MODELLING VOLATILITY AND FINANCIAL MARKET RISKS OF SHARES ON THE JOHANNESBURG STOCK EXCHANGE

by

MAKHWITING MONNYE RHODA

DISSERTATION Submitted in fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

in

STATISTICS

in the

FACULTY OF SCIENCE AND AGRICULTURE
(School of Mathematical and Computer Sciences)

at the

UNIVERSITY OF LIMPOPO

SUPERVISOR: Prof. M Lesaoana

CO-SUPERVISOR: Mr. C Sigauke

2014

DECLARATION

I declare that the dissertation hereby submitted to the University of Limpopo, for the degree of Master of Science in Statistics has not previously been submitted by me for a degree at this or any other university; that it is my work in design and in execution, and that all material contained herein has been duly acknowledged.

_____   _____

Makhwiting, MR (Ms)   Date

DEDICATION

This dissertation is dedicated to my parents, my younger sister Jacksinah and my two cousins Sheidah and Thato.

ACKNOWLEDGEMENT

ABSTRACT

A number of previous research studies have investigated volatility and financial risks in the ermeging markets. This dissertation investigates stock returns volatility and financial risks in the Johannesburg Stock Exchange (JSE). The investigation is conducted in modelling volatility using Autoregressive Moving Average-Generalised Autoregressive Conditional Heteroskedastic (ARMA-GARCH)-type models. Daily data of the log returns at the JSE over the period 08 January, 2002 to 30 December, 2011 is used. The results suggest that daily returns can be characterised by an ARMA (1, 0) process. Empirical results show that ARMA (1, 0)-GARCH (1, 1) model achieves the most accurate volatility forecast. Modelling tail behaviour of rare and extreme events is an important issue in the risk management of a financial portfolio. Extreme Value Theory (EVT) is applied to quantify upper extreme returns. Generalised Extreme Value (GEV) distribution is used to model the behaviour of extreme returns. Empirical results show that the Weibull distribution can be used to model stock returns on the JSE. In using the Generalised Pareto Distribution (GPD), the modelling framework used accommodates ARMA and GARCH models. The GPD is applied to ARMA-GARCH filtered returns series and the model is referred to as the ARMA-GARCH-GPD model. The threshold value is estimated using Pareto quantile plot while peak-over-threshold approach is used to model the upper extreme returns. In general, the ARMA-GARCH-GPD model produces more accurate estimates of extreme returns than the ARMA-GARCH model. The out of sample forecast indicates that the ARMA (1, 3)-GARCH (1, 1) model provides the most accurate results.

Contents

List of Figures

List of Tables

List of Acronyms

ACF	Autocorrelation Function
ACGARCG	Asymmetric Component Generalised Autoregressive Conditional Heteroskedasticity
ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criterion
ALSI	All Share Price Index
AR	Autoregressive
ARCH	Autoregressive Conditional Heteroskedasticity
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
AV	Autoregressive Volatility
CGARCH	Component Generalised Autoregressive Conditional Heteroskedasticity
EGARCH	Exponential Generalised Autoregressive Conditional Heteroskedasticity
EVI	Extreme Value Index
EVT	Extreme Value Theory
FIAPARCH	Fractionally Integrated Asymmetric Power Autoregressive Conditional Heteroskedasticity
GARCH	Generalised Autoregressive Conditional Heteroskedasticity
GARCH-M	Generalised Autoregressive Conditional Heteroskedasticity in Mean
GDP	Gross Domestic Product
GEV	Generalised Extreme Value
GJR	Glosten Jagannathan Runkle
GPD	Generalised Pareto Distribution
i.i.d.	independent and identically distributed
JB	Jaque-Bera
JSE	Johannesburg Stock Exchange

KSE	Khartoum Stock Exchange
LM	Lagrange Multiplier
Log	Natural Logarithm
MA	Moving Average
MAE	Mean Absolute Error
MAPE	Mean Absolute Percent Error
MLE	Maximum Likelihood Estimator
MLq	Maximum Linear Quadratic
MSE	Mean Square Error
PACF	Partial Autocorrelation Function
PGARCH	Power Generalised Autoregressive Conditional Heteroskedasticity
POT	Peak Over Threshold
QML	Quasi Maximum Likelihood
Q-Q	Quantile-Quantile
RMSE	Root Mean Square Error
RSE	Root Square Error
TASE	Tel Aviv Stock Exchange
TGARCH	Threshold Generalised Autoregressive Conditional Heteroskedasticity
VaR	Value-at-Risk
ZSE	Zimbabwe Stock Exchange

Research Outputs

Peer Reviewed Journal Publications

1. Makhwiting MR., Lesaoana M. and Sigauke C. 2012. Modelling volatility and financial market risk of shares on the Johannesburg Stock Exchange. African Journal of Business Management 6(27): 8065-8070.

2. Makhwiting MR., Sigauke C. and Lesaoana M. Modelling Tail Behavior of Returns using the Generalised Extreme Value Distribution. 2014. Economics, Management, and Financial Markets 9(1): 41-52.

Accepted for Publication

1. Makhwiting MR., Sigauke C. and Lesaoana M. Modelling Conditional Heteroskedasticity in JSE stock returns using Generalised Pareto Distribution. African Review of Economics and Finance.

Conferences Attended

1. Makhwiting MR., Lesaoana M., and Sigauke C. Volatility modelling of heteroskedastic shares using Generalised Pareto Distribution. 54th Annual Conference of the South African Statistical Association, 5-9 November 2012, Nelson Mandela Metropolitan University, PE, South Africa. `http://www.sasa2012.co.za/Images/Uploads/SASA_2012_program.pdf`

2. Makhwiting MR., Lesaoana M., and Sigauke C. Modelling Volatility and financial market risks of shares on the JSE. 53th Annual Conference of the South African Statistical Association, 31 October to 4 November, 2011, CSIR, Pretoria, South Africa. `http://www.sastat.org.za/sasa2011/images/SASA_2011_PROGRAM.PDF`

3. Makhwiting MR., Lesaoana M., and Sigauke C. Modelling Volatility and financial market risks of shares on the JSE. 40th Annual Conference of the Operations Research Society of South Africa, 18-21 September 2011, Elephant Hills Hotel, Victoria Falls, Zimbabwe. `http://www.orssa.org.za/wiki/uploads/Conf/2011ORSSAConferenceProgramme.pdf`

Chapter 1

Introduction and Background

1.1 Introduction

Stock market performance is crucial to businesses in South Africa. Studying stock market performance requires assessing constituents such as volatility modelling and financial market risks. It is therefore important to examine the Johannesburg Stock Exchange (JSE). The JSE is a full service securities exchange which provides trading, clearing and settlement of equities, derivatives, interest rate products and other associated instruments. The stock exchange was established in 1887 and has been experiencing numerous changes since then. The JSE is licensed as an exchange under the Securities Services Act, 2004 and Africa's premier exchange. It has operated as a market place for the trading of financial products for nearly 120 years.

The JSE does not only channel funds into the economy, but also provides investors with returns on investments in the form of dividends, analyses business information to identify areas of risk and makes recommendations on profitability models. The exchange is fulfilling its main function by rechanneling cash resources into the productive economic activity, thus building the economy while enhancing job opportunities and wealth creation. The exchange functions inside a proper regulatory framework that is adhered to by all market players and is carefully enforced by a regulatory act. The JSE has been based on self-regulation with rules and directives to protect the interests of the general public who are buying and selling shares.

The JSE All Share Top 40 Companies Index is an equity index intended to reflect the performance of the South African ordinary share market as a whole. The purpose of the index is to have a tool that will be able to describe the market at a given point in time in terms of price levels, dividend yield and earnings yields. The All Share Price Index (ALSI) is an equity index which mirrors the performance of the South African ordinary share market. The ALSI is benchmarked against global standards and is basically an indicator of the general mood of the market. The ALSI measures the performance of the overall market. A relatively big proportion of the total number of securities listed on the JSE are incorporated into the index, on the basis that movements in the share prices of those constituent companies can be said to represent the market as a whole.

JSE limited released results that showed flexibility of the stock exchange even with difficulties faced by all stock markets in 2008 which include increased volatility and declining investor sentiments. Despite the fact that stock markets globally faced extraordinary tests in 2008, the JSE performed well. The rising trend was a result of increased purchase of shares and competitive share prices. However, at the end of 2008, the market cap fell due to falling share prices and this was a signal of a recession. The share prices started to rise in the last quarter of 2009 and this was a sign that the world recession was easing.

Investing in financial markets is challenging because prices are highly volatile and exhibit extreme price movement of magnitude. Volatility is defined as the statistical measure of the dispersion of returns for a security or market index within a specific time horison. It can either be measured by using the standard deviation or variance between returns from that same security or market index. It is used to quantify the risk of the financial instrument over the specified time period. Risk is the possibility of losing some or all of the original investment. Generalised Autoregressive Conditional Heteroskedastic (GARCH) is a time series technique used to model the serial dependence of volatility. This study models volatility and financial market risks of

shares using GARCH-type models.

Extreme Value Theory (EVT) is the theory of measuring and modelling extreme events. It is especially well suited to describe the tails of the profits and losses distributions typically found in stock returns. Two main distributions for EVT analysis are Generalised Extreme Value (GEV) distribution and Generalised Pareto Distribution (GPD). Nowadays EVT has experienced a boom in the financial field, especially with respect to risk managements. In this study the GEV distribution has been applied to model tail behaviour of returns while GPD is used to model conditional heteroskedasticity in the JSE stock returns.

1.2   Research problem

Volatility has become a topic of enormous importance to almost anyone who is involved in financial markets and has been one of the most active and successful areas of research in time series econometrics and economic forecasting in recent decades. Financial markets across the world have seen increased volatility in recent periods. The problem is that investors and financial analysts are concerned about uncertainty of returns on their investment assets, caused by variability in speculative market prices and the instability of business performance. Investors are interested in the direct impact of time varying volatility on the pricing and hedging derivatives. In order to address the above problem, GARCH-type models and EVT distributions are used to examine time varying volatility of ALSI on the JSE.

1.3   Purpose of the study

1.3.1   Aim

The aim of this study is to examine the use of GARCH-type models for modelling volatility and financial market risks of shares on the JSE. The study will detect the best fit estimation model for conditional variance on GARCH-type models including GARCH (p, q), Generalised Autoregressive Conditional Heteroskedastic in

Mean (GARCH-M), Threshold Generalised Autoregressive Conditional Heteroskedastic (TGARCH), Exponential Generalised Autoregressive Conditional Heteroskedastic (EGARCH) and EVT distributions.

1.3.2   Objectives of the study

The objectives of the study are to:

a. Develop symmetric and asymmetric GARCH-type models,

b. Use EVT for modelling extreme market risk for the ALSI on the JSE, and

c. Identify areas for further study.

1.4   Significance of the study

Investors and financial analysts are concerned about the uncertainty of the returns on their investment assets caused by variability in speculative market prices and the instability of business performance. When stock market risk increases, risk starts to averse investors, who in turn tend to reduce their holding of equities relative to safe assets such as Treasury bills. As a proxy of risk, volatility is not only of great concern to investors but also to policy makers. On the other hand, policy makers are mainly focused on the effect of volatility on the stability of financial markets in particular, and the whole economy in general. Finally, volatility estimation is essential in many Value-at-Risk (VaR) models.

1.5   Organisation of the study

This study is divided into five chapters. Following this introductory chapter, chapter 2 reviews empirical literature on modelling volatility and financial market risks. Chapter 3 provides methods that are used to model volatility and financial market risks. Analysis of the data used in this study is discussed in chapter 4. Finally, chapter 5 concludes the study, makes recommendations and identifies areas for further study.

Chapter 2

Literature review

2.1 Introduction

This chapter analyses the theoretical and empirical literature related to modelling volatility and financial market risks. Following this introductory part, section 2.2 analyses volatility using GARCH-type models. In section 2.3 modelling of returns is measured using GEV distribution, while in section 2.4 volatility modelling is measured using GPD. Finally, section 2.5 summarises the chapter.

2.2 Volatility modelling using GARCH-type models

Mala and Reddy (2007) conducted a study on measuring stock market volatility in emerging economies. They used both Autoregressive Conditional Heteroskedastic (ARCH) and GARCH models to find the presence of the stock market volatility on Fiji's stock market. The analysis was done using time series data for the period 2001 to 2005. The volatility of stock returns were then regressed against the interest rates and the results showed that the interest rates changes have a significant effect on stock market volatility. The test for the presence of the volatility was carried for each specific firm listed on the stock market but results revealed that of the 16 listed companies, only 7 firms were volatile.

Alberg *et al.* (2008) examined a comprehensive empirical analysis of the return and conditional variance of the Tel Aviv Stock Exchange (TASE) indices in Israel using

GARCH models. They compared the forecasting performance of several GARCH models using different distributions for two TASE index returns and parameters were estimated using Quasi Maximum Likelihood (QML) technique. The Akaike Information Criterion (AIC) and the log-likelihood values indicated that the EGARCH model and Asymmetric Power Autoregressive Conditional Heteroskedastic model estimate the series better than the traditional GARCH. The results further showed that EGARCH model using a skewed student-t distribution is the most successful in forecasting the TASE indices. The results indicated that asymmetric GARCH models improve the forecasting performance.

Floros (2008) examined the use of GARCH-type models for modelling volatility and explaining financial market risk. Daily data from Egypt and Israel were used. The study employed various GARCH-type models such as simple Component GARCH (CGARCH), EGARCH, TGARCH, Asymmetric Component GARCH (ACGARCH) and Power GARCH (PGARCH) models. The results showed strong evidence that daily returns can be characterised by GARCH models. The findings are that for both Egypt and Israel markets, increased risk will not necessarily lead to a rise in the returns. The research on examining financial returns has raised the question of whether GARCH-family models are able to capture volatility clustering.

Hein (2008) examined stock return volatility in Vietnam stock market. The empirical investigation was conducted by means of GARCH models including both symmetric and asymmetric models. The data set used was from Vietnam index over a six-year period from March, 2002 to March, 2008. The findings presented the inappropriateness of asymmetric GARCH in modelling Vietnam stock return volatility. The excess kurtosis and skewness in residual series of Vietnam stock return still revealed even with the best performing GARCH models. Empirical results suggested the sufficiency of GARCH (1, 1) and GARCH (2, 1) models in capturing properties of conditional variance in Vietnam stock market.

Aktan *et al.* (2010) addressed the issue of conditional volatility modelling by using

symmetric and asymmetric GARCH-type models. The study examined the characteristics of conditional volatility in the three Baltic stock markets (Estonia, Latvia and Lithuania) by using a broad range of GARCH volatility models. The results showed that there is strong evidence that daily returns from Baltic stock markets can be successfully modelled by GARCH-type models. For all Baltic markets, they concluded that increased risk will not necessarily lead to a rise in the returns. All of the analysed indices exhibited complex time series characteristics involving asymmetric GARCH model, long tails and complex autoregressions in the returns.

Emenike (2010) investigated the volatility of stock market returns in Nigeria using GARCH (1, 1) and the Glosten Jagannathan Runkle (GJR)-GARCH (1, 1) models. Volatility clustering, leptokurtosis and leverage effects were examined for the Nigerian Stock Exchange returns series from January, 1985 to December, 2008. The results from GARCH (1, 1) model showed that volatility of stock returns was persistent in the return series data. The results of GJR-GARCH (1, 1) model showed the existence of leverage effects in the stock returns data. Overall results from this study provided evidence to show volatility persistence, fat-tail distribution and leverage effects for the Nigerian stock returns data.

Hamadu and Ibiwoye (2010) studied the volatility behaviour of the Nigerian insurance stock price. Several deviations of heteroskedastic conditional volatility models were evaluated using model evaluation performance metrics. The post estimation revealed that most of the models studied were competitive. However, the results showed that the EGARCH is a more preferred modelling framework for evaluating risk volatility of Nigerian insurance stocks. The findings were validated by using AIC, Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) evaluation information measures.

Sigauke *et al.* (2010) investigated the use of Autoregressive Integrated Moving Average (ARIMA)-GARCH-type models for modelling volatility in a hyperinflationary economic environment. Their study used monthly stock prices of eight counters listed

on the Zimbabwe Stock Exchange (ZSE) over the period 1993 to 2004. The results suggested that the monthly returns are characterised by an ARMA (0, 1)-GARCH (1, 1) model and that increased risk does not necessarily imply an increase in returns. The out of sample forecasting evaluation indicated that ARMA (0, 1)-TGARCH (1, 1) model achieves the most accurate volatility forecast followed by ARMA (0, 1)-GARCH (1, 1) model. The study showed that supply and demand theory which was used at the ZSE underestimated the value of assets.

Wang (2010) investigated the link between the volatility of China's stock market and macroeconomic variables such as the real Gross Domestic Product (GDP), inflation, and interest rate for the period from 1992 to 2008 using monthly data. The study implemented two steps in its investigation. In stage one, the volatility for each variable using EGARCH model was estimated and an examination of the causal relationship between the volatility of the stock prices and the macroeconomic variables was found. Results revealed that there was no causal relationship between stock market volatility and economic activity measured by real GDP. This implied that stock prices were not significant in explaining economic activity. Results showed unidirectional causal relationship between stock market volatility and interest rate volatility running from stock prices to the interest rate.

Ahmed and Suliman (2011) estimated volatility in the daily returns of the principal stock exchange of Sudan using GARCH models. In their study, they used the Khartoum Stock Exchange (KSE) over the period from January, 2006 to November, 2010. The models used include both symmetric and asymmetric GARCH that capture the most common stylised facts about index returns such as volatility clustering and leverage effect. The empirical results showed that the conditional variance process is persistent. The results also provided evidence on the existence of risk premium for the KSE index return series which support the positive correlation between volatility and the expected stock returns. The findings of their study explained that all GARCH specifications applied clarify that explosive volatility process is present in the KSE index returns over the sampled period.

Angabini and Wasiuzzaman (2011) examined different GARCH models to investigate and quantify the changes in volatility of the Malaysian stock market with respect to the global financial crisis in 2007/2008. The unit root test was applied to check for stationarity and series were found to be stationary. Conditional mean was then modelled using Autoregressive Moving Average (ARMA) models and the Autoregressive (AR) (4) model was selected as the best model. The GARCH models were estimated using QML assuming the Gaussian normal distribution. Different lags were examined for each model and the GARCH (1, 1), EGARCH (1, 1) and GJR-GARCH (1, 1) were found to be the most successful models. AR (4)-GARCH (1, 1), AR (4)-EGARCH (1, 1) and AR (4)-GJR-GARCH (1, 1) were the final models to describe the process.

Li and Hong (2011) examined and demonstrated the ability and superiority of price range estimators to forecast future volatility by comparing with the GARCH volatility. The study employed data from United States stock market. They adopted Autoregressive Volatility (AV) model in order to model dynamics of volatility process. Two types of volatility models discussed and estimated were return based GARCH model and range-based AV model. The comparison study included out of sample forecasting performance as well as in sample comparison. The results from both out of sample and in sample forecasts showed that the range based AV model successfully captured the dynamics of the volatility. Furthermore the results showed that AV gained good performance relative to the GARCH model.

2.3 Modelling returns using GEV distribution

Bali (2007) proposed an extreme value distribution to estimate interest rate during high volatility period on Federal Reserve in the United States. The study provided strong evidence for the presence of serial correlation and extreme conditionally heteroskedastic volatility effects. The study showed that the tails of the empirical distribution are much thicker than the tails of the normal distribution. The empirical results indicated that the volatilities of the maximum and minimum changes in the interest rate decline as time to maturity arises. The study also introduced a closed

form option pricing model based on the GEV distribution, and showed that the newly proposed model provides more accurate predictions.

Hasan *et al.* (2011) studied extreme share returns in Malaysia. The monthly, quarterly, half yearly and yearly maximum returns are fitted to the GEV distribution. Maximum Likelihood Estimation (MLE) is used to estimate the parameter while L-moment estimate is used to initialise the MLE optimisation routine for the stationary model. Likelihood ratio test is performed to determine the best model and return levels are then estimated for prediction and planning purpose. The results showed that maximum returns for all selection periods are stationary. The results also concluded that yearly maximum is better for the convergence to the GEV distribution especially if longer records are available.

Zhao *et al.* (2011) developed a GEV-GARCH model by applying a conditional autoregressive heteroskedastic structure to the classical GEV distribution on high frequency data of Dow Jones stock returns. A simulation study and real data showed that the GEV-GARCH model can capture the dynamic of conditional variance of extremes and model the tail behaviour of the underlying variables. The study indicated that model identification is problematic if the tail behaviour changes as fast as the volatilities. Results also indicated that the model can be used in estimation of tail related risk for heteroskedastic time series. Further results demonstrated that model identification and parameter estimation complications arise when considering a time varying shape parameter with similar GARCH structure.

Huang *et al.* (2012) studied the applications of EVT on analysis for closing price data of Dow Jones industrial index and Danish fire insurance claims data. The study proposed the hypothesis testing problem for the extreme value index based on a new test statistics, the Linear Quadratic Regulation statistics. The proposed test statistics is obtained by using maximum Lq-likelihood (MLq) method. The focus for the study is only on the testing problem for GEV distribution. The results showed that MLq method performs well on testing problems for GEV distribution.

2.4   Volatility modelling using GPD

Bystrom (2005) investigated electricity prices quoted on Nord Pool, the first multinational exchange for electricity trading. Initially, return series was filtered with an AR-GARCH model and then applied results from EVT to the residuals. The study revealed that price changes in the market are not only volatile but their empirical distributions are also highly non-normal. Results showed a good fit of the GPD to AR-GARCH filtered price change series. The results also indicated that Peak Over Threshold (POT) method of modelling the extreme tails changes with high accuracy.

Chan and Gray (2006) examined a number of approaches to forecast VaR for electricity markets. The daily data of aggregated electricity spot prices from five international power markets (Victoria, Nord Pool, Alberta, Hayward and PJM) are examined. Leverage effects in conditional volatility are modelled with an EGARCH specification. Model residuals are standardised to produce near independently and identically distributed observations, and EVT is applied to the standardised residuals to forecast the tail quantiles. The results showed that AR-EGARCH model with EVT method produces the most accurate forecast of VaR than pure AR-EGARCH model.

Cotter (2007) examined extreme risk in Asian markets. EVT model and Gaussian distribution were applied to daily log returns of Asian equity market. POT of EVT generated the risk measures where tail returns are modelled with a fat-tailed distribution. The GPD parameters are estimated by maximum likelihood method. It is established that EVT distribution support modelling of tail returns in an unconditional setting. The results also revealed that distribution measures are much smaller if Gaussian distribution is applied compared to GPD estimates.

Magheyereh and Al-Zoubi (2008) investigated the tail behaviour of daily stock returns for the three emerging stock in the Gulf region (Bahrain, Oman and Saudi Arabia) over the period 1998 to 2005. In order to model extreme returns and to estimate tail quantiles, the study filtered the return series of the Gulf markets with the

skewed student's t distribution of AR (1)-Fractionally Integrated Asymmetric Power Autoregressive Conditional Heteroskedastic (FIAPARCH) (1, d, 1) model and then applied results from EVT to the standardised residuals. The results revealed that POT method of modelling extreme tail quantiles is more accurate than conventional methodologies in estimating the tail behaviour of the Gulf market returns. The results also showed that extreme values are located further out in the right tail than in the left tail and tail parameters differ within the region.

Djakovic *et al.* (2011) investigated the performance of EVT with daily stock index returns of four different emerging markets. The research covered the sample representing the Serbian, Croatian, Slovenian and Hungarian stock indexes using data from January, 2006 to September, 2009. Performance test was carried out for the success of application of the EVT in estimating the tails of daily return distribution. Research results according to estimated GPD parameters indicated the necessity of applying market risk estimation method. The results also indicated that the GPD fits the tails of the return distribution in selected emerging markets well, and that the daily return distributions have different characteristics at the left and right tails.

Guru (2012) applied extreme value theoretic technique to the NSE Nifty index to quantify the tail risks in the index for the time period 1995 to 2011. POT method of GPD was fitted to the data. The positive value of the shape parameter of the distribution indicated fat-tailed nature of the return series. The results showed features of financial asset returns. It was also found that extreme value based modelling of tails of the distribution provides more accurate measures of risk compared to estimates based on normal assumption.

Lee (2012) focused on modelling and estimating tail parameters of loss distribution from Taiwanese commercial fire loss severity. The GPD was employed and compared with standard parametric modelling based on Lognormal, Exponential, Weibull and Gamma distribution. In the results, parametric curve-fitting method is described for modelling extreme historical losses using mean excess function plot. Optional thresh-

old is determined and parameter value of GPD is modelled using a Hill plot and mean excess function plot. The results showed that GPD can be fitted to commercial fire insurance loss security.

Sigauke *et al.* (2012) conducted a study on modelling of intraday increases in peak electricity demand using ARMA-EGARCH-generalised single Pareto model. The developed model is used for extreme tail quantile estimation using daily peak electricity demand data from South Africa during the period from 2002 to 2011. The advantage of the modelling approach used lies in its ability to capture conditional heteroskedasticity in the data through EGARCH framework. It was found that ARMA-EGARCH-generalised single Pareto model produces more accurate estimate of extreme tails than a pure ARMA-EGARCH model. The modelling approach can be applied in any study for modelling conditional heteroskedasticity.

Song and Song (2012) proposed a new, fast and stable parameter estimation method for extreme quantiles of heavy-tailed distributions with massive data. The method employed the POT with GPD that is commonly used to estimate extreme quantiles and parameter estimation. The results demonstrated that parameter estimation method has a smaller Mean Square Error (MSE) than other common methods when the shape parameter of GPD is at least zero. The estimated quantiles also show the best performance in terms of RMSE and absolute relative bias for heavy-tailed distribution.

2.5   Conclusion

This chapter has presented the previous work done on volatility modelling in financial markets. Results from literature show that GARCH-type models can be used to model volatility especially in stock markets. Empirical evidence reveals that extreme value distributions like GEV distribution and GPD can be measured to model extreme events and reduce risks in the markets.

Chapter 3

Methodology

3.1 Preliminaries

The chapter discusses the methods and techniques which will be used in analysing the data. This study employs a quantitative research paradigm since it is much more focused on the collection and analysis of the numerical data and statistics. The study follows a model testing research design for modelling volatility and financial market risks of shares on the JSE. Secondary data from the officials of the JSE is measured. The JSE daily data of all share price index from 07 January, 2002 to 30 December, 2011 is used.

3.2 Data analysis techniques

The study uses Time series, Autoregressive Moving Average (ARMA) model, Generalised Autoregressive Conditional Heteroskedasticity (GARCH)-type models and Extreme Value Theory (EVT) distributions which include Generalised Extreme Value (GEV) distribution and Generalised Pareto Distribution (GPD).

3.3 Time series analysis

Time series is a sequence of data points, measured typically at successive time instants spaced at uniform time intervals. Time series analysis is used for identifying and characterising the nature of the analysed variables and predicting future values of

the time series variable. A time series plot is used to display the time variation of one or more scalar data sets associated with a mesh or grid at observation points in observation coverage. This is basically a plot of the response or variable of interest (Z_t) against time.

3.3.1 Components of time series

A time series can be decomposed into the following components:

1. Trend: A trend exists when there is a long-term increase or decrease in the data.

2. Seasonal component: A seasonal pattern exists when a series is influenced by seasonal factors (e.g., the quarter of the year, the month, or day of the week). Seasonality is always of a fixed and known period.

3. Cyclic component: A cyclic pattern exists when data exhibit rises and falls that are not of fixed period. The duration of these fluctuations is usually of at least 2 years.

4. Irregular components: These are erratic movements in a time series that follow no recognisable pattern.

3.4 Stationarity

The basic idea of stationarity is that the probability laws that govern the behaviour of process do not change over time. In a sense, the process is in statistical equilibrium. Specifically, a process Y_t is said to be strictly stationary if the joint distribution of $Y_{t_1}, Y_{t_2}, ..., Y_{t_n}$ is the same as the joint distribution of $Y_{t_{1-k}}, Y_{t_{2-k}}, ..., Y_{t_{n-k}}$ for all choices of time points $t_1, t_2, ..., t_n$ and all choices of time lag k. The time series Y_t is said to be second order stationary if the mean is constant for all t and if for any t and k the covariance between Y_t and Y_{t+k} only depends on the lag difference k. In other words there exists a function $c : \mathbb{Z} \to \mathbb{R}$ such that for all t and k we have $c(k) = cov(X_t, X_{t+k})$. The various techniques which can be used to standardise the data include differencing, logarithm transformation, square root transformation,

power transformation, etc. Differencing and logarithm transformation techniques are applied in this study.

3.4.1 Stationary through differencing

Models that are not stationary when subjected to differencing often yield stationary process. Thus, a time series of a differenced data can be denoted by: $BZ_t = Z_t - Z_{t-1}$, where B denotes backshift operator. In some instances, differencing once may not yield a stationary process. In that regard we continue to difference the dataset until it is stationary.

3.4.2 Stationary through logarithm transformation

The logarithm transformation is normally used when a serious problem is encounted where increased dispersion seems to be associated with increased levels of the series, implying that the larger the series the more the variation there is around that level and conversely.

3.5 Augmented Dickey-Fuller test

In testing for a unit root, the Augmented Dickey-Fuller (ADF) test can be used. It is an augmented version of the Dickey-Fuller test for a larger and more complicated set of time series models. The testing procedure for the ADF is the same as for Dickey-Fuller test but applied to the model:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \cdots + \delta_{p-1} \Delta y_{t-p+1} + \epsilon_t \tag{3.1}$$

where Δ is a difference operator, α is a constant, β is the coefficient on a time trend and γ is the lag of the autoregressive process. Imposing the constraints $\alpha = 0$ and $\beta = 0$ correspond to modelling a random walk and using the constraint $\beta = 0$ corresponds to modelling a random walk with a drift. By including lags of the order p, the ADF formulation allows for higher order autoregressive processes. This means that the lag length p has to be determined when applying the test. The unit root test is carried out as follows:

$H_0 : \gamma = 0$ (The data needs to be differenced to make it stationary or there is a unit root).

$H_1 : \gamma < 0$ (The data is stationary and does not need to be differenced or there is no unit root).

The value for the test statistic $DF_T = \frac{\hat{\gamma}}{SE\hat{\gamma}}$ can be compared to the relevant critical value for the ADF test.

3.6   Test of normality

3.6.1   Skewness and Kurtosis

Skewness is the third standardised moment and it describes the asymmetry of a distribution. Positive value of skewness indicates a long right tail while negative value of skewness indicates a long left tail. Zero skewness indicates symmetry around the mean. The skewness (S) of a random variable X is given by:

$$S = \frac{E(X - \mu)^3}{\sigma^3} \tag{3.2}$$

where μ and σ are the mean and standard deviation of X.

The kurtosis is the fourth standardised moment of the distribution and is a measure of flatness of a distribution. The kurtosis for the normal distribution is exactly three. Kurtosis higher than three indicates that the distribution has heavy tails and peaks close to its mean. The sample data has a flatter distribution than the normal if the value of kurtosis is less than three. The kurtosis (K) of a random variable X is given by:

$$K = \frac{E(X - \mu)^4}{\sigma^4} \tag{3.3}$$

where μ and σ are the mean and standard deviation of X.

3.6.2   Jarque-Bera test

The Jaque-Bera (JB) is a popular test of normality that incorporates both skewness and kurtosis. It is given by:

$$\text{JB} = \frac{n}{6} \left[S^2 + \frac{(K - 3)^2}{4} \right] \tag{3.4}$$

where n is the sample size, S is the skewness and K is the sample excess kurtosis. The JB test has an asymptotic χ^2 (chi-square) distribution with two degrees of freedom and can be used to test the null hypothesis that the data are from a normal distribution. The null hypothesis of normality is rejected if the calculated test statistic exceeds a critical value from the χ^2 (2) distribution. The critical value of 5.99 is corresponding to 5% level of significance. The JB test for normality is presented as:

H_0 : Normal distribution, skewness is zero and excess kurtosis is zero.
H_1 : Non-normal distribution.

3.7   Histogram and Kernel Density estimation

Histogram is a graphical representation of the frequency distribution of a data set. It provides insight into skewness, tail behaviour and outliers. The basic procedure of constructing a histogram consists of dividing the interval covered by the data set into length of sub-intervals known as bins. Histogram is a good starting point for analysing the shape and location of the data distribution. However, some of its properties such as non-smooth can be unsatisfactory. The unsatisfactory of non-smooth can be alleviated by Kernel density estimation. The Kernel density estimation is not only a popular tool for visualising the distribution of data but also a well-known nonparametric estimator of univariate or multivariate densities. Let $X_1, X_2, ..., X_n$ denote a sample of size n from a random variable with density f. The Kernel density estimate of f at the point x is given by:

$$f_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{n}\right)$$

(3.5)

where the Kernel K satisfies $\int_{\infty}^{\infty} K\left(x\right) dx = 1$ and the smoothing parameter h is known as the bandwidth. The function $\int_{\infty}^{\infty} f_h\left(x\right) dx = 1$ is due to the division of the sum by nh.

3.8   Diagnostic test

The JB test statistic is the test for normally under the null hypothesis that the coefficients of skewness and kurtosis are equal to zero and three respectively. If the critical value of JB test is greater than 5.99, it indicates statistical significance at the 5% level. In this dissertation we employ four main GARCH-type models and carry out the diagnosis tests to determine whether the models are adequate and which model is the best fit for the data. The estimated GARCH model which provides a good fit should capture all dynamic aspects of the model of the mean and the model of the variance.

3.9   Model building strategy

The model building strategy developed in this study uses Box and Jenkins (1976) technique. Box and Jenkins (1976) propose a practical three-stage procedure for finding a good model.

Stage 1: Model identification
The identification stage uses two graphical devices to measure the correlation between the observations within a single data series. These devices are called an estimated autocorrelation function (ACF) and estimated partial autocorrelation function (PACF). The next step at the identification stage is to summarise the statistical relationships within the data series in a more compact way than is done by the estimated ACF and PACF. The estimated ACF and PACF are used as guide in choosing one or more autoregressive integrated moving average (ARIMA) models that seem appropriate.

Stage 2: Model estimation
The precise estimates of the coefficients of the model chosen at the identification stage are discovered. This stage provides some warning signals about the adequacy of the model. In particular, if the estimated coefficients do not satisfy certain mathematical inequality conditions, that model is rejected.

Stage 3: Diagnostic checking

Box and Jenkins (1976) suggest some diagnostic checks to help determine if an estimated model is statistically adequate. The results at this stage may also indicate how a model could be improved and this leads to the identification stage. The cycle of identification, estimation and diagnostic checking is repeated until a good model is found.

3.10 Autocorrelation and Partial Autocorrelation plots

After a time series has been stationarised by differencing, the next step in fitting an ARMA model is to determine whether AR or MA terms are needed to correct any autocorrelation that remains in the differenced series. ACF plot is merely a bar chart of the coefficients of correlation between a time series and lags of itself. The PACF plot is a plot of the partial correlation coefficients between the series and lags of itself. In general, the partial correlation between two variables is the amount of correlation between them which is not explained by their mutual correlations with a specified set of other variables. A partial autocorrelation is the amount of correlation between a variable and a lag of itself that is not explained by correlations at all lower-order-lags.

3.11 Autocorrelation and Partial Autocorrelation functions

Box and Jenkins (1976) suggest the number of lags to be no more than $\left(\frac{n}{4}\right)$ autocorrelations. The autocorrelation coefficient measures the correlation between a set of observations and a lagged set of observations in a time series. The autocorrelation between Z_t and Z_{t+k} measures the correlation between pairs $(Z_1, Z_{1+k}), (Z_2, Z_{2+k}), ..., (Z_n, Z_{n+k})$. The sample autocorrelation coefficient r_k is an estimate of γ_k, defined as follows:

$$r_k = \frac{\sum (Z_t - \overline{Z})(Z_{t+k} - \overline{Z})}{\sum (Z_t - \overline{Z})^2} \tag{3.6}$$

where Z_t is the data from the stationary time series, Z_{t+k} is the data from k time period ahead of t and $\overline{Z}$ is the mean of the stationary time series. The estimated ACF and PACF are used as a guide in choosing one or more ARIMA models that might fit the available data. The idea of partial autocorrelation analysis is to measure

how $\widehat{Z}_t$ and $\widehat{Z}_{t+k}$ are related. The equation that gives a good estimate of the partial autocorrelation is:

$$\widehat{\varphi}_{kk} = \frac{r_k - \sum_{j=1}^{k-1} \widehat{\varphi}_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} \widehat{\varphi}_{k-1,j} r_{k-j}}, k = 2, 3, ... \tag{3.7}$$

where $\widehat{\varphi}_{k-1,j}$ for $k = 2, 3, ...$ and $j = 1, ..., k-1$

3.12   Autoregressive model

The notation AR (p) refers to the autoregressive model of order p. AR (p) model is written as:

$$X_t = c + \sum_{i=1}^{p} \phi_i X_{t-i} + \varepsilon_t \tag{3.8}$$

where X_t is the time series, $\phi_1, ..., \phi_p$ are parameters, c is a constant and ε_t is a sequence of independent random variables with mean 0 and variance σ^2, assuming that X_t is stationary.

3.13   Moving Average model

The notation MA (q) refers to the moving average of order q. MA (q) model is written as:

$$X_t = \mu + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{3.9}$$

where X_t is a stationary time series, $\theta_1, ..., \theta_q$ are the parameters of the model, μ is the expectation of X_t (often assumed to be equal to zero) and ε_t is a sequence of independent random variables with mean 0 and variance σ^2.

3.14   Autoregressive Moving Average model

The notation ARMA (p, q) refers to the model with p autoregressive terms and q moving average terms. This model contains the AR (p) and MA (q) models:

$$X_t = c + \varepsilon_t + \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{3.10}$$

where c is a constant and ε_t is a sequence of independent random variables with mean 0 and variance σ^2.

3.15   Behaviour of ACF and PACF

Table 3.1 defines the behaviour of the ACF and PACF plots on models AR(p), MA(q) and ARMA(p, q).

Table 3.1: Behaviour of the ACF and PACF

Model	ACF	PACF
AR (p)	Tails off gradually from lag p and may contain damped fluctuations	Cuts off after p lags
MA (q)	Cuts off after q lags	Tails off gradually from lag q and may contain damped fluctuations
ARMA (p,q)	Both decay exponentially from max(p,q) and contain damped fluctuations	Both decay exponentially from max(p,q) and contain damped fluctuations

3.16   Mean equation

There are a large number of previous studies modelling the volatility which is approximated by the variance of the error terms when estimating expected returns. Literature suggests that the main focus in stock returns analysis is the variance equation rather than the mean equation. Thus an unpredictable part of stock returns should be concentrated on. The unpredictable part of stock returns, error or innovation terms, is obtained by autoregressive regression which removes the predictable part of returns. The mean equation is specified as:

$$r_t = \varepsilon_t + c \tag{3.11}$$

where r_t is the returns, $\varepsilon_t \sim N(0, \sigma_t^2)$ and c is a constant. The number of AR and MA terms are selected based on the AC and PAC plots. The best model is then estimated

by the smaller Akaike Information Criterion (AIC) and the Durban-Watson statistics which is two or closer. Durban-Watson statistics is used to detect the presence of autocorrelation. The value of Durban-Watson statistics is always between zero and four and there is no autocorrelation in the sample if the value of Durban-Watson statistics is two.

AIC is the most widely used model selection among researchers. Shawky and Abu-Zinadah (2008) indicate that AIC is a good model selection for a large sample. The AIC can be calculated as:

$$\texttt{AIC} = 2L - K\log(n) \tag{3.12}$$

where L is the maximised value of the log likelihood function for the estimated model, K is the number of independently estimated parameters in the model and n is the count of data points in the estimated dataset. Provided the ARMA specification displays no sign of autocorrelation, the number of lags with the lowest AIC is selected.

3.17 Autoregressive Conditional Heteroskedastic model

A more sophisticated volatility model is the Autoregressive Conditional Heteroskedastic, ARCH (q) model suggested by Engle (1982). It is unlikely in financial time series that the error terms will be constant overtime, therefore allowing for conditional heteroskedasticity in stock returns analysis is reasonable. Unlike historical estimation process using sample standard deviations, the ARCH model constructs the conditional variance σ_t^2 of asset returns with maximum likelihood method. The ARCH (q) model by Engle (1982) formulates volatility as follows:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_t^2 \tag{3.13}$$

The time varying volatility is captured by allowing volatility to depend on the lagged values of the innovation terms ε_t and q chosen such that the residuals of the variance equation are white noise. All of the coefficients in the conditional variance equation are required to be non-negative, thus $\omega > 0, \alpha_i \geq 0$, for $i = 1, ..., q$. In particular, if $q = 1$, the conditions become $\omega > 0$, and $\alpha_1 \geq 0$. The ARCH effect is exhibited by α_1

to capture the short-run persistence. The ARCH model is simple to apply but many parameters are required to estimate the volatility of stock returns. The problem of parsimony among the other problems of ARCH model such as how to specify the value of p and the violation of non-negativity constraints led to more general framework GARCH (p, q) proposed by Bollerslev (1986).

3.18 Lagrange Multiplier test and Ljung-box statistics

Lagrange multiplier (LM) test for autoregressive conditional heteroskedasticity must be tested on estimated equation. The ARCH LM test is the most commonly utilised method within a GARCH framework, and it is therefore used in this study. The null hypothesis is rejected if the test statistics which follow a chi-squared distribution is significant and concludes that there is evidence of ARCH effects in the data and a GARCH model is appropriate. Engle's LM can be used to test the null hypothesis of no remaining ARCH effects. LM test for autoregressive conditional heteroskedasticity must be tested on estimated equation.

Ljung-box statistic can be used to test the null of no autocorrelation up to a specific lag. The Ljung-box statistic is defined as:

$$Q = n\,(n+2) \sum_{j=1}^{h} \frac{\hat{\rho}_j^2}{n-j} \tag{3.14}$$

where n is the sample size, $\hat{\rho}_j$ is the sample autocorrelation at lag j and h is the number of lags being tested. The critical region for the significance level α is rejected if $Q > \chi^2_{1-\alpha,h}$ where $\chi^2_{1-\alpha,h}$ is the α-quantile of the chi-square distribution with h degrees of freedom. In our study we use Ljung-box test from one up to twentieth order autocorrelation of residuals of the fitted ARMA (p, q) model ε_t and squared residuals of the fitted ARMA (p, q) model ε_t^2.

3.19 Volatility measurement

Volatility is the spread of all likely outcomes of an uncertain variable. In financial markets, the spread of asset returns is of concern to risk managers and stock brokers.

Volatility is associated with the sample standard deviation of returns over some period of time. It is computed using the equation:

$$\hat{\sigma} = \sqrt{\frac{1}{T-1} \sum_{t=1}^{T} (r_t - \mu)^2} \qquad (3.15)$$

where r_t is the return on day t and μ is the average return over the $T - day$ period. The return r_t is defined as:

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) \qquad (3.16)$$

where P_t denotes the current stock price on day t and P_{t-1} denotes one lagged stock price on day $t-1$. The variance, σ^2 could also be used as a measure of volatility. Modelling volatility using variance and standard deviation is less common because of their simple relationship. In this study, variance is used as a measure of volatility. Volatility is a qualified measure of market risk. Volatility is related to risk, but it is not exactly the same. Risk is associated with undesirable outcome, whereas volatility strictly measure uncertainty.

3.20 GARCH-type models

GARCH model is used to estimate the serial dependence of volatility. Financial time series modelling has been a subject of considerable research both in theoretical and empirical statistics and econometrics. Numerous parametric specifications of ARCH models have been considered for the description of the characteristics of financial markets. ARCH model was introduced by Engle (1982) for modelling financial time series, while Bollerslev (1986) came up with the GARCH to parsimoniously represent the higher order ARCH model. A GARCH model for the conditional variance process extends the simple ARCH model by assuming that $\varepsilon_t = z_t \sigma_t$ where z_t are innovations, that is z_t are standardised residuals since $z_t = \frac{\varepsilon_t}{\sigma_t}$. The innovations z_t are such that $E(z_t) = 0$ and $Var(z_t) = 1$. We can assume these innovations to be conditionally normally distributed, Student-t distributed and Generalised error distributed. GARCH-M model was considered by Floros (2008) and Nelson (1991) introduced the EGARCH model to capture the asymmetric effect. Other specifications of the GARCH models include TGARCH model introduced independently by

Glosten *et al.* (1993) and Zakoian (1994) to capture the asymmetric effect.

3.20.1   GARCH (p, q) model

Following the natural extension of the ARMA process as a parsimonious representation of a higher order AR process, Bollerslev (1986) extended the work of Engle (1982) to the GARCH process. The GARCH (p, q) process is defined as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_t^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \qquad (3.17)$$

for $\omega > 0, \alpha_i \geq 0, \beta_j \geq 0$ and σ_t^2 is the conditional variance, which is a linear function of q lags of the squares of the error terms ε_t^2 or the ARCH terms and p lags of the past value of the conditional variances σ_t^2 or the GARCH terms, and the constants α_i, β_j and ω.

The most widely used model in practice for many financial time series is GARCH (1, 1) which contains only three parameters in the conditional variance equation. The model is very parsimonious and shown to be sufficient to capture the volatility clustering in data without the requirement of higher order models. GARCH (1, 1) is specified as:

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \qquad (3.18)$$

where σ_t^2 is the conditional variance and ε_t is the residual at time t. The conditional variance equation models the time varying nature of volatility of the residuals generated from the mean equation. The study employs GARCH (1, 1) and higher order GARCH models to verify the best fit model for JSE returns.

3.20.2   GARCH-M (p, q) model

The GARCH-M model is an extension of the basic GARCH framework which allows the conditional mean of a sequence to depend on its conditional variance or standard deviation. GARCH-M model is often used for prediction of the risk of a portfolio at a given point in time. The model is useful in presenting the time-varying risk premium in explaining excess returns. This model is considered by Floros (2008) for checking

if increase in variance result in a higher expected returns. The standard GARCH-M model is given by:

$$r_t = \mu + \beta\sigma_t^2 + \varepsilon_t, \varepsilon_t \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{3.19}$$

In GARCH-M (1, 1), when $(\alpha+\beta)$ approach unity the persistence of shocks to volatility is greater. If shocks to volatility persist over a long time, the effect of volatility on stock prices can be significant. The parameter β is called the risk premium parameter. If risk premium parameter is greater than zero and statistically significant, the model indicates that the return is positively related to its volatility. In other words, a rise in mean return is caused by an increase in conditional variance as a proxy of increased risk.

3.20.3 EGARCH (p, q) model

EGARCH model is designed to capture the leverage effect noted in Black (1976) while Nelson (1991) developed this model. A simple variance specification of EGARCH (p, q) model is given by:

$$\log\sigma_t^2 = \omega + \sum_{j=1}^{q}\beta_j\log\sigma_{t-j}^2 + \sum_{i=1}^{p}\alpha_i\left(\left|\frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right| + \rho_i\frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right) \tag{3.20}$$

The α parameter represents a magnitude effect or the symmetric effect of the model, the "GARCH" effect. For stock prices, negative shocks (bad news) generally have large impacts on their volatility than positive shocks (good news). The presence of leverage effect can be tested by the hypothesis that $\rho < 0$. The leverage effect is asymmetric if the parameter $\rho \neq 0$. The term $\frac{\varepsilon_{t-i}}{\sigma_{t-i}}$ in equation (3.20) represents the asymmetric effect of shocks. A special variation of the EGARCH (p, q) model is the EGARCH (1, 1) model given by:

$$\log\sigma_t^2 = \omega + \beta\log\sigma_{t-1}^2 + \alpha\left(\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| + \rho\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right) \tag{3.21}$$

The logarithmic form of the conditional variance implies that the leverage effect is exponential. The exponential nature of the EGARCH model guarantees that the

conditional variance is always positive even if the coefficients are negative. The persistence in conditional volatility is captured by the parameter β. When β is relatively large, then volatility takes a long time to die out following a crisis in the market, Alexander (2008).

3.20.4   TGARCH (p, q) model

TGARCH model was introduced independently by Glosten *et al.* (1993) and Zakoian (1994) to capture the asymmetric effect. The TGARCH (p, q) model specification for the conditional variance is given by:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \gamma_i \varepsilon_{t-i}^2 d_{t-i} + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{3.22}$$

where $d_t = 1$ if $\varepsilon_t < 0$ and $d_t = 0$ otherwise. In this model, good news ($\varepsilon_t > 0$) and bad news ($\varepsilon_t < 0$) have differential effects on the conditional variance. Good news has impact of α, while bad news has impact of $\alpha + \gamma$. If $\gamma > 0$ then the leverage effect exists and bad news increases volatility, while if $\gamma \neq 0$ the news impact is asymmetric, Hill *et al.* (2007). The specification of the conditional variance of TGARCH (1, 1) model is given as:

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \gamma d_{t-1} \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{3.23}$$

where the coefficient γ is known as the asymmetric term. When $\gamma = 0$, the model collapse to the standard GARCH forms.

3.21   Residual analysis

Residuals can be used to assess if the ARMA model is adequate and if the parameter estimates are close to the true values. The residual $\hat{Z}_t$ can be defined as:

$$\hat{Z}_t = X_t - \hat{X}_t \tag{3.24}$$

where X_t is the actual value and $\hat{X}_t$ is the fitted value. Model adequacy is checked by assessing whether the model assumptions are satisfied. The basic assumption is that the a_t are white noise. That is, they possess the properties of independence, identically and normally distributed random variables with mean zero and variance

σ_a^2. A model is good if it has residuals that are independent, normally distributed and constant variance.

(i) Test for Independence

A test for independence can be performed by examining ACF of the residuals. Residuals are independent if they do not form any pattern and are statistically insignificant. The Ljung-Box test can be used for randomness and the hypothesis is as follows:

H_0: The data are independently distributed.
H_1: The data are not independently distributed.

The Ljung-Box test statistic is given as:

$$Q_{LB} = n\,(n+2) \sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{n-k} \tag{3.25}$$

where n is the sample size, ρ_k^2 is the sample autocorrelation at lag k and m is the number of lags being tested. For significance level α, the critical region for rejection of the hypothesis of randomness is $Q_{LB} > \chi_{\alpha,h}^2$, which is the α-quantile of the chi-squared distribution with h degrees of freedom. Because the test is applied to residuals, the degrees of freedom must account for the estimated model parameters so that h = m - p - q, where p and q indicate the number of parameters from the ARMA (p, q) model fit to the data.

(ii) Test for Normality

In reality, residuals will never be perfectly normally distributed. If the departure from normality is extreme, then the test statistics do not have a t-distribution. Test for normality can be performed by constructing a histogram. A histogram of normally distributed residuals should approximately be asymmetric and bell shaped. A plot of the tth ordered data values versus the corresponding normal scores should fall approximately on a straight line. Normal score correlation test is based on the sample correlation coefficient between the residuals a_t and the corresponding normal scores S_t.

(iii) Test for Homoskedasticity

Test of constant variance can be inspected by plotting the residuals and plotting the residuals against fitted value. The model is adequate if the plot suggest a rectangular scatter around a zero horizontal level with no trends.

3.22   Forecasting and Evaluation measures

Evaluating the performance of different forecasting models plays a critical role in choosing the most accurate models. A huge number of papers has studied the construction of modelling and forecasting volatility, a few of them focus on the volatility forecasting evaluation. The economic loss function is usually unavailable because it requires the specific details of the investors' decision process and the cost or benefits that result from using these forecasts. Therefore, the statistical loss function is utilised in practice instead of economic loss function. The most widely used evaluation measures are Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percent Error (MAPE). The common way to solve the problem is to carry out the average figures of some statistical measures and then compare the forecast models based on the parameter obtained.

The RMSE for the conditional mean is calculated as:

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{n}(r_{at} - r_{ft})^2}{n}} \tag{3.26}$$

where n is the number of out of sample forecast data points, with $r_{at} - r_{ft}$ being the forecast errors. The terms r_{at} and r_{ft} are the actual returns and its future forecast, respectively.

The conditional volatility model for RMSE is calculated as:

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{n}(\sigma_{at}^2 - \sigma_{ft}^2)^2}{n}} \tag{3.27}$$

where σ_{at}^2 and σ_{ft}^2 are realised and forecasts of volatility, respectively.

MAE is also dependent on the scale of the dependent variable but it is less sensitive to large deviations than the usual squared loss. The conditional volatility model for MAE is calculated as:

$$\texttt{MAE} = \frac{1}{n} \sum_{t=1}^{n} \left| \sigma_{at}^2 - \sigma_{ft}^2 \right| \tag{3.28}$$

Another popular accuracy measure is the MAPE, which is scale independent. However, MAPE was criticised for the problem of asymmetry and instability when the original value is small. MAPE as accuracy measure is affected by large percentage errors that occur when the value of the original series is small and MAPE cannot be compared directly with simple models such as random walk. The conditional volatility model for MAPE is calculated as:

$$\texttt{MAPE} = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{\sigma_{at}^2 - \sigma_{ft}^2}{\sigma_{ft}^2} \right| \tag{3.29}$$

3.23 Extreme Value Theory

Extreme Value Theory (EVT) is a branch of statistics that studies "rare" or extreme events. It has been established for modelling catastrophic events in insurance and finance. It is especially well suited to describe the "fat-tails" of the profit and loss distributions typically found in stock returns. Nowadays EVT has experienced a boom in the financial field, especially with respect to risk management, particularly with its application to VaR estimation. The EVT method applied in this study assumes that the returns series are independent and identically distributed (i.i.d.) random variables. The i.i.d. random variables do not hold in market data returns as they present stylised facts such as heavy tails and clustered extremes. This feature can be corrected by filtering the returns series using time series analysis to get i.i.d. variables and then apply EVT method. In this study EVT methods are used in order to model the tails of the returns distributions of the stock indices. The two most widely used distributions for EVT analysis are Generalised Extreme Value (GEV) distribution and Generalised Pareto distribution (GPD). Both distributions are limiting distribution of extremes, but they differ in their definition of extremes.

3.23.1 Generalised Extreme Value Distribution

The Generalised Extreme Value (GEV) distribution is a motivated method for describing the distribution of the maxima and minima, Beirlant *et al.* (2004). The GEV distribution is a family of continuous probability distributions developed within the EVT to combine the Gumbel, Frechet and Weibull families. The GEV distribution has cumulative distribution function:

$$
F(x; \mu, \sigma, \xi) = \begin{cases} \exp\left\{-[1 + \xi(\frac{x-\mu}{\sigma})]^{\frac{-1}{\xi}}\right\}, 1 + \xi(\frac{x-\mu}{\sigma}) > 0 \text{ and } \xi \neq 0 \\ \exp\left(-e^{\frac{-(x-\mu)}{\sigma}}\right), \xi = 0 \end{cases} \tag{3.30}
$$

where ξ is an extreme value index, also known as shape parameter. For $\xi > 0$, $\xi < 0$ and $\xi = 0$ we obtain the Frechet, Weibull and Gumbel families, respectively. The Frechet distribution is fat-tailed as its tail decays slowly. The Weibull distribution (upper bounded) is a thin-tailed distribution with finite tail. The Gumbel distribution is thin-tailed for which all moments are finite and whose cumulative distribution function declines exponentially in the tails. The shape parameter ξ, governs the tail behaviour of the distribution. The scale parameter σ and location parameter μ represent the dispersion and average of the extreme observations respectively.

The quantile function for the GEV distribution is used to predict the probability of exceedances levels and estimate high quantiles. The quantile function of the GEV distribution is given by:

$$
x_p = \begin{cases} \mu + \frac{\sigma}{\xi}\left\{[\texttt{-log}(1-p)]^{-}\xi - 1\right\}, \xi \neq 0 \\ \mu - \sigma\texttt{log}[\texttt{-log}(1-p)], \xi = 0 \end{cases} \tag{3.31}
$$

The quantity x_p satisfy $P(X > x_p) = 1 - F(x; \mu, \sigma, \xi) = p$, for $0 < p < 1$, where x_p is the return level associated with the return period $\frac{1}{p}$. A derivation of the quantile function is given in the Appendix A1. The parameters μ, σ and ξ can be estimated using the maximum likelihood method. The log-likelihood function L is given as:

$$
L(x; \mu, \sigma, \xi) = \prod_{i=1}^{n} (f_{x;\mu,\sigma,\xi}) \tag{3.32}
$$

32

where the probability density function $f_{x;\mu,\sigma,\xi}$ is given as:

$$f(x;\mu,\sigma,\xi) = \frac{1}{\sigma}\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{\frac{-1-\xi}{\xi}}\exp\left\{-[1+\xi(\frac{x-\mu}{\sigma})]^{\frac{-1}{\xi}}\right\} \qquad (3.33)$$

for $1+\xi\left(\frac{x-\mu}{\sigma}\right) > 0$ and $\xi \neq 0$.

If $\xi = 0$ the pdf $f(x;\mu,\sigma,\xi)$ is given as:

$$f(x;\mu,\sigma,\xi) = \frac{1}{\sigma}\exp\left(\frac{-x-\mu}{\sigma}\right)\exp\left[-\exp\left(\frac{-x-\mu}{\sigma}\right)\right] \qquad (3.34)$$

3.23.2 Probability-Probability plot

Let $x_1, x_2, ..., x_n$ be a random sample from the cumulative distribution function (cdf) F and $\hat{F}$ be the estimated cdf. Then a plot of $\hat{F}_{x_i,n}$ against $p_{i,n}; i = 1, 2, ..., n$ is called the Probability-Probability plot, where $x_{i,n}$ is the ith order statistics and $p_{i,n}$ is the plotting position, which is defined as:

$$p_{i,n} = \frac{i-\alpha}{n+\beta}; i = 1, 2, ..., n \qquad (3.35)$$

The parameter $\alpha, \beta \geq 0$ can be chosen empirically based on the behaviour of the data, the type of distribution and the estimation method used to estimate the parameters. If the model fits the data well, then the pattern of points will be very close to the 45-degree line.

3.23.3 Quantile-Quantile plot

The Quantile-Quantile (Q-Q) plot is a probability plot which is a graphic in which the empirical order statistics on the Y-axis are compared to expected values of some theoretical order statistics located on the X-axis. The idea of Q-Q plot has emerged from the observation that for important classes of distributions, the Q-Q plots are linearly related to the corresponding quantiles of a standard example from the class. Let $\hat{F}_x$ be the estimate of the distribution function F. The Q-Q plot consider the estimate of the inverse of cdf such that $\hat{F}^{-1}_{p_i,n}$ versus $x_i, n, ; i = 1, 2, ..., n$. The model will fit the data well if the points of the scatter plot are very close to 45-degree line.

3.23.4 Return Level plot

The return level can be defined as the level which is expected to be exceeded once every $\frac{1}{p}$ period, which is known as return period. The return level, say x_T, exceeded on average once in T days can be written as:

$$T = \frac{1}{P}$$

$$P\left(X > x\right) = 1 - F(x) = \frac{1}{T}$$

$$X_T = F^{-1}\left(1 - \frac{1}{T}\right) \tag{3.36}$$

which is given by the quantile function F^{-1}.

3.23.5 Generalised Pareto Distribution

The use of GPD is suggested by the results of Pickands (1975) for limiting distribution for the excesses over a sufficiently high threshold. Different methods can be applied under GEV distribution to perform statistical inferences on extreme value. The block maxima are the mostly used method on GEV distribution but its weakness is that it cannot peak observations over a selected high threshold. The Peak Over Threshold (POT) method of GPD is one of the most widely used to estimate sufficiently high threshold. The threshold for GPD can be selected using graphical techniques like quantile plots and mean excess plots. This study applies Pareto quantile plot to select a threshold for GPD.

The cumulative distribution function associated with GPD denoted as W_ξ is defined as:

$$W_\xi(x) = \begin{cases} 1 - \left(1 + \frac{\xi(x-\tau)}{\sigma}\right)^{\frac{-1}{\xi}} & \text{and } x - \tau > 0 \\ 1 - \exp\left(\frac{-x-\tau}{\sigma}\right), \xi = 0 \text{ and } x - \tau > 0 \\ 1 - \left(1 + \frac{\xi(x-\tau)}{\sigma}\right)^{\frac{-1}{\xi}}, \xi < 0 \text{ and } 0 < x - \tau < \frac{-\sigma}{\xi} \end{cases} \tag{3.37}$$

where ξ is called the shape parameter or extreme value index which takes a negative, a positive or a zero. The parameters τ and σ are location and scale parameters respectively. The tail index ξ gives an indication of the heaviness of the tail. The larger

ξ, the heavier the tail. Distributions with shape parameter $\xi \neq 0$ are suited to model financial returns. If $\xi > 0$ then $W_\xi(x)$ is the Pareto distribution, $\xi = 0$ correspond to the Exponential distribution and $\xi < 0$ is known as Beta distribution.

The survival function of the GPD is calulated as $P(X > x | X > \tau) = 1 - W_\xi(x)$. This equation can be expanded as:

$$1 - W_\xi(x) = \begin{cases} \left(1 + \frac{\xi(x-\tau)}{\sigma}\right)^{\frac{-1}{\xi}} \text{ and } x - \tau > 0 \\ \exp\left(\frac{-x-\tau}{\sigma}\right), \xi = 0 \text{ and } x - \tau > 0 \\ \left(1 + \frac{\xi(x-\tau)}{\sigma}\right)^{\frac{-1}{\xi}}, \xi < 0 \text{ and } 0 < x - \tau < \frac{-\sigma}{\xi} \end{cases} \qquad (3.38)$$

The quantile function of the GPD is given by:

$$x_p = \begin{cases} \tau + \frac{\sigma}{\xi}\left[(p)^-\xi - 1\right], \xi \neq 0 \\ \tau - \sigma\log(p), \xi = 0 \end{cases} \qquad (3.39)$$

where p is the tail probability. The maximum likelihood of GPD is based on the value of threshold τ. A derivation of the quantile function given in equation (3.39) is presented in Appendix A2.

The log likelihood function for a GDP random variables is given by:

$$\log L(\xi, \sigma | \tau) = -m\log\sigma - \left(1 + \frac{1}{\xi}\right)\sum_{i=1}^{m}\log\left[1 + \xi\left(\frac{x_i - \tau}{\sigma}\right)\right] \qquad (3.40)$$

with $1 + \xi\left(\frac{x_i - \tau}{\sigma}\right) > 0, \xi \neq 0$ and $i = 1, ..., m$. The estimator holds the usual properties of consistency, asymptotic efficiency and asymptotic normality when $\xi > -0.5$.

3.23.6 Peak Over Threshold method

A threshold is chosen as defining the start point of the tail and the POT method then estimates the distribution of the excess beyond the threshold. The distribution of excesses over a sufficient high threshold τ on the underlying return distribution F is defined as:

$$F_\tau(y) = P_r\{X - \tau = y | X > \tau\} \qquad (3.41)$$

where τ is the threshold. The POT method is based on what is called the GPD in the following manner: It has been shown Pickands (1975) that asymptotically, the excess values above a high level will follow a GPD if and only if the parent distribution belongs to the domain of attraction of one of the extreme value distributions. The POT method involves the selection of τ, the exceedances above threshold. This method consists of fitting the GPD to the distribution of excesses over a threshold τ. The choice of threshold is an important practical problem, which is mainly based on a compromise between bias and variance. Pareto quantile plot is seen in this study as a useful graphical technique in selecting a sufficiently high threshold. Selecting a threshold is bound on a statement that says: above a threshold τ at which the GPD is a valid approximation of the excess distribution, the Pareto quantile plot should be approximately linear.

3.24 Pareto quantile plot

A Pareto quantile plot is a graphical method for inspecting the parameters of a Pareto distribution. A Pareto model holds if there exist a linear relationship between the logarithms of the observed values and the quantiles of the standard exponential distribution. The logarithms of the observed values, $\log(x_i), i = 1, ..., n$ are plotted against the theoretical quantiles. The Pareto quantile plot is generalised by using the theoretical quantiles:

$$-\log\left(1 - \frac{\displaystyle\sum_{j=1}^{i} w_j}{\displaystyle\sum_{j=1}^{n} w_j} \frac{n}{(n+1)}\right) \quad i = 1, ..., n \tag{3.42}$$

where the correlation factor $\frac{n}{(n+1)}$ ensures that the quantiles reduce to $-\log\left(1 - \frac{i}{n+1}\right)$ if all sample weights are equal.

3.25 Conclusion

This chapter has presented the techniques, methods and steps to be applied when modelling volatility and financial market risks. The data is analysed using Time

series, GARCH-type models and EVT distributions.

Chapter 4

Data analysis

4.1 Introduction

This chapter sets out to achieve the major objectives of the study. The statistical packages used for data analysis are Eviews and R. The chapter is divided into ten sections. Section 4.2 expresses the data used. Section 4.3 analyses the behaviour of the dataset. Section 4.4 discusses ARMA (p, q) model applied and section 4.5 discusses residual analysis. Section 4.6 discusses the symmetric and asymmetric GARCH-type models. Section 4.7 models the tail behaviour of returns using GEV distribution while section 4.8 models conditional heteroskedasticity on the returns using GPD. Section 4.9 evaluates out of sample predictions and finally, section 4.10 concludes the chapter. Appendix B presents R codes used to generate some of the results obtained in this chapter.

4.2 Data collection

The study uses secondary data provided by the JSE. The data used consists of daily log returns on JSE return index (in percentages) over the period 08 January, 2002 to 30 December, 2011, generating 2495 observations. The inspiration behind choosing daily data is that the stock markets react very quickly to new information.

4.3   Behaviour of the data

Figure 4.1 shows that all share JSE index is not stationary, meaning that variance is not stable. Some transformation is needed since the index series is not stationary. Figure 4.1 also indicates the fall (deflation) of shares towards the end of 2008 that might be caused by investors who rush and sell their shares in a panic due to the world recession. This created some excellent buying opportunities for well-informed traders and investors.
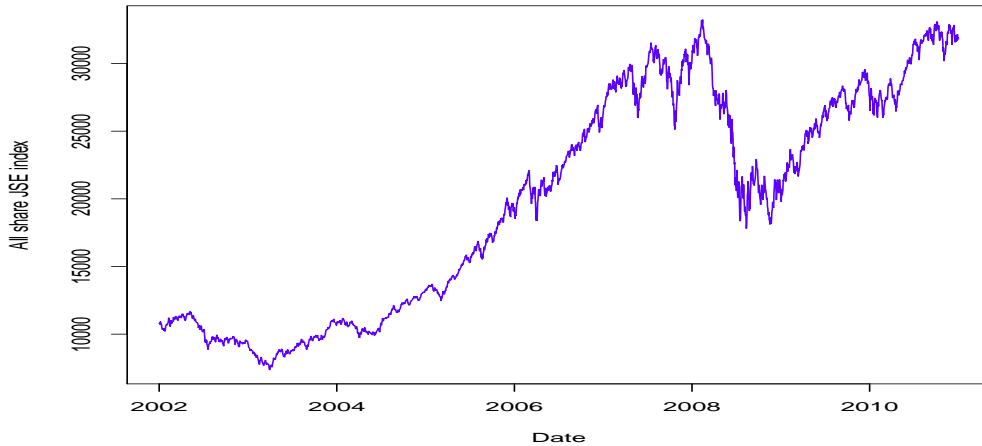
Figure 4.1: Plot of daily prices for JSE price index (2002-2011)

Formal unit root tests were carried out using the ADF test as shown in Table 4.1. Since the computed ADF test statistics (-0.7316) is greater than the critical values at 1%, 5% and 10% significant levels, we fail to reject the null hypothesis that there is a unit root and the series needs to be differenced to make it stationary.

Table 4.1: Augmented Dickey-Fuller test of the JSE price index

ADF Test Statistic = -0.7316	1% Critical Value	-3.4360
	5% Critical Value	-2.8632
	10% Critical Value	-2.5677

Figure 4.2 shows that returns series has a constant mean and constant variance which implies the first difference series of JSE index achieves stationarity. The returns are

multiplied by 100 to generate percentage changes in price. The multiplication also reduces the errors as the raw returns could be very small and produce large rounding errors in some calculations. Based on the stationarity requirements, percentage logarithmic returns are calculated as:

$$
\begin{aligned}
r_t &= 100\nabla \mathrm{log} P_t \\
&= 100(1 - B)\mathrm{log} P_t \\
&= 100(\mathrm{log} P_t - \mathrm{log} P_{t-1}) \\
&= 100\mathrm{log}\left(\frac{P_t}{P_{t-1}}\right)
\end{aligned}
\tag{4.1}
$$

where P_t denotes the current stock price on day t and P_{t-1} denotes one lagged stock price on day $t - 1$.

Figure 4.2 also displays volatility clustering and shows that volatility occurs in bursts which gives a hint that the returns may not be independent and identically distributed (i.i.d.). Returns show the extent of the day-to-day change in price where a positive spike represents a large daily gain and a negative spike indicates a significant daily loss. The series seems to be stationary since the data is fluctuating around zero. Figure 4.2 exhibits the defining characteristics of financial markets which are high volatility, occasional extreme movements and volatility clustering.
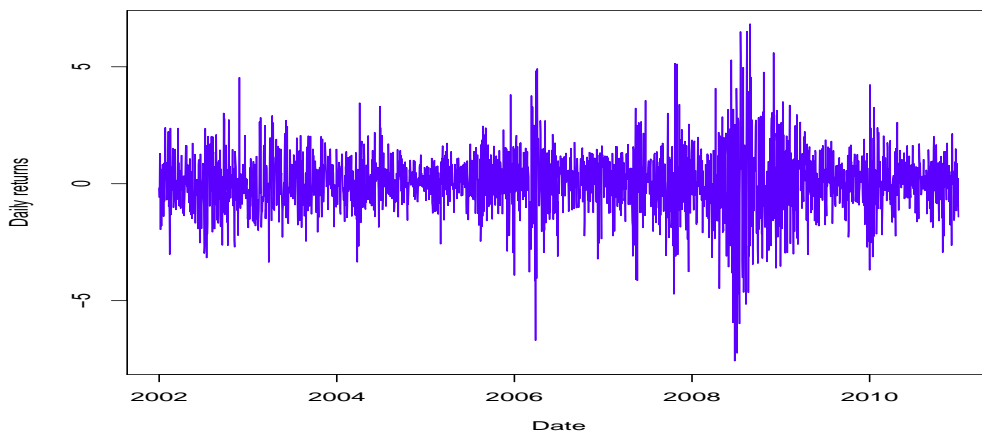

Figure 4.2: Plot of daily returns for JSE index (2002-2011)

Table 4.2 shows that the absolute computed ADF test statistics of -83.60657 is much smaller than the critical values at 1%, 5% and 10% significant level, thus we can reject the null hypothesis and conclude that JSE index series is a non-stationary series, but after taking the 1st-difference stationary is achieved.

Table 4.2: Augmented Dickey-Fuller test of the JSE returns

ADF Test Statistic = -83.60657	1% Critical Value	-3.4360
	5% Critical Value	-2.8632
	10% Critical Value	-2.5677

Table 4.3 shows descriptive statistics for prices index and returns. The mean and median are positive, suggesting that stock prices in general increase slightly overtime. The skewness measures the asymmetric and kurtosis measures the peakedness of the probability distribution. The coefficient of skewness indicates that both price index and returns have asymmetric distribution skewed to the left. The kurtosis of returns is 5.9695 which is greater than 3 indicating that the distribution of the returns is leptokurtic, that is, it is fat-tailed. This shows that the returns series exhibit financial characteristics of leptokurtosis and volatility clustering. The implication of non-normality is supported by the Jarque-Bera test statistics for both prices index and returns which show that the null hypothesis of normality is rejected at the 5% level of significance. The conclusion that returns series observed in the JSE market have non-normality distribution is reasonable as it is common phenomenon in emerging stock markets.

Table 4.3: Summary statistics of the price index and returns

Measure	Prices	Returns
Mean	20485.70	0.0435
Median	21264.36	0.0886
Minimum	7361.150	-7.5807
Maximum	33232.89	6.8339
Std. deviation	8340.726	1.3322
Skewness	-0.0972	-0.1316
Kurtosis	1.4776	5.9695
Jarque-Bera	254.3360 (0.0000)	923.8839 (0.0000)

The Q-Q plot for the returns shown in Figure 4.3 falls nearly on a straight line except at the beginning, where the plot goes up marginally. Q-Q plots that fall on a straight line in the middle but curve upward at the beginning indicate that the distribution is leptokurtic and has a thicker tail than the normal distribution. The reason that the distribution would not be exactly normal is because of volatility clustering.
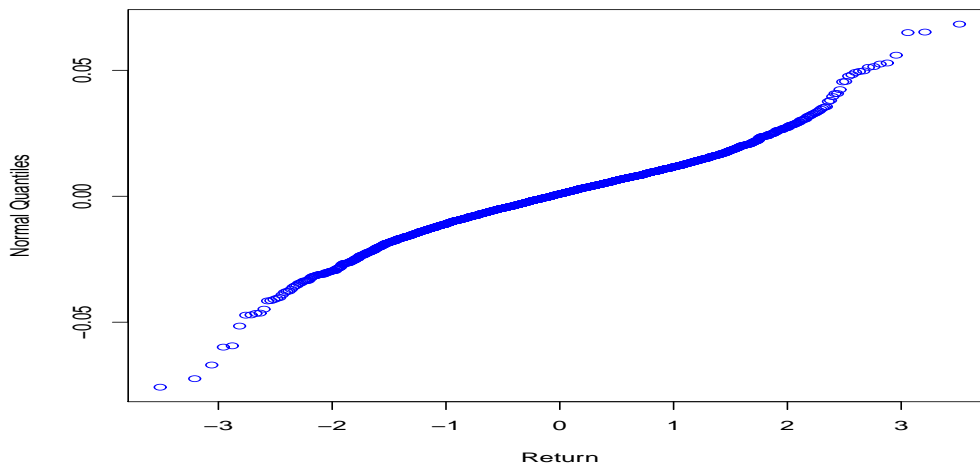

Figure 4.3: Q-Q plot of the returns series

The Kernel density of the returns series given in Figure 4.4 shows that the distribution of the data is non-normal. The density is estimated using Kernel density estimation, Silverman (1986).
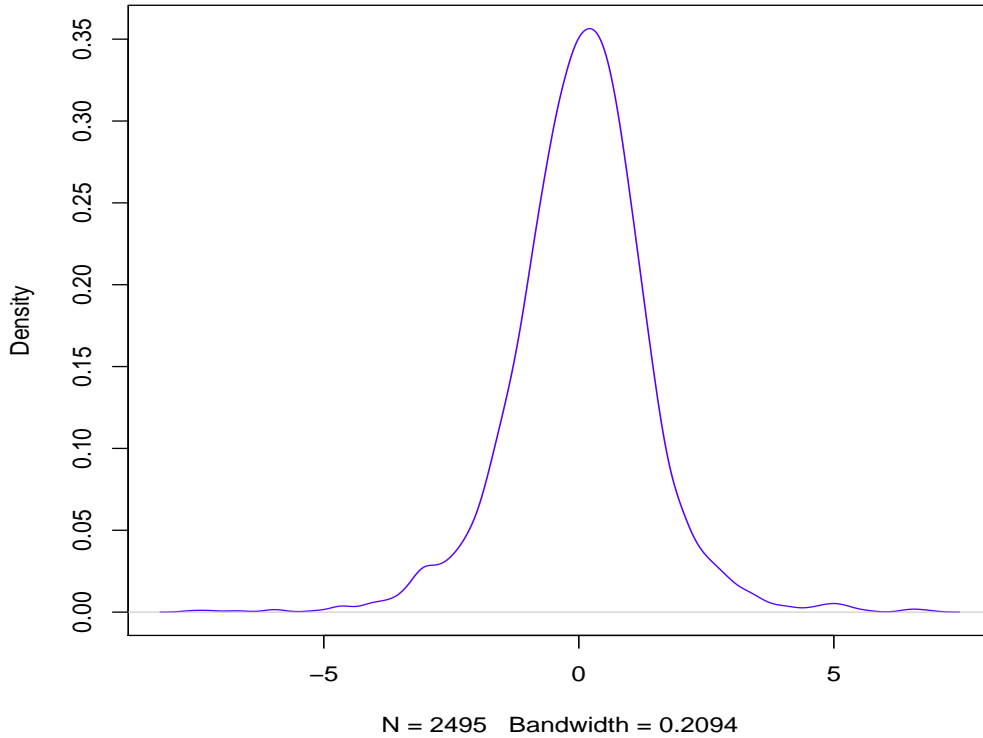
Figure 4.4: Kernel density of the returns series

The red line shown on Figure 4.5 represent a normal curve. The non-normality is driven by the presence of an outlier. Figure 4.5 shows that JSE returns distribution is peaked, confirming the evidence of non-normal distribution in Table 4.3. Peaked distribution is a sign of recurrent wide changes, which is an indication of uncertainty in the price discovery process. The suggestion is that heteroskedasticity issue needs to be taken into account. It is reasonable to use GARCH-type models to acquire heteroskedasticity.
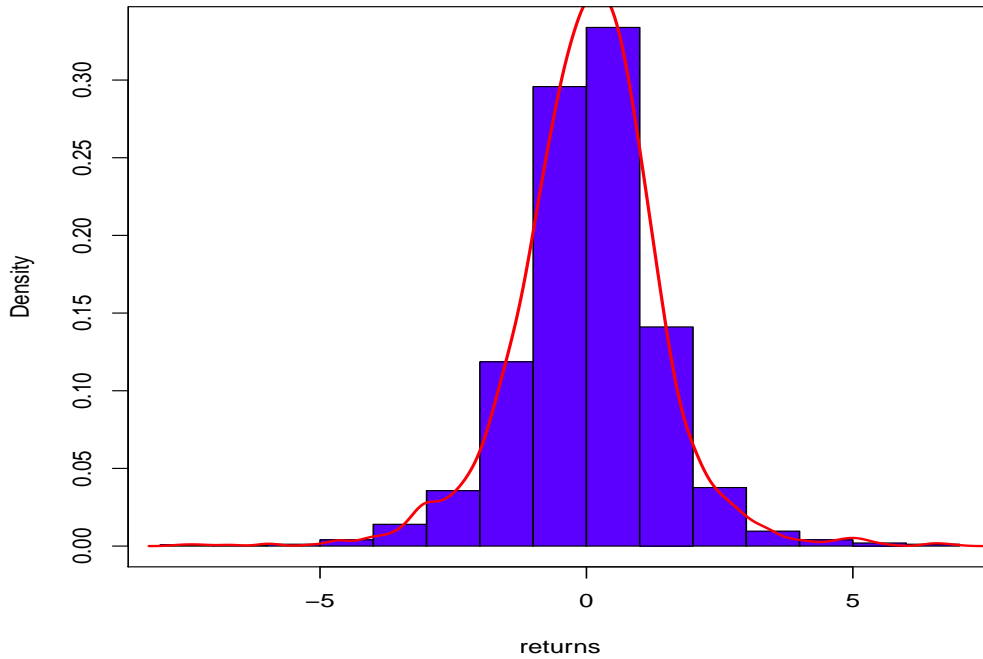
Figure 4.5: Histogram of daily returns

4.4 Model specification

The ARMA (p, q) models provide a flexible and parsimonious approximation to conditional mean dynamics. Autocorrelation and partial autocorrelation plots are used to determine the order of ARMA (p, q) models. The ACF and PACF plots shown in Figure 4.6 and Figure 4.7 respectively, give a clearer picture about the nature of the model. The result for ACF plot decays exponentially while PACF shows only one lag strongly significant at the first strike. It is clear that an ARMA (1, 0) model is appropriate and yet can be applied in the returns data.
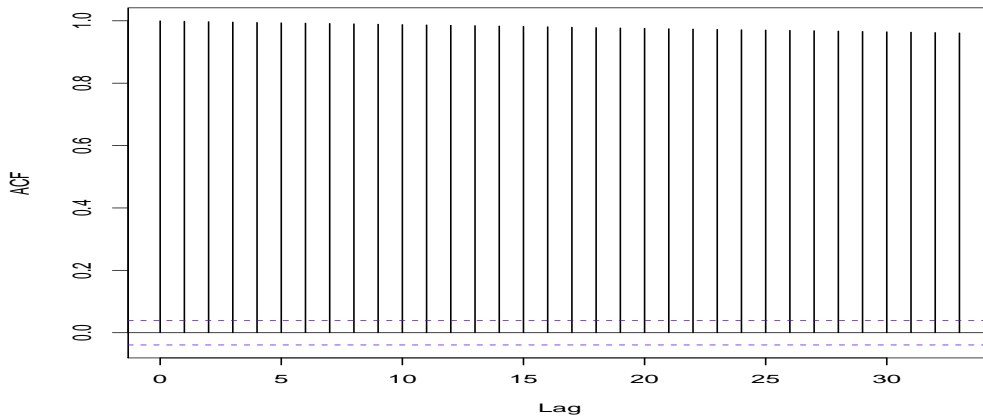
Figure 4.6: Autocorrelation plot
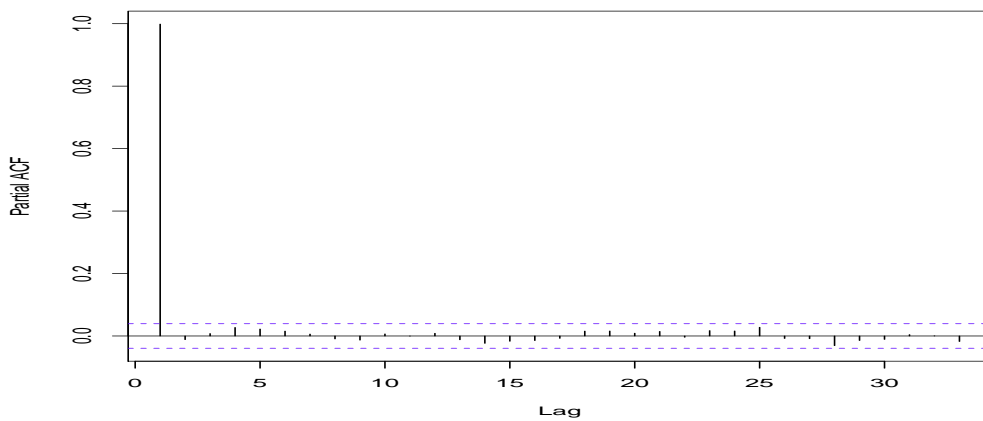
Figure 4.7: Partial autocorrelation plot

According to the ACF and PACF, ARMA (1, 0) model is found to be the best fitting model. The ARMA (1, 1) model is also tested and the better model was found to be AR (1, 0) based on the AIC. Table 4.4 demonstrates the ARMA (1, 0) model summary and its test score. The returns can be written as rt = 0.0436 + [AR (1) = 0.0389] with p-values in the parenthesis. The value of the Durban-Watson statistics is almost 2, implying that there is no autocorrelation in the sample.

Table 4.4: ARMA (1, 0) model summary and test score

Statistics	Estimation
Constant	0.0436 (0.1165)
AR(1)	0.0389 (0.0522)
AIC	3.4116
Durban-Watson statistics	1.9991
Log likelihood	-4253.2830

4.5 Residual analysis

4.5.1 Test of independence

The best test for residual autocorrelation is to examine the autocorrelation plot of the residuals as shown in Figure 4.8. Ideally, most of the residual autocorrelations should fall within 95% confidence bands around zero, which are located at roughly $\frac{2}{\sqrt{n}}$ where n is the sample size. The residuals seem to be statistically insignificant since most of the residual autocorrelations are located roughly at $\frac{2}{\sqrt{2495}} = 0.0400$.
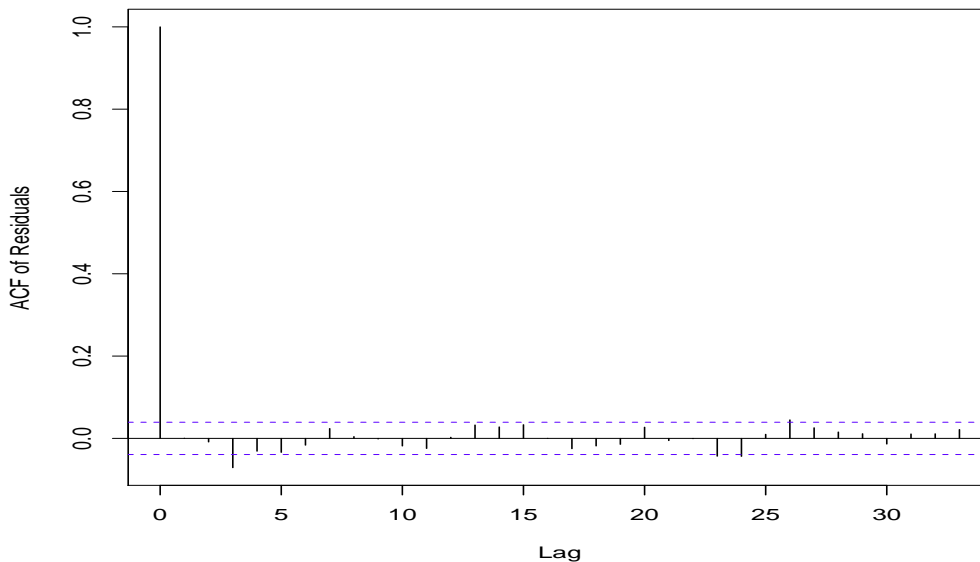
Figure 4.8: Autocorrelation of the residuals

The Q-statistic of Ljung-Box test for randomness at lag 5 is 18.2130 (0.0010) as shown in Table 4.5. The $\chi^2_{\alpha,h} = \chi^2_{0.05,4} = 9.49$ and $Q_{LB} > \chi^2_{\alpha,h}$, thus the null hypothesis is rejected and that the residuals are not independent.

Table 4.5: Correlogram of the residuals

Lag	AC	PAC	Q-stat	P-value
1	0.0000	0.0000	0.0004	0.0000
2	-0.0071	-0.0090	0.1928	0.6610
3	-0.0710	-0.0710	12.8920	0.0020
4	-0.0310	-0.0310	15.3210	0.0020
5	**-0.0340**	**-0.0360**	**18.2130**	**0.0010**

4.5.2   Test of normality

Figure 4.9 is a histogram of the residuals and it shows that residuals are approximately normal but negatively skewed. The figure is bell shaped and it also shows that residuals are symmetric about the mean.
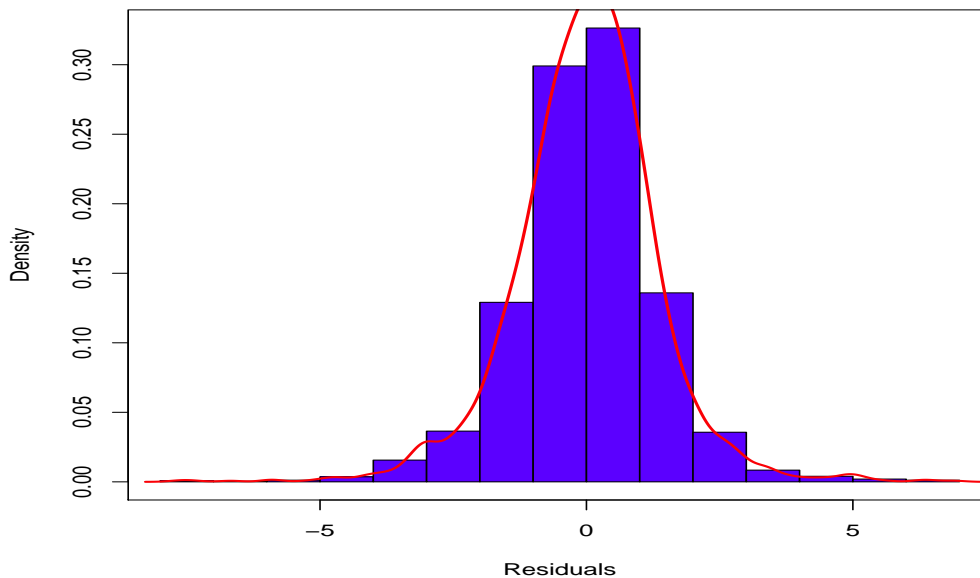
Figure 4.9: Histogram of the residuals

4.5.3   Test of constant variance

Figure 4.10 shows the residual plot that is used to test for constant variance. The figure shows that there is no trend; therefore the conclusion is that the ARMA (1, 0) model is adequate.
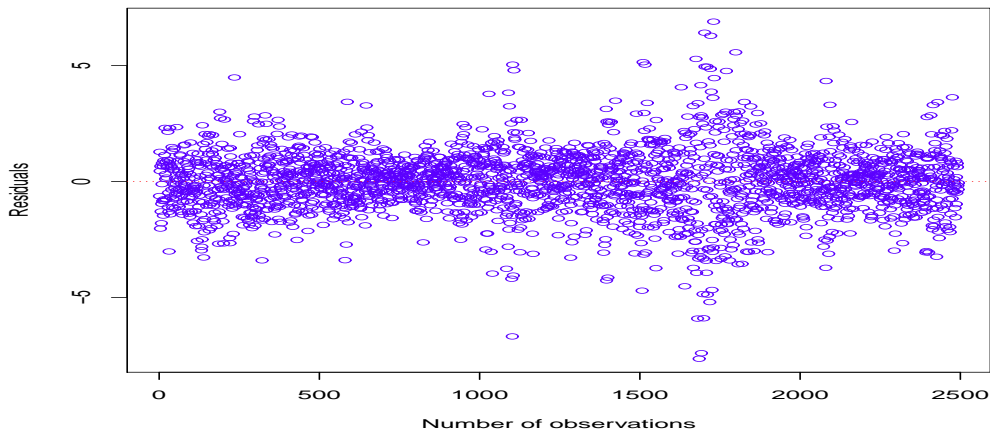
Figure 4.10: Residuals plot

The results in Figure 4.11 show that there is no trend. The graph of residual against fitted values suggests that ARMA (1, 0) model is adequate. In general, residual plot and residual against fitted plot recommend ARMA (1, 0) model.
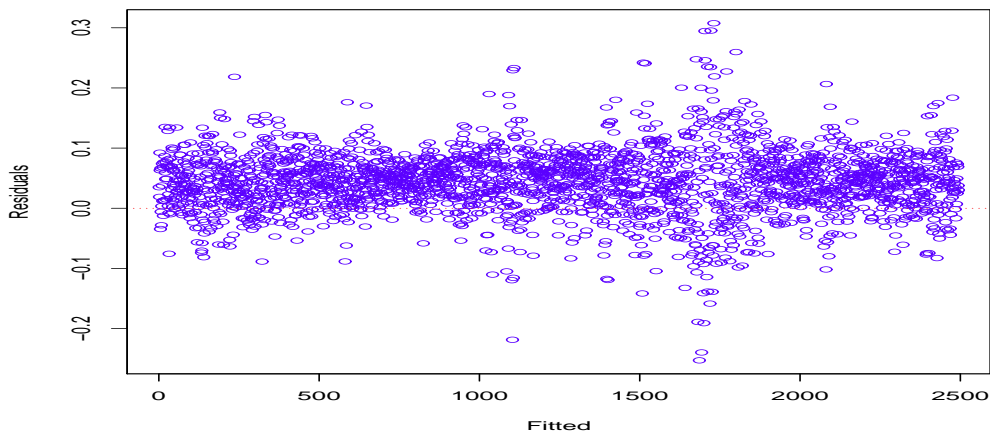
Figure 4.11: Residual against Fitted plot

4.6 GARCH-type models

The returns series data is not constant over time, which suggests that heteroskedasticity issue needs to be taken into account. It is reasonable to use GARCH-type models to acquire heteroscedasticity in the JSE returns data. There are fundamental distinctions between the symmetric GARCH-type models that are used to model ordinary volatility clustering and asymmetric GARCH-type models that are designed to capture leverage effects. GARCH (1, 1) model is chosen because it is the simplest specification and the most widely used in the literature. This study uses GARCH (1, 1) model and GARCH (1, 1)-M model to resolve ordinary volatility clustering. EGARCH (1, 1) model and TGARCH (1, 1) model are designed to capture leverage effects. GARCH-type models are used with ARMA (1, 0) model as mean equation.

4.6.1 ARMA (1, 0)-GARCH (1, 1) model

Results of the ARMA (1, 0)-GARCH (1, 1) model for returns are shown in Table 4.6. The mean equation can be written as rt = 0.0854 + [AR (1) = 0.0497], as shown in Table 4.6. The estimate of ϕ is statistically significant at 5% level supporting the use of ARMA (1, 0) model for the returns. Theory expects parameters ϕ and α to be higher than zero and β to be positive to ensure that the conditional variance (σ_t^2) is non-negative. Table 4.6 shows that ϕ and α are greater than zero and significant at 5% level. Thus, GARCH (1, 1) model seems to be quite good for explaining the behaviour of stock returns volatility in JSE.

Table 4.6: ARMA (1, 0)-GARCH (1, 1) models for returns

Mean Equation			
Coefficient	Std. Error	z-Statistics	P-value
$\mu = 0.0854$	0.0218	3.9089	0.0001
$\phi = 0.0497$	0.0497	2.4290	0.0151
Variance Equation			
Coefficient	Std. Error	z-Statistics	P-value
$\omega = 0.0228$	0.0062	3.6673	0.0002
$\alpha = 0.0938$	0.0128	7.3211	0.0000
$\beta = 0.8933$	0.0132	67.5847	0.0000

Volatility shocks are persistent since the sum of the ARCH and GARCH coefficients are very close to one as shown in Table 4.7. The Ljung-Box Q-statistics of order 20 computed on the standardised residuals and squared standardised residuals are presented by $Q(20)$ and $Q^2(20)$, respectively. The Ljung-Box $Q(20)$ and $Q^2(20)$ statistics up to twentieth order autocorrelation are both less than $\chi^2_{20,0.05} = 31.410$ suggesting that the hypothesis of independence in daily returns should not be rejected. Engle's LM test indicates that there are no more ARCH effects left up to lag 10. In Table 4.7, the values in parenthesis denote the standard errors.

Table 4.7: ARMA (1, 0)-GARCH (1, 1) model diagnostic

$\alpha + \beta$	0.9871
$Q(20)$	17.2110 (0.5760)
$Q^2(20)$	16.4260 (0.6290)
ARCH (10)	0.0297 (0.1655)
AIC	3.1451
Durban-Watson statistics	2.0186

4.6.2 ARMA (0, 1)-GARCH (1, 1)-M model

Table 4.8 summarises the parameter estimates for the ARMA (0, 1)-GARCH (1, 1)-M model. The coefficient denoted by β_1 is positive and insignificant, meaning that

increased risk does not necessarily imply higher returns. The coefficient of α is significantly positive and less than one, indicating that the impact of "old" news on volatility is significant with persistent shocks.

Table 4.8: ARMA (1, 0)-GARCH (1, 1)-M models for returns

Mean Equation			
Coefficient	Std. Error	z-Statistics	P-value
$\mu = 0.0321$	0.07626	0.4207	0.6740
$\phi = 0.0508$	0.0223	2.2820	0.0225
$\beta_1 = 0.0502$	0.0719	0.6983	0.4850
Variance Equation			
Coefficient	Std. Error	z-Statistics	P-value
$\omega = 0.0246$	0.0075	3.2966	0.0010
$\alpha = 0.0926$	0.0116	7.9817	0.0000
$\beta = 0.8935$	0.0133	67.1677	0.0000

Table 4.9 shows that volatility shocks are persistent since the sum of the ARCH and GARCH coefficients are very close to one. The Ljung-Box Q-statistics of order 20 computed on the standardised residuals and squared standardised residuals are presented by $Q(20)$ and $Q^2(20)$ respectively. The Ljung-Box $Q(20)$ and $Q^2(20)$ statistics up to twentieth order autocorrelation are both less than $\chi^2_{20,0.05} = 31.410$ suggesting that the hypothesis of independence in daily returns should not be rejected. Engle's LM test indicates that there are no more ARCH effects left up to lag 10. In choosing the symmetric model, it is accepted that GARCH (1, 1) model is the best due to the values of the AIC while Durban-Watson statistics is almost the same in these two models.

Table 4.9: ARMA (1, 0)-GARCH (1, 1)-M model diagnostic

$\alpha + \beta$	0.9861
$Q(20)$	17.1090 (0.5820)
$Q^2(20)$	16.8940 (0.5970)
ARCH (10)	0.0297 (0.1398)
AIC	3.1457
Durban-Watson statistics	2.0187

4.6.3   ARMA (1, 0)-EGARCH (1, 1) model

The skewness statistics of the returns series generated in Table 4.3 imply the asymmetry in the data set. This study employs EGARCH and TGARCH models for asymmetric persistence.

Table 4.10 presents results for the ARMA (1, 0)-EGARCH (1, 1) model. The leverage effect term, γ is negative, indicating the existence of the leverage effect in future returns during the sampling period. Since $\gamma \neq 0$ the news impact is asymmetric, supporting the use of the skewed Student-t distribution for z_t (the standardised residuals). The parameter β measures the persistence in conditional volatility irrespective of anything happening in the market. The value of parameter β is 0.9819 which is closer to 1, implying that volatility will take long time to die in the Johannesburg stock market.

Table 4.10: ARMA (1, 0)-EGARCH (1, 1) model for returns

Mean Equation			
Coefficient	Std. Error	z-Statistics	P-value
$\mu = 0.0365$	0.0217	1.6814	0.0927
$\phi = 0.0459$	0.0202	2.2758	0.0229
Variance Equation			
Coefficient	Std. Error	z-Statistics	P-value
$\omega = -0.0896$	0.0158	-5.6599	0.0000
$\alpha = 0.1189$	0.0205	5.7989	0.0000
$\beta = 0.9819$	0.0034	272.0490	0.0000
$\gamma = -0.0961$	0.0146	-6.5809	0.0000

In Table 4.11, the sum of α and β is above one, suggesting that shocks to the conditional variance are highly persistent. This implies that large changes in returns tend to be followed by large changes, and small changes tend to be followed by small changes which confirm that volatility clustering is observed in JSE returns series. The Ljung-Box statistics up to twentieth order autocorrelation are less than the critical value from the χ^2 distribution at 5% level of significance suggesting that shocks to the conditional variance are highly persistent. Engle's LM test indicates that there are no more ARCH effects left up to lag 10. The values in parenthesis in Table 4.11 denote the standard errors.

Table 4.11: ARMA (1, 0)-EGARCH (1, 1) model diagnostic

$\alpha + \beta$	1.1008
$Q(20)$	17.3940 (0.5630)
$Q^2(20)$	23.5110 (0.1440)
ARCH (10)	0.0293 (0.1958)
AIC	3.1205
Durban-Watson statistics	2.0131

4.6.4   ARMA (1, 0)-TGARCH (1, 1) model

The ARMA (1, 0)-TGARCH (1, 1) model for returns shown in Table 4.12 indicates that the news impact is asymmetric since $\gamma \neq 0$. The parameter γ is significantly positive which indicates that the leverage effect exists, which in turn implies that bad news will increase volatility.

Table 4.12: ARMA (1, 0)-TGARCH (1, 1) model for returns

Mean Equation			
Coefficient	Std. Error	z-Statistics	P-value
$\mu = 0.0359$	0.0219	1.6383	0.1014
$\phi = 0.0494$	0.0201	2.4569	0.0140)
Variance Equation			
Coefficient	Std. Error	z-Statistics	P-value
$\omega = 0.0311$	0.0068	4.5909	0.0000
$\alpha = -0.0022$	0.0156	-0.1415	0.8875
$\beta = 0.9062$	0.0141	64.3148	0.0000
$\gamma = 0.1482$	00.0213	6.9614	0.0000

Table 4.13 shows that persistence in volatility shocks is evident as the sum of the ARCH and GARCH terms makes 0.9040 which is close to one. The TGARCH model captures asymmetric effect better than EGARCH model, and it is therefore regarded as the best model for asymmetry in the JSE return series. The Ljung-Box $Q(20)$ and $Q^2(20)$ statistics up to twentieth order autocorrelation are both less than $\chi^2_{20,0.05} = 31.410$ suggesting that the hypothesis of independence in daily returns not be rejected. Engle's LM test indicates that there are no more ARCH effects left up to lag 10. In Table 4.13, the values in parenthesis denote the standard errors.

Table 4.13: ARMA (1, 0)-TGARCH (1, 1) model diagnostic

$\alpha + \beta$	0.9040
Q (20)	18.2840 (0.5040)
$Q^2(20)$	25.3430 (0.1500)
ARCH (10)	0.0311 (0.1645)
AIC	3.1209
Durban-Watson statistics	2.0199

4.7 Generalised Extreme Value Distribution

Financial data is well known to be heavy tailed and Extreme Value Theory (EVT) has been shown to be a very useful tool in estimating and predicting the extreme behaviour of financial products. EVT has arisen as a new methodology to analyse the tail behaviour of stock returns. The application of the EVT to statistics allows us to fit models to data from the upper and lower tails and provides a firm theoretical foundation on which we can build statistical models describing extreme events. Extreme share returns on stock markets are of particular interest to fund managers, investment analysts and financial regulators. This is because extreme returns affect the whole stock market and may dramatically reduce the benefits of risk variation.

An extreme share price change can significantly affect the performance of an investment over a long time period (e.g. a year) or even threaten the stability of the financial system. A popular assumption usually made is that financial logarithmic returns follow a normal distribution. There is empirical evidence that distributions of returns can possess fat or heavy tails so that a careful analysis of returns is required. The Generalised Extreme Value (GEV) distribution has been successfully used to model the extreme returns events for many countries and regions. The importance of GEV distribution arises from the fact that it is the limit distribution of the maxima of a sequence of independent and identically distributed random variables. This section seeks to establish whether or not, the returns on the JSE price index follow a heavy tailed distribution.

Figure 4.12 shows the dataset of positive and negative returns at the JSE from the period 08 January, 2002 to 30 December, 2011. In all the 2495 returns, there are 1321 positive returns and 1174 negative returns as displayed in Figure 4.12. This shows that there was more chance for investors to sell their shares at profit during the given time interval.
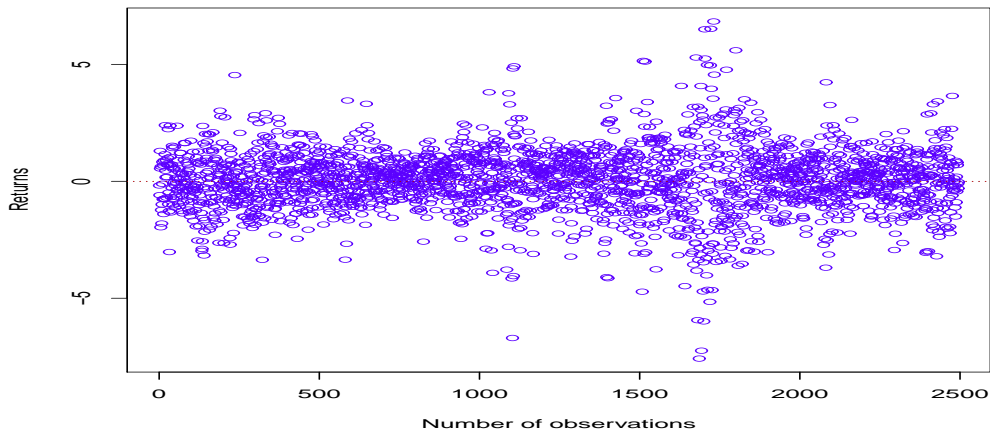

Figure 4.12: Scatter plot of the returns

4.7.1 Tail quantiles

Maximum likelihood fitted parameter values and their corresponding standard errors are summarised in Table 4.14. The parameter estimate of $\hat{\xi}$ obtained (-0.1897) implies that the underlying daily returns can be modelled using the Weibull class of distributions, since $\hat{\xi} < 0$.

Table 4.14: GEV parameter estimates from fitting daily returns

Parameter	Estimate	Standard Error
Location ($\hat{\mu}$)	-0.4917	0.0309
Scale ($\hat{\sigma}$)	1.4299	0.0185
Shape ($\hat{\xi}$)	-0.1897	0.0034

4.7.2 Model checking

The diagnostic plot shown in Figure 4.13 is used to assess the accuracy of the GEV distribution for the returns. The diagnostic plot generated in Figure 4.13 uses ismev package of R established by Coles (2001). The probability plot and quantile plot provide techniques for assessing whether or not a data set follows a given distribution. The data is plotted against a theoretical distribution in such a way that the points form a straight line. Departures from this straight line indicate departures from the specified distribution. The return level plot shows the return period against the return level, and shows an estimated 95% confidence interval. The return level is the level that is expected to be exceeded. The return period is the amount of time expected to wait to exceed a particular return level. Thus, all these diagnostic plots support the fitted model.
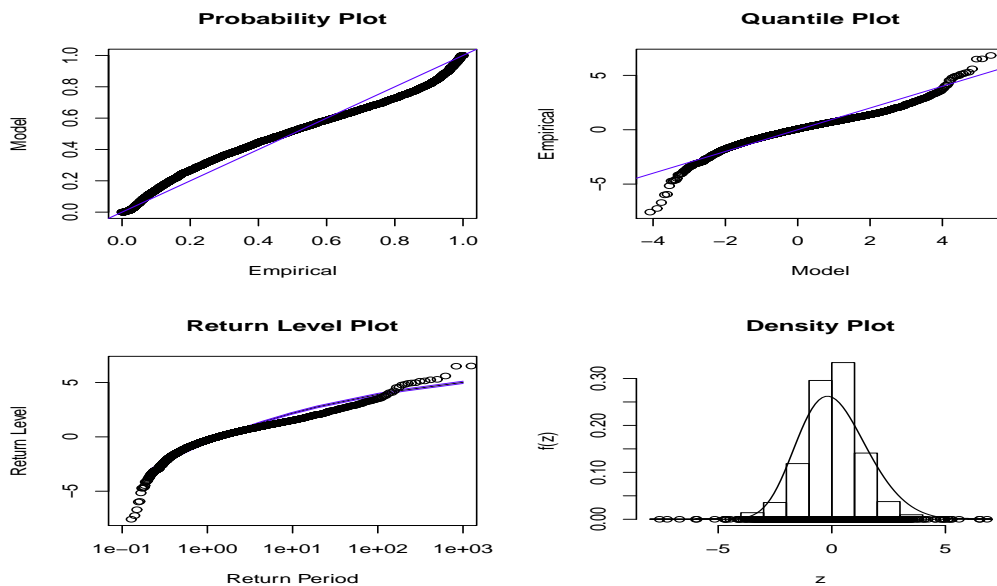

Figure 4.13: Diagnostic plot illustrating the fit of the data to the GEV distribution

Table 4.15 shows the number of exceedances above the extreme tail quantiles. The theoretical number of exceedances of 95% tail probability over a 125 return is calculated as 0.05 * 2495. The number of exceedances decreases as the quantiles increase. The Conditional GEV is computed using quantile function. In place of 95% tail probability $x_{0.05} = -0.4917 - \frac{1.4299}{0.1897} \left\{ \left[-\log(1 - 0.05) \right]^{0.1897} - 1 \right\} = 2.7552$ and the number

57

of exceedances above 2.7552 is 54.

Table 4.15: Estimated tail quantiles at different probabilities (Number of exceedances)

Tail probability (p)	Expected observation	Conditional GEV
0.1	250	109
0.05	125	54
0.01	25	18
0.005	12	14
0.001	2	7

Figure 4.14 shows that the GEV distribution predicts accurately the number of exceedances above very high quantiles. The expected observations are represented by a solid line while the dashed line represents GEV distribution.
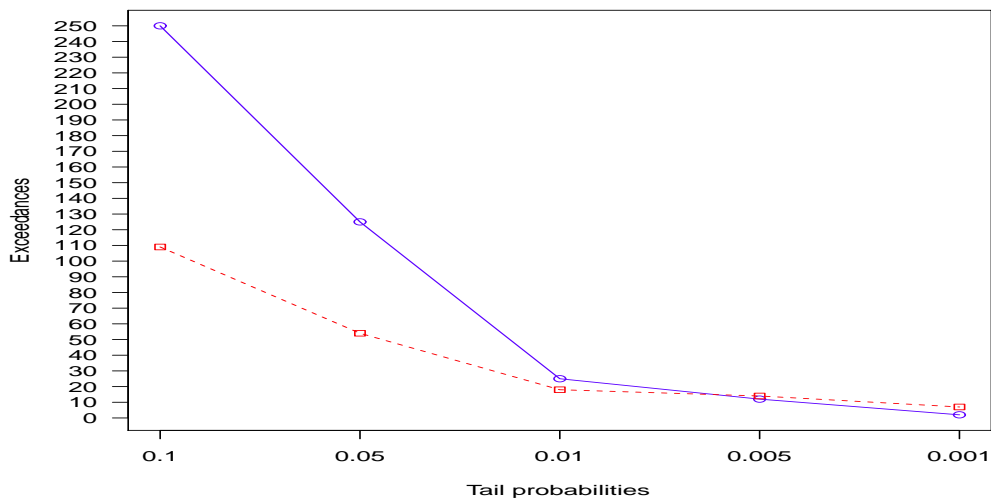
Figure 4.14: Plot of exceedances against tail probabilities

4.7.3 Monthly frequency analysis

The monthly frequency analysis of observations above 95th quantile is carried out. Table 4.16 shows the frequency of occurrence of return values above the 95th quantile. This indicates that May and October have the highest returns over the sampling period (2002-2011). The investigation provides an important implication to

investors and risk managers when modelling extreme events in the JSE. The extreme exceedances are also presented by bar-plot in Figure 4.15.

Table 4.16: Monthly frequency analysis of observations above 95th quantile

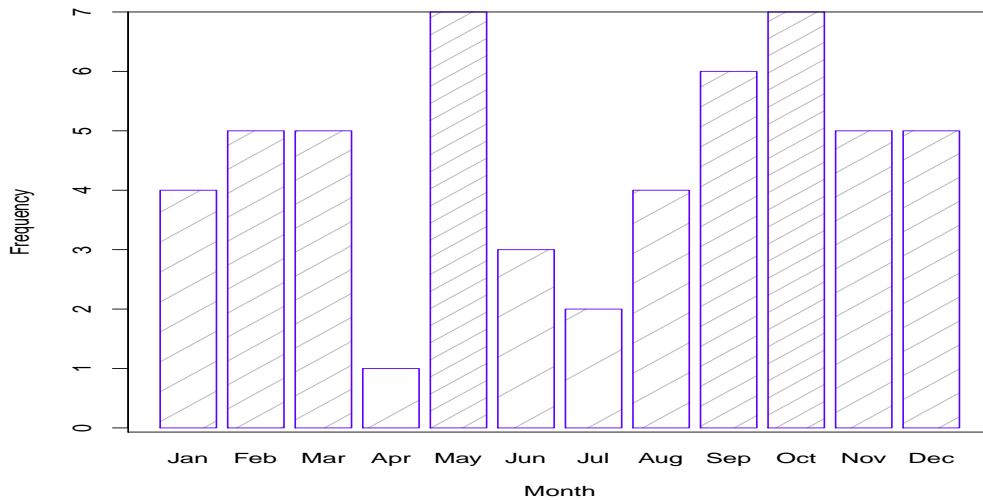| Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Freq	4	5	5	1	7	3	2	4	6	7	5	5

Figure 4.15: Bar chart of the monthly frequency above 95th quantile.

4.8   Generalised Pareto Distribution

Accurate modelling of extreme returns is vital to financial risk management. Risk management gained importance in the last decade due to the increase in the volatility in financial markets. The common assumption in finance theory is that financial returns are normally distributed. Conversely, several tail studies indicate that most financial time series are fat tailed. Investing in financial markets is challenging because prices are highly volatile and exhibit extreme price movement of magnitudes.

It is important to note that EVT relies on an assumption of independent and identically distributed (i.i.d.) observations. There is empirical evidence that distribution of

59

returns can possess fat or heavy tails so that a careful analysis of returns is required. The GPD was first introduced by Pickands (1975) in the extreme value framework as a distribution of the sample excesses (or exceedances) above a sufficient high threshold. The POT method models a distribution of excess over a given threshold. The importance of the GPD in the extreme value theory is dominant and extensively used in various practical situations.

4.8.1   Model selection

The key assumption in EVT is that extreme returns are i.i.d. series. Hence, before using an EVT method we may want to check that observations are approximately i.i.d. In case the observations do not fulfil this hypothesis, it is important to apply a filter before fitting the EVT model. In particular it is very common to find heteroskedasticity in financial returns series. According to the AIC, the ARMA (1, 0)-EGARCH (1, 1) is a better model but it does not fall in the constraint $\alpha + \beta = 1$. The fitted model for ARMA (1, 0)-GARCH (1, 1) model shows that $\alpha + \beta = 0.9871$, which is closer to 1. ARMA (1, 0)-GARCH (1, 1) model is estimated in a view of filtering the return series to obtain nearly i.i.d. residuals.

The results of fitting an ARMA (1, 0)-GARCH (1, 1) model for the JSE returns series are presented in Table 4.17. The estimate of ϕ is significant in supporting the use of ARMA (1, 0) model for the returns. Volatility shocks are persistent since the sum of the ARCH and GARCH coefficients are very close to one. The estimates for α and β are highly significant. The Box-Pierce Q-statistics is insignificant up to lag 20 indicating that there is no excessive autocorrelation left in the residuals. The critical value $\chi^2(20)$ distribution is 31.41. In all cases 5% level of significance is used and p-values are shown in the parentheses. In modelling volatility of heteroskedastic shares, EVT is applied to standardised residuals from this model. The model diagnostic results support the use of the ARMA (1, 0)-GARCH (1, 1) model which is applied to filter returns in obtaining nearly i.i.d.

Table 4.17: ARMA (1, 0)-GARCH (1, 1) model for returns

Mean Equation
$\mu = 0.0854(0.0001)$
$\phi = 0.0497(0.0151)$
Variance Equation
$\omega = 0.00228(0.0002)$
$\alpha = 0.0938(0.0000)$
$\beta = 0.8933(0.0000)$
Model Diagnostics
$\alpha + \beta = 0.9871$
$Q(20) = 17.2110(0.5760)$
$Q^2(20) = 16.4260(0.6290)$
$ARCH(10) = 0.098048(0.0002)$

4.8.2 The GPD fit

The Pareto quantile plot is a graphical method for inspecting the parameters of Pareto distribution. The logarithm of the observed positive residuals is plotted against the theoretical quantiles. The Pareto quantile plot shown in Figure 4.16 is illustrated by using the laeken package of R established by Beirlant *et al.* (1996) to predict a threshold. Figure 4.16 displays the Pareto quantile plot for the positive residual data. The threshold is $\tau = \exp(0.9634) = 2.6206$. There are 58 observations above the threshold as shown in Figure 4.17. The maximum likelihood estimation is used for the determination of the GPD parameters from 58 exceedances. The estimated GPD parameters are $\hat{\sigma} = 1.0896$ (0.2442) and $\hat{\xi} = $ -0.0321 (0.1824), with standard errors in the parenthesis. The results show that residuals can be modelled using the Weibull class of distributions, since $\xi < 0$.
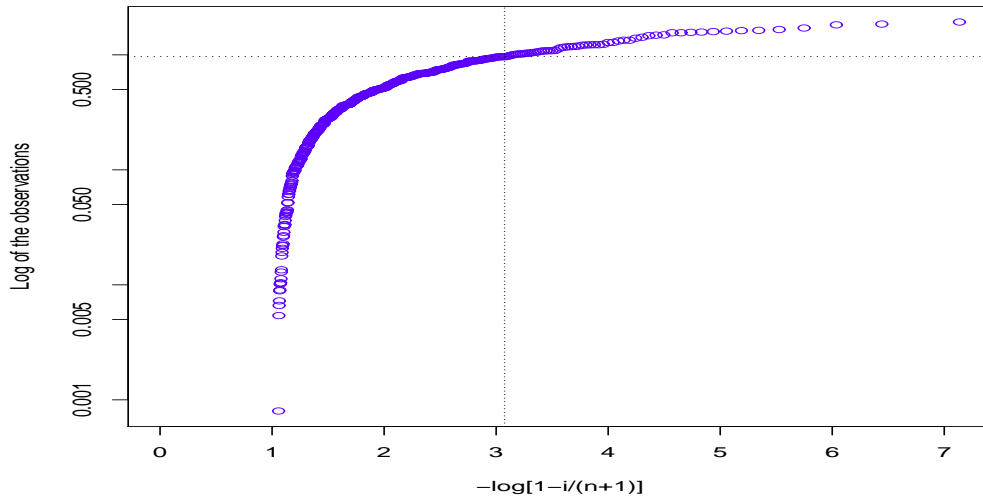
Figure 4.16: Pareto quantile plot

Figure 4.17 shows the plot of all 1255 positive residuals. The residuals above a horizontal line are those that are above a threshold $\tau = 2.6206$. The observations above the threshold are assumed to follow GPD.
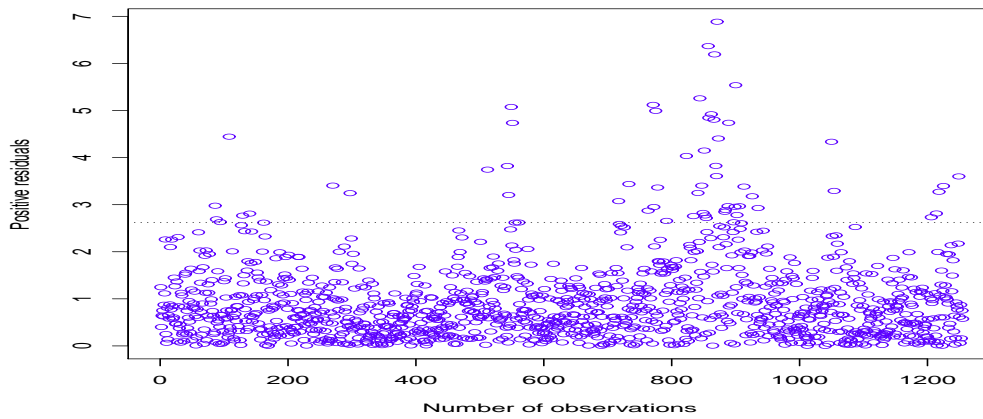

Figure 4.17: Scatter plot of positive residuals

Diagnostic plots for the fitted GPD are shown in Figure 4.18. The GPD model is adequate since the probability and quantile plots consist of points close to the unit diagonal, indicating a good fit to the data.
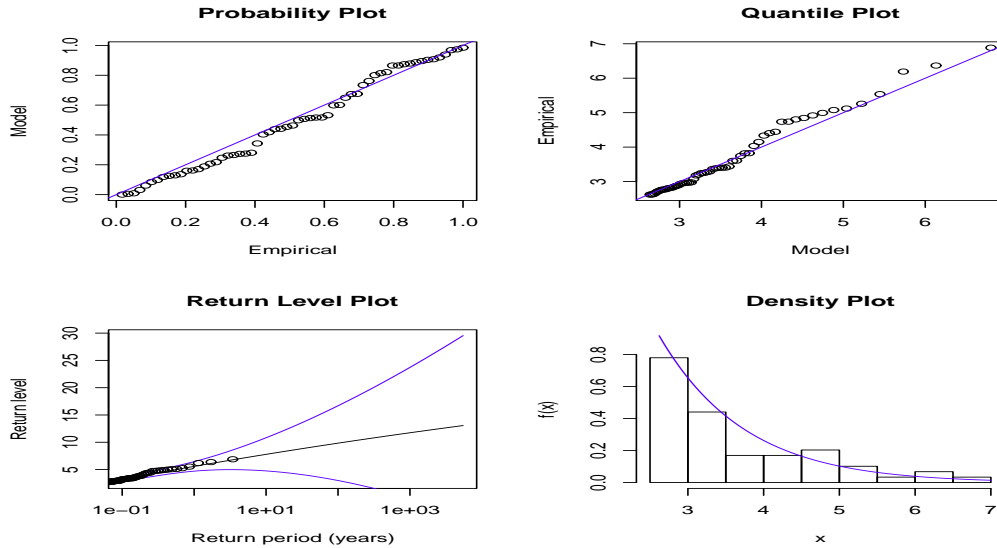

Figure 4.18: Diagnostic plot illustrating the fit of the data to the GPD

4.8.3 Estimation of extreme quantiles

Table 4.18 shows the number of the exceedances related to the corresponding tail probabilities. It is discovered that for both expected observations and conditional GPD, the number of exceedances decrease as the tail probability increase. The method to calculate quantiles works well if the observed number of exceedances is close to the number of expected observations. The number of expected observations can be calculated by multiplying the number of residuals by tail quantiles, Bystrom (2005). The theoretical number of exceedances of a 95% tail quantile over positive residuals of 1255 is (0.05*1255) = 63. ARMA-GARCH model is presented with conditionally normally distributed standardised residuals.

Table 4.18: Estimated tail quantiles at different probabilities (Number of exceedances)

Tail probability (p)	Expected	ARMA-EGARCH	Conditional GPD
0.1	126	161	76
0.05	63	107	34
0.01	13	53	12
0.005	6	40	3
0.001	1	17	0

The conditional GPD is attained by applying the quantile function and count the number of observations that are larger than the estimated tail quantile. The ARMA-GARCH model significantly underestimates all tail quantiles resulting in an excessive number of exceedances. The conditional GPD distribution produces better forecasting results, thus the ARMA-GARCH-GPD model yields more accurate estimates of extreme tail quantile.

The findings of this study also show that the ARMA-GARCH-GPD model performs well, especially in financial markets where the distribution of returns exhibits large movements. Plot of exceedances against tail probabilities for expected observations is given in Figure 4.19. The solid line represents the expected observations, the blue dashed line is the conditional GPD and the red dashed line represents the ARMA-GARCH model.
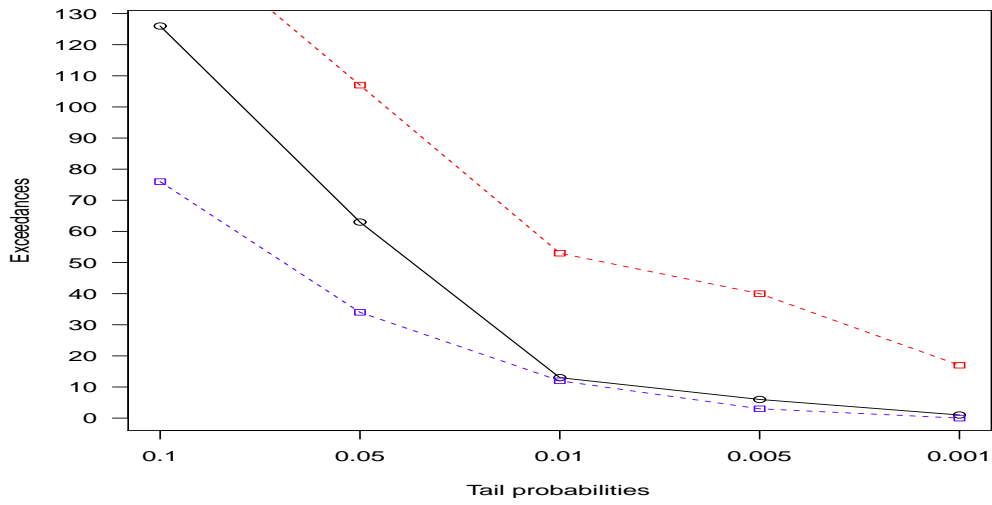
Figure 4.19: Plot of exceedances against tail probabilities

Figure 4.20 shows the bar chart of the monthly frequency of occurrence of exceedances. All months have the highest frequencies above the threshold except in April and July. October is the month with highest exceedances above threshold. This investigation provides an important implication to investors and risk managers when modelling extreme events in the JSE.

Figure 4.20: Bar chart of the monthly frequency of occurrence of exceedances

The bar chart of the yearly frequency of occurrence of exceedances is given in Figure 4.21. The number of exceedances were high in 2008, which was caused by the 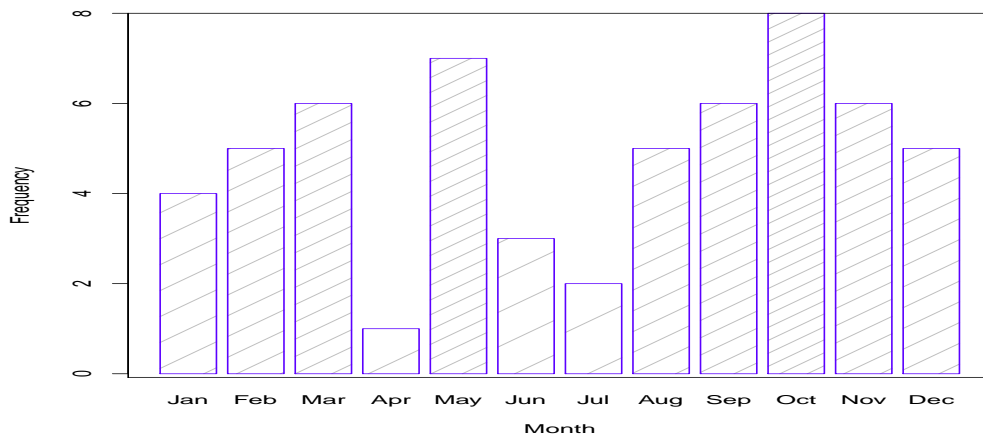global recession of 2008/2009, resulting in a large increase in unemployment and a deflationary scare in many countries.
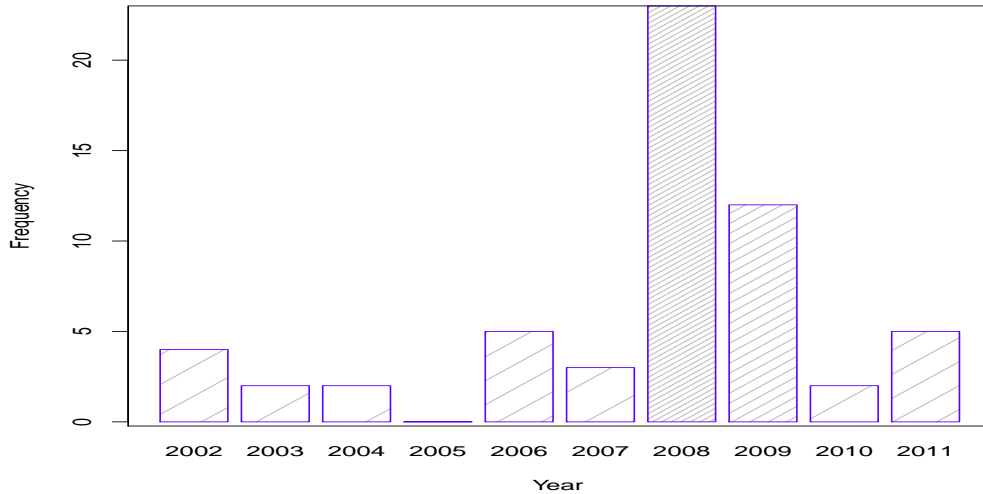

Figure 4.21: Bar chart of the yearly frequency of occurrence of exceedances

4.9 Out of sample forecasting

A good forecast capability of volatility models provide a practical tool for stock market analysis and enable investors to give more appropriate securities pricing strategies. The forecast error statistics used in the evaluation of the out of sample predictions takes the period from 10 January, 2012 to 28 December, 2012 generating 244 observations. The Root Mean Square Error (RMSE) statistics and Mean Absolute Error (MAE) statistics are used to rank the models based on their out of sample forecasting accuracy. The alternative GARCH models and some diagnostic checking are performed to compare volatility specifications. Different ARMA (p, q)-GARCH (1, 1) models are selected as the representative models to compare out of sample performance. The ACF and PACF are used to test all possible models. The out of sample forecasting for the possible ARMA (p, q)-GARCH (1, 1) models assume three different distributions namely: Normal distribution, Student-t distribution and Generalised

error distribution. The four models used are ARMA (1, 0)-GARCH (1, 1), ARMA (1, 1)-GARCH (1, 1), ARMA (1, 2)-GARCH (1, 1) and ARMA (1, 3)-GARCH (1, 1).

The RMSE statistics and MAE statistics for the four ARMA (p, q)-GARCH (1, 1) are reported in Table 4.19. The RMSE statistics indicates that ARMA (1, 1)-GARCH (1, 1) model provides the most accurate forecast, followed by ARMA (1, 0)-GARCH (1, 1) model and ARMA (1, 3)-GARCH (1, 1) model. The MAE statistics indicates that ARMA (1, 1)-GARCH (1, 1) model provides the most accurate forecast, followed by ARMA (1, 3)-GARCH (1, 1) model.

Table 4.19: Comparison of out of sample forecast accuracy models

Model	RMSE	MAE
ARMA(1, 0)-GARCH(1, 1)	0.7056	0.5624
ARMA(1, 1)-GARCH(1, 1)	0.7045	0.5604
ARMA(1, 2)-GARCH(1, 1)	0.7057	0.5625
ARMA(1, 3)-GARCH(1, 1)	0.7056	0.5622

The error statistics in Table 4.19 are unable to provide the obvious distinction between the three performing models. The data justification is performed to provide the best accurate model. Table 4.20 shows results for data validation for the out of sample forecast which is performed using data from 07 January, 2013 to 02 August, 2013. The out of sample forecast for data validation assumes that the models follow Normal distribution, Student-t distribution and Generalised error distribution. The AIC and MAE indicate that the ARMA (1, 3)-GARCH (1, 1) model provides the most accurate forecast.

Table 4.20: Data validation

Model	AIC	RMSE	MAE
ARMA(1, 0)-GARCH(1, 1)	11.9497	89.0309	25.5215
ARMA(1, 1)-GARCH(1, 1)	11.8412	88.5014	19.6507
ARMA(1, 2)-GARCH(1, 1)	11.5865	88.5256	20.4412
ARMA(1, 3)-GARCH(1, 1)	9.5721	88.7082	8.2686

4.10 Conclusion

Chapter 4 has presented and analysed the results pertaining to volatility and financial market risk at the JSE. The behaviour of the data and the stationarity tests on the data were first presented. The basic picture revealed the properties of financial data such as excess volatility, volatility clustering, excess kurtosis and non-normality. Results show that returns are characterised by an ARMA (1, 0) process. The symmetric and asymmetric GARCH-type models are used to model volatility and financial risks at the JSE. The ARMA (1, 0)-GARCH (1, 1)-M model shows that increased risk does not necessarily imply an increase in returns. The TGARCH model captures asymmetric effect better that the EGARCH model. The out of sample forecasting evaluations indicate that ARMA (1, 0)-GARCH (1, 1) model achieves most accurate volatility forecast. In modelling the extreme returns, the two most EVT distributions (GEV distribution and GPD) are used. The GEV distribution shows that daily returns for the JSE follow the Weibull class. The results show that the ARMA-GARCH-GPD model yields more accurate estimates of extreme returns than the ARMA-GARCH model. The out of sample forecast indicates that the ARMA (1, 3)-GARCH (1, 1) model provides the most accurate results.

Chapter 5

Conclusion

5.1 Introduction

This chapter gives a report of the conclusion related to the objectives and research problem of the study. The chapter is divided into five sections. Following this introductory part, section 5.2 discusses research findings and section 5.3 states the limitations and outlines some recommendations of the study. Section 5.4 provides areas for further study and section 5.5 concludes the chapter.

5.2 Research findings

The study investigates the estimation and forecast ability of ARMA (p, q) model and four GARCH-type models. EVT is used in the study for the asymptotic behaviour of extreme observations. The JSE returns series exhibits the defining characteristics of financial markets, which are high volatility, occasional extreme movements and volatility clustering. The conclusion that daily returns series observed in JSE have non-normality distribution is reasonable as it is common phenomenon in datasets of emerging markets. Overall results provide evidence to show volatility clustering, leptokurtic distribution and leverage effects for the Johannesburg stock returns data. The high values of kurtosis for the returns suggest that extreme price changes occurred more frequently during the sample period, 08 January, 2002 to 30 December, 2011.

The JSE price index series achieved stationarity after taking the 1st-difference. Results show that returns are characterised by an ARMA (1, 0) model. The results indicate that the daily returns can be characterised by the GARCH-type models. The Ljung-Box Q-statistics is tested up to the twentieth order autocorrelation which indicated that the null hypothesis of independence in daily returns should be rejected in all GARCH-type models. The financial theory that suggests that an increase in variance results in a higher expected return does not hold in the JSE returns data. The Ljung-Box Q-statistics justified the use of all GARCH-type models in modelling volatility at the JSE. The out of sample forecasting evaluations indicate that the ARMA (1, 0)-GARCH (1, 1) model achieve the most accurate volatility forecast.

The GEV distribution illustrates that daily returns for the JSE follows the Weibull class. In summary, the probability plot and the quantile plot suggest that the GEV distribution is a good fit to the data. The exceedances above the 95th quantile indicate that May and October have the highest returns over the sampling period. The investigation provides an important implication to investors and risk managers when modelling extreme events in the JSE. In modelling heteroskedasticity using GPD, the ARMA-GARCH model was applied in stage one with a view of filtering the return series to obtain nearly i.i.d. residuals. In stage two, the EVT framework is applied to the standardised residuals from ARMA-GARCH model. The results show that residuals can be modelled using the Weibull class of distributions.

The ARMA-GARCH model overestimates all tail quantiles and thus the distribution disables to model the positive tail accurately. The ARMA-GARCH-GPD model yields more accurate estimates of extreme returns than the ARMA-GARCH model. In summary the results of this study support the combination of ARMA-GARCH model with EVT for estimating upper extreme quantiles. In particular, the results show that the participants in the JSE market can rely on EVT-based models like GPD when modelling conditional heteroskedasticity of extreme events. This study also provides an important implication to investors and risk managers. The study by Bystrom (2005) indicates that VaR performance under a GARCH-EVT framework is

superior to a number of parametric approaches. The findings of this study also show that the ARMA-GARCH-GPD model performs well, especially in financial markets where the distribution of returns exhibits large movement of magnitude.

5.3   Limitations and Recommendations

One limitation of this study is that it is confined to secondary data only. All limitations of GARCH-type models and EVT distribution are applicable to this study. GARCH-type models are recommended to be used to model and forecast volatility at the JSE, more specially the ARMA (1, 0)-GARCH (1, 1) model. The participants in the JSE market can rely on EVT-based models like GEV distribution and GPD when modelling volatility of extreme events.

5.4   Areas for further study

Future research should look at forecasting volatility of daily data using Markov regime switching GARCH models. Bayesian GARCH approach can be used in the estimation of the volatility of the residual returns. The impact of GEV quantile estimators has yet to be assessed and remains the focus of further research. This can have severe significances for risk management tools like VaR, as it could leave a financial market inadequately protected against extreme risk. Further combination of models of the overall time series with models of the tails appears to be an interesting issue to address in future work on extremes in financial markets.

5.5   Conclusion

This chapter presented the summarised results found in attaining the objectives and research problems of this study. The findings of this study have important implications for investors and risk managers at the stock markets. It may be worthwhile to investors to invest in industries and sectors that are generally more stable or less volatile. The findings are important to investors and risk managers who are concerned with assessing risk portfolios.

Bibliography

[1] Ahmed, A.E.M. and Suliman, S.Z. 2011. *Modeling stock market volatility using GARCH models evidence from Sudan.* International Journal of Business and Social Science 2(23): 114-128.

[2] Aktan, B., Korsakiene, R. and Smaliukiene, R. 2010. *Time-varying volatility modelling of Baltic stock markets.* Journal of Business Economics and Management 11(3): 511-532.

[3] Alberg, D., Shalit, H. and Yosef, R. 2008. *Estimating stock market volatility using asymmetric GARCH Models.* Applied Financial Economics 18: 1201-1208.

[4] Alexander, C. 2008. *Market risk analysis volume II: Practical Financial Econometrics.* John Wiley and Sons Ltd, Chichester.

[5] Angabini, A. and Wasiuzzaman, S. 2011. *GARCH models and the financial crisis: A study of the Malaysian stock market.* The International Journal of Applied Economics and Finance 5(3): 226-236.

[6] Bali, T.G. 2007. *The generalized extreme value approach to financial risk measurement.* Journal of Money, Credit and Banking 39(7): 1613-1649.

[7] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. 2004. *Statistics of Extremes: Theory and Applications.* Wiley Series in Probability and Statistics, England.

[8] Beirlant, J., Vynckier, P., and Teugels, J. 1996. *Tail index estimation, Pareto quantile plot, and regression diagnostics.* Journal of American Statistical Association 91(436): 1659-1667.

[9] Black, F. 1976. *Studies in stock price volatility changes, proceedings of the 1976 meeting of the business and economic statistics section.* American Statistical Association 177-181.

[10] Bollerslev, T. 1986. *Generalized autoregressive conditional heteroskedasticity.* Journal of Econometrics 31: 307-327.

[11] Box Jenkins methodology. *2006 Statistical Forecasting.* `http://www.statisticalforecasting.com/box-jenkins-methodology.php` (2012-11-12)

[12] Box, G.E.P. and Jenkins, G.M. 1976. *Time series analysis: forecasting and control, fourth edition.* Holden-day.

[13] Bystrom, H.N.E. 2005. *Extreme value theory and extremely large electricity price change.* International Review of Economics and Finance 14(1): 41-55.

[14] Chan, K.F. and Gray, P. 2006. *Using extreme value theory to measure value-at-risk for daily electricity spot prices.* International Journal of Forecasting 22(2): 283-300.

[15] Coles, S. 2001. *An Introduction to Statistical Modeling of Extreme Values.* Springer Series in Statistics, London.

[16] Cotter, J. 2007. *Extreme risk in Asian equity markets.* MPRA paper 3536. University Library of Munich, Germany.

[17] Djakovic, V., Andjelic, G. and Borocki, J. 2011. *Performance of extreme value theory in emerging markets: An empirical treatment.* African Journal of Business Management 5(2): 340-369.

[18] Emenike, K.O. 2010. *Modelling stock returns volatility in Nigeria using GARCH models.* Proceeding of International Conference on Management and Enterprice Development 1(4): 5-11.

[19] Engle, R.F. 1982. *Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation.* Econometrica 50(4): 987-1007.

73

[20] Floros, C. 2008. *Modelling volatility using GARCH models: Evidence from Egypt and Israel.* Middle Eastern Finance and Economics 2: 31-41.

[21] Glosten, L.R., Jaganathan, R. and Runkle, D. 1993. *Relationship between expected value and volatility of the nominal excess returns on stocks.* Journal of Finance 48(5): 1779-1801.

[22] Guru, A. 2012. *Extreme movements in the Indian equity markets: An extreme value theory approach.* International Journal of Social Sciences and Interdisciplinary Research 1(2): 6-18.

[23] Hamadu, D. and Ibiwoye, A. 2010. *Modelling and Forecasting the volatility of the daily returns of Nigerian Insurance Stocks.* International Business Research 3(2): 106-116.

[24] Hasan, H., Radi, N.F.A. and Kassim, S. 2011. *Modelling the distribution of extreme share return in Malaysia using Generalized Extreme Value (GEV) distribution.* AIP conference proceedings 1450: 82-89.

[25] Hein, M. 2008. *Modelling and Forecasting Volatility by Generalized Autoregressive Conditional Heteroskedastic-type models: the case of Vietnam Stock Exchange.* MA Dissertation in Finance and Investment. University of Nottingham, United Kingdom.

[26] Hill, R.C., Griffiths, W.E. and Lim, G.C. 2007. *Principles of Econometrics, 4th Edition.* John Wiley and Sons, New York.

[27] Huang, C., Lin, J. and Ren, Y.Y. 2012. *Testing for the shape parameter of Generalized Extreme Value distribution based on Lq-likelihood ratio statistic.* Springer-Verlag 76(5): 641-671.

[28] Johannesburg Stock Exchange. *Market, regulation and trade operations.* FTSE/JSE Africa Index series. `http://www.jse.co.za` (2011-21-02)

[29] Lee, W.C. 2012. *Fitting the Generalized Pareto distribution to commercial fire loss severity: evidence from Taiwan.* Journal of Risk 14(3): 63-80.

[30] Li, H. and Hong, Y. 2011. *Financial Volatility Forecasting with range-based Autoregressive Volatility Model.* Finance Research Letters 8(2): 69-76.

[31] Maghyereh, A.I. and Al-Zoubi, H.A. 2008. *The tail behaviour of extreme stock returns in the Gulf emerging markets: An implication for financial risk management.* Studies in Economics and Finance 25(1): 21-37.

[32] Mala, R. and Reddy, M. 2007. *Measuring stock market volatility in an emerging economy.* International Research Journal of Finance and Economics 8: 126-133.

[33] Mcleod, Y., Yu, H. and Mahdi, E. 2012. *Time series analysis with R.* A handbook of Statistics 30: 661-712. `http://www.jstor.org` (2012-01-09)

[34] Nelson, D.B. 1991. *Conditional heteroscedasticity in asset returns: A New Approach.* Econometrica 59(2): 347-370.

[35] Pickands, J. 1975. *Statistical Inference using extreme order statistics.* The Annals of Statistics 3(1): 119-131.

[36] Securities Services Act 36. 2004. *Financial Services Laws General Amendment Act.* Published in Government Gazette 475 (27190). Republic of South Africa. `http://www.jse.co.za` (2011-21-02)

[37] Shawky, A.I. and Abu-Zinadah, H.H. 2008. *Characterizations of the Exponentiated Pareto Distribution based on record values.* Applied Mathematical Sciences 2(26): 1283-1290.

[38] Shumway, R.H. and Stoffer, D.S. 2011. *Springer test in Statistics.* Time series analysis and its application, ARIMA models 83-171. `http://www.springerlink.com` (2011-05-09)

[39] Sigauke, C., Maposa, D., Mudimu, E. and Nyamugure P. 2010. *Volatility modelling using ARIMA-GARCH models in a hyperinflationary economic environment.* Proceedings of the 52nd annual conference of the South African Statistical Association for 2010. North-West University, Potchefstroom Campus.

[40] Sigauke, C., Verster, A. and Chikobvu, D. 2012. *Tail quantile estimation of heteroskedastic intraday increases in peak electricity demand.* Open Journal of Statistics 2(4): 435-443.

[41] Silverman, B.W. 1986. *Density estimation for statistics and data analysis.* Monographs of statistics and Applied probability. London, Chapman and Hall.

[42] Song, J. and Song, S. 2012. *A quantile estimation for massive data with Generalized Pareto distribution.* Computational Statistics and Data Analysis 56(1): 143-150.

[43] Tsay, R.S. 2011. *Analysis of financial time series, second edition.* Wiley series in probability and statistics. Financial time series and their characteristics 26-48. `http://www.sciencedirect.com` (2011-05-02)

[44] Wang, X. 2010. *The relationship between stock market volatility and macroeconomic volatility: Evidence from China.* Journal of Chinese Economic and Finance 2: 67-77.

[45] Zakoian, J.M. 1994. *Threshold heteroskedastic models.* Journal of Dynamics and Control 18(5): 931-955.

[46] Zhao, X., Scarrott, C.J., Oxley, L. and Reale, M. 2011. *GARCH dependence in extreme value models with Bayesian inference.* Mathematics and Computers in Simulation 81: 1430-1440.

Appendix A1

Derivations of the quantile function of the GEV distribution.

The cumulative distribution function of GEV distribution is given by:

$$F(x; \mu, \sigma, \xi) = \begin{cases} \exp\left\{-[1 + \xi(\frac{x-\mu}{\sigma})]^{\frac{-1}{\xi}}\right\}, 1 + \xi(\frac{x-\mu}{\sigma}) > 0 \text{ and } \xi \neq 0 \\ \exp\left(-e^{\frac{-(x-\mu)}{\sigma}}\right), \xi = 0 \end{cases}$$

The survival function of the GEV distribution is $P(X > x) = 1 - G_{x;\mu,\sigma,\xi}$

$$P(X > x) = \exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{\frac{-1}{\xi}}\right\}, \text{ for } 1 + \xi\left(\frac{x-\mu}{\sigma}\right) > 0 \text{ and } \xi \neq 0$$

Let $p = P(X > x)$

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

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

$$\log(1-p)^{-\xi} = -\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]$$

$$x_p = \mu + \frac{\sigma}{\xi}\left[-\log(1-p)^{-\xi} - 1\right]$$

Similarly when $\xi = 0$ we have $x_p = \mu\log\left[-\log(1-p)\right]$

77

Appendix A2

Derivation of the quantile function of the GPD.

The cumulative distribution function of GPD is given by:

$$W\xi(x) = \begin{cases} 1 - [1 + \xi(\frac{x-\tau}{\sigma})]^{\frac{-1}{\xi}}, \xi \neq 0 \\ 1 - \exp\left(\frac{-x-\tau}{\sigma}\right), \xi = 0 \end{cases}$$

The survival function of the GPD is $P(X > x) = 1 - W_{\xi(x)}$

$$P(X > x) = \begin{cases} 1 - [1 + \xi(\frac{x-\tau}{\sigma})]^{\frac{-1}{\xi}}, \xi \neq 0 \\ 1 - \exp\left(\frac{-x-\tau}{\sigma}\right), \xi = 0 \end{cases}$$

Let $p = P(X > x)$

$$p = 1 - [1 + \xi(\frac{x-\tau}{\sigma})]^{\frac{-1}{\xi}}, \text{ for } \xi \neq 0$$

$$(1 - p)^{\xi} = [1 + \xi(\frac{x-\tau}{\sigma})]$$

$$x_p = \tau + \frac{\sigma}{\xi}\left[(p)^{-\xi} - 1\right]$$

where x_p is the quantile function.

Similarly when $\xi = 0$ we have $x_p = \tau - \sigma\log(p)$

Appendix B

This Appendix contains some of the R codes used to generate the results and plots in Chapter 4

FITTING TIME SERIES PLOT
attach(Prices)
win.graph()
Prices.ts=ts(Prices,start=2002,end=2011,freq=260)
plot(Prices.ts,xlab="Date",ylab="All share JSE index",col="blue",main="")

attach(Returns)
win.graph()
Returns.ts=ts(Returns,start=2002,end=2011,freq=260)
plot(Returns.ts,xlab="Date",ylab="Daily returns",col="blue",main="")

FITTING Q-Q PLOT, DENSITY PLOT AND HISTOGRAM
attach(Returns)
head(Returns)
win.graph()
qqnorm(RT)

attach(Returns)
head(Returns)
win.graph()
plot(density(Returns)),col="blue",main="")
box()

attach(Returns)
head(Returns)

```
win.graph()
hist(returns,col="blue",main="")
box()
```

FITTING ACF AND PACF PLOTS SIMULTANEOUSLY
```
attach(Price)
win.graph()
par(mfrow=c(2,1))
acf(price,main="")
acf(price,type="p",main="")
box(ity="solid")
win.graph()
```

FITTING RESIDUALS
```
attach(residual)
head(residual)
win.graph()
win.graph()
acf(residual,ylab="ACF of Residual",main="")
box()
```

```
attach(res)
head(res)
win.graph()
hist(residual,xlab="Residual",col="blue",main="")
box()
```

```
attach(res)
win.graph() residual.plot=plot(residual,start=2002,end=2011,freq=260)
plot(residual,xlab="Number of observations",ylab="Residual",col="blue",main="")
box()
```

```
plot(residual,col="blue",main="")
attach(fitted)
win.graph()
plot(Fitted,xlab="Fitted",ylab="Residuals",col="blue",main="")
abline(h = 0.0, lty = 3,col="red")
box()
```

FITTING GENERALISED EXTREME VALUE DISTRIBUTION

```
attach(return)
library(ismev)
sharefit=gev.fit(return)
win.graph()
gev.diag(sharefit)

win.graph()
GEV=c(109,54,18,14,7)
EO=c(250,125,25,12,2)
g-range=range(0,260)
plot(EO,type="o",axes=F,col="blue", ylab="Exceedances", xlab="Tail probabili-
ties")
lines(GEV,type="o",pch=22,lty=2,col="red")
axis(1,at=1:5, lab=c("0.1","0.05", "0.01","0.005","0.001"))
axis(2, las=1, at=10*0:g-range[2])
box()

attach(Exceedance)
head(Exceedance)
win.graph()
barplot(Frequency, main="", xlab="Month",ylab="Frequency",
names=c("Jan","Feb","Mar","Apr","May","Jun","Jul","Aug","Sep","Oct","Nov",
```

"Dec"),border="blue",

density=c(4,5,5,1,7,3,2,4,6,7,5,5))

box()

FITTING GENERALISED PARETO DISTRIBUTION

attach(lnresidual)

head(lnresidual)

win.graph()

library(laeken)

paretoQPlot(lnresidual,w = NULL, xlab ="-log[1-i/(n+1)]", ylab ="Log of the ob-
servations",interactive = TRUE,main="",col="blue")

attach(residual)

head(residual)

win.graph()

plot(Residual,xlab="Number of observations",ylab="Positive residuals",col="blue",

xlim=c(0,1250))

abline(h = 2.6206, lty = 3)

attach(residual)

head(residual)

library(ismev)

sharefit=gpd.fit(Residual,2.6206)

win.graph()

gpd.diag(sharefit)

win.graph()

AEG=c(161,107,53,40,17)

GPD=c(76,34,12,3,0)

EO=c(126,63,13,6,1)

g-range=range(0,170)

```
plot(AEG,type="o",axes=F,col="blue", ylab="Exceedances", xlab="Tail probabil-
ities")
lines(EO,type="o",pch=22,lty=2,col="red")
lines(GPD,type="o",pch=22,lty=2,col="black")
axis(1,at=1:5, lab=c("0.1","0.05", "0.01","0.005", "0.001"))
axis(2, las=1, at=10*0:g-range[2])
box()


attach(Exceedance)
head(Exceedance)
win.graph()
barplot(Frequency, main="", xlab="Month",ylab="Frequency",
names=c("Jan","Feb","Mar","Apr","May","Jun","Jul","Aug","Sep","Oct","Nov",
"Dec"),border="blue",
density=c(4,5,6,1,7,3,2,5,6,8,6,5))
box()


attach(Exceedance)
head(Exceedance)
win.graph()
barplot(Frequency, main="", xlab="Year",ylab="Frequency",
names=c("2002","2003","2004","2005","2006","2007","2008","2009","2010","2011"),
border="blue",
density=c(4,2,2,0,5,3,23,12,2,5))
box()
```