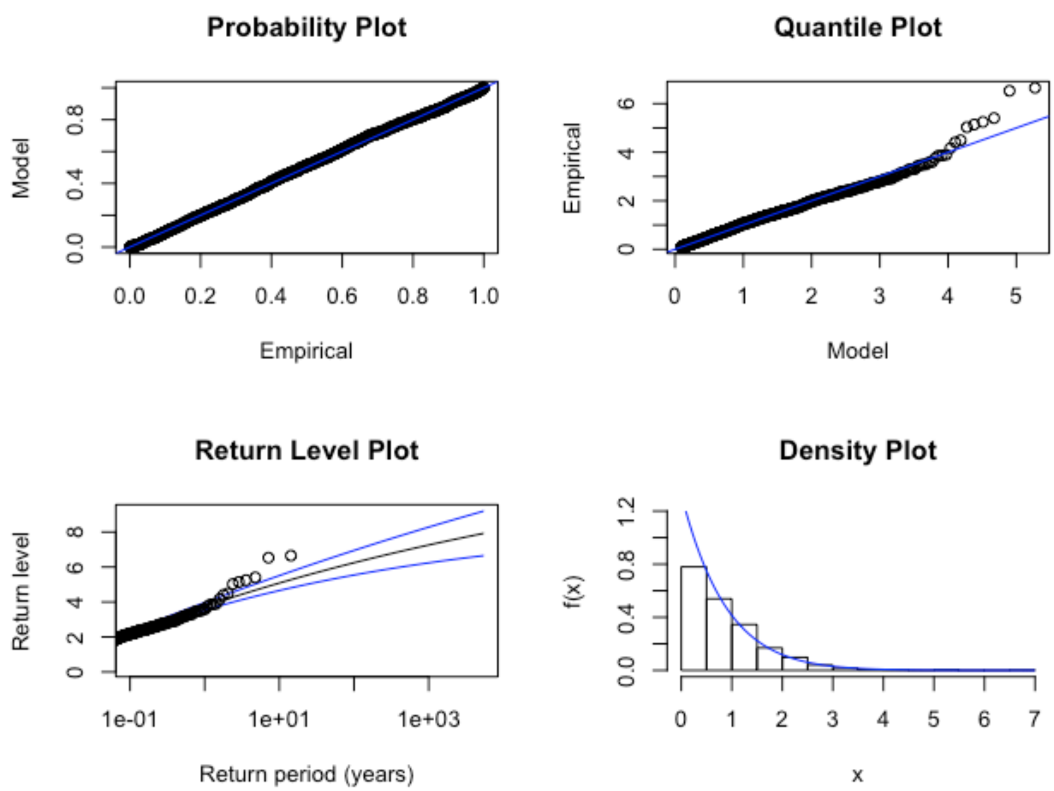

Figure 30: GPD diagnostic plots for S&P 500 data

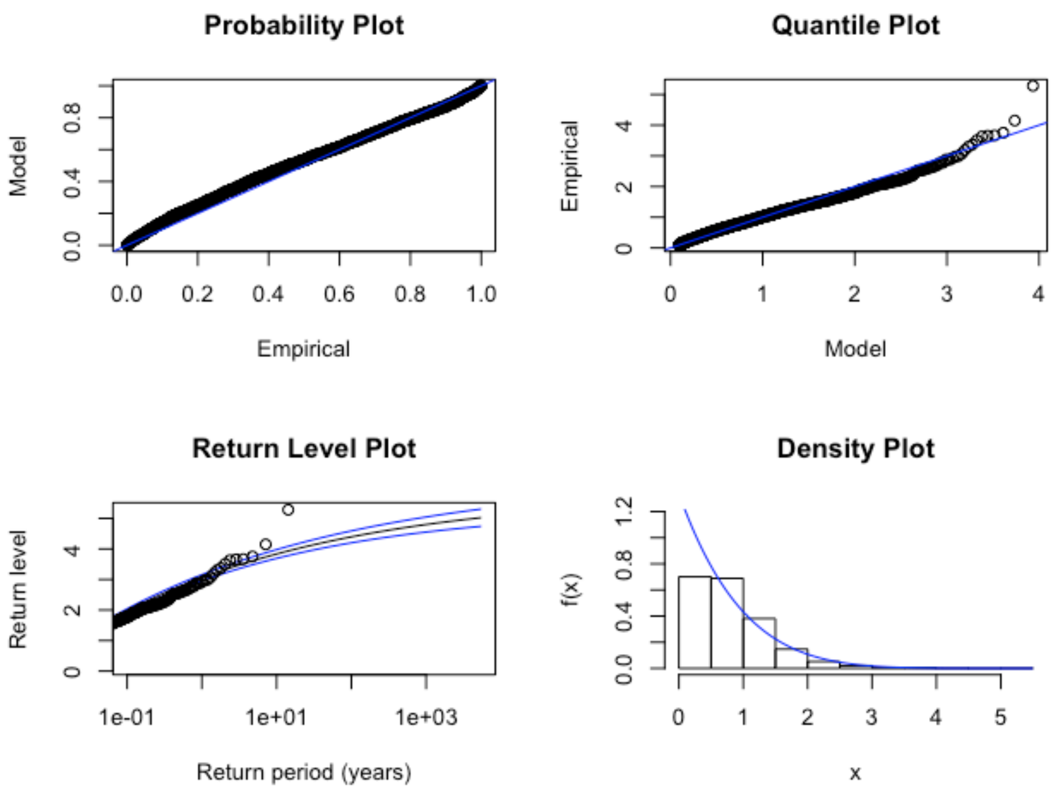

Figure 31: GPD diagnostic plots for JSE All Share Index data

Parameter Estimates for Generalized Pareto Distribution

	xi	**beta**
S&P 500	- 0.1424366	0.5537540
JSE All Share Index	0.02725862	0.46077408

Table 15: Parameter estimates for Generalized Pareto Distribution using Maximum Likelihood parameter estimation

As can be seen from Table 15, the shape parameter (xi) is negative when fitted using the S&P 500 data. When the Generalized Pareto Distribution is fitted to the JSE All Share Index, the shape parameter in negative. The beta represents the scale parameter of the Generalized Pareto Distribution. Once the data has been fitted to the Generalized Pareto Distribution, the Value at Risk and Expected Shortfall are calculated. This done at the 99th , 97.5, and 95th percentile.

Value at Risk and Expected Shortfall summary for GPD (p = 0.99)

	Value at Risk		**Expected Shortfall**	
	S&P 500	**JSE**	**S&P 500**	**JSE**
Mean	- 0.012267	- 0.011336	- 0.014968	- 0.013577
Median	- 0.010447	- 0.010418	- 0.012425	- 0.012593
Maximum	- 0.068507	- 0.037029	- 0.085438	- 0.042760
Minimum	- 0.004768	- 0.005085	- 0.005558	- 0.005977
1st Quantile	- 0.013929	- 0.012975	- 0.017434	- 0.015731
3rd Quantile	- 0.007936	- 0.008509	- 0.009411	- 0.010063

Table 16: Value at Risk and Expected Shortfall summary for GPD (p = 0.99)

Value at Risk and Expected Shortfall summary for GPD (p = 0.975)

	Value at Risk		**Expected Shortfall**	
	S&P 500	**JSE**	**S&P 500**	**JSE**
Mean	- 0.009769	- 0.009278	- 0.012489	- 0.011522
Median	- 0.008399	- 0.008465	- 0.010566	- 0.010618
Maximum	- 0.053896	- 0.031138	- 0.070209	- 0.037310

Minimum	- 0.003936	- 0.004209	- 0.004813	- 0.005141
1st Quantile	- 0.011003	- 0.010532	- 0.014308	- 0.013242
3rd Quantile	- 0.006389	- 0.007022	- 0.008002	- 0.008619

Table 17: Value at Risk and Expected Shortfall summary for GPD (p = 0.975)

Value at Risk and Expected Shortfall summary for GPD (p = 0.95)

	Value at Risk		**Expected Shortfall**	
	S&P 500	**JSE**	**S&P 500**	**JSE**
Mean	- 0.007815	- 0.007712	- 0.010588	- 0.009964
Median	- 0.006717	- 0.007019	- 0.009023	- 0.009117
Maximum	- 0.043199	- 0.026384	- 0.059059	- 0.032910
Minimum	- 0.003246	- 0.003520	- 0.004191	- 0.004484
1st Quantile	- 0.008822	- 0.008696	- 0.012025	- 0.011370
3rd Quantile	- 0.005112	- 0.005880	- 0.006837	- 0.007494

Table 18: Value at Risk and Expected Shortfall summary for GPD (p = 0.95)

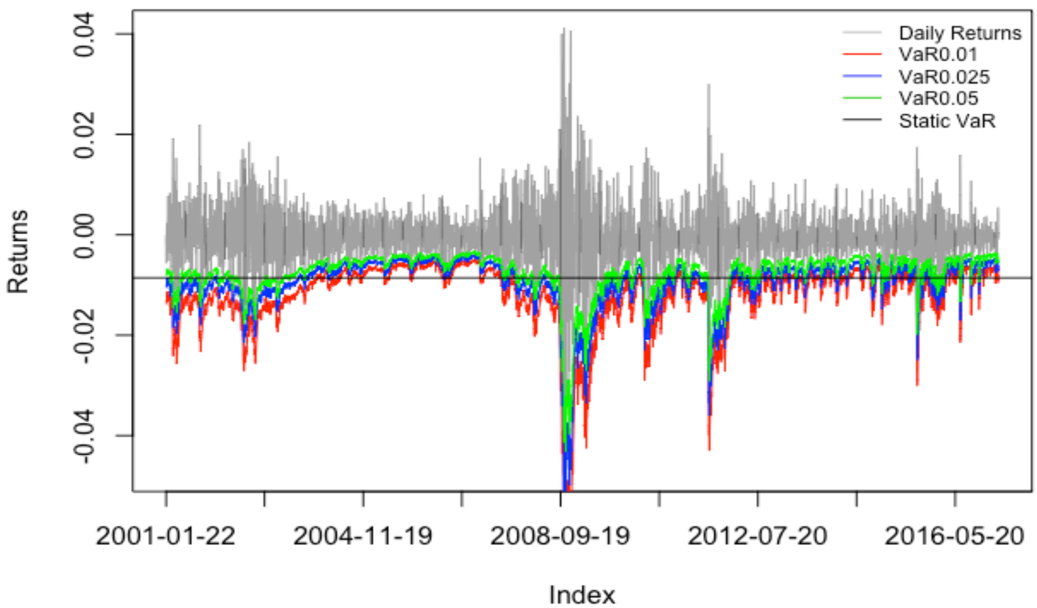

Figure 32: Value at Risk for S&P 500 (using Peaks over Threshold Method)

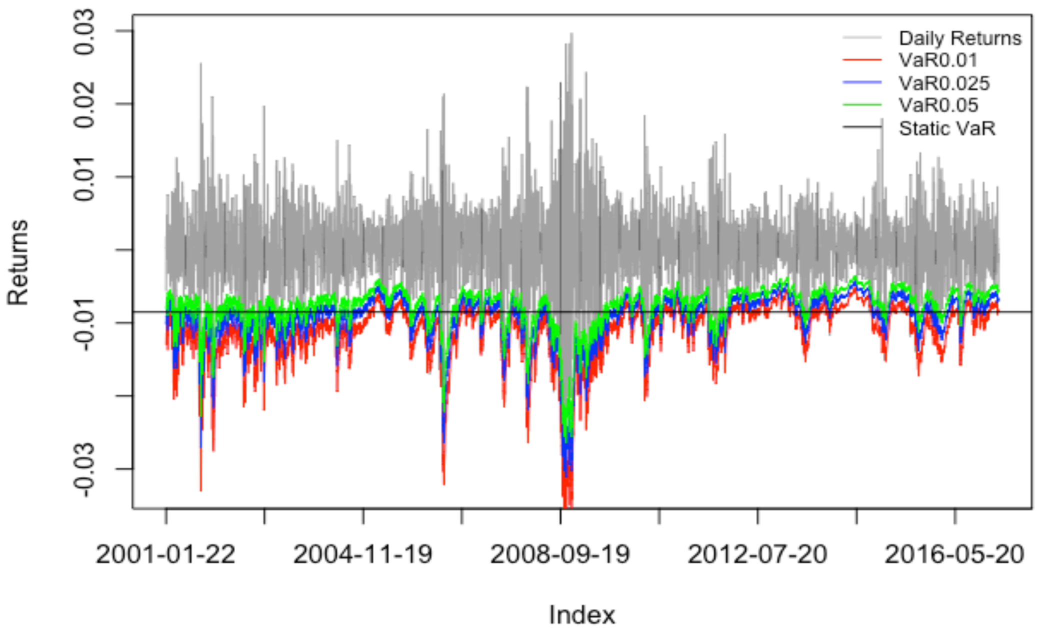

Figure 33: Value at Risk for JSE All Share Index (using Peaks over Threshold Method)

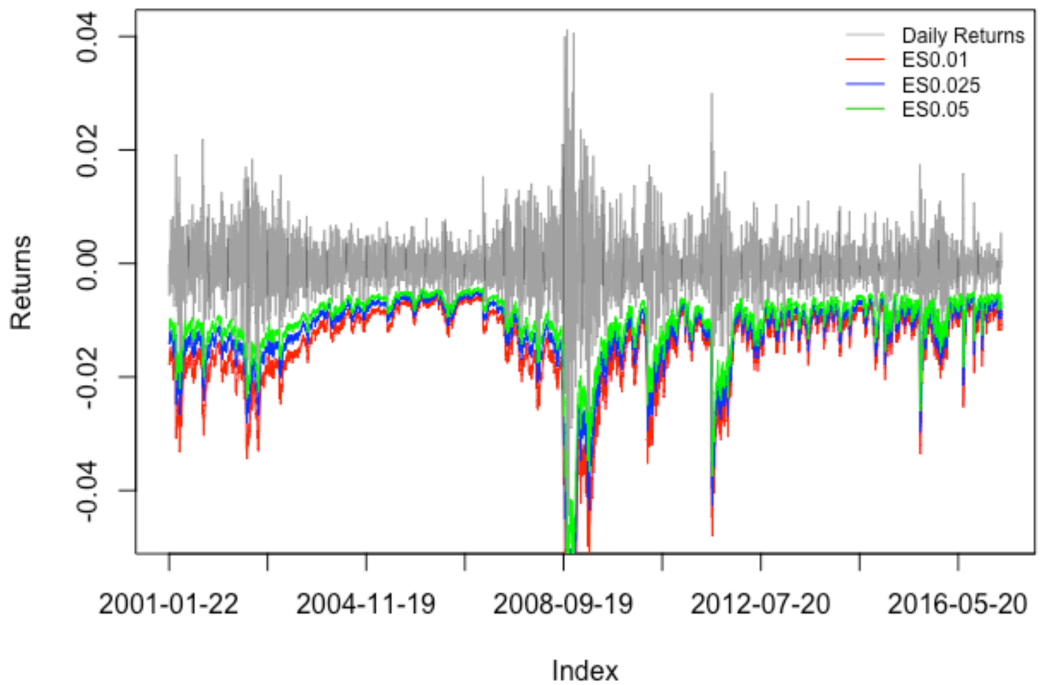

Figure 34: Expected Shortfall for S&P 500 (using Peaks over Threshold Method)

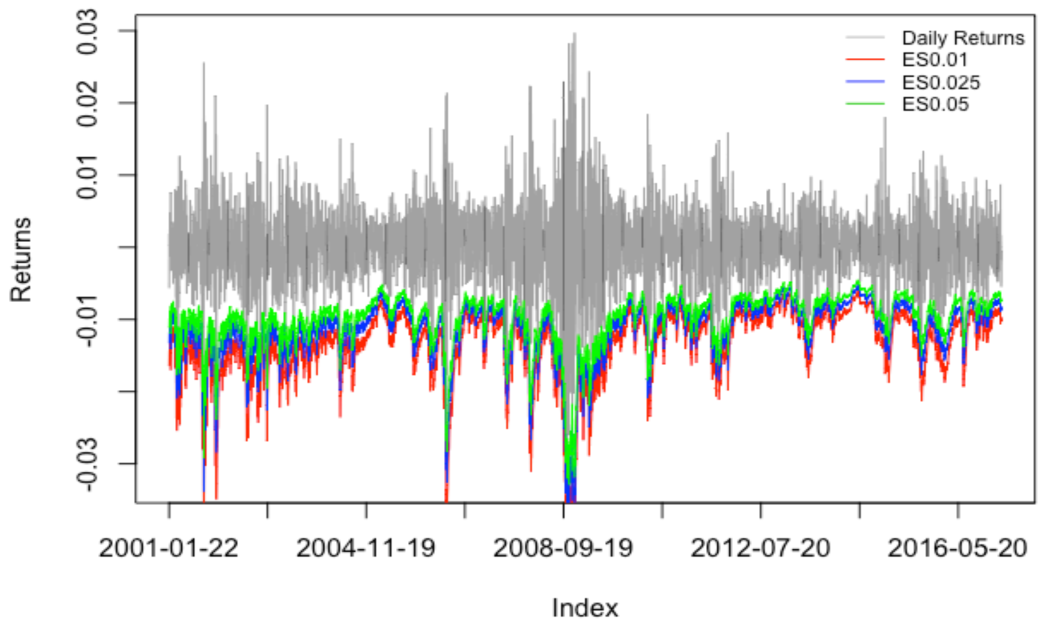

Figure 35: Expected Shortfall for JSE All Share Index (using Peaks over Threshold Method)

The tables above show the Value at Risk and Expected Shortfall estimates for both the S&P 500 and the JSE All Share Index when calculated using the peaks over threshold method. The Value at Risk and Expected Shortfall estimate for the Generalized Pareto Distribution are very similar to those of the Generalized Extreme Value Distribution, and are summarised below.

Table 16 show that at the 99th quantile, the average losses expected to be 0.012% for the S&P 500, and 0.011% for the JSE All Share Index. At the 97.5th quantile the average losses are expected to be 0.009% for both the S&P 500 and the JSE All Share Index, and the maximum expected losses are 0.053% and 0.031% for the S&P 500 and the JSE All Share Index respectively, as seen in Table 17. In Table 18 the expected losses at the 95th quantile are displayed. The average losses for the S&P 500 and JSE All Share index is 0.07%, and the maximum expected losses are 0.043% for the S&P 500 and 0.026% for the JSE All Share Index.

The changes in the Value at Risk and Expected Shortfall over time can be seen in Figure 32, Figure 33, Figure 34, and Figure 35.

7.4 Backtesting Results

The following four null hypotheses were tested for the Generalized Extreme Value Distribution and the Generalized Pareto Distribution for both the S&P 500 and the JSE All Share index.

The following four hypotheses where tested:

H_0: Exceedances with correct probability – the Value at Risk is tested using Christofferson's Test for Conditional Convergence

H_0: Exceedances with the correct probability and independent – the Value at Risk is tested using Christofferson's Test for Conditional Convergence

H_0: The mean of excess violations of Value at Risk is not greater than zero – Expected Shortfall is tested

7.4.1 Generalized Extreme Value Distribution Backtesting

The Value at Risk estimates are then back tested using the Christoffersen's test to assess the performance of the model (Ardia, 2008).

The backtesting results for the Generalized Extreme Value Distribution (using the Block Maxima Method) for the S&P 500:

	$\alpha = 0{,}01$	$\alpha = 0.025$	$\alpha = 0.05$
Expected	42	105	210
Actual	45	117	218
Unconditional convergence p-value	0.667	0.264	NaN
Conditional convergence p-value	0.732	0.347	NaN
Expected Shortfall p-value	0	0	0

Table 19: Value at Risk results for Generalized Extreme Value Distribution using S&P 500 data.

The backtesting results for the Generalized Extreme Value Distribution (using the Block Maxima Method) for the JSE All Share Index:

	$\alpha = 0,01$	$\alpha = 0.025$	$\alpha = 0.05$
Expected	42	105	210
Actual	35	95	177
Unconditional convergence p-value	0.252	0.294	0.014
Conditional convergence p-value	0.387	0.065	0.014
Expected Shortfall p-value	0	0	0

Table 20: Value at Risk results for Generalized Extreme Value Distribution using JSE All Share Index data.

The expected value is the number of exceedances which are expected at the given confidence interval. For the purposes of this study, only the left tail is considered since this is associated with large losses. The results for the S&P 500 are summarised in Table 19. Therefore, at a 99% quantile, the Value at Risk condition is expected to be violated 42 times, when using the S&P 500 data set. The actual value is the number of Value at Risk violations which actually occur when the data is fitted to the Generalized Extreme Value Distribution using the Block Maxima Method. In this case the actual value is 45 Value at Risk violations. Similarly, at a 97.5% quantile there are 105 expected violations, but only 117 actual violations. Lastly, at a 95% quantile there is an expected number of 210 violations, and only 218 actual Value at Risk violations.

The South African market is considered to be an emerging market, so the JSE All Share Index has different characteristics to a more developed market such as the S&P 500. This can be seen in the results in Table 20. The expected number of exceedances remains the same as the S&P 500 for the JSE All Share index. At the 99% quantile level there are 35 actual exceedances, when the expected number of exceedances is 42. Next, at the 97.5% quantile level there are 95 actual exceedances. Lastly, at the 95% quantile level there are 177 actual exceedances.

Based on the results, it can be seen that the Generalized Extreme Value distribution captures the behaviour of the tails, and gives relatively good Value at Risk predictions. Next the results of the Generalized Extreme Value Distribution are compared to that of the Generalized Pareto Distribution.

7.4.2 Generalized Pareto Distribution Backtesting

The results for the Generalized Pareto Distribution (using the Peaks over Threshold Method) for the S&P 500:

	$\alpha = 0{,}01$	$\alpha = 0.025$	$\alpha = 0.05$
Expected	42	105	210
Actual	37	98	202
Unconditional convergence p-value	0.379	0.191	NaN
Conditional convergence p-value	0.581	0.084	NaN
Expected Shortfall p-value	0.474	0.870	0.451

Table 21: Results for Value at Risk of Generalized Pareto Distribution using S&P 500 data.

The results for the Generalized Pareto Distribution (using the Peaks over Threshold Method) for the JSE All Share Index:

	$\alpha = 0{,}01$	$\alpha = 0.025$	$\alpha = 0.05$
Expected	42	105	210
Actual	40	102	199
Unconditional convergence p-value	0.732	0.730	NaN
Conditional convergence p-value	0.643	0.075	NaN
Expected Shortfall p-value	0.403	0.473	0.324

Table 22: Results for Value at Risk of Generalized Pareto Distribution using JSE All Share Index data.

As expected, the Generalized Extreme Value Distribution doesn't perform as well as the Generalized Pareto Distribution. This is largely due to the fact that the Generalized Pareto Distribution better captures the shape of the tails.

The results for the S&P 500 are summarised in Table 21. At the 99% quantile, the Value at Risk is expected to be violated 42 times, the actual value is 45 Value at Risk violations. Similarly, at the 97.5% quantile there are 105 expected violations, but only 117 actual violations. Lastly, at the 95% quantile there is an expected number of 210 violations, and only 218 actual Value at Risk violations.

When the JSE All Share Index data is used, the expected number of exceedances remains the same. At the 99% quantile level there are 35 actual exceedances. Next, at the 97.5% quantile level there are 95 actual exceedances. Lastly, at the 95% quantile level there are 177 actual exceedances.

Chapter 8 – Impact of Estimation Frequency

The previous chapter focused on taking an Extreme Value Theory approach to the daily returns of S&P 500 and the JSE All Share Index. This was achieved using the Generalized Extreme Value Distribution and the Generalized Pareto Distribution to model the rare events which occur in the tails of the distribution.

This chapter looks at the effects of updating the GARCH model's parameters less frequently, when taking an extreme value approach to calculating Value at Risk and Expected Shortfall. The work of Ardia & Hoogerheide (2014) is extended using an extreme value approach, as done in Mc Neil & Frey (2000). This allows the effects of estimation frequency to be analysed while focusing on the extreme events which occur in the tails of the distribution of the returns. This is important for large risk management systems, such as in banking and investments.

In the study by Ardia & Hoogerheide (2014), the performance of two different GARCH models were compared when the parameters of the model were updated at different frequencies. The models used in this study are the GARCH(1,1) model and the asymmetric GARCH model also known as the GJR(1,1) model.

The GARCH models were updated daily, weekly (5 days), monthly (20 days), and quarterly (60 days) (Ardia & Hoogerheide, 2014). Twelve years of S&P 500 data was used from 2000 to 2012. Both the Value at Risk and the Expected Shortfall were calculated, and then compared between the different update frequencies (Ardia & Hoogerheide, 2014).

One of the interesting results from their study, was that more advanced models, which are updated less frequently result in much better Value at Risk and Expected Shortfall forecasts, than simpler models which are updated daily (Ardia & Hoogerheide, 2014).

They found that, the impact of the update frequency had a relatively small impact on both the value at risk and expected shortfall forecasts (Ardia & Hoogerheide, 2014). With only marginal differences in the performance of the model (Ardia & Hoogerheide, 2014).

For all of the analysis in the chapter the following colours will be used to represent different update frequencies:

Update frequency colour key

Blue – The model is updated daily
Orange – The model is updated every 5 days (weekly)
Green – The model is updated every 10 days (fortnightly)
Purple – The model is updated every 20 days (monthly)

8.1 GARCH Model Results

As done in the previous chapter, the daily returns are fitted to the GARCH (1,1) model, but in this chapter the performance of the model is compared when it is update less frequently. Initially the GARCH model is updated daily, and then every 5 day, then every 10 day, and lastly every 20 days. In each case, the model is still fed all of the data points, but at less frequent intervals.

Error Analysis of GARCH model for S&P 500

Parameter	Estimate	Std Error	t-value	p-value
Mu	- 0.0002760432 - 0.000279710 - 0.0002833329 - 0.0002943486	0.0000933345 0.000092947 0.0000929277 0.0000927860	- 2.958 - 3.009 - 3.049 - 3.172	0.00310 0.00262 0.00230 0.00151
AR	- 0.0512979314 - 0.052534062 - 0.0520447474 - 0.0482307320	0.0353386833 0.035367345 0.0352351769 0.0351673816	- 1.452 - 1.485 - 1.477 - 1.371	0.14661 0.13744 0.13966 0.17023
Omega	0.0000012523 0.000001245 0.0000012597 0.0000012381	0.0000004244 0.000000424 0.0000004236 0.0000004224	2.951 2.937 2.974 2.931	0.00317 0.00331 0.00294 0.00337
Alpha	0.1919482572 0.195757108 0.1961084117 0.1967960197	0.0483107792 0.048714095 0.0483344336 0.0476842131	3.973 4.018 4.057 4.127	0.0000709 0.0000586 0.0000496 0.0000367

Beta	0.7023315048 0.699401275 0.6979330078 0.6996055492	0.0642810161 0.064261333 0.0636128117 0.0622584752	10.926 10.884 10.972 11.237	$< 0{,}1 \times 10^{-8}$ $< 0{,}1 \times 10^{-8}$ $< 0{,}1 \times 10^{-8}$ $< 0{,}1 \times 10^{-8}$

Table 23: Error Analysis of GARCH model for S&P 500

Error Analysis of GARCH model for JSE All Share Index

Parameter	**Estimate**	**Std Error**	**t-value**	**p-value**
Mu	0.0001510824 0.0001525832 0.0001483257 0.0001613052	0.0001168444 0.0001170665 0.0001163881 0.0001168609	1.293 1.303 1.274 1.380	0.1960 0.192 0.2025 0.167
AR	0.0016163734 0.0009531905 0.0031011836 0.0038940862	0.0327798524 0.0327033385 0.0327085190 0.0325824626	0.049 0.029 0.095 0.120	0.9607 0.977 0.9245 0.905
Omega	0.0000003495 0.0000003440 0.0000003361 0.0000003480	0.0000001648 0.0000001613 0.0000001611 0.0000001677	2.121 2.133 2.087 2.075	0.0339 0.033 0.0369 0.038
Alpha	0.0725531599 0.0717391193 0.0737542333	0.0143388762 0.0139317141 0.0143165587	5.060 5.149 5.152	0.000000419 0.000000261 0.000000258

	0.0747545448	0.0145109074	5.152	0.000000258
Beta	0.9076154771 0.9087371055 0.9074645720 0.9062332741	0.0174375317 0.0168729333 0.0171787131 0.0176163739	52.050 53.858 52.825 51.443	$< 0{,}1 \times 10^{-8}$ $< 0{,}1 \times 10^{-8}$ $< 0{,}1 \times 10^{-8}$ $< 0{,}1 \times 10^{-8}$

Table 24: Error Analysis of GARCH model for the JSE All Share Index

The above tables give the Maximum Likelihood parameter estimates and the standard error estimates for the GARCH (1,1) model. Table 23 shows the impact of estimation frequency when using the MLE to estimate the parameters of the GARCH model for the S&P 500 data. Similarly, Table 24 shows the parameter estimates of the GARCH model for the JSE All Share Index.

Standardised Residuals Tests for S&P 500

			Statistic	**p-value**
Ljung-Box Test	**R**	**Q(10)**	7.273701 7.055216 6.347582 5.434805	0.6993778 0.7202222 0.7852647 0.8603073
Ljung-Box Test	**R**	**Q(15)**	14.97824 14.36018 13.59909 14.00753	0.4529858 0.4984059 0.5561261 0.5249583
Ljung-Box Test	**R**	**Q(20)**	19.39037	0.4965971

			19.51938 18.31093 18.22365	0.4883321 0.5669342 0.572678
Ljung-Box Test	**R^2**	**Q(10)**	2.592301 2.597662 2.554426 2.546894	0.9960048 0.9893748 0.9900586 0.9901746
Ljung-Box Test	**R^2**	**Q(15)**	4.41948 4.603049 4.525937 4.69067	0.9960048 0.9949871 0.9954361 0.9944372
Ljung-Box Test	**R^2**	**Q(20)**	5.731233 5.958879 5.879367 5.980551	0.9992148 0.9989519 0.9990509 0.9989235
LM Arch Test	**R**	**TR^2**	3.163374 3.047823 2.921633 2.946854	0.9942765 0.9951969 0.9960725 0.9959077

Table 25: Standardised Residuals Tests for the S&P 500

Standardised Residuals Tests for JSE All Share Index

			Statistic	p-value
Ljung-Box Test	**R**	**Q(10)**	13.55057 13.25874 14.37348 12.85973	0.1944959 0.2095628 0.1566276 0.2316206
Ljung-Box Test	**R**	**Q(15)**	16.40169 16.2648 17.10368 15.48349	0.3558689 0.3646747 0.3127045 0.4171833
Ljung-Box Test	**R**	**Q(20)**	19.11157 18.58467 19.01497 17.72962	0.5145849 0.5489516 0.5208533 0.6052141
Ljung-Box Test	R^2	**Q(10)**	9.189734 9.624591 10.15174 10.52536	0.5141969 0.4740229 0.427283 0.3956705
Ljung-Box Test	R^2	**Q(15)**	14.90868 14.56932 14.89091 15.45906	0.4580151 0.482861 0.4593032 0.4188835

Ljung-Box Test	**R^2**	**Q(20)**	22.91403 22.56234 22.97199 22.75328	0.2930356 0.3107916 0.2901722 0.3010708
LM Arch Test	**R**	**TR^2**	13.00044 12.57503 13.70715 13.40736	0.3690085 0.40067 0.3197986 0.3401399

Table 26: Standardised Residuals Tests for the JSE All Share Index

Table 25 and Table 26 show the Ljung-Box statistic and the ARCH Lagrange multiplier. $Q(L)$ is used to test for auto-correlation, $Q^2(L)$ and TR^2 tests for higher order heteroskedasticity between the residuals. Since the p-value is greater than 0.05 in all cases, the null hypothesis of the data being independently distributed is not rejected. This indicates that the GARCH (1,1) model successfully captures the volatility clustering.

The $Q(L)$ corresponds to what can be seen in the Autocorrelation Function (ACF) plots and $Q^2(L)$ and TR^2 corresponds to the ACF squared plots. The ACF is used as a graphical representation to compare the standardised residuals against the lags. For complete details of the Ljung-Box Test and the ARCH Lagrange multiplier test see Appendix 3.

Model Selection Tests

	AIC	BIC	Log likelihood
S&P 500	- 8.683944 - 8.687514 - 8.685282 - 8.681965	- 8.659406 - 8.662976 - 8.660743 - 8.657426	4346.972 4348.757 4347.641 4345.983
JSE	- 8.227759 - 8.223627 - 8.224688 - 8.220552	- 8.203220 - 8.199088 - 8.200150 - 8.196014	4118.88 4116.814 4117.344 4115.276

Table 27: GARCH Model Information Criterion Statistics

Table 27 illustrates that the GARCH (1, 1) model performs better when fitted to the S&P 500 as opposed to the JSE All Share Index. The GARCH model performs well regardless of how frequently the model is update. It should be noted that these statistics on their own do not give much information about the goodness of fit for the GARCH Model.

Jarque-Bera Test

		Statistic	p-value
S&P 500	**Chi^2**	164.6061 165.7962 164.9758 166.3091	0 0 0 0
JSE All Share Index	**Chi^2**	43.54943 44.68929 43.65776 45.63366	0 0 0 0

Table 28: Jarque-Bera Test

Table 28 displays the statistic is for the Jarque-Bera Test, which tests for goodness of fit against the normal distribution. As can be seen in the above table, both the S&P 500 and the JSE All Share index have a p-value less than 0.05 for the null hypothesis for normality is rejected at a 5% significance level. This confirms the notion that neither the S&P 500 or the JSE All Share Index follow a normal distribution. The residuals are still heavy-tailed, and therefore do not follow a normal distribution. See Appendix 3 for complete details on the Jarque-Bera Test.

8.2 Generalized Extreme Value Distribution

Once the daily returns have been fitted to the GARCH model, the residuals are extracted and fitted to either the Generalized Extreme Value Distribution or the Generalized Pareto Distribution. First, the results will be looked at when fitted to the Generalized Extreme Value Distribution.

Parameter Estimates for Generalized Extreme Value Distribution

	xi	sigma	mu
S&P 500	- 0.005802828	0.738623422	1.266321726
	- 0.03394043	0.74078216	1.34141930
	- 0.003323827	0.746484749	1.209516548
	- 0.002860601	0.747568767	1.225695532
JSE`	- 0.07941019	0.52832616	1.23506914
	- 0.09476781	0.54451321	1.19982862
	- 0.0726628	0.5021547	1.2479568
	- 0.07413556	0.50359340	1.24439713

Table 29: Parameter estimates for Generalized Extreme Value Distribution using Maximum Likelihood parameter estimation

As can be seen from Table 29, the shape parameter (xi) is negative regardless of how frequently the model is updated. This suggests the standardised residuals for both the S&P 500 and the JSE All Share index have left tails which can be described by a Weibull distribution.

Value at Risk and Expected Shortfall summary for GEV (p = 0.99)

	Value at Risk		Expected Shortfall	
	S&P 500	**JSE**	**S&P 500**	**JSE**
Mean	- 0.012295 - 0.012252 - 0.012178 - 0.012062	- 0.011508 - 0.011569 - 0.011499 - 0.011473	- 0.00001691 - 0.00008158 - 0.00002659 - 0.00000196	- 0.0019745 - 0.0017084 - 0.0018508 - 0.0018683
Maximum	- 0.067559 - 0.064526 - 0.061665 - 0.061665	- 0.037230 - 0.035925 - 0.033761 - 0.031562	- 0.01210814 - 0.00992371 - 0.00859808 - 0.00859807	- 0.0109319 - 0.0109319 - 0.0109319 - 0.0109105

Table 30: Value at Risk and Expected Shortfall summary for GEV (p = 0.99)

Value at Risk and Expected Shortfall summary for GEV (p = 0.975)

	Value at Risk		Expected Shortfall	
	S&P 500	**JSE**	**S&P 500**	**JSE**
Mean	- 0.009738 - 0.009686 - 0.009582 - 0.009491	- 0.009532 - 0.009592 - 0.009538 - 0.009515	- 0.0001298 - 0.00003244 - 0.00008266 - 0.0001096	- 0.0020605 - 0.0018008 - 0.0019401 - 0.0019569
Maximum	- 0. 053052 - 0.050804	- 0. 031679 - 0.030643	- 0. 0124652 - 0.01029696	- 0. 0110996 - 0.0110996

	- 0.047567 - 0.047567	- 0.028303 - 0.026513	- 0.00893186 - 0.0089319	- 0.0110996 - 0.0110574

Table 31: Value at Risk and Expected Shortfall summary for GEV (p = 0.975)

Value at Risk and Expected Shortfall summary for GEV (p = 0.95)

	Value at Risk		**Expected Shortfall**	
	S&P 500	**JSE**	**S&P 500**	**JSE**
Mean	- 0.007708 - 0.007650 - 0.007541 - 0.007471	- 0.007952 - 0.007986 - 0.007956 - 0.007938	- 0.0003032 - 0.0002074 - 0.0002507 - 0.0002751	- 0.0021930 - 0.0019429 - 0.0020773 - 0.0020931
Maximum	- 0. 042620 - 0.040735 - 0.037508 - 0.037508	- 0. 027179 - 0.026221 - 0.024199 - 0.022758	- 0. 0130272 - 0.0108824 - 0.0094580 - 0.0094580	- 0. 0113615 - 0.0113615 - 0.0113615 - 0.0112873

Table 32: Value at Risk and Expected Shortfall summary for GEV (p = 0.95)

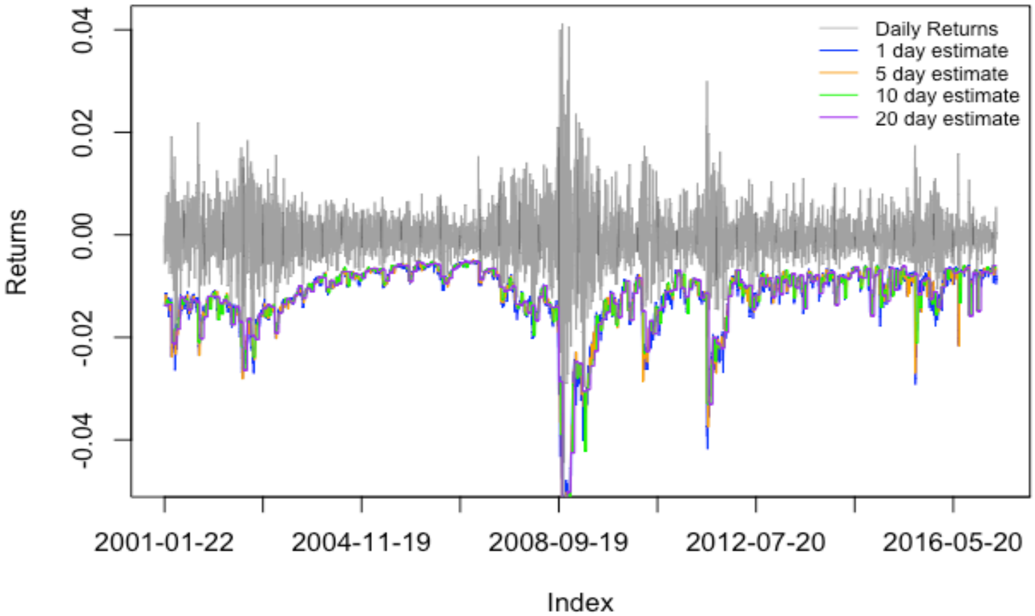

Figure 36: GEV Value at Risk estimations at 99th quantile using S&P 500 data

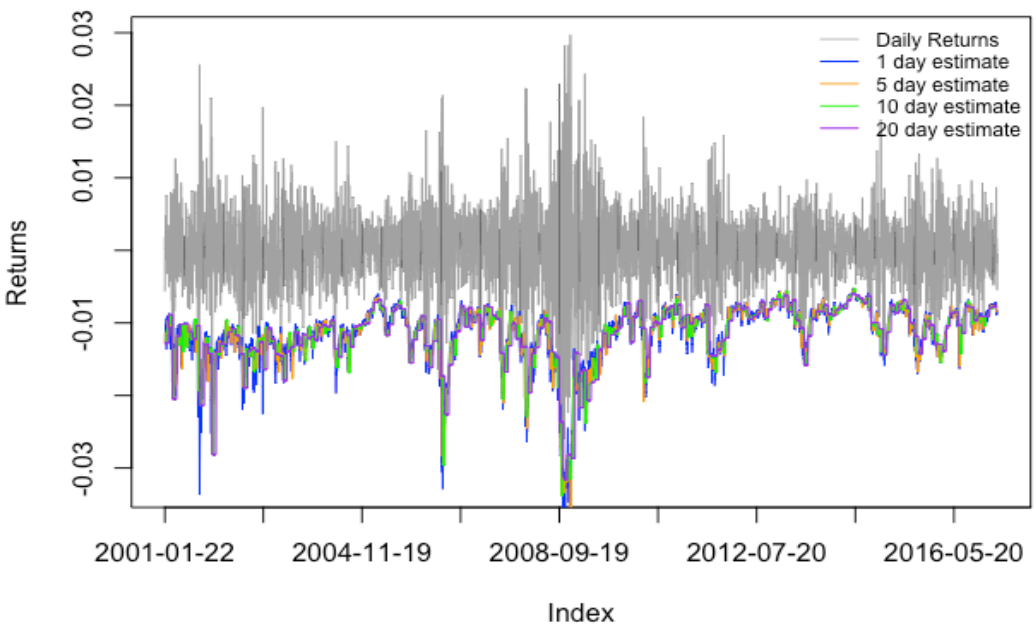

Figure 37: GEV Value at Risk estimations at 99th quantile using JSE All Share Index data

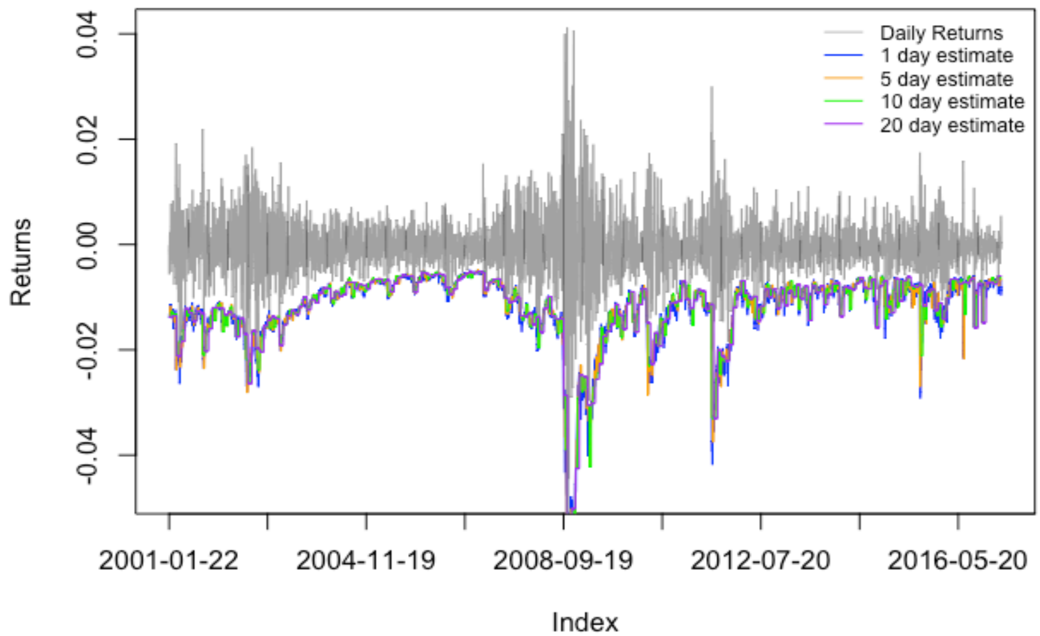

Figure 38: GEV Expected Shortfall estimations at 99th quantile using S&P 500 data

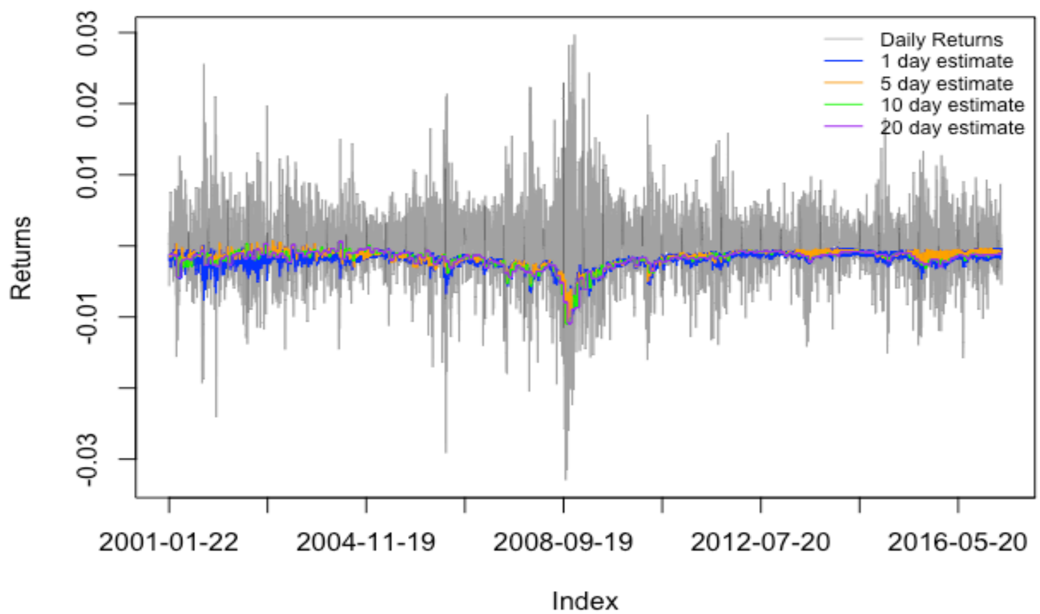

Figure 39: GEV Expected Shortfall estimations at 99th quantile using JSE All Share Index data

The tables above show the Value at Risk and Expected Shortfall estimates for both the S&P 500 and the JSE All Share Index when calculated using the block maxima method at different estimation frequencies.

Table 30Table 12 show that at the 99th quantile, the average losses expected to be 0.012% for the S&P 500, and 0.011% for the JSE All Share Index regardless of how frequently the model is update. Table 31 shows the 97.5th quantile, for both the S&P 500 and the JSE All Share Index. In both cases there is little difference between when the model is updated daily or monthly (every 20 days). At the 95th quantile is where there is a marked difference when the results are updated less frequently.

The changes in the Value at Risk over time can be seen in Figure 36: GEV Value at Risk estimations at 99th quantile using S&P 500 data and Figure 37: GEV Value at Risk estimations at 99th quantile using JSE All Share Index data. In both graphs, the Value at Risk estimates are very similar regardless of how frequently the model is updated. Figure 38: GEV Expected Shortfall estimations at 99th quantile using S&P 500 data and Figure 39: GEV Expected Shortfall estimations at 99th quantile using JSE All Share Index data, show the changes to Expected Shortfall overtime. Again, the Expected Shortfall estimates are very similar regardless of how frequently the model is updated.

8.2.1 Generalized Extreme Value Distribution Backtesting Results

Backtesting results for the Generalized Extreme Value Distribution (using the Block Maxima Method) for the S&P 500 can be seen in the following table:

Quantiles	**0,99**	**0.975**	**0.95**
Expected	42	105	210
Actual 1 day	45	117	218
Unconditional convergence p-value	0.667	0.264	NaN
Conditional convergence p-value	0.732	0.347	NaN
Expected Shortfall p-value	0	0	0
Actual 5 day	41	110	219
Unconditional convergence p-value	0.852	0.659	NaN
Conditional convergence p-value	0.708	0.093	NaN
Expected Shortfall p-value	0	0	0
Actual 10 day	49	124	236
Unconditional convergence p-value	0.305	0.076	NaN
Conditional convergence p-value	0.040	0.010	NaN
Expected Shortfall p-value	0	0	0
Actual 20 day	61	136	250
Unconditional convergence p-value	0.006	0.004	NaN
Conditional convergence p-value	0.005	0.002	NaN
Expected Shortfall p-value	0	0	0

Table 33: Results for Value at Risk of Generalized Extreme Value Distribution using S&P 500 data.

Table 33 contains a summary of the estimated tail quantiles at different update frequencies for the Generalized Extreme Value distribution for the S&P 500 data. At the 99th quantile there are 42 expected exceedances. When the model is updated daily there were 45 exceedances, 41 exceedances when the model is updated weekly, 49 exceedances when the model is updated every two weeks, and 61 exceedances when the model is updated monthly.

At the 97,5th quantile there are 105 expected exceedances. On daily updates there were 117 exceedances, when the model is updated once a week there were 110 exceedances. When the model is updated fortnightly, there are 124 exceedances, and 136 exceedances when the model is updated monthly.

Lastly, at the 95th quantile, there are 210 expected exceedances. When the model is updated daily there are 218 exceedances, 219 exceedances when the model is updated weekly, 236 exceedances when the model is updated every two weeks, and 250 exceedances when the model is updated once a month.

Backtesting results for the Generalized Extreme Value Distribution (using the Block Maxima Method) for the JSE All Share Index can be seen in the following table:

Quantiles	**0,99**	**0.975**	**0.95**
Expected	42	105	210
Actual 1 day	35	95	177
Unconditional convergence p-value	0.252	0.294	0.014
Conditional convergence p-value	0.387	0.065	0.014
Expected Shortfall p-value	0	0	0
Actual 5 day	31	78	155
Unconditional convergence p-value	0.069	0.004	0.009
Conditional convergence p-value	0.152	0.016	0.016
Expected Shortfall p-value	0	0	0
Actual 10 day	42	93	173
Unconditional convergence p-value	0.975	0.209	0.006
Conditional convergence p-value	0.655	0.321	0.017
Expected Shortfall p-value	0	0	0
Actual 20 day	43	103	193
Unconditional convergence p-value	0.902	0.804	NaN
Conditional convergence p-value	0.637	0.654	NaN
Expected Shortfall p-value	0	0	0

Table 34: Results for Value at Risk of Generalized Extreme Value Distribution using JSE All Share Index data.

Table 34 contains a summary of the estimated tail quantiles at different update frequencies for the Generalized Extreme Value distribution for the JSE All Share

Index. At the 99th quantile there are 42 expected exceedances. When the model is updated daily there were 35 exceedances, 31 exceedances when the model is updated weekly, 42 exceedances when the model is updated every two weeks, and 43 exceedances when the model is updated monthly.

At the 97,5th quantile there are 105 expected exceedances. On daily updates there were 95 exceedances, when the model is updated once a week there were 78 exceedances. When the model is updated fortnightly, there are 93 exceedances, and 103 exceedances when the model is updated monthly.

Lastly, at the 95th quantile, there are 210 expected exceedances. When the model is updated daily there are 177 exceedances, 155 exceedances when the model is updated weekly, 173 exceedances when the model is updated every two weeks, and 193 exceedances when the model is updated once a month.

8.3 Generalized Pareto Distribution

According to Coles (2004) the peaks-over-threshold method with the Generalized Pareto Distribution is a better alternative than the block maxima method since it uses more information. In this section the residuals from the GARCH models are extracted and fitted to the Generalized Pareto Distribution.

Parameter Estimates for Generalized Pareto Distribution

	xi	**beta**
S&P 500	- 0.05394545	0.77770225
	- 0.05813888	0.78998869
	- 0.04050225	0.75884109
	- 0.04125494	0.76029277
JSE All Share Index	- 0.02701273	0.50105354
	- 0.02012265	0.49475982
	- 0.01664065	0.47976237
	- 0.02758233	0.49047831

Table 35: Parameter estimates for Generalized Pareto Distribution using Maximum Likelihood parameter estimation

Table 35 shows the maximum likelihood parameter estimates for the Generalized Pareto Distribution. The shape parameter (xi) is negative when fitted using the S&P 500 and the JSE All Share Index data. The beta represents the scale parameter of the Generalized Pareto Distribution.

Value at Risk and Expected Shortfall summary for GPD (p = 0.99)

	Value at Risk		**Expected Shortfall**	
	S&P 500	**JSE**	**S&P 500**	**JSE**
Mean	- 0.012267 - 0.012252 - 0.012209 - 0.012092	- 0.011336 - 0.011346 - 0.011270 - 0.011248	- 0.014968 - 0.014952 - 0.014900 - 0.014763	- 0.013577 - 0.013585 - 0.013495 - 0.013469
Maximum	- 0. 068507 - 0.065369 - 0.062194 - 0.062194	- 0. 037029 - 0.035876 - 0.033789 - 0.031577	- 0. 085438 - 0.081528 - 0.078156 - 0.078156	- 0. 042760 - 0.041070 - 0.039160 - 0.036617

Table 36: Value at Risk and Expected Shortfall summary for GPD (p = 0.99)

Value at Risk and Expected Shortfall summary for GPD (p = 0.975)

	Value at Risk		**Expected Shortfall**	
	S&P 500	**JSE**	**S&P 500**	**JSE**
Mean	- 0.009769 - 0.009755 - 0.009720 - 0.009627	- 0.009278 - 0.009288 - 0.009223 - 0.009206	- 0.012489 - 0.012473 - 0.012430 - 0.012313	- 0.011522 - 0.011532 - 0.011454 - 0.011432
Maximum	- 0. 053896 - 0.051475	- 0. 031138 - 0.030255	- 0. 070209 - 0.067009	- 0. 037310 - 0.036043

	- 0.048479 - 0.048479	- 0.028254 - 0.026488	- 0.063818 - 0.063818	- 0.034048 - 0.031852

Table 37: Value at Risk and Expected Shortfall summary for GPD (p = 0.975)

Value at Risk and Expected Shortfall summary for GPD (p = 0.95)

	Value at Risk		**Expected Shortfall**	
	S&P 500	**JSE**	**S&P 500**	**JSE**
Mean	- 0.007815 - 0.007802 - 0.007773 - 0.007700	- 0.007712 - 0.007722 - 0.007666 - 0.007653	- 0.010588 - 0.010573 - 0.010535 - 0.010436	- 0.009964 - 0.009974 - 0.009905 - 0.009886
Maximum	- 0.043199 - 0.041322 - 0.038462 - 0.038462	- 0. 026384 - 0.025578 - 0.023778 - 0.022423	- 0. 059059 - 0.056400 - 0.053345 - 0.053345	- 0. 032910 - 0.031862 - 0.029914 - 0.028045

Table 38: Value at Risk and Expected Shortfall summary for GPD (p = 0.95)

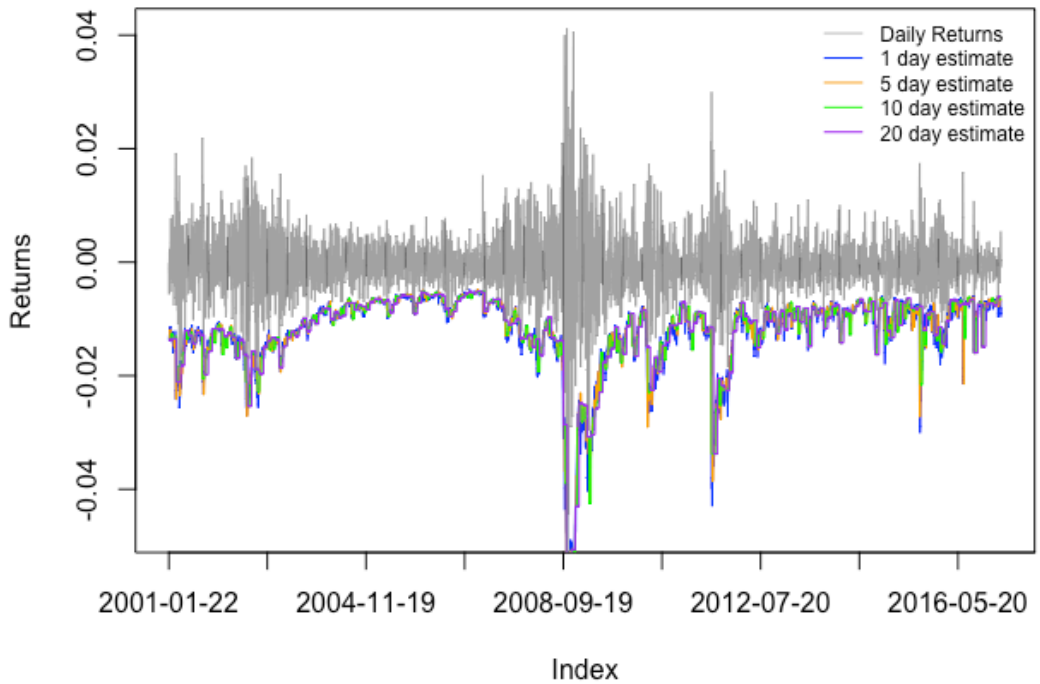

Figure 40: GPD Value at Risk estimations at 99th quantile using the S&P 500 data

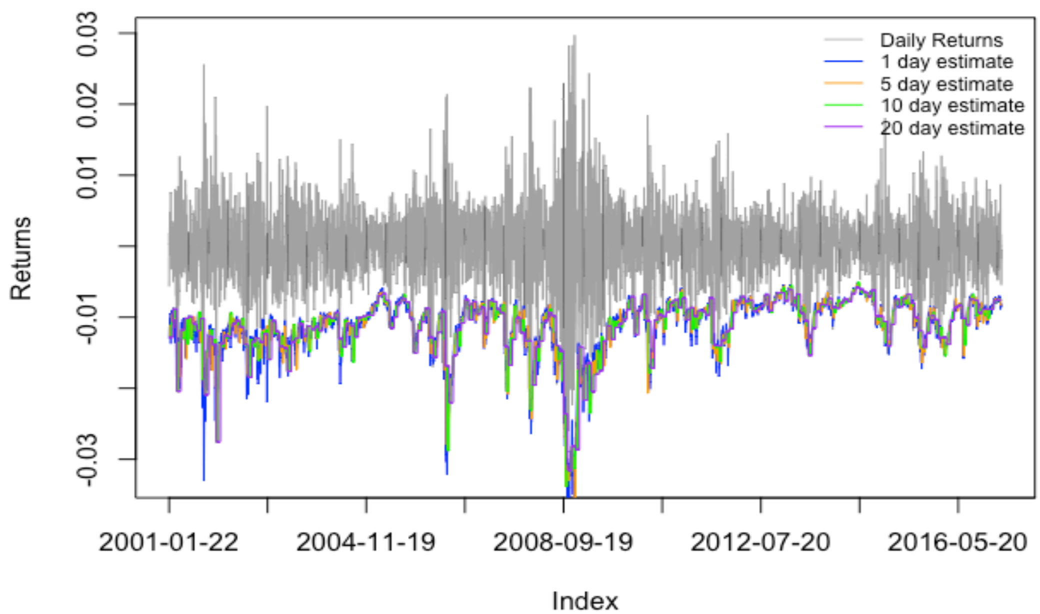

Figure 41: GPD Value at Risk estimations at 99th quantile using the JSE All Share Index data

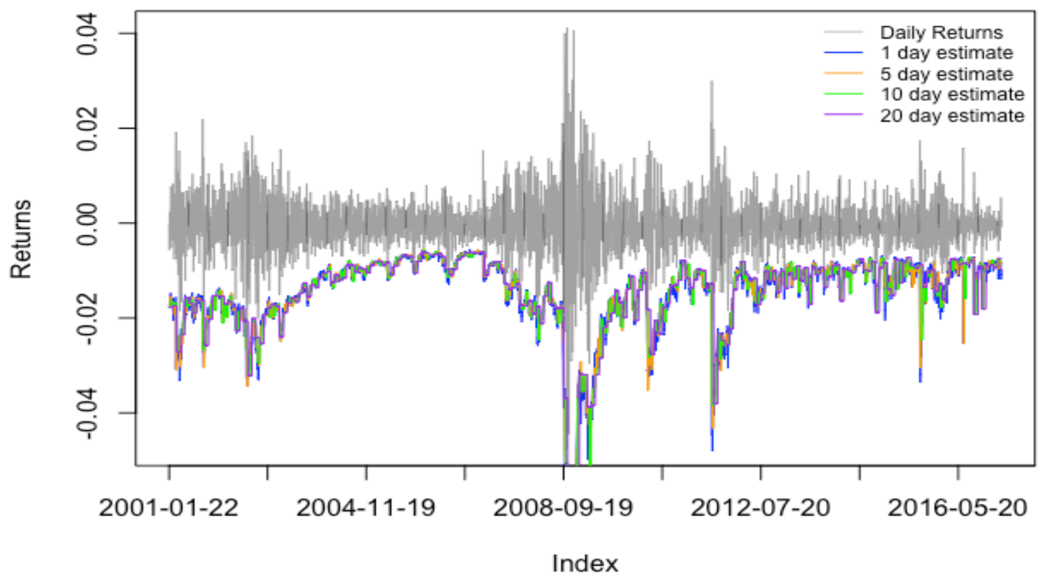

Figure 42: GPD Expected Shortfall estimations at 99th quantile using the S&P 500 data

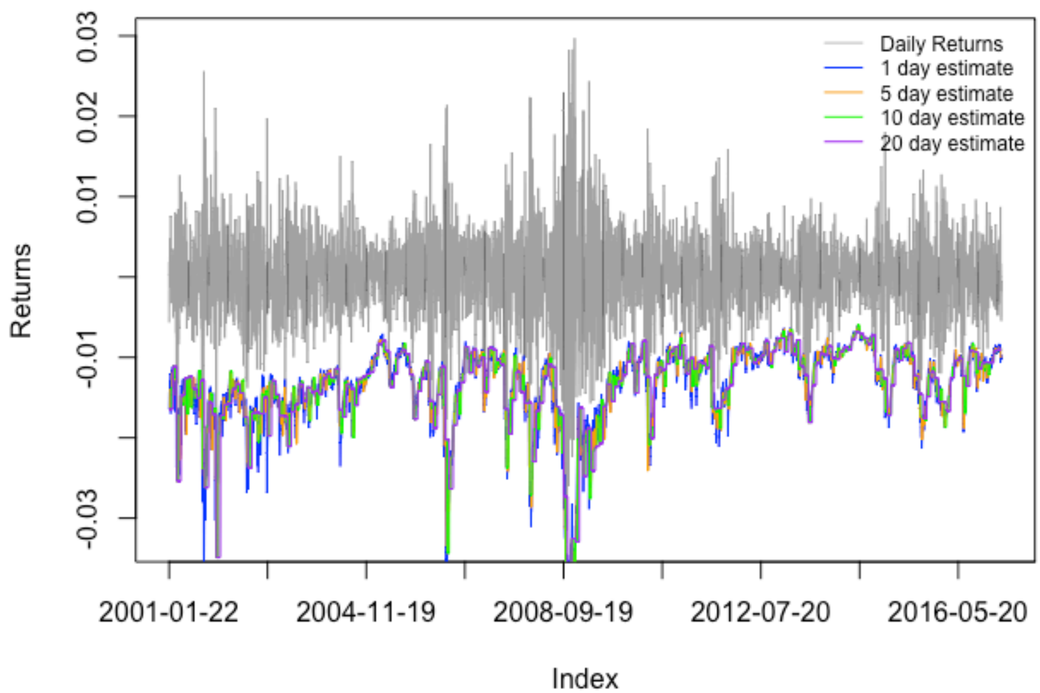

Figure 43: GPD Expected Shortfall estimations at 99th quantile using the JSE All Share Index data

The Value at Risk and Expected Shortfall estimates for the different update frequencies can be seen in the tables above. These estimates are calculated using the peaks over threshold method on both the S&P 500 and the JSE All Share Index data sets.

Table 36 show that at the 99th quantile, the average losses expected to be 0.012% for the S&P 500, and 0.011% for the JSE All Share Index regardless of how frequently the model is update. Table 37 shows very similar results at the 97.5th quantile, for both the S&P 500 and the JSE All Share Index. In both cases there is little difference between when the model is updated daily or monthly (every 20 days). As with the Generalized Extreme Value distribution, at the 95th quantile is where there is a noticeable difference in the Value at Risk and Expected Shortfall estimates when the model is updated less frequently.

The changes in the Value at Risk over time can be seen in Figure 40 and Figure 41. In both graphs, the Value at Risk estimates are very similar regardless of how frequently the model is updated. Figure 42 and Figure 43, show the changes to Expected Shortfall overtime. As with the Generalized Extreme Value Distribution, the Expected Shortfall estimates are very similar regardless of how frequently the model is updated.

8.3.1 Generalized Pareto Distribution Backtesting Results

Backtesting results for the Generalized Pareto Distribution (using the Peaks over threshold method) for the S&P 500 data can be seen in the following table:

Quantiles	**0,99**	**0.975**	**0.95**
Expected	42	105	210
Actual 1 day	48	119	207
Unconditional convergence p-value	0.379	0.191	NaN
Conditional convergence p-value	0.581	0.084	NaN
Expected Shortfall p-value	0.474	0.870	0.450
Actual 5 day	42	111	204
Unconditional convergence p-value	0.975	0.174	NaN
Conditional convergence p-value	0.741	0.037	NaN
Expected Shortfall p-value	0.850	0.932	0.853
Actual 10 day	48	121	221
Unconditional convergence p-value	0.380	0.135	NaN
Conditional convergence p-value	0.042	0.012	NaN
Expected Shortfall p-value	0.401	0.605	0.523
Actual 20 day	61	132	235
Unconditional convergence p-value	0.006	0.012	NaN
Conditional convergence p-value	0.005	0.004	NaN
Expected Shortfall p-value	0.007	0.037	0.035

Table 39: Results for Value at Risk of Generalized Pareto Distribution using the S&P 500 data.

Table 39 contains a summary of the estimated tail quantiles at different update frequencies for the Generalized Pareto Distribution for the S&P 500 data. At the 99th quantile there are 42 expected exceedances. When the model is updated daily there were 48 exceedances, 42 exceedances when the model is updated weekly, 48 exceedances when the model is updated every two weeks, and 61 exceedances when the model is updated monthly.

At the 97,5th quantile there are 105 expected exceedances. On daily updates there were 119 exceedances, when the model is updated once a week there were 111 exceedances. When the model is updated fortnightly, there are 121 exceedances, and 132 exceedances when the model is updated monthly.

Lastly, at the 95th quantile, there are 210 expected exceedances. When the model is updated daily there are 207 exceedances, 204 exceedances when the model is updated weekly, 221 exceedances when the model is updated every two weeks, and 235 exceedances when the model is updated once a month.

Backtesting results for the Generalized Pareto Distribution (using the Peaks over threshold method) for the JSE All Share Index can be seen in the following table:

Quantiles	**0,99**	**0.975**	**0.95**
Expected	42	105	210
Actual 1 day	40	102	199
Unconditional convergence p-value	0.732	0.730	NaN
Conditional convergence p-value	0.643	0.075	NaN
Expected Shortfall p-value	0.403	0.473	0.324
Actual 5 day	35	90	174
Unconditional convergence p-value	0.251	0.117	0.007
Conditional convergence p-value	0.386	0.222	0.025
Expected Shortfall p-value	0	0	0
Actual 10 day	45	106	196
Unconditional convergence p-value	0.668	0.961	NaN
Conditional convergence p-value	0.562	0.909	NaN
Expected Shortfall p-value	0	0	0
Actual 20 day	47	114	216
Unconditional convergence p-value	0.466	0.408	NaN
Conditional convergence p-value	0.451	0.622	NaN
Expected Shortfall p-value	0	0	0

Table 40: Results for Value at Risk of Generalized Pareto Distribution using JSE All Share Index data.

Table 40 contains a summary of the estimated tail quantiles at different update frequencies for the Generalized Pareto Distribution for the JSE All Share Index. At the 99th quantile there are 42 expected exceedances. When the model is updated daily there were 40 exceedances, 35 exceedances when the model is updated weekly, 45 exceedances when the model is updated every two weeks, and 47 exceedances when the model is updated monthly.

At the 97,5th quantile there are 105 expected exceedances. On daily updates there were 102 exceedances, when the model is updated once a week there were 90 exceedances. When the model is updated fortnightly, there are 106 exceedances, and 114 exceedances when the model is updated monthly.

Lastly, at the 95th quantile, there are 210 expected exceedances. When the model is updated daily there are 199 exceedances, 174 exceedances when the model is updated weekly, 196 exceedances when the model is updated every two weeks, and 216 exceedances when the model is updated once a month.

Chapter 9 – Discussion and Conclusion

This dissertation has focused on the application of Extreme Value Theory to both the S&P 500 and the JSE All Share index. The daily returns from the two different data sets is fitted to the GARCH (1,1) model, and the residuals are extracted. From here the residuals are fitted to one of two heavy tailed distributions, using the two-step process suggested by McNeil & Frey (2000).

Once the residuals have been fitted to either the Generalized Extreme Value Distribution or the Generalized Pareto Distributions, the Value at Risk and Expected Shortfall estimates are calculated. Next, the Value at Risk and Expected Shortfall estimates are backtested to determine how well the extreme value models performed. Finally, the models are examined at different update frequencies.

One of the key finding is that the model which is updated every 5 days performs just as well, and sometimes better than the model which is updated daily. This corresponds to the results from Ardia & Hoogerheide (2013). This could be cause by some underlying weekly trend. Further analysis would be needed to determine exactly how frequently a model needs to be updated.

Only the GARCH (1,1) model was considered in this study. Using a more advanced GARCH model could lead to better prediction. As well as looking different window sizes. In this study, a rolling window of 1000 days is used. The window size might also effect the results, since when the model is updated monthly, it might require a wider window to produce better results. It would be interesting to determine the optimal window size for each different update frequency.

Further investigation will be needed to determine exactly how frequently a model needs to be updated, depending on the volatility of the markets at the time. Ideally, if the model could change its update frequency depending on how volatile the markets are. When the markets are fairly calm, updating the model every 20 days produces sufficient results. Conversely, when the markets are experiencing high volatility, the model needs to be updated more frequently since.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 716-723.

Angus, J. E. (1993). Asymptotic theory for bootstrapping the extremes. *Communication in Statistics - Theory and Methods*, 15-30.

Ardia, D. (2008). *Financial Risk Management with Bayesian Estimation of GARCH Models.* Berlin: Springer.

Ardia, D., & Hoogerheide, L. (2014). GARCH models for daily stock returns: Impact of estimation frequency on Value-at-Risk and Expected Shortfall forecasts. *Tinbergen Institute Discussion Paper*, 1-16.

Bailey, D. H., Borwein, J. M., Lopez de Prado, M., & Jim Zhu, Q. (2014). Pseudo-Masthematics and Financial Charlatanism: The Effects of Backtest Overfitting on Out-of-Sample Performance. *Notices of the AMS*, 458-471.

Bali, T. G. (2007). A Generalized Extreme Value Approach. *Journal of Money, Credit and Banking, 39*(7), 1613-1649.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Economics*, 307-327.

Broberg, K. B. (1997). Encounters with Professor Waloddi Weibull. In H. P. Rossmanith, & G. R. Irwin, *Fracture Research in Retrospect: An anniversary volume in honour of G.R. Irwin's 90th birthday* (pp. 141-147). Rotterdam: CRC Press.

Brooks, C. (2008). *Introductory Econometricsfor Finance.* Cambridge: Cambridge University Press.

Bru, B., & Hertz, S. (2001). Maurice Frechet. In C. C. Heyde, & E. Seneta, *Statisticians of the centuries* (pp. 331-334). New york: Springer.

Bystrom, H. (2004). Managing extreme risks in tranquil and volatile markets using conditional extreme value theory. *International Review of Financial Analysis*, 133-152.

Chan, N. H., Deng, S.-J., Peng, L., & Xia, Z. (2007). Interval estimation of Value at Risk based on GARCH models with heavy tailed innovations. *Journal of Econometrics*, 556-576.

Chang, C.-L., Jimenez-Martin, J.-A., Maasoumi, E., Mc Aleer, M., & Perez-Amaral, T. (2015). Choosing Expected Shortfall over VaR in Basel III Using Stochastic Dominance . *Tinbergen Institute Discussion Paper* , 1-40.

Canty, A., & Ripley, B. (2019, April 2). Bootstrap Functions.

Charpentier, A. (2015). *Computational Actuarial Science with R.* Florida: CRC Press.

Chernick, M. R., & LaBudde, R. A. (2011). *An introduction to bootstrap methods with applications to R.* Hoboken: John Wiley & Sons, Inc.

Chinhamu, K., Huang, C.-K., Huang, C.-S., & Chikobvu, D. (2015, Februarg). Extreme Risk, Value-At-Risk and Expected Shortfall in the Gold Market. *International Business & Economics Research Journal, 14*(1), 107-122.

Christoffersen, P. F. (1998). Evaluating interval forcasts. *International Economic Review, 39*(4), 841-862.

Christoffersen, P., & Pelletier, D. (2004). Backtesting Value-at-Risk: ADuration-Based Approach. J*ournal of Financial Econometrics*, 84-108.

Coles, S. (2004). *An Introduction to Statistical Modeling of Extreme Values.* Bristol: Springer.

Davenport, T. (1991). *South Africa: A Modern History.* London: Macmillan Press LTD.

David, H. A., & Edwards, A. W. (2001). *Annotated Readings in the History of Statistics.* New York: Springer.

Dowd, K. (2005). R-version of Kevin Dowd's MATLAB Toolbox from book "Measuring Market Risk".

Efron, B. (1979). Bootstrap Methods: Another look at the Jackknife. *The Annals of Statistics*, 1-26.

Engle, R. (1982). A general approach to lagrange multiplier model diagnostics. *Journal of Econometrics*, 83-104.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 987-1008.

Feinstein, C. H. (2005). *An Economic History of South Africa.* Cambridge: Cambridge University Press.

Feng, L., & Shi, Y. (2017). A simulation study on the distributions of disturbances in the GARCH model. *Cogent Economics & Finance*, 1-19.

Francq, C., & Zakoian, J.-M. (2010). *GARCH Models: Structure, Statistical Inference and Financial Applications.* Sussex: John Wiley & Sons Ltd.

Fréchet, M. R. (1927). Sur la loi de probabilité de l'écart maximum. *Annales de la Societe Polonaise de Mathematique*, 93 - 116.

de Haan, L., & Ferreira, A. (2006). *Extreme Value Theory: An Introduction.* New York: Springer.

Hamann, R., Khagram, S., & Rohan, S. (2008). South Africa's charter approach to socioeconomic transformation: Collaborative governance or hardball bargaining? . *Journal of Southern African Studies*, 21-37.

Hamdi, Y., Bardet, L., Duluc, C.-M., & Rebour, V. (2015). Use of historical information in extreme-surge frequency estimation: the case of marine flooding on the La Rochelle site in France. *Natural hazards and earth system sciences*, 1515-1531.

Harmantzis, F., Miaso, L., & Chien, Y. (2006). Empirical study of value-at-risk and expected shortfall models with heavy tails. *The Journal of Risk Finance, 7*(2), 117-135.

Hertz, S. (2001). Emil Julius Gumbel. In J. M. Gani, C. C. Heyde, & E. Seneta, *Statisticians of the Centuries* (pp. 406-410). New York: Springer.

Hu, W., & Kercheval, A. (2007). Risk Management with Generalized HyperbolicDistribution. *FEA '07 Proceedings of the Fourth IASTED International* (pp. 19-24). California: Conference on Financial Engineering and Applications.

Jorion, P. (2007). *Value at Risk: The new benchmark for managing financial risk.* New York: The McGraw-Hill Companies, Inc.

Künsch, H. R. (1989). The Jackknife And The Bootstrap For General Stationary Observations. *The Annals of Statistics*, 1217-1241.

Kupiec, P. H. (1995). Techniques for Verifying the Accuracy of Risk Measurement Models. *The Journal of Derivatives*, 73-84.

Laker, I., Huang, C.-K., & Clark, A. E. (2017). Dependent bootstrapping for value-at-risk and expected shortfall. *Risk Management*, 301-322.

Ljung , G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, 297-303.

Mc Grattan, E. R., & Waddle, A. (2017). The Impact of Brexit on Foreign Investment and Production. *National Bureau of Economic Research*, 1-32.

McNeil, A., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approuch. *Journal of Emperical Finance, 7*, 271-300.

Miller, M. B. (2012). *Mathematics and Statistics for Financial Risk Management.* New Jersey: John Wiley & Sons, Inc.

Milwidsky, C., & Mate, E. (2010). Value at Risk in the South African Equity Market: A view from the tails. *SAJEMS NS*, 345-361.

Moody's Analystics. (2013). *Moody's Analystics.* Retrieved from Moody's Analytics, Inc. All Rights Reserved.: https://www.moodysanalytics.com/-/media/article/2013/2013-18-10-basel-iii-capital-and-liquidity-standards-faq.pdf

Nadarajah, S., Chan, S., & Afuecheta, E. (2013). *VaRES: Computes value at risk and expected shortfall for over 100 parametric distributions*. Retrieved from VaRES: Computes value at risk and expected shortfall for over 100 parametric distributions: https://CRAN.R-project.org/package=VaRES

Peters, G., & Shevchenko, P. (2015). *Advances in Heavy Tailed Risk Modeling: A Handbook of Operational Risk.* New Jersey: John Wiley & Sons.

Pfaff, B., & Mc Neil, A. (2012). *evir: Extreme Values in R*. Retrieved from evir: Extreme Values in R: https://CRAN.R-project.org/package=evir

Pfaaff, B. (2016). *Financial Risk Modelling and Portfolio Optimization with R.* Chichester: John Wiley & Sons, Ltd.

Qi, Y. (2008). Bootstrap and empirical likelihood methods in extremes. *Extremes*, 81-97.

R Core Team. (2017). R: A Language and Environment for Statistical Computing. *https://www.R-project.org/*. Vienna, Austria: R Foundation for Statistical Computing.

Ruppert, D. (2004). *Statistics and Finance: An Introduction.* New York: Springer.

Schmitt, T., Chetalova, D., Schafer, R., & Guhr, T. (2013, September). Non–Stationarity in Financial Time Series and Generic Features. *EPL (Europhysics Letters), 103*(5), 58003.

Singh, A., Allen, D., & Robert, P. (2013). Extreme market risk and extreme value theory. *Mathematics and Computers in Simulation*, 310-328.

Tsay, R. S. (2013). *An Introduction to Analysis of Financial Data with R.* New Jersey: John Wiley and Sons.

Tsay, R. S. (2002). *Analysis of Financial Time Series.* New York: John Wiley & Sons, Inc.

Van De Venter, T. (2000). *Backtesting Historical Simulation Value-at-Risk for a selected portfolio of South African bonds.* University of Pretoria. Pretoria: University of Pretoria.

Writer, S. (2017, February 28). *Why Brexit is a problem for South Africa.* Retrieved from Bisiness Tech: https://businesstech.co.za/news/finance/160959/why-brexit-is-a-problem-for-south-africa/

Wuertz, D., Setz, T., Chalabi, Y., Boudt, C., Chausse, P., & Miklovac, M. (2017). *fGarch: Rmetrics - Autoregressive Conditional Heteroskedastic Modelling*.

Retrieved from fGarch: Rmetrics - Autoregressive Conditional Heteroskedastic Modelling: https://CRAN.R-project.org/package=fGarch

Yang, X., Chaptpatanasiri, R., & Sattayatham, P. (2016). Dynamic Risk Measurement of Financial Time Series with Heavey-Tailed: A New Hybrid Approach. *Thai Journal of Mathematical*, 269.

Appendix 1

Extreme Value Theorem (Fisher-Tippett-Gnedenko theorem)

The class of extreme value distributions is:

$$G(z,\mu,\sigma,\xi) = \begin{cases} exp\left(-\left(1+\xi\left(\frac{z-\mu}{\sigma}\right)^{-1/\xi}\right)\right), & \xi \neq 0 \\ exp\left(-exp\left(-\frac{z-\mu}{\sigma}\right)\right), & \xi = 0 \end{cases}$$

Proof:

Let us consider the class of limit functions D where, $\lim_{t\to\infty} \frac{U(tz)-b(t)}{a(t)} = D(z)$, for each $z > 0$ continuity point of $D(z) = G \leftarrow \left(e^{-\frac{1}{z}}\right)$, $a(t) := a_{[t]}$, and $b(t) := b_{[t]}$.

First suppose that 1 is a continuity point of D. Then note that for continuity points for $x > 0$,

$$\lim_{t\to\infty} \frac{U(tz)-U(t)}{a(t)} = D(x) - D(1) =: E(x).$$

Take $y > 0$ and write

$$\frac{U(tzy)-U(t)}{a(t)} = \frac{U(tzy)-U(ty)}{a(ty)}\frac{a(ty)}{a(t)} + \frac{U(ty)-U(t)}{a(t)}.$$

We claim that $\lim_{t\to\infty} \frac{U(ty)-U(t)}{a(t)}$ and $\lim_{t\to\infty} \frac{a(ty)}{a(t)}$ exist. Suppose these limits do exist.

Then there are A_1, A_2, B_1, B_2 with $A_1 \neq A_2$ or $B_1 \neq B_2$, where B_i are limit points of $\frac{U(ty)-U(t)}{a(t)}$ and A_i are limit points of $\frac{a(ty)}{a(t)}$, $i = 1, 2$, as $t \to \infty$.

We find from

$$\frac{U(tzy) - U(t)}{a(t)} = \frac{U(tzy) - U(ty)}{a(ty)} \frac{a(ty)}{a(t)} + \frac{U(ty) - U(t)}{a(t)}$$

that $E(zy) = E(z)\, A_i + B_i$, $i = 1, 2$, for all z continuity points of $E(\cdot)$ and $E(\cdot\, y)$.
For an arbitrary z take a sequence of continuity points z_n with $z_n \uparrow z (n \to \infty)$.

Then $E(z_n y) \to E(zy)$ and $E(z_n) \to E(z)$ since E is left continuous for $i = 1, 2$ from each other one obtains $E(z)\, (A_1 - A_2) = B_2 - B_1$ for all $z > 0$. Since E cannot be constant (G in nondegenerate) we must have $A_1 = A_2$ and hence also $B_1 = B_2$.

In conclusion:

$A(y) := \lim_{t \to \infty} \frac{a(ty)}{a(t)}$ exists for $y > 0$, and for $x, y > 0$,

$$E(zy) = E(z)\, A(y) + E(y).$$

Hence for $s := \log(z)$, $t := \log(y)$, $(z, y \neq 1)$, and $H(z) := E(e^z)$, we have

$$H(t + s) = H(s) A(e^t) + H(t),$$

Since $H(0) = 0$, the equation can be written as:

$$\frac{H(t + s) - H(t)}{s} = \frac{H(s) - H(0)}{s} A(e^t).$$

There is certainly one t at which H is differentiable (since H is monotone);

hence by

$$\frac{H(t+s)-H(t)}{s} = \frac{H(s)-H(0)}{s} A(e^t)$$

H is differentiable everywhere and $H^{'}(t) = H^{'}(0)\, A(e^t)$.

Write $Q(t) := \frac{H(t)}{H'(0)}$. Note that $H^{'}(0)$ cannot be zero : H cannot be constant since G is nondegenerate.

Then $Q(0) = 0$, $Q'(0) = 1$.

By $H(t+s) = H(s)A(e^t) + H(t)$:

$$Q(t+s) - Q(t) = Q(s)A(e^t)$$

And by $H^{'}(t) = H^{'}(0)\, A(e^t)$:

$$Q(t+s) - Q(t) = Q(s)Q'(t).$$

Subtracting the same expression with t and s interchanged we get:

$$Q(t)\, \frac{Q'(s)-1}{s} = \frac{Q(s)}{s}\, (Q'(t)-1),$$

Hence (let $s \to 0$)

$$Q(t)\, Q''(0) = Q'(t) - 1 .$$

It follows that Q is twice differentiable, and by differentiation,

$$Q^{''}(0)\, Q^{'}(t) = Q^{''}(t).$$

Hence,

$$(\log(Q'))'(t) = Q''(0) =: \xi \in \mathbb{R}, \forall\, t.$$

Since $Q'(0) = 1$, it follows that

$$Q'(t) = e^{\xi\, t}$$

And since $Q(0) = 0$,

$$Q(t) = \int_0^t e^{\xi\, t}\, ds.$$

This means that

$$H(t) = H'(0)\frac{e^{\xi\, t} - 1}{\xi}$$

And

$$D(t) = D(1) + H'(0)\ \frac{t^{\xi} - 1}{\xi}.$$

Hence,

$$D\ \leftarrow (z) = \left(1 + \ \xi\ \frac{z - D(1)}{H'(0)}\right)^{1/\xi}.$$

Now $D(z) = G\ \leftarrow \left(e^{-1/z}\right)$, and hence

$$D \leftarrow (z) = \frac{1}{-\log\big(G(z)\big)}.$$

Finally, combining $D \leftarrow (z) = \left(1 + \xi \frac{z - D(1)}{H'(0)}\right)^{1/\xi}$ and $D \leftarrow (z) = \frac{1}{-\log(G(z))}$, we obtain the statement of the theorem.

For complete details of the proof, please see de Haan and Ferreira (2006, pp. 6-8).

Jarque-Bera Test

The Jarque-Bera test is a goodness of fit test developed by Carlos Jarque and Anil Bera. It tests whether the sample data have skewness and kurtosis corresponding to a normal distribution.

The test statistic is defined as follows:

$$JB = \frac{n-k+1}{6}\left(S^2 + \frac{1}{4}(C-3)^2\right)$$

Where n is the sample size.

S represents the skewness and is defined as:

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{n}\sum_{i-1}^{n}(x_i - \overline{x})^3}{\left(\frac{1}{n}\sum_{i-1}^{n}(x_i - \overline{x})^2\right)^{\frac{3}{2}}}$$

C is the kurtosis and is defined as:

$$C = \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{n}\sum_{i-1}^{n}(x_i - \overline{x})^4}{\left(\frac{1}{n}\sum_{i-1}^{n}(x_i - \overline{x})^2\right)^{2}}$$

and k is the number of independent variables.

Ljung-Box Test

The Ljung-Box Test was developed by Greta Ljung and George Box. It is used to test for autocorrelations in a time series. This determines if the time series is Independently Distributed.

The test statistic is defined as follows:

$$Q(L) = n(n+2)\sum_{k=1}^{L}\frac{\hat{\rho}_k^2}{n-k}$$

Where n is the sample size, L is the number of lags being tested, and $\hat{\rho}_k$ is the sample autocorrelations at lag L.

Under the null hypothesis the test statistic $Q(L)$ follows a χ_h^2 distribution.

ARCH Langrange Multiplier Test

The ARCH Langrange multiplier is used to test for autoregressive conditional heteroskedasticity in the standardized residuals. This is done using linear regression as follows:

$$LM = TR^2$$

Where T represents the sample size and R represents the sample multiple correlation coefficient obtained from:

$$z_t^2 = \alpha_0 + \alpha_1 z_{t-1}^2 + \cdots + \alpha_m z_{t-m}^2 + e_t$$

Where $t = m + 1, \ldots, T$.

Under the null hypothesis the test statistic follows a χ_m^2 distribution.

Akaike Information Criterion

The AIC measures how well the model fits the data relative to other models (Akaike, 1974). Each GARCH model is fitted to the S&P 500 or JSE All Share Index using the quasi maximum likelihood estimator, and then the AIC is computed for each model.

The AIC can be computed as follows:

$$AIC = 2k - 2\ln(\hat{L})$$

Where

$$k = \ number\ of\ parameters\ in\ the\ model$$

and

$$\hat{L} = \ the\ maximum\ value\ of\ the\ likelihood\ function\ for\ the\ model.$$

The model with the smallest AIC is considered to be the best model of the data (Tsay, 2013).

Bayesian Information Criterion

The Bayesian Information Criterion is closely related to the AIC. It is defined as follows:

$$BIC = \ln(n)\, k - 2 \ln\,(\hat{L})$$

Where

$$k = \ \ number\ of\ parameters\ in\ the\ model$$

and

$$\hat{L} = \ the\ maximum\ value\ of\ the\ likelihood\ function\ for\ the\ model$$

and

$$n = \ sample\ size.$$

Expected Shortfall (GARCH (1,1))

Let $\{X_t\}$ be a GARCH (1,1) process such that

$$X_t = X_{t-h} + r_t$$

Where $r_t = \overline{r_t} + \mu$, $\overline{r_t} = \sigma_t e_t$, $\sigma_t = \sqrt{\alpha_0 + \alpha_1 \bar{r}_{t-h}^2 + \beta_1 \sigma_{t-h}^2}$, and e_t are independent standard normal variables.

Let $\hat{\mu}$, $\hat{\alpha}_0$, $\hat{\alpha}_1$, and $\hat{\beta}_1$ denote the maximum likelihood estimates.

Let $\hat{\mu}^k = k\mu$ and $\hat{\sigma}^k = \hat{\sigma}(t,t)$ where $\hat{\sigma}^k = \hat{\sigma}(t,t)$ is defined as

$$\hat{\sigma}^2(t^*,t) = \hat{\alpha}_0 + \hat{\alpha}_1 \left(r_{t^*}^k - \hat{\mu}^k\right)^2 + \hat{\beta}_1 \hat{\sigma}^2 \ (t^* - kh, t)$$

$$\hat{\sigma}^2(t - nkh, t) = \frac{k}{nk-1} \sum_{i=0}^{nk-1} (r_{t-ih} - \hat{\mu})^2$$

For $t^* = t - (n-1)kh, \dots, \ t - kh, t$; where n is the number of k-period returns.

The Expected Shortfall can be estimated as follows

$$ES_p(x) = -1 + \frac{1}{p} \int_0^p \exp\left\{\hat{\mu}^k + \hat{\sigma}^k x_{\hat{v}^k}^q\right\} \ dq$$

Where x_v^q denotes the qth quantile of the Student's t distribution with v degrees of freedom.

Properties of Generalized Extreme Value Distribution

The Generalized Extreme Value distribution has the following probability density function:

$$f(z,\mu,\sigma,\xi) = \frac{1}{\sigma}\begin{cases} \left(1+\xi\left(\frac{z-\mu}{\sigma}\right)\right)^{-1/\xi-1} exp\left(-\left(1+\xi\left(\frac{z-\mu}{\sigma}\right)\right)^{-1/\xi}\right), & \xi \neq 0 \\ exp\left(-\frac{z-\mu}{\sigma}\right) exp\left(-exp\left(-\frac{z-\mu}{\sigma}\right)\right), & \xi = 0 \end{cases}$$

The quantiles associated with the upper tails of the GEV distribution can be obtained by inverting the Cumulative Distribution Function equation:

$F(z,\mu,\sigma,\xi) = \begin{cases} exp\left(-(1+\xi\left(\frac{z-\mu}{\sigma}\right)^{-1/\xi}\right), & \xi \neq 0 \\ exp\left(-\,exp\left(-\,\frac{z-\mu}{\sigma}\right)\right), & \xi = 0 \end{cases}$ and solving for $F(z_p) = 1-p$.

Solving this equation, it can easily be shown that:

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi}\left[1-\{-\ln(1-p)\}^{-\xi}\right], & for\ \xi \neq 0 \\ \mu - \sigma\ln\{-ln(1-p)\}, & for\ \xi = 0 \end{cases}.$$

The quantiles z_p, represent the daily return level for the corresponding return period $\frac{1}{p}$ and is associated with the tails of the distribution (Coles, 2004).

Properties of Generalized Pareto Distribution

The Generalized Pareto Distribution has the following probability density function:

$$f(z,\mu,\sigma,\xi) = \frac{1}{\sigma}\begin{cases} \left(1+\xi\left(\frac{z-\mu}{\sigma}\right)\right)^{-1/\xi-1} exp\left(-\left(1+\xi\left(\frac{z-\mu}{\sigma}\right)\right)^{-1/\xi}\right), & \xi \neq 0 \\ exp\left(-\frac{z-\mu}{\sigma}\right) exp\left(-exp\left(-\frac{z-\mu}{\sigma}\right)\right) & , \quad \xi = 0 \end{cases}$$

The quantiles associated with the upper tails of the GPD can be obtained by inverting the Cumulative Distribution Function equation:

$$F(z,\mu,\sigma,\xi) = \begin{cases} exp\left(-(1+\xi\left(\frac{z-\mu}{\sigma}\right)^{-1/\xi}\right), & \xi \neq 0 \\ exp\left(-exp\left(-\frac{z-\mu}{\sigma}\right)\right) & , \xi = 0 \end{cases}$$ and solving for

$$F(z_p) = 1-p.$$

Solving this equation, it can easily be shown that:

$$z_p = \begin{cases} \mathrm{u} - \frac{\sigma}{\xi}\left[\left(\frac{1-p}{N_u/n}\right)^{-\xi} - 1\right], & for\ \xi \neq 0 \\ \mathrm{u} - \sigma\log\left(\frac{1-p}{N_u/n}\right), & for\ \xi = 0 \end{cases}.$$

The quantiles z_p, represent the daily return level for the corresponding return period $\frac{1}{p}$ and is associated with the tails of the distribution. Where N_u represents the number of excesses, and u is the threshold (Coles, 2004).

Update frequency colour key

Blue – The model is updated daily
Orange – The model is updated every 5 days (weekly)
Green – The model is updated every 10 days (fortnightly)
Purple – The model is updated every 20 days (monthly)