STABLE DISTRIBUTIONS WITH APPLICATIONS TO SOUTH AFRICAN FINANCIAL DATA

UNIVERSITY OF KWAZULU - NATAL

INYUVESI YAKWAZULU-NATALI

Kimera Naradh (211501459)

July, 2024

Stable distributions with applications to South African financial data

by

Kimera Naradh (211501459)

A thesis submitted to the
University of KwaZulu-Natal
in fulfilment of the requirements for the degree
of
DOCTOR OF PHILOSOPHY
in
STATISTICS

Thesis Supervisor: Professor Knowledge Chinhamu

Thesis Co-supervisor: Professor Retius Chifurira

UNIVERSITY OF KWAZULU-NATAL
SCHOOL OF MATHEMATICS, STATISTICS AND COMPUTER SCIENCE
WESTVILLE CAMPUS, DURBAN, SOUTH AFRICA

Declaration - Plagiarism

I, Kimera Naradh (211501459), declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.

2. This thesis has not been submitted for any degree or examination at any other university.

3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then

   (a) their words have been rewritten but the general information attributed to them has been referenced, or

   (b) where their exact words have been used, then their writing has been placed in italics and referenced.

5. This thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source being detailed in the thesis and in the reference sections.

Kimera Naradh (211501459) (Student) Date

Professor Knowledge Chinhamu (Supervisor) Date

Professor Retius Chifurira (Co-supervisor) Date

Disclaimer

This document describes work undertaken as a PhD programme of study at the University of KwaZulu-Natal (UKZN). All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institution.

ABSTRACT

In recent times, researchers, analysts and statisticians have shown a keen interest in studying Extreme Value Theory (EVT), particularly with the application to mixture models in the medical and financial sectors. This study aims to validate the use of stable distributions in modelling three Johannesburg Stock Exchange (JSE) market indices, namely the All Share Index (ALSI), Banks Index and the Mining Index, as well as the United States of American Dollar (USD) to South African Rand (ZAR) exchange rate. This study leverages the unique properties of stable distributions when modelling heavy-tailed data. Nolan's S_0-parameterization stable distribution (SD) was fitted to the returns of the three FTSE/JSE indices and USD/ZAR exchange rate and a hybrid Generalized Autoregressive Conditional Heteroskedasticity (GARCH)-type model combined with stable distributions was fitted to each return series. The two-tailed mixture model of the Generalized Pareto Distribution (GPD), stable distribution, Generalized Pareto Distribution referred to as GSG, as well as the Stable-Normal-Stable (SNS) and Stable-KDE-Stable (SKS) was fitted to evaluate its relative performance in modelling financial data. Results show that the S_0-parameterization SD fits the South African financial returns well. The hybrid GARCH (1,1)-SD model competes favourably with the GARCH-GPD model in estimating Value-at-Risk (VaR) for FTSE/JSE Banks Index, FTSE/JSE Mining Index and the USD/ZAR exchange rate returns. The hybrid EGARCH (1,1)-SD competes well against the GARCH-GPD model for the FTSE/JSE ALSI returns. Inconclusive results are observed for the short position of the fitted GKG models; however, in the long position, an appropriate fit of the GPD-KDE-GPD (GKG) model, where KDE is the kernel density estimator, is emphasised for all four return series. The proposed mixture models, GSG, SNS and SKS models, are found to be a good alternative in fitting South African financial data to the commonly used GPD-Normal-GPD (GNG) mixture model. The results of this study are important to financial practitioners, risk managers and researchers as the proposed mixture models add more value to the literature on the applications of extreme mixture models.

Keywords: Stable distributions, Nolan's S_0-parameterization, mixture models, GPD-Normal-GPD, GPD-Stable-GPD, Stable-Normal-Stable, Stable-KDE-Stable, Kolmogorov-Smirnov test, Anderson-Darling test, VaR, Kupiec likelihood ratio

DEDICATION

To my beloved father, Mr. Prithiraj (Raj) Naradh, whose belief in my dreams never wavered. Your presence is deeply missed but your spirit lives on in every word of this thesis.

ACKNOWLEDGEMENTS

First and foremost, I offer my heartfelt gratitude to God for the guidance, blessings, and strength throughout this PhD journey. Divine grace has been a source of comfort and perseverance during both challenging and joyful times.

As I reach the end of this PhD journey, I am filled with gratitude for the many people who have supported and inspired me along the way.

I am deeply indebted to my supervisors, Professor Knowledge Chinhamu and Professor Retius Chifurira, for their unwavering support, expert guidance and encouragement. Their constructive criticism and praise have been transformative in shaping me into a more resilient PhD candidate and a more compassionate human being. I am privileged to be under their supervision and mentorship. Their invaluable inputs played a pivotal role in shaping this thesis.

I am eternally grateful to the University of KwaZulu-Natal for providing an excellent academic environment and the necessary resources for my research. A special thank you to Vanessa-Lynn Neophytou for her prompt and thorough editing assistance. Her insights significantly enhanced the overall quality of this thesis.

My utmost appreciation and respect goes out to my parents Raj and Nisha for their efforts, support, sacrifices and instilling the values of hard work and perseverance in me. To my sister, Yoshka, I am infinitely thankful for you always being my cheerleader and for the countless moments of laughter and support, your belief in me, even when I doubted myself made this journey more bearable. To my little pug, Kairo, your loyal companionship and presence provided me with a sense of calm with your playful demeanour reminding me to find joy in the little things even in the midst of long working hours. My family's unconditional love, encouragement and unwavering faith in my dreams is the cornerstone of my success.

Lastly, I would also like to take a moment to acknowledge and thank myself for the perseverance and dedication which have been essential in completing this PhD journey. This journey was challenging but my determination to overcome obstacles and stay focused has brought me to this achievement. I am grateful for my strength, resilience and hard work that have made this thesis possible.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AD	Anderson Darling
ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criterion
ALSI	All Share Index
ARCH	Autoregressive Conditional Heteroskedasticity
APARCH	Asymmetric Power Autoregressive Conditional Heteroskedasticity
BIC	Bayesian Information Criterion
CDF	Cumulative Density Function
DF	Dickey-Fuller
DIC	Deviance Information Criterion
ECF	Empirical Characteristic Function
EGARCH	Exponential Generalized Autoregressive Conditional Heteroskedasticity
EM	Expectation Maximization
EVT	Extreme Value Theory
FTSE/JSE	Financial Times Stock Exchange/ Johannesburg Stock Exchange
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GNG	GPD-Normal-GPD
GKG	GPD Kernel Density Estimator GPD
GPD	Generalized Pareto Distribution
GSG	GPD Stable Distribution GPD
iid	Independent and identically distributed
JB	Jarque-Bera
KDE	Kernel Density Estimator
KPSS	Kwiatkowski-Phillips-Schmidt-Shin
K-S	Kolmogorov-Smirnov
LM	Lagrange Multiplier
MCMC	Monte Carlo Markov Chain
ML	Maximum Likelihood
MSE	Mean Square Error
MRL	Mean Residual Life
PDF	Probability Distribution Function
POF	Proportion of Failure
P-P	Philips-Perron
Q-Q	Quantile to Quantile
RMSE	Root Mean Square Error

ABBREVIATIONS

SD	Stable Distribution
SNS	Stable Normal Stable
TGARCH	Threshold Generalized Autoregressive Conditional Heteroskedasticity
UC	Unconditional Coverage
USD	United States of American Dollar
VaR	Value-at-Risk
ZAR	South African Rand

RESEARCH OUTPUTS

Journals for publications were recommended by the University of KwaZulu-Natal. The following papers have been published with the core research contributions derived from this thesis.

1. Naradh, K., Chinhamu, K., and Chifurira, R. (2021). Estimating the value-at-risk of JSE indices and South African exchange rate with Generalized Pareto and stable distributions. *Investment Management and Financial Innovations*, 18(3),151-165.

2. Naradh, K., Chifurira, R., and Chinhamu, K. (2022). Analysis of stock exchange risk and currency in South African Financial Markets using stable parameter estimation. *International Journal of Finance and Banking Studies*, 11(1),120-131.

3. Naradh, K., Chinhamu, K., and Chifurira, R. (2023). Investigating risk within South African Financial Markets using Extreme Value Mixture Models. *Journal of Statistics Applications and Probability.* 12(3),1061-1072.

4. Naradh, K., Chinhamu, K., and Chifurira, R. (2024). Extreme Value Stable Mixture Modelling with applications to South African stock market indices and exchange rate.*Journal of Statistics, Optimization and Information Computing.*[Accepted].

CHAPTER 1

INTRODUCTION

This chapter outlines the rationale of this study, related literature, statement of the problem, aim and objectives, significance and scientific contribution of the study. It also covers the research layout of the thesis.

1.1 Rationale of the study

The theory used to describe the likely occurrence of rare events or unusual behaviour is referred to as Extreme Value Theory (EVT). This theory is commonly used to create statistically sound models that can make valid inferences and in recent times a plethora of extreme value mixture models are explored. It is well-known that in the last decade South Africa has experienced bleak economic growth, high levels of crime, corruption and unemployment. These societal concerns have intensified with the unprecedented Coronavirus disease 2019 (COVID-19) global pandemic, the Russia-Ukraine war and the Israel-Palestine conflict. The requirement for reliable models that monitor the movement of volatile indices and exchange rates during globally disruptive events is of major importance in curtailing risk and implementing relevant macroeconomic structural changes and governance essential for financial stability.

Literature references that financial data displays heavy tails as well as skewness and a possible solution to dealing with this problem is to recommend a volatility model that adequately describes the features of financial data. Noteworthy work by Zhao (2010) proposed an extreme value mixture model that fits the Generalized Pareto Distribution (GPD) model at the upper and lower tail, and the Normal Distribution is fitted as the bulk model between the two tails. The mixture model created by Zhao (2010) is an extension of literature by McNeil and Frey (2000) where a two-stage model is developed where, at the initial stage, a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model is fitted to capture volatility clustering and the latter stage fits a GPD to the tails.

Stable distributions are a versatile four-parameter family of probability laws that offer a reliable fit to the suggested model. Nolan (2003) highlights the application of stable distributions in financial modelling due to the fact that stable distributions are a generalised form of the Normal or Gaussian distribution and allows for heavy tails and skewness, which are characteristics very prominent in financial data. De Wet et al. (2007) and Chinhamu et al. (2015) investigated the use of stable distributions to South African financial indices or precious commodities, with results showing stable distributions to be robust in modelling extreme risk in South African financial data returns.

EVT has garnered significant interest in several industries, particularly the financial industry. Based on the above, this study aims to explore extreme value mixture models with the inclusion of stable distributions. The main contribution of this thesis is to propose extreme value mixture models, namely, the GPD-Stable-GPD (GSG), Stable-Normal-Stable (SNS) and Stable-KDE-Stable (SKS), where KDE represents the Kernel density estimator, models to three FTSE/JSE stock market indices and the United States of American Dollar to the South African Rand. Distribution density plots of the fitted models will validate the adequacy of the models on each of the return series investigated in this study. This study highlights the usefulness of extreme value mixture models in the context and application of a South African financial perspective and suggests the stable mixture models as recommended distributions to consider when conducting financial analyses.

1.2 Literature Review

Nolan (2003) explored the application of stable distributions with exchange rate data, specifically evaluating the exchange rate dynamics between the British Pound and the German Mark. The data comprised of the daily exchange rate values spanning from 2 January 1990 to 21 May 1996. Returns were computed for the data series, and parameters were estimated using the Maximum Likelihood (ML) method. The data was analysed using the fitted stable model and the suggested Normal distribution model. Nolan (2003) also investigated the monthly exchange rates between the United States (U.S.) Dollar relative to the Tanzanian Shilling. The data covered the period from January 1975 to September 1997. The returns of the data were calculated and the parameter estimation was conducted using the ML method. The study observed that the exchange rate of Tanzanian Shilling showed heightened variability.

McCulloch (1997) investigated the appropriateness of stable distributions using stock market data, specifically the stock price data referred to as the Centre for Research in Security Prices (CRSP). The analysis covered a period of forty years, from January 1953 to December 1992. The study calculated ML estimates as well as quantile estimates. Goodness-of-fit was assessed using graphical methods, which involved analysing the Probability-Probability (P-P) plot and the stable density plot. The diagnostics indicated an accurate fit.

Chinhamu et al. (2015) studied the preferred generalized hyperbolic distribution for modelling gold price returns and compared it with fitted stable distributions. The suitability of the distributions was evaluated using the Anderson-Darling test, Bayesian information criterion (BIC), Akaike information criterion (AIC), and by backtesting VaR estimates. It was discovered that the most suitable model for gold returns varies depending on the VaR level, with both the stable distribution and the generalized hyperbolic distribution effectively capturing extreme risk in gold returns.

Kallah-Dagadu (2013) assessed three methodologies for estimating α-stable distributions. The study utilised the ML method, empirical characteristic function, and sample quantile methods to estimate parameters for stable distributions, as well as parameters for Normal and Cauchy distributions. This analysis was conducted using data from the Ghana Stock Exchange All-Shares Index, the U.S. Dollar to Ghana Cedi (USD/GHC) exchange rate, the British Pound Sterling to Ghana Cedi (GBP/GHC) exchange rate, and the European Euro to Ghana Cedi (EUR/GHC) exchange rate. The data from 2000 to 2011 was examined, showing that weekly returns in Ghanaian financial data exhibit heavy tails and asymmetry. The study concluded that the maximum likelihood method yielded the most accurate and efficient estimates for fitting stable distributions to the data.

Naradh (2016) applied stable distributions to model exchange rates between each of the BRICS (Brazil, Russia, India, China, and South Africa) countries and the U.S. Dollar, examining both the univariate and multivariate scenarios. The dataset comprises exchange rate data recorded between January 2011 and January 2016. Both the Kolmogorov-Smirnov test and the Anderson-Darling test indicate that stable distributions provide a good fit for the returns of BRICS financial data. Value-at-Risk (VaR) calculations and in-sample VaR backtesting were performed following the International Basel Regulatory standards. The robustness of each model describing the financial data was assessed using the Kupiec likelihood ratio test and Christoffersen's conditional coverage test. This study demonstrated the efficacy of stable distributions in analyzing BRICS financial data.

Kateregga et al. (2017) discussed various parameter estimation techniques, namely, the quantiles, logarithmic moments method, ML and the empirical characteristics function (ECF) method on the α-stable distribution. Results from this study suggest the ECF method performs better than the ML method over several values of the shape parameter α and skewness parameter β. The empirical analysis was based on crude oil, natural gas, gasoline, corn, wheat, gold, silver and platinum commodities data. A graphical analysis of the density and Quantile-Quantile (Q-Q) plots analysed the fit of the fitted stable model. Kateregga (2017) explored a risk management topic and applies the Malliavin calculus to derive a Bismut-Elworthy-Li (BEL) representation formula for computing financial Greeks under the framework of subordinated Brownian motion by an inverse α-stable process. It is noted that the BEL model performs well and is recommended to investors in emerging markets to construct hedge fund portfolios.

An extreme value mixture model, where the bulk segment is a parametric distribution, and the GPD is modelled at the tails, was studied by Behrens et al. (2004). The two distributions are spliced at some point known as the threshold that is treated as a parameter to be estimated. This analysis used a truncated gamma distribution as the bulk model.

Frigessi et al. (2002) introduced a dynamically weighted mixture model that combines the GPD for the tail model with the Weibull distribution for the bulk model. It is assumed that the model exhibits light-tailed behavior, where the GPD predominantly governs the upper tail. The parameters were estimated using the maximum likelihood method, and this model was applied to both simulated data and the Danish fire loss data. The methodology presented in this work offers value for unsupervised tail estimation in scenarios characterised by heavy-tailed distributions.

Do Nascimento et al. (2012) adopted a semi-parametric Bayesian approach that includes for the bulk distribution, a mixture of weighted gamma densities and the GPD model for the tail. BIC or DIC statistics is used to determine the number of gamma components in the bulk model.

Like Do Nascimento et al. (2012) and extending on the work of Behrens et al. (2004), Lee et al. (2012) proposed a mixture model comprising a mixture of exponential distribution components below a threshold and the GPD is fitted for the threshold excess. The EM algorithm is utilised for parameter estimation where a mixture of two exponential distributions is used to model the Danish fire and medical claim data sets.

De Melo Mendes and Lopes (2004) evaluated a data-driven approach and proposed a two-tailed mixture model. The GPD model is suggested to fit the tails and a Normal distribution is used as the fitted model for the bulk density. The estimation process involves the ordering and standardisation of data. Hu (2013) summarises the 6-step estimation procedure.

A model similar to the work of De Melo Mendes and Lopes (2004) is investigated by Zhao (2010), where Bayesian inference is used as the method of estimation. The Normal distribution is recommended as the bulk distribution and GPD for the lower and upper tails. A novel finding using this method is the bulk distribution does not influence the fitting of the tail model. A potential limitation highlighted by Hu (2013) deals with the misfit of the model, specifically the bulk model. Zhao (2010) applies extreme value theory to financial data to investigate financial risks and extreme events in markets. The GNG model was also implemented in neonatal data where low birth weights of infants and days spent in the neonatal intensive care unit were recorded. The usefulness of the GNG model is highlighted in both financial risk management and critical health planning as well as medical staff resource allocation.

A flexible extreme value mixture modelling framework was defined by MacDonald et al. (2011), where the standard kernel density estimator is the bulk model and the GPD is the tail models. The performance of the proposed two-tailed mixture extreme value model was evaluated by empirical analysis and simulation study to determine normal physiological measurements for pre-mature infants.

Hu (2013) reviewed existing extreme value models and created a package in the statistical programming language R. This study allowed for the easy application of extreme value mixture models, which is not commonly available in any other software application. Hu (2013) noted the various forms of mixture models, that is, parametric, semi-parametric and non-parametric models and provided an automated way to determine thresholds in mixture models. A simulation study carried out in this study investigated the performance of various extreme value mixture models. Results showed that the KDE model based on a non-parametric form of a mixture model provided a good tail fit, whereas a reasonable fit is noticed with parametric and semi-parametric form mixture models.

Recent research by Coulibaly et al. (2024), Teimouri (2020), Liu and Shi (2022), Zhao et al. (2010) and Qiu (2024) further emphasise the significance and applicability of stable distributions and mixture models in financial modelling.

Table 1.1 provides a summary of the related literature on stable distributions and mixture model applications.

Table 1.1: Summary of related literature.

Authors	Title	Models	Findings	Criticism/Gap
Nolan (2003)	Modelling financial data with stable distributions	Stable model with ML estimation	Financial data exhibiting heavy-tails are effectively characterised by stable distributions.	Certainty cannot be made on tail probabilities.
McCulloch (1997)	Measuring Tail Thickness to Estimate the Stable Index α: A Critique	Stable distribution (ML and quantile estimation)	Many methods are unreliable in estimating α and result in biased estimates.	Alternative models are suggested that are not sample size sensitive. Formal goodness-of-fit tests are needed to be implemented.
Chinhamu et al. (2015)	Evaluating risk in gold prices with generalized hyperbolic and stable distributions	Generalized Hyperbolic and Stable Distributions	Generalized Hyperbolic Distribution is most adequate for modelling gold returns with the stable distribution out-performing at some VaR levels.	Gap for implementation of stepwise functions and mixture models.
Kallah-Dagadu (2013)	Modelling Ghana stock exchange Indices and exchange rates with stable distributions	α-stable distribution	ML estimation method is the best in providing accurate stable estimates.	Results can be skewed for large data sets where the Central Limit Theorem can be applied.
Naradh (2016)	Multivariate elliptically contoured stable distributions with applications to BRICS financial data	Nolan's S_0-stable model and GARCH (1,1)-Stable	Stable distributions are an apt choice to fit BRICS data where skewness and heavy-tailed phenomenon are evident.	Scope to evaluate VaR estimates and model performance in the multivariate stable case.
Kateregga et al. (2017)	Parameter estimation for stable distributions with application to commodity futures log-returns	Stable model	ECF outperforms the ML method of estimation	Formal model diagnostics evaluating the fit of stable models
Kateregga (2017)	Stable Processes: Theory and Applications in Finance	Stable and BEL model	The BEL model fits well to various energy,grain and metal commodities.	Explore a practical application of the BEL model.
Behrens et al. (2004)	Bayesian analysis of extreme events with threshold estimation	truncated gamma-GPD mixture model	A Bayesian threshold methodology is proposed for analysing extreme events.	Subjectivity of the choice of prior models and model mis-specification.
Frigessi et al. (2002)	A dynamic mixture model for unsupervised tail estimation without threshold selection	Webull-GPD mixture model	An unsupervised learning framework (no prior knowledge required) is introduced by a dynamic mixture model where no separate threshold estimation is required.	Statistical practitioners identify model fitting and interpretation complexities.
Do Nascimento et al. (2012)	A semi-parametric Bayesian approach to extreme value estimation	weighted gamma-GPD mixture model	The flexibility of a semi-parametric modelling approach is shown to produce robust models by considering the vast characteristics of extreme data sets.	Hybrid modelling has computational complexities and is difficult to implement.
Lee et al. (2012)	Modelling insurance claims via a mixture exponential model combined with peaks-over-threshold approach	exponential GPD mixed model	An exponential mixture model is proposed representing the claim size of insurance claims. The EM algorithms used to ensure a good fit to insurance data through an iterative procedure.	An iterative EM estimation approach is computationally intense for practitioners.
De Melo Mendes and Lopes (2004)	Data-driven estimates for mixtures	weighted EM mixture model	The EM algorithm is used to support a modelling approach that relies on the underlying data set rather than subject matter assumptions.	The model estimation method can be computationally demanding and challenging to implement.
Zhao (2010)	Extreme value modelling with application in finance and neonatal research	GNG model by a Bayesian approach	Propose the novel Bayesian GNG model with applications to finance and neonatal health.	Model mis-specification especially the bulk model is a concern.
MacDonald et al. (2011)	A flexible extreme value mixture model	GKG	A new proposed mixture model with GPD tails and KDE bulk model.	MCMC in this model has a long computational time.
Hu (2013)	Extreme Value Mixture Modelling with Simulation Study and Applications in Finance and Insurance	various existing mixture models	KDE mixture models provide a good fit and a R package (evmix) created.	consider alternative non-parametric bulk model- the logspline density estimator

Given the recent literature presented in Table 1.1, there is limited research on mixture models involving stable distributions. This study aimed at modelling financial data using a mixture of stable distributions. This study extends the work of Naradh (2016), Zhao (2010) and MacDonald et al. (2011) by proposing the GSG, SNS and SKS mixture models which investigates the suitability of modelling South African financial data.

1.3 Statement of the problem

A plethora of extreme value mixture models have been developed by analytical practitioners in recent times. There seems to be a vested interest in researching the properties and applications of extreme value mixture models in various sectors, particularly the stock market as well as the medical and financial service sectors. A robust model of financial data should be able to capture the empirical properties of data. There is empirical evidence that the tail distribution of financial data is heavier than the Normal distribution. Nolan's (2003) study showed the stable distribution to be suitable for heavy-tailed financial data. There is a need to explore the suitability of extreme value mixture models involving the stable distribution in capturing the tail distribution of financial data.

1.4 Aim and Objectives of the Study

The main aim of this study is to investigate the suitability of extreme value mixture models that include a component of stable distributions in either the tail or bulk component of a mixture model in modelling South African financial data. This is achieved by:

(i) Fitting the Nolan's S_0-parameterization stable model on South African financial data;

(ii) Estimating VaR values of South African financial data using hybrid GARCH-type combined with a stable distribution;

(iii) Proposing a GPD-Stable-GPD model and investigating its suitability in modelling South African financial data;

(iv) Proposing Stable-Normal-Stable and Stable-KDE-Stable mixture models for South African financial data and evaluating model fit to South African financial data.

1.5 Significance of the study

Many extreme value modelling frameworks have become frequently utilised to evaluate risk in several industries, particularly the financial industry. The study of extreme events is deemed important for many financial practitioners where EVT methods are evaluated to provide statistically sound models for tail distributions. These models ideally should provide reliable extrapolations that project beyond the observed data range.

The occurrence of heavy tails in financial data is well-known in literature and EVT methods as well as stable distributions are considered when examining financial data. There is no general framework for defining extreme value mixture models and the applications of such models are often not that straightforward. Therefore, the enhancements of new extreme value mixture models from traditional models need to be considered, especially in finance, where risk is of great concern.

This thesis contributes to a framework to understand extreme value stable distribution mixture models, specifically the GPD-Stable-GPD (GSG), Stable-Normal-Stable (SNS) and Stable-KDE-Stable (SKS) models, using Nolan's S_0-parameterization for the respective bulk or tail models. The application of stable distributions in extreme value theory and, furthermore in mixture models are relatively understudied or under-explored. This thesis may fill the gaps in knowledge of existing literature and propose a robust two-tailed model, suitable for financial applications, for dealing with the fat tail phenomenon in financial data. Financial analysts, risk specialists and academics interested in extreme value mixture modelling in South Africa are likely to benefit from the insights in this study.

1.6 Scientific contribution of the study

The major contribution of this study is the application of stable distributions in modelling South African financial data. The scientific contributions are as follows:

(i) Evaluate the usefulness of fitting the Nolan's S_0-parameterization stable model to South African indices and exchange rate data.

(ii) Combining GARCH-type models with stable distribution covering the innovations.

(iii) Exploration of the GPD-Normal-GPD and the GPD-KDE-GPD mixture models in modelling South African financial indices and exchange rate data

(iv) Proposing the GPD-Stable-GPD, Stable-Normal-Stable and Stable-KDE-Stable models for modelling South African financial data.

1.7 Thesis structure

This thesis consists of a total of seven chapters. Following this introductory chapter, Chapter 2 discusses the relevant research methodology in this study. Chapter 3 provides insights into the JSE stock market indices and United States of American Dollar to South African Rand data sets by a statistical data analysis. The background material and theoretical concepts of stable distributions are covered in Chapter 4. Chapter 5 discusses the hybrid GARCH-type and stable distribution compared to the GARCH type-GPD hybrid model. Chapter 6 develops the main scientific contribution of this study by implementing the proposed two-tailed stable distribution mixture models. Lastly, Chapter 7 summarises the findings and concludes the study.

CHAPTER 2

METHODOLOGY

This chapter discusses the relevant research methodology and statistical tests used in the study.

2.1 Tests for normality

2.1.1 Q-Q plots

The Quantile-Quantile (Q-Q) plot is a probability plot used to check if a data set comes from a theoretical distribution. Quantiles of two probability distributions are plotted together as a graphical method for comparison. In most cases, the Q-Q plot is used to determine whether or not a data set follows a normal distribution.

As a rule of thumb, data points on the Q-Q plot that lie on the straight diagonal line indicate normally distributed data. Alternatively, less normally distributed data is indicated by data points that deviate from the straight line. It is important to note that Q-Q plots are not a formal test for normality but merely provides an indication if a data set roughly follows a normal distribution (Bobbitt, 2024).

2.1.2 Jacque-Bera test

The Jarque-Bera test is a common test to evaluate whether a sample of n data points follows a normal distribution. The test procedure is based on a joint statistic that comprises of skewness and kurtosis coefficients, and the test statistic is given by:

$$JB = n\left[\frac{\left(b_1^{1/2}\right)^2}{6} + \frac{(b_2-3)^2}{24}\right]. \tag{2.1}$$

If Y_i is independent and identically distributed (i.i.d) as well normally distributed, then:

$$\sqrt{n}\left(b_1^{1/2}-0\right) \xrightarrow{\text{D}} N\left(0,6\right), \sqrt{n}\left(b_2-3\right) \xrightarrow{\text{D}} N\left(0,24\right). \tag{2.2}$$

With standardisation:

$$Z_1 = \frac{\sqrt{n}\left(b_1^{1/2} - 0\right)}{\sqrt{6}} \xrightarrow{\text{D}} N\left(0,1\right), Z_2 = \frac{\sqrt{n}\left(b_2 - 3\right)}{\sqrt{24}} \xrightarrow{\text{D}} N\left(0,1\right). \tag{2.3}$$

Therefore,

$$Z_1^2 + Z_2^2 \xrightarrow{\text{D}} \chi_2^2. \tag{2.4}$$

Under the null hypothesis of normality where skewness is 0 and kurtosis is 3, the Jarque-Bera (JB) test statistic is asymptotically chi-squared distributed with 2 degrees of freedom as the JB statistic is the sum of squares for 2 asymptotic standardised Normal distributions. This suggests H_0 is to be rejected at α level of significance if $JB \geq \chi^2_{1-\alpha,2}$.

2.2 Tests for autocorrelation

2.2.1 Autocorrelation

Autocorrelation measures the degree of similarity between successive observations in a time-ordered series. In the classical linear regression model, it is assumed that the disturbances u_i exhibit no autocorrelation. That is,

$$E(u_i u_j) = 0,\ i \neq j. \tag{2.5}$$

2.2.2 Autocorrelation function

For a stationary process Y_t with mean $E(Y_t) = \mu$ and constant variance $\text{Var}(Y_t) = E(Y_t - \mu)^2 = \sigma$. The covariance between Y_t and Y_{t+k} is expressed as:

$$\gamma_k = \text{Cov}(Y_t, Y_{t+k}) = E(Y_t - \mu)(Y_{t+k} - \mu), \tag{2.6}$$

and the autocorrelation function (ACF) is expressed as:

$$\rho_k = \frac{\text{Cov}(Y_t, Y_{t+k})}{\sqrt{\text{Var}(Y_t)}\sqrt{\text{Var}(Y_{t+k})}} = \frac{\gamma_k}{\gamma_0}. \tag{2.7}$$

where k denotes a separation by k lags (Wei, 2006).

2.2.3 Partial autocorrelation function (PACF)

After accounting for the mutual dependence of the intervening variables $Y_{t+1}, Y_{t+2}, \ldots, Y_{t+k-1}$, the correlation between Y_t and Y_{t+k} can be investigated. The conditional correlation

$$\mathrm{Corr}(Y_t, Y_{t+k}|Y_{t+1}, \ldots, Y_{t+k-1}), \tag{2.8}$$

is known as partial correlation which for convenience is denoted by ϕ_{kk}. Consider a regression model that includes k lagged variables, that is

$$Y_{t+k} = \phi_{k1}Y_{t+k-1} + \phi_{k2}Y_{t+k-2} + \ldots \phi_{kk}Y_t + e_{t+k},$$

where ϕ_{ki} denotes the ith regression parameter and e_{t+k} is the error term. It can be shown, by using Cramer's rule, that the PACF, as denoted by Wei (2006):

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{vmatrix}}. \tag{2.9}$$

2.3 Ljung-Box test

The Ljung-Box test is a frequently used statistical test used to determine autocorrelation in a time series. The test statistic for the Ljung-Box test is represented by:

$$Q = n\,(n+2)\sum_{k=1}^{h}\frac{r_k^2}{n-k}, \tag{2.10}$$

where n denotes the sample size, h is the number of lags being tested and r_k is the sample autocorrelation at lag k. The null hypothesis, which assumes no autocorrelation, suggests the test statistic Q follows a chi-squared distribution with h degrees of freedom $\left(\chi^2_{(h)}\right)$.

2.4 Tests for ARCH Effects

2.4.1 ARCH LM test

The Autoregressive Conditional Heteroskedasticity Lagrange Multiplier (ARCH LM) statistical test of Engle (1982) and described by Tsay (2005), aids with the detection of ARCH effects in a time series. ARCH effects highlight the variability of a time series; that is, periods of high volatility are followed by high volatility and vice versa for low volatility, where periods of low volatility are followed by low volatility. The ARCH LM test statistic is defined as:

$$LM = n \times R^2, \tag{2.11}$$

where n is the sample size and R^2 is the coefficient of determination of the squared residuals from the regression.

2.5 Tests for unit root and stationarity

2.5.1 The unit root test

The unit root stochastic process is described by:

$$Y_t = \rho Y_{t-1} + u_t, \qquad -1 \leq \rho \leq 1, \tag{2.12}$$

where u_t indicates a white noise error term. When $\rho = 1$, in the case of the unit root, the equation above transforms into a random walk model without drift, which is a non-stationary stochastic process. Y_t is regressed on the lagged value Y_{t-1} to verify if the estimated ρ is statistically equal to 1. If this criterion is met, Y_t is deemed non-stationary. This concept forms the basis of unit root tests for stationarity. The equation above is transformed by subtracting Y_{t-1} from both sides to yield

$$Y_t - Y_{t-1} = \rho Y_{t-1} - Y_{t-1} + u_t,$$

$$= (\rho - 1)Y_{t-1} + u_t.$$

Alternatively, this can be expressed as

$$\triangle Y_t = \delta Y_{t-1} + u_t. \tag{2.13}$$

where $\delta = (\rho - 1)$ and $\triangle$ signifies the first difference operator. Equation (2.13) is estimated and the null hypothesis is tested as follows:

$$H_0 : \delta = 0.$$

If $\delta = 0$, then $\rho = 1$. This indicates a unit root, indicating that the time series under consideration is non-stationary. If $\delta = 0$, then equation (2.13) simplifies to

$$\triangle Y_t = (Y_t - Y_{t-1}) = u_t, \tag{2.14}$$

where u_t denotes the white noise error term which is stationary, implying that the first differences of a random walk time series are also stationary. To estimate equation (2.14), the first differences of Y_t are computed and regressed Y_{t-1} to assess whether the estimated slope coefficient in this regression ($= \hat{\delta}$) is equal to zero or not. If $= \hat{\delta}$ is zero, it is inferred that Y_t is non-stationary. However, if $= \hat{\delta}$ is negative, it is inferred that Y_t is stationary.

Several considerations must be made when implementing the Dickey-Fuller (DF) test procedure. A random walk process can exhibit no drift, drift, or a combination of both deterministic and stochastic trends.
The DF test is estimated under three different null hypotheses:

1. Y_t is a random walk: $\triangle Y_t = \delta Y_{t-1} + u_t.$

2. Y_t is a random walk with drift: $\triangle Y_t = \beta_1 + \delta Y_{t-1} + u_t.$

3. Y_t is a random walk with drift around a stochastic trend: $\triangle Y_t = \beta_1 + \beta_2 t \delta Y_{t-1} + u_t,$

where t represents the time or trend variable. The null hypothesis in each of the above cases is, $H_0 : \delta = 0$, that is, there exists a unit root and the time series is non-stationary. The alternative hypothesis states that: $H_1 : \delta < 0$ which implies the time series is stationary. If H_0 is rejected, then in the first scenario, Y_t is a stationary time series with a zero mean. In the second case, it is suggested that Y_t is stationary with non-zero mean $\left(\frac{\beta_1}{(1-\rho)}\right)$ and lastly, Y_t exhibits stationarity around a deterministic trend. The estimation process involves using ordinary least squares (OLS) method, where the estimated coefficient of Y_t is divided by ts standard error to calculate the τ statistic. Referring to the DF tables or any other statistical package, if the computed absolute value of the τ statistic exceeds the DF critical τ values , we fail to reject the null hypothesis. In such a scenario, the time series is non-stationary (Gujarati and Porter, 2009).

2.5.2 Augmented Dickey-Fuller test (ADF)

When the error terms are correlated, Dickey and Fuller developed an alternative test called the Augmented Dickey-Fuller (ADF) test. The ADF test is performed by augmenting the the DF test by adding the lagged values of the dependent variable $\triangle Y_t$. The following regression equation is estimated

$$\triangle Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^{m} \alpha_i \triangle Y_{t-i} + \epsilon_t, \tag{2.15}$$

where ϵ_t denotes a pure white noise error term and $\triangle Y_{t-1} = (Y_{t-1} - Y_{t-2}), \triangle Y_{t-2} = (Y_{t-2} - Y_{t-3})$ and so on. The number of lagged differences to include is determined empirically, with the goal of ensuring that the error term becomes serially uncorrelated. In the ADF test, $\delta = 0$ is tested. The same critical values are utilised because the ADF test adheres to the same asymptotic distribution as the DF statistic.(Gujarati and Porter, 2009).

2.5.3 Philips-Perron test

The Phillips-Perron (P-P) test provides a more comprehensive approach to testing unit root non-stationarity. This test is similar to the ADF test but incorporates an automatic correction to the DF test to allow for autocorrelated residuals.
The Phillips-Perron test fits the following regression

$$Y_t = \alpha + \rho Y_{t-1} + \epsilon_t, \tag{2.16}$$

where the inclusion or exclusion of a constant term or trend term is considered. Z_ρ and Z_τ are two statistics that are calculated as follows:

$$Z_\rho = n(\hat{\rho_n} - 1) - \frac{1}{2}\frac{n^2\hat{\sigma}^2}{s_n^2}\left(\hat{\lambda}_n^2 - \hat{\gamma}_{0.n}\right), \tag{2.17}$$

$$Z_\tau = \sqrt{\frac{\hat{\gamma}_{0,n}}{\hat{\lambda}_n^2}}\frac{\hat{\rho}_n - 1}{\hat{\sigma}} - \frac{1}{2}\left(\hat{\lambda}_n^2 - \hat{\gamma}_{0.n}\right)\frac{1}{\hat{\lambda}_n}\frac{n\hat{\sigma}}{s_n}, \tag{2.18}$$

$$\hat{\gamma}_{j,n} = \frac{1}{n}\sum_{t=j+1}^{n} \hat{u}_i \hat{u}_{t-j},$$

$$\hat{\lambda}_n^2 = \hat{\lambda}_{0,n} + 2\sum_{j=1}^{q}\left(1 - \frac{j}{q+1}\right)\hat{\lambda}_{j,n},$$

$$s_n^2 = \frac{1}{n-k}\sum_{t=1}^{m} \hat{u}_i^2,$$

where u_t describes the OLS residual, k represents the number of covariates in the regression model, q denotes the number of lags to use in calculating $\hat{\lambda}_n^2$ and $\hat{\sigma}$ represents the OLS standard error of $\hat{\rho}$. s_n^2 is an unbiased OLS estimator of the variance of the error terms.

The regression involves Y regressed on lagged Y rather than differenced Y regressed on lagged Y. Z_τ is the adjusted *t*-statistic, similar to the DF test:

$$\frac{\hat{\rho}_n - 1}{\hat{\sigma}}. \tag{2.19}$$

The ML estimate of the variance of error terms is assumed when $j = 0$,

$$\hat{\lambda}_n^2 = \hat{\gamma}_{0,n} + 2\sum_{j=1}^{q}\left(1 - \frac{j}{q+1}\right)\hat{\gamma}_{j,n},$$

where q represents the number of lagged covariances. When the covariances are zero, it implies that the autocorrelation between error terms $\hat{\gamma}_{j,n}$ is zero for $j > 0$. Therefore, the second term in the equation is removed, and $\hat{\lambda}_n^2 = \hat{\gamma}_{0,n}$. A replacement for Z_τ yields:

$$Z_\tau = \sqrt{\frac{\hat{\gamma}_{0,n}}{\hat{\lambda}_n^2}}\frac{\hat{\rho}_n - 1}{\hat{\sigma}} - \frac{1}{2}\left(\hat{\lambda}_n^2 - \hat{\gamma}_{0,n}\right)\frac{1}{\hat{\lambda}_n}\frac{n\hat{\sigma}}{s_n}. \tag{2.20}$$

In this scenario, $\hat{\lambda}_n^2 - \hat{\gamma}_{0,n} = 0$, and the second term disappears. $\frac{\hat{\gamma}_{0,n}}{\hat{\lambda}_n^2} = 1$ hence the term $\sqrt{\frac{\hat{\gamma}_{0,n}}{\hat{\lambda}_n^2}}\frac{\hat{\rho}_n-1}{\hat{\sigma}}$ reduces to $\frac{\hat{\rho}_n-1}{\hat{\sigma}}$ and therefore $Z_\tau = \frac{\hat{\rho}_n-1}{\hat{\sigma}}$. We notice the similarity as in the standard DF test. When there is no autocorrelation between the error terms, this aspect of the P-P test aligns with the DF test. The P-P test adjusts the DF test for autocorrelation among error terms in a non-parametric regression framework. The critical values follow the same distribution as those of the DF statistic.
If there is no autocorrelation among error terms, the covariances are equal leading to the second term in the other P-P statistic becoming zero, where $\hat{\lambda}_n^2 = \hat{\gamma}_{0,n}$,

$$Z_\rho = n(\hat{\rho}_n - 1) - \frac{1}{2}\frac{n^2\hat{\sigma}^2}{s_n^2}\left(\hat{\lambda}_n^2 - \hat{\gamma}_{0,n}\right). \tag{2.21}$$

In the above scenario, $Z_\rho = n(\hat{\rho}_n - 1)$ which corresponds to the DF test (StataCorp, 2015).

2.5.4 Kwiatkowski-Phillips-Schmidt-Shin test

The Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test evaluates whether a time series exhibits stationarity around a mean or a linear trend as per the null hypothesis, or displays non-stationarity attributable to the existence of a unit root under the alternative hypothesis. The KPSS model consists of a series of observations that can be described as the combination of three components: a deterministic trend, a random walk, and a stationary error term. The model is formulated as follows:

$$Y_t = \xi t + r_t + \epsilon_t, \tag{2.22}$$

$$r_t = r_{t-1} + u_t, \tag{2.23}$$

where $Y_t, t = 1, 2, \ldots, T$ denotes a series of observations of the variable of interest, t is the deterministic trend, r_t is a random walk process and ϵ_t is the error term, which by assumption is stationary. u_t is the error term in the second equation above and is a series of i.i.d. with mean 0 and variance $\hat{\sigma}_u^2$ The null hypothesis of stationarity is equivalent to the assumption that $\hat{\sigma}_u^2$ of the random walk process r_t equals to zero. When $\xi = 0$, the null hypothesis implies Y_t is stationary around r_0. Conversely, if $\xi \neq 0$, then this suggests Y_t is stationary around a linear trend. If $\hat{\sigma}_u^2 > 0$, then Y_t is non-stationary due to the presence of a unit root.
By subtracting Y_t from both sides of the first equation above, we have

$$\triangle Y_t = \xi + u_t + \triangle\epsilon_t = \xi + w_t, \tag{2.24}$$

where w_t by the assumption that the error terms (ϵ_t and u_t) are i.i.d. is generated by an AR (1) process where: $w_t = v_t + \theta v_{t-1}$. The KPSS model may be described as:

$$Y_t = \xi + \beta Y_{t-1} + w_t, \tag{2.25}$$

$$w_t = v_t + \theta v_{t-1},\ \beta = 1. \tag{2.26}$$

Equations (2.25) and (2.26) show a link between the KPSS and the DF test. In the DF test, the parameter β is set to 1 assuming that θ, nuisance parameter is 0. The assumption that β serves as the nuisance parameter and, therefore, tests if θ is equal to -1. Under the assumption $\beta = 0$, a one-side Lagrange Multiplier (LM) test is conducted with the null hypothesis, $H_0 : v\sigma_u^2 = 0$. Here, u_t is normally distributed and, ϵ_t are iid random variables with a mean of 0 and a constant variance σ_ϵ^2 (Kwiatkowski et al., 1992)

The statistic used in the KPSS test is as follows:

(a) Testing a null hypothesis of stationarity around a linear trend versus the alternative hypothesis that a unit root is present. Let $e_t, t = 1, 2, 3, \ldots, T$ represent the estimated errors obtained from a regression on Y_t. The estimated variance denoted by $\hat{\sigma}_t^2$ is calculated as the sum of squared errors divided by the number of observations T. Partial sums of errors are calculated from:

$$S_t = \sum_{t=1}^{t} e_t \text{ for } t = 1, 2, \ldots, T.$$

The LM test statistic is subsequently defined as:

$$\text{LM} = \sum_{t=1}^{T} \frac{S_t^2}{\sigma_\epsilon^2}. \tag{2.27}$$

(b) Testing the null hypothesis of stationarity around the mean versus an alternative hypothesis of the presence of a unit root, where e_t represents the estimated errors computed as residuals from the regression on Y_t; specifically, $e_t = Y_t - \bar{Y}$. The remaining definitions are left unchanged. The long-run variance is:

$$\sigma^2 = \lim T^{-1} E\left[S_T^2\right], \tag{2.28}$$

The long-run variance is encountered when defining the asymptotic distribution of a test statistic. Kwiatkowski et al. (1992) presents a consistent estimate of the long-run variance, defined as

$$s^2(k) = T^{-1} \sum_{t=1}^{T} e_t^2 + 2T^{-1} \sum_{j=1}^{k} w(j,k) \sum_{t=s+1}^{T} e_t e_{t-1}, \tag{2.29}$$

where $w(j,k)$ denote weights determined by the chosen spectral window. The Bartlett window is utilised by Kwiatkowski et al. (1992) by defining the weights as $w(j,k) = 1 - \frac{j}{k+1}$. This ensures that $s^2(k)$ is non-negative. It is argued that for quarterly data, lag $k = 8$ is considered optimal. Choosing $k < 8$, distorts the test size and if $k > 8$ the power is reduced. Here, the KPSS test statistic is determined by dividing the sum of squared partial sums by the estimated long-run variance.

$$\hat{\eta} = T^{-2} \frac{\sum S_t^2}{s^2(k)}. \tag{2.30}$$

The symbols $\hat{\eta}_\mu$ and $\hat{\eta}_t$ represent the KPSS test statistic for testing stationarity around a mean and around a trend, respectively. The asymptotic distribution

of the KPSS test statistic is unconventional and converges to a higher-order Brownian bridge.
The $\hat{\eta}_\mu$ statistic, testing for stationarity around mean converges to

$$\hat{\eta}_\mu \rightarrow \int_0^1 V(r)^2 dr.$$

The Brownian bridge is denoted by: $V(r) = W(r) - rW(1)$ which is defined in terms of a standard Wiener process $W(r)$. When $\xi \neq 0$, the KPSS test statistic, $\hat{\eta}_t$ for testing stationarity around a trend converges to a second-order Brownian bridge, $V_2(r)$, which is described as :

$$V_2(r)_2 = W(r) + (2r - 3r^2)W(1) + (-6r + 6r^2) \int_0^1 W(s) ds.$$

The test statistic $V_2(r)_2$ converges weakly to the limit:

$$\hat{\eta}_t \rightarrow \int_0^1 V_2(r)^2 dr.$$

To summarize, the KPSS test is carried out as follows:

(i) We test the null hypotheses where H_0 : proposes stationarity around a mean or trend versus the alternative hypothesis which suggests non-stationarity of a time series due to the presence of a unit root.

(ii) Calculate the test statistic value.

(iii) If the calculated value exceeds the critical value at a specified level of significance, we reject the null hypothesis of stationarity (Syczewska, 2010).

2.6 Goodness-of-Fit tests

2.6.1 Anderson-Darling (AD) goodness-of-fit test

The Anderson-Darling test statistic A^2 is defined as:

$$A_n^2 = -\sum_{i=1}^{n} \frac{2i-1}{n} \Big(\ln(\hat{F}(y_{(i)})) + \ln(1 - \hat{F}(y_{(n+1-i)})) \Big) - n, \tag{2.31}$$

where $y_{(1)} < \ldots < y_{(m)}$ represents the ordered sample size m arranged from from smallest to largest and $F(y)$ is the theoretical cumulative distribution against which the sample is compared (Anderson and Darling, 1954). The null hypothesis $y_{(1)} < \ldots < y_{(m)}$ stems from the underlying distribution $F(y)$. The null hypothesis is rejected at an α level of significance, if the test statistic A^2 exceeds the critical value from a table of critical values corresponding to different sample sizes. Typically, the critical values of the Anderson-Darling test statistic vary depending on the distribution under examination. The Anderson-Darling test (Anderson and Darling, 1954) is commonly used to assess the goodness-of-fit for various distributions.

2.6.2 Kolmogorov-Smirnov (K-S) test

The K-S goodness-of-fit test compares a fitted CDF $\hat{F}(y)$ with an empirical CDF $F_n(y)$ in order to evaluate the adequacy of the fit. The empirical CDF, $F_n(y)$, is defined as the proportion of the observations $Y_1, Y_2, \ldots, Y_n$ that are less than or equal to y. $F_n(y)$ is expressed as:

$$F_n(y) = \frac{I(y)}{n},$$

where n represents the sample size and $I(y)$ denotes the number of Y_i's $\leq y$ (Evans et al., 2008). D_n, the K-S statistic is the maximum vertical distance between $\hat{F}(y)$ and $F_n(y)$ across all values of y

$$D_n = \sup_y |F_n(y) - \hat{F}(y)|. \tag{2.32}$$

The D_n statistic is computed from

$$D_n^+ = \max_{i=1,2,\ldots,n} \left\{ \frac{i}{n} - \hat{F}(Y_{(i)}) \right\},\ D_n^- = \max_{i=1,2,\ldots,n} \left\{ \hat{F}(Y_{(i)}) - \frac{i-1}{n} \right\}, \tag{2.33}$$

where $Y_{(i)}$ describes the statistic of i^{th} order and letting

$$D_n = \max\{D_n^+, D_n^-\}. \tag{2.34}$$

2.7 Model selection

In this study, Akaike information criterion (AIC) and Bayesian information criterion (BIC) are used for model selection.

2.7.1 Akaike information criterion (AIC)

The Akaike Information Criterion (AIC) is a commonly used statistical metric developed by Akaike (1974) for model selection. The AIC formula is calculated as:

$$AIC = 2k - 2ln(L), \tag{2.35}$$

where k is the number of parameters in the model and L is the maximum likelihood of the model. When comparing models, a lower AIC is preferred and indicates a good model fit (Burnham and Anderson, 2004).

2.7.2 Bayesian information criterion (BIC)

The Bayesian information criterion (BIC), also referred to as the Schwarz Criterion, was introduced by Schwarz (1978) as a metric for model selection. The BIC is calculated as:

$$BIC = kln(n) - 2ln(L), \tag{2.36}$$

where k is the number of parameters in the model, n is the sample size and L is the maximum likelihood of the model.

2.8 Tests for asymmetry

2.8.1 Sign bias test

The sign bias test of Engle and Ng (1993) tests the presence of leverage effects in standardised residuals to identify a possible mis-specification of the GARCH model. The squared standardised residuals on lagged positive and negative shocks is noted as:

$$z_t^2 = c_0 + c_1 \cdot I_{\epsilon_{t-1}<0} + c_2 \cdot I_{\epsilon_{t-1}<0}\epsilon_{t-1} + c_3 \cdot I_{\epsilon_{t-1}\geq 0}\epsilon_{t-1} + u_t, \tag{2.37}$$

where I is the indicator function and ϵ_t is the estimated GARCH residuals. The null hypothesis is $H_0 = c_1 = c_2 = c_3$ or jointly as $H_0 = c_i$ for $i = 1, 2, 3$. The indicator variable $I_{\epsilon_{t-1}<0}$ takes the value of 1 when the estimated residuals ϵ_{t-1} is negative and 0 otherwise. The sign bias test evaluates the impact of both the positive and negative shocks on volatility. The negative sign bias test considers the $I_{\epsilon_{t-1}<0}$ variable and focuses on the impact of only large or small negative shocks on volatility. The positive sign bias test uses the $I_{\epsilon_{t-1}\geq 0}$ and solely focuses on the impact positive shocks have on volatility (Ghalanos et al., 2018).

2.9 Forecasting evaluation metrics

From a technical standpoint, measuring the performance of a fitted model is highly important. There are different types of evaluation metrics that deal with the performance of a time series model. In this research, the mean square error (MSE) and the root mean square error (RMSE) evaluate the forecasting performance of the fitted models.

2.9.1 Mean square error (MSE)

Mean square error (MSE) is the average of error squares and is defined as:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \left(Y_i - \hat{Y}i \right)^2 , \tag{2.38}$$

where Y_i is the actual value, $\hat{Y}_i$ is the predicted value and n is the sample size. MSE is mostly positive and lower values are favoured. With the square term, larger errors are penalised.

2.9.2 Root mean square error (RMSE)

Root mean square error (RMSE) is an extension of MSE and is the square root of the Mean Square Error. The RMSE test statistic is defined as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(Y_i - \hat{Y}i \right)^2}, \tag{2.39}$$

The RMSE statistic is always positive and lower values indicate better performance. In this case, the root term penalises error terms more.

2.10 Value-at-Risk and backtesting

2.10.1 Value-at-Risk

The Basel Committee on Banking Supervision adopts Value-at-Risk (VaR) as the standard benchmark measure for evaluating market risk (BIS, 2016). The capital requirements of financial institutions are derived from VaR estimates; therefore, tests for assessing the out-of-sample forecast accuracy of VaR models through backtesting procedures is essential (Escanciano and Olmo, 2010). VaR intends to estimate the maximum potential loss for a portfolio over a specified period, where VaR estimations focus on the tails of a distribution and robustness testing procedures for a model. For a random variable Z, describing the log return of a risky financial instrument with distribution function F over a specified period, the VaR at a given probability p is defined as the p^{th} quantile of distribution F, specifically,

$$\mathrm{VaR_p} = F^{-1}(1-p), \tag{2.40}$$

where, F^{-1} denotes the quantile function.

Various backtesting procedures are employed to assess the effectiveness and reliability of VaR estimates.

2.10.2 Kupiec likelihood ratio test

The Kupiec test, also known as the proportion of failures (POF) test, proposed by Kupiec (1995), is widely recognised and utilised. The Kupiec test examines the unconditional coverage (UC) property. The accuracy of the VaR model is assessed by observing the failure rate, which is the proportion of instances where VaR is exceeded in a given sample. The number of exceptions is represented by (y) and N represents the number of exceptions by and the total number of observations the rate of failure is expressed as $\frac{y}{N}$. If a VaR value is reported at the confidence level c, an exception is defined as when the realized loss surpasses this VaR value. Therefore, the expected number of exceptions y from N, the total observations expressed as $(1-c)N$ (Katsenga, 2013). The number of exceptions will vary from $(1-c)N$ within an acceptable range. The backtesting method allows for the calculation of the range of y. Hence, acceptance or rejection of the VaR model may occur (Campbell, 2006).

The parameters needed for backtesting the VaR model using the Kupiec test include, y, the number of exceptions, (N), the total number of observations and the confidence level (c). The null hypothesis H_0 is defined as:

$$H_0 : \frac{y}{N} = \frac{y^*}{N}, \tag{2.41}$$

where $\frac{y}{N}$ represents the expected failure rate at a given confidence level c and $\frac{y^*}{N}$ denotes the observed failure rate. The Kupiec test is conducted similarly to a likelihood ratio (LR) test, expressed as:

$$L_{UC} = -2\ln\left[(1-p)^{(N-y)}\,p^y\right] + 2\ln\left[\left(1-\frac{y}{N}\right)^{N-y}\left(\frac{y}{N}\right)^y\right], \tag{2.42}$$

where $p = (1-c)$. Under the null hypothesis H_0, the test statistic conforms to a chi-squared distribution with 1 degree of freedom, $\chi^2(1)$. If the value of L_{UC} statistic is lower than the critical value $\chi^2(1)$, then the model is considered adequate. Values above the critical region indicate inaccuracies in the model and should consequently be rejected (Katsenga, 2013).

CHAPTER 3

DATA AND EXPLORATORY DATA ANALYSIS

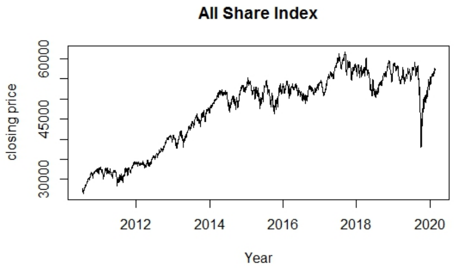
This chapter discusses data sources and presents the results of the exploratory data analysis.

3.1 Data source

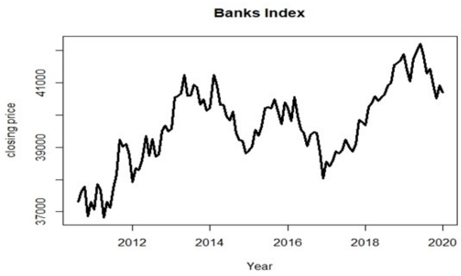
The FTSE/JSE stock market indices and exchange rate data, namely the All Share Index (ALSI), Banks Index and the Mining Index, and United States of American Dollar (USD) to South African Rand (ZAR) exchange rate was obtained from McGregor IRESS (Bureau for Financial Analysis). This research selected the time period between 13 August 2010 and 14 August 2020.

3.1.1 Exploratory data analysis

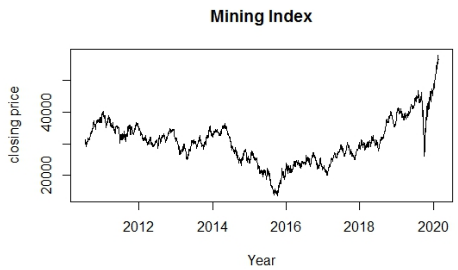
We explore the data sets used in this study. Figure 3.1 shows the time series plot of the daily closing prices of three FTSE/JSE market indices and the USD/ZAR exchange rate. Top left of the plots shown in Figure 3.1 is the FTSE/JSE ALSI, the top right is the FTSE/JSE Banks Index, the bottom left is the FTSE/JSE Mining Index and the bottom right is the USD/ZAR exchange rate.



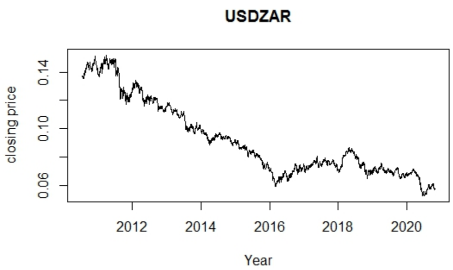
(a) FTSE/JSE ALSI



(b) FTSE/JSE Banks Index



(c) FTSE/JSE Mining Index



(d) USD/ZAR

Figure 3.1: Time series plots of daily closing FTSE/JSE indices and USD/ZAR exchange rate.

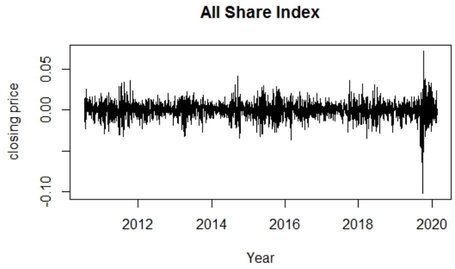
The time series plots shown on the left of Figure 3.1 indicate non-stationarity. Over time, a general upward trend is observed for the FTSE/JSE All Share Index and FTSE/JSE Bank Index. A general decreasing trend followed by an increasing trend is observed for the FTSE/JSE Mining Index. An overall decreasing trend is observed for the USD/ZAR closing prices.

The return series for each stock market indices and exchange rate are calculated as the first backward differences as the index values natural algorithm. For day t, the daily log return r_t is defined as:

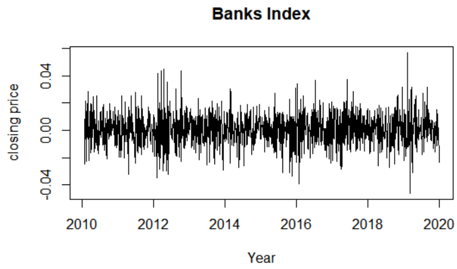
$$r_t = ln(P_t) - ln(P_{t-1}),$$

where P_t is the closing price and day t and P_t.

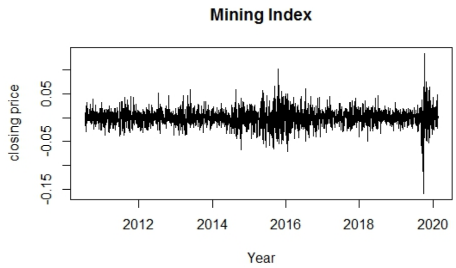
Figure 3.2 shows the time series plots of the log-returns of each FTSE/JSE stock market indices and USD/ZAR exchange rate.



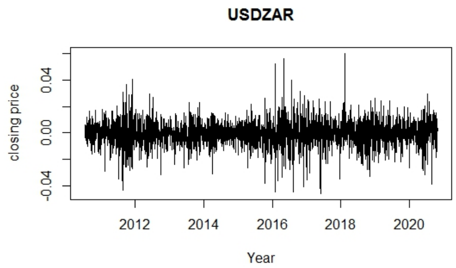
(a) FTSE/JSE ALSI returns



(b) FTSE/JSE Banks Index returns



(c) FTSE/JSE Mining Index returns



(d) USD/ZAR returns

Figure 3.2: Time series plots of daily closing FTSE/JSE indices and USD/ZAR exchange rate returns.

Figure 3.2 shows the time series plot of the log returns. The log returns are now stationary as the mean fluctuates around 0. However, heteroskedasticity and volatility clustering are evident due to the varying variance over time. Isolated instances of extreme returns triggered by financial market shocks are apparent, such as the 2015 stock market crash and the global COVID-19 pandemic.

Table 3.1 shows the descriptive statistics of the log returns of the four data sets analysed in this study. From Panel A, the FTSE/JSE ALSI and FTSE/JSE Mining Index have a positive mean, indicating the daily stock market indices have an upward trend, whereas the negative mean of the FTSE/JSE Banks Index and USD/ZAR indicated a negative trend over time. The excess kurtosis value signifies the leptokurtic nature of the return series. This means that the empirical distribution of the daily returns is much heavier than the well-known normal distribution. Significant skewness and excess kurtosis values are observed in the Banks Index returns, potentially reflecting the underperformance of the South African economy, high unemployment rates, and the adverse impact of government support for the state power utility, Eskom Holdings SOC Ltd which has proven to have negative effects for South African Bank Index stocks (Changole, 2019).

In this study, the results of the normality and autocorrelation tests in Table 3.1 are discussed in conjunction with the Q-Q and ACF and PACF plots from Figure 3.3 and Figure 3.4-Figure 3.7, respectively.

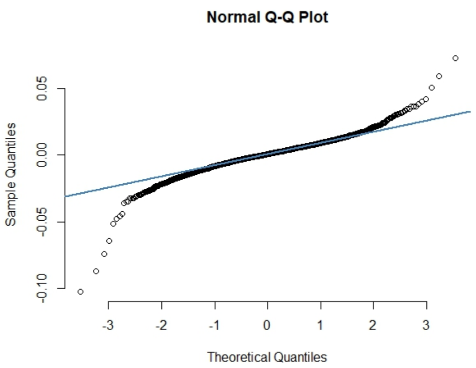
Observing the findings from Panel C, indicate that at a 5% level of significance, the null hypothesis of a unit root is rejected, suggesting that all return series are stationary. According to the KPSS test results, all returns are stationary since all p-values are 0.1, which is exceed 0.05; therefore the null hypothesis of stationarity is rejected.

Table 3.1: Descriptive summary statistics of daily return of financial stock market indices and exchange rate.

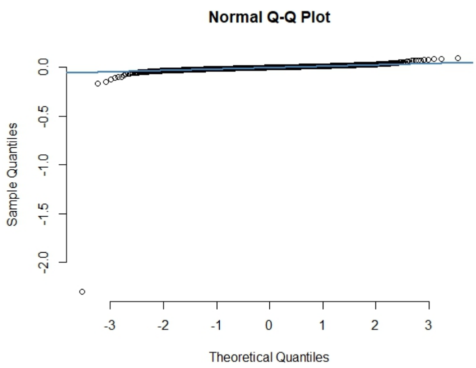
	FTSE/JSE ALSI		FTSE/JSE Banks Index		FTSE/JSE Mining Index		USDZAR	
Panel A: Descriptive Statistics								
Number of observations	2499.00		2499.00		2499.00		2675.00	
Minimum	0.1023		2.3021		0.1589		-0.0460	
Maximum	0.0726		0.0991		0.1346		0.0603	
Mean	0.0003		-0.0008		0.0002		-0.0003	
Skewness	-0.7310		-41.0875		-0.1605		-0.1671	
Excess Kurtosis	8.8822		1919.8771		6.0443		2.7766	
Panel B: Testing for normality, autocorrelation and heteroscedasticity								
Test	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
Jarque-Bera	8455.2667	$< 0.0001^{***}$	385117.1320	$< 0.0001^{***}$	3823.9300	$< 0.0001^{***}$	874.3992	$< 0.0001^{***}$
Ljung-Box $Q(10)$	40.6628	$< 0.0001^{***}$	2.0700	0.9958	22.3538	0.0134	3.8590	0.9535
Ljung-Box $Q(10)^2$	2551.8000	$< 0.0001^{***}$	0.0044	1.0000	1305.7000	$< 0.0001^{***}$	127.1600	$< 0.0001^{***}$
ARCH LM Test	936.0966	$< 0.0001^{***}$	0.0000	0.9900	50.0300	0.0000	15.4100	0.0000
Testing for unit root and stationarity								
Unit root test	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
ADF Test	–13.6259	0.0100	–12.9576	0.0100	–13.5509	0.0100	–14.7402	$< 0.0001^{***}$
PP Test	-2586.908	0.0100	-2548.2810	0.0100	-2430.2470	0.0100	-2688.0280	0.0100
KPSS Test	0.1303	0.1000	0.0914	0.1000	0.2394	0.1000	0.0431	0.1000

Note. *** symbolises a very small p-value

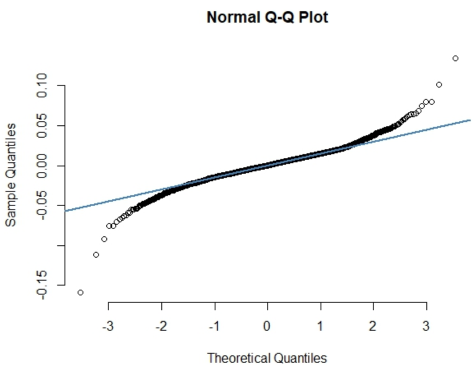
The normal Q-Q plots for the three FTSE/JSE stock market indices and USD/ZAR exchange rate are shown in Figure 3.3. It is evident that there is a lack of fit for extreme values for all returns, as the normal Q-Q plots show the tails for each return series to be heavier than the normal distribution. The Jarque-Bera test results in Table 3.1 for normality indicates a p-value less than 0.0001 for all four returns, thus rejecting the normality assumption at all levels of significance. The null hypothesis of normality for the Jarque-Bera test is rejected at 5% level of significance for all stock and currency returns. This presumes considering heavy-tailed models when analysing the returns series. The rejection of the normality assumption by the Jarque-Bera test is further confirmed by the results from the normal Q-Q plots.



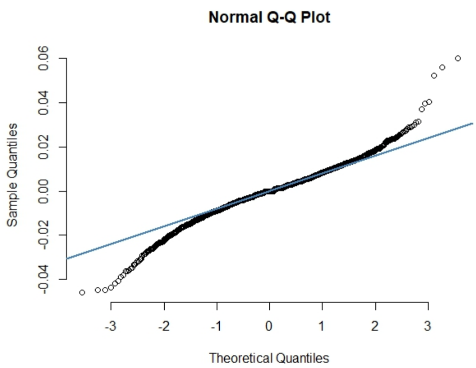
(a) FTSE/JSE ALSI returns



(b) FTSE/JSE Banks Index returns



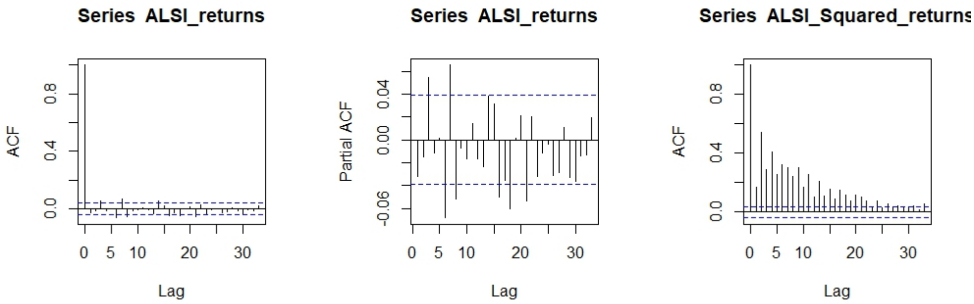
(c) FTSE/JSE Mining Index returns



(d) USD/ZAR returns

Figure 3.3: Normal Q-Q plots of daily closing FTSE/JSE indices and USD/ZAR exchange rate returns.

The ACF and PACF plots for each return series of the data are shown in Figures 3.4, 3.5, 3.6 and 3.7. The ACF and PACF plots for the FTSE/JSE ALSI and FTSE/JSE Mining Index show serial correlation in the returns data, whereas the FTSE/JSE Banks Index and USD/ZAR suggest no serial correlation in the return series data. Clearly, the ACF plots of the squared returns for the FTSE/JSE ALSI and FTSE/JSE Mining Index have serial correlation, and no correlation is observed for FTSE/JSE Banks Index, whereas some correlation is seen in the squared returns of the USD/ZAR return series. The significant p-values of the Ljung-Box test in Table 3.1 for FTSE/JSE Banks Index and USD/ZAR exchange rate fail to reject the null hypothesis of no autocorrelation. On the contrary, the null hypothesis for FTSE/JSE ALSI and the FTSE/JSE Mining Index is rejected, implying that the return series shows serial correlation. The ACF and PACF plots confirm the results of the Ljung-Box test in Table 3.1.

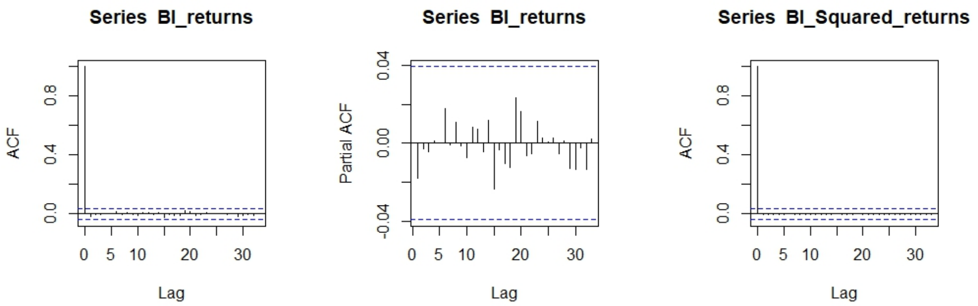


(a) ACF of FTSE/JSE ALSI returns

(b) PACF of FTSE/JSE ALSI returns

(c) ACF of FTSE/JSE ALSI squared returns

Figure 3.4: ACF and PACF plots of FTSE/JSE ALSI returns.

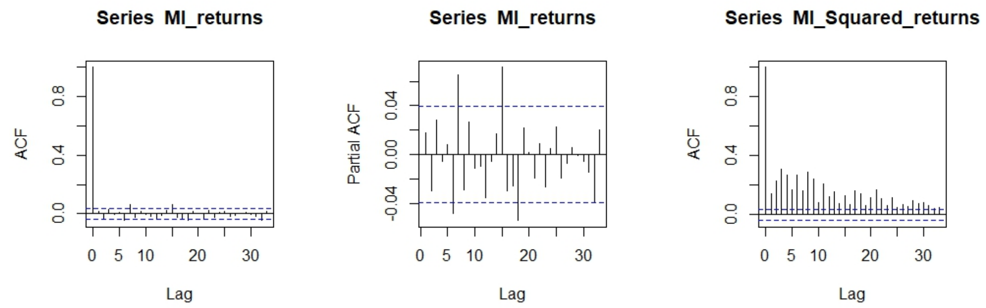


(a) ACF of FTSE/JSE Banks Index returns

(b) PACF of FTSE/JSE Banks Index returns

(c) ACF of FTSE/JSE Banks Index squared returns

Figure 3.5: ACF and PACF plots of FTSE/JSE Banks Index returns.

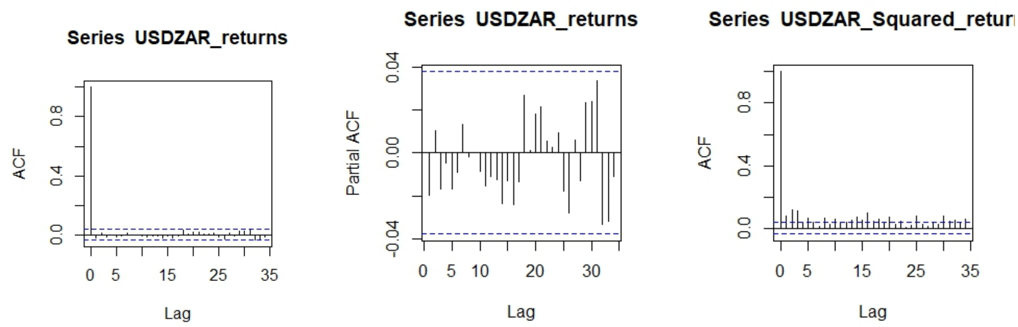


(a) ACF of FTSE/JSE Mining Index returns

(b) PACF of FTSE/JSE Mining Index returns

(c) ACF of FTSE/JSE Mining Index squared returns

Figure 3.6: ACF and PACF plots of FTSE/JSE Mining Index returns.



(a) ACF of USD/ZAR exchange rate returns

(b) PACF of USD/ZAR exchange rate returns

(c) ACF of USD/ZAR exchange rate squared returns

Figure 3.7: ACF and PACF plots of USD/ZAR returns.

3.2 Summary

In this chapter, the empirical properties of the daily closing prices of three FTSE/JSE stock markets and USD/ZAR exchange rate returns were evaluated. The following stylised characteristics of financial data were observed:

- The returns data in this study are characterised by heavy tails.
- Excess kurtosis is observed, implying the data series returns are leptokurtic.
- Serial autocorrelation is observed for FTSE/JSE ALSI and FTSE/JSE Mining Index; however, no autocorrelation is seen in the FTSE/JSE Banks Index and USD/ZAR.

- Volatility clustering is observed over time.

These stylised facts can be captured by appropriate statistical models or distributions. These suggested models/distributions are:

- Generalized Pareto distribution (GPD).
- Stable distributions (SD).
- GARCH-type model with heavy tail innovations such as the hybrid GARCH-GPD or GARCH-SD models.
- Extreme value mixture models (parametric, semi-parametric or non-parametric models).

3.3 Statistical software packages

In this study, we used the following R packages for the exploratory data analysis: stats by Team et al. (2018), methods by Chambers (2008), fbasics by Wuertz et al. (2017), tseries by Trapletti et al. (2015), TSA by Chan et al. (2022) and graphics by R Core Team (2016).

CHAPTER 4

UNIVARIATE STABLE DISTRIBUTIONS

4.1 Introduction

Stable distributions are a unique class of models in the field of probability and statistics. Stable models generalise the normal distribution and allow for heavy tails therefore taking into account the modelling of rare phenomena or extreme events. The study of stable distribution applications to real-world data by Nolan (2001) shows the maximum likelihood (ML) method to be feasible in estimating stable parameters to exchange rate, stock price, radar noise and ocean wave data sets. Xu et al. (2011) demonstrated that for Chinese stock returns data, the fitted α-stable model is better suited than the classical Black–Scholes model. Naradh (2016) highlighted the usefulness of stable models for Brazil, Russia, India, China and South Africa (BRICS) country exchange rates against the American Dollar in the univariate and multivariate case. This chapter provides an introduction to univariate stable distributions and builds on the work by Naradh (2016) by fitting univariate stable distributions to South African financial data.

4.2 Stable distributions

Stable distributions, also commonly referred to as α-stable or Lévy α-stable, are a subset of probability distributions that have several unique and interesting mathematical properties. Stable distributions are particularly important in statistics, finance and signal processing. This class is a four-parameter family of models that is well known for heavy tails, which is a favourable characteristic for modelling extreme events, financial returns and risk analysis (Nolan, 2014). The theory of stable distributions stems from the novel work of sums on independent identically distributed variables by Paul Lévy in the 1920s.

4.2.1 Definition of stable

Definition 4.1.1 (Nolan, 2003)

(i) The addition of two random variables that follow a normal distribution results in another random variable that is also normally distributed. If Y follows a normal distribution, then Y_1 and Y_2 are independent and follows the same distribution to Y with any positive constants a and b.

$$aY_1 + bY_2 \stackrel{\text{d}}{=} cY + d, \tag{4.1}$$

for $c \geq 0$ and $d \in \mathbb{R}$ where $\stackrel{\text{d}}{=}$ denotes being equal in distribution.

(ii) A random variable is considered symmetrically stable if it is both stable and symmetrically distributed around 0, meaning $Y \stackrel{\text{d}}{=} -Y$.

(iii) A random variable is classified as strictly stable when $d = 0$.

When adding independent normally distributed random variables, the sum has a mean that is the sum of the individual means, and a variance that is the sum of the individual variances. If Y follows a normal distribution with mean μ and variance σ, then the expressions $N(a\mu, a\sigma^2)$ and $N(b\mu, b\sigma^2)$ appear on the left-hand-side of equation (4.1). On the right-hand-side, the expression is denoted as $N(c\mu + d, c\sigma^2)$. Applying the addition rule, we derive that $c^2 = a^2 + b^2$ and $d = (a + b - c)\mu$. The term "stable" is aptly used because the shape remains consistent even after addition, as shown in equation (4.1). In the literature, the term sum stable is used to further reinforce the addition property in equation (4.1). The term "stable" should also be differentiated from other distributions such as max-stable, min-stable, multiplication stable, and geometric stable distributions. In older literature, different terms were used. "Stable" was referred to as what is now termed "strictly stable", while "quasi-stable" corresponds to what we now call "stable". Two random variables, Y and Z are considered to be similar if there exists constants $A > 0$ and $B \in \mathbb{R}$ where, $Y \stackrel{\text{d}}{=} AZ + B$. Therefore, the definition of stable can be rephrased as $aY_1 + bY_2$ and of the same type as Y.

Stable distributions are theoretically appealing but challenging to apply in practice. There are three specific cases with closed-form density functions that are known to exhibit stable properties. The family of α-stable distributions is a rich class and includes the normal, Cauchy and Lévy distributions as subclasses, which are described below by their respective density functions. The stable parameters α, β, γ and δ are further defined in the subsequent section.

(i) Normal/Gaussian distribution $Y \sim N(\mu, \sigma^2)$

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2 \right], \ -\infty < y < \infty. \tag{4.2}$$

The normal distribution is stable with $\alpha = 2$ and skewness $\beta = 0$.

(ii) Cauchy distribution $Y \sim$ Cauchy(γ, δ)

$$f(y) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (y-\delta)^2}, \ -\infty < y < \infty. \tag{4.3}$$

The Cauchy laws are stable with $\alpha = 1$ and $\beta = 0$.

(iii) Lévy distribution $Y \sim$ Lévy(γ, δ)

$$f(y) = \sqrt{\frac{\gamma}{2\pi}} \frac{1}{(y-\delta)^{3/2}} \exp - \left[\frac{\gamma}{2(y-\delta)} \right], \delta < y < \infty. \tag{4.4}$$

The Lévy distributions are stable with $\alpha = \frac{1}{2}$ and $\beta = 1$.

The normal and Cauchy distributions are both characterised as symmetric and bell-shaped curves. The key difference between the two lies in the fact that the Cauchy distribution exhibits heavier tails. However, the Lévy distribution is characterised by skewness and has heavier tails in contrast to the Cauchy distribution (Nolan, 2015).

The normal distribution is widely utilised in financial modelling due to its advantageous analytical properties, which are also shared by other members of the stable distribution family (Yang, 2012a).
The reasons for the popularity of the Normal distribution in financial modelling include:

- A straightforward and practical distribution that allows for the implementation of numerical methods.
- The Central Limit Theorem and the Law of Large Numbers simplify statistical problems by approximating distributions to be normal.
- Normally distributed random variables cluster around the mean, with decreasing odds of deviation from the mean as one moves away from it.

Prominent financial frameworks relying on the normal distribution include (Stoyanov et al., 2011):

(a) The Black-Scholes option pricing model.

(b) The Capital Asset Pricing Model (CAPM).

(c) Markowitz's Modern Portfolio Theory.

4.2.2 Alternative definitions of stability

Definition 4.1.2 (Nolan, 2003)
A non-degenerate Y is stable if and only if for every, $n > 1$, there exists constants $c_n > 0$ and $d_n \in \mathbb{R}$ such that

$$Y_1 + ... + Y_n \stackrel{\mathrm{d}}{=} c_n Y + d_n, \tag{4.5}$$

where $Y_1, \ldots, Y_n$ are independent and identical copies of Y. Y is strictly stable if $d_n = 0$ for all n. The constant c_n must take the form $c_n = n^{1/\alpha}$ where α is in the interval $(0, 2]$. The definitions above rely on the distributional properties of Y. Another way to characterize the distribution is through the Generalized Central Limit Theorem. Stable distributions are best described using either their characteristic function or Fourier transform. For a random variable Y with distribution function $F(y)$, the characteristic function is defined as:

$$\phi(t) = E\left(e^{itY}\right) = \int_{-\infty}^{\infty} e^{itY} dF(y), \tag{4.6}$$

where $\phi(t)$ specifies the distribution of Y and the sign function is defined as:

$$\text{sign } y = \begin{cases} -1, & y < 0, \\ 0, & y = 0, \\ 1, & y > 0. \end{cases}$$

Definition 4.1.3 (Nguyen and Sampson, 1991)
A distribution function $F(y)$ is considered to be univariate stable if for every $b_1 > 0$, $b_2 > 0$, real c_1, c_2, and there exists a positive number *b* and a real number *c* such that for every scalar *y*, where $-\infty < y < \infty$,

$$F\left(\frac{y - c_1}{b_1}\right) * F\left(\frac{y - c_2}{b_2}\right) = F\left(\frac{y - c}{b}\right), \tag{4.7}$$

where $*$ represents the convolution operation.

The characteristic function ϕ of a univariate stable distribution is defined by:

$$\phi\left(t\right) = iut - \gamma \left|t\right|^{\alpha} \left[1 + i\beta \frac{t}{\left|t\right|} \omega\left(t, \alpha\right)\right], \tag{4.8}$$

where $-\infty < t < \infty$, with $-\infty < \mu < \infty$, $-1 \leq \beta \leq 1, 0 < \alpha < 2$, $\dfrac{t}{|t|}$ at $t = 0$ and for all t :

$$\omega(t, \alpha) = \begin{cases} \tan\frac{\pi\alpha}{2}, & \alpha \neq 1, \\ \frac{2}{\pi}\ln(|t|), & \alpha = 1. \end{cases}$$

A random variable with a stable distribution can be characterised by having the same distribution as a linear combination of n independent copies of that random variable. Its characterisation also hinges on the relationships among the coefficients of this linear combination.

Definition 4.1.4 (Nolan, 2015)
A random variable Y is considered stable if and only if $Y \overset{\mathrm{d}}{=} aZ + b$, where $0 < \alpha \leq 2, -1 \leq \beta \leq 1$, a is not equal to 0 , $b \in \mathbb{R}$ and Z is a random variable with characteristic function

$$E\left(e^{itZ}\right) = \begin{cases} \exp(-\left|t\right|^{\alpha}\left[1 - i\beta\tan\frac{\pi\alpha}{2}(\text{sign}(t))\right]), & \alpha \neq 1, \\ \exp(-\left|t\right|\left[1 + i\beta\frac{2}{\pi}(\text{sign}(t))\log\left|t\right|\right]), & \alpha = 1. \end{cases} \tag{4.9}$$

The distributions are symmetric when $\beta = 0$ and $b = 0$. In this case, the characteristic function of aZ take the form $\phi(t) = e^{-a^{\alpha}|t|^{\alpha}}$.

4.2.3 Characterisation and parameterization of stable distributions

Stable distributions are characterised by four parameters:α, β, γ, and δ. The index of stability, index of law or characteristic of the exponent is explained by the parameter α where $0 < \alpha < 2$. The skewness parameter is represented by β where $-1 < \beta < 1$. The distribution is symmetric when $\beta = 0$. If $\beta > 0$, distribution exhibits right skewness and for $\beta < 0$, it shows left skewness. The shape of the distribution is defined by α and β. $\gamma > 0$ describes the scale parameter. The rightward or leftward shift of the distribution is denoted by δ and is called the location parameter. The distribution has a leftward shift when $\delta < 0$. Alternatively, the distribution has a rightward shift if $\delta > 0$. Various parameterizations are employed to describe stable distributions. This arises from historical evolution and the numerous issues encountered during the analysis of stable distributions. When working with data fitting or numerical methods, the first parameterization is typically preferred. However, if one prefers to work with simple algebraic structures, a different parameterization is recommended. For studying the analytical properties of strictly stable distributions, yet another parameterization would be beneficial. The notation $S\left(\alpha,\beta,\gamma,\delta\ k\right)$ is used to denote the class of stable distributions. The four parameters α,β,γ and δ are unknown and require estimation. The integer k serves to differentiate between various parameterizations. (Nolan, 2015).

Definition 4.2.1 (Nolan, 2015)

Nolan's S_0-parametrization A random variable Y is $S\left(\alpha,\beta,\gamma,\delta;0\right)$ if

$$Y \overset{\mathrm{d}}{=} \begin{cases} \gamma\left(Z-\beta\tan\frac{\pi\alpha}{2}\right)+\delta, & \alpha\neq 1, \\ \gamma Z+\delta, & \alpha=1, \end{cases} \tag{4.10}$$

where $Z\equiv Z(\alpha,\beta)$ has characteristic function (4.9). In this case, Y has characteristic function:

$$E\left(e^{itY}\right)=\begin{cases} \exp(-\gamma^{\alpha}\left|t\right|^{\alpha}\left[1+i\beta\left(\tan\frac{\pi\alpha}{2}\right)(\mathrm{sign}(t))\times(\left|\gamma t\right|^{1-\alpha}-1)\right]+i\delta t), & \alpha\neq 1, \\ \exp(-\gamma\left|t\right|\left[1+i\beta\frac{2}{\pi}(\mathrm{sign}(t))\times\log(\gamma\left|t\right|)\right]+i\delta t), & \alpha=1. \end{cases} \tag{4.11}$$

Nolan (2014) recommends using the S_0-parameterization for statistical inferences and numerical applications because it offers the simplest form of the characteristic function that remains continuous across all four parameters. The S_0-parameterization defines a location-scale family. If $Z\sim S(\alpha,\beta,\gamma,\delta;0)$, then for $\alpha\neq 0$, $b\in\mathbb{R}$, $aZ+b\sim S(\alpha,\mathrm{sign}(\alpha)\beta,\left|a\right|\gamma,a\delta+b;0)$.

Definition 4.2.2 Nolan's S_1-parametrization (Nolan, 2015)

A random variable Y is $S\left(\alpha,\beta,\gamma,\delta;1\right)$ if

$$Y \overset{\mathrm{d}}{=} \begin{cases} \gamma Z+\delta & \alpha\neq 1, \\ \gamma Z+\left(\delta+\beta\frac{2}{\pi}\gamma\log\gamma\right) & \alpha=1, \end{cases} \tag{4.12}$$

where $Z\equiv Z(\alpha,\beta)$ has characteristic function (4.9). In this scenario, Y has characteristic function:

$$E\left(e^{itY}\right)=\begin{cases} \exp(-\gamma^{\alpha}\left|t\right|^{\alpha}\left[1-i\beta(\tan\frac{\pi\alpha}{2})(\mathrm{sign}(t))\right]+i\delta t) & \alpha\neq 1, \\ \exp(-\gamma\left|t\right|\left[1+i\beta\frac{2}{\pi}(\mathrm{sign}(t))\log(\gamma\left|t\right|)\right]+i\delta t), & \alpha=1. \end{cases} \tag{4.13}$$

Yang (2012a) discusses the following parameterizations by Zolotarev.

Definition 4.2.3 Zolotrev A-parameterization

A random variable Y is $S\left(\alpha,\beta,\gamma,\delta;A\right)$ if the characteristic function can be expressed as:

$$E\left(e^{i}tY\right)=\begin{cases} \exp(\gamma[it\delta-\left|t\right|^{\alpha}+it\left|t\right|^{\alpha-1}\beta\tan\frac{\pi\alpha}{2}]) & \alpha\neq 1, \\ \exp(\gamma[it\delta-\left|t\right|^{\alpha}-i\beta\frac{2}{\pi}t\log\left|\mathrm{t}\right|), & \alpha=1. \end{cases} \tag{4.14}$$

The characteristic functions in (4.14) exhibit discontinuities in the parameters that define them. These discontinuities occur at all points of the form $\alpha = 1$ and $\beta \neq 0$. Taking limits $\alpha^\star \to 1\,(\alpha^\star \neq 1), \beta^\star \to \beta \neq 0, \gamma^\star \to \gamma$ and $\delta^\star \to \delta$, does not result in a stable distribution with parameters $\alpha = 1, \beta, \gamma$ and δ; instead, it leads to an improper distribution where the entire measure tends to infinity. Adding a shift to the location parameter, $-\beta \tan\frac{\pi\alpha}{2}$, removes this discontinuity.

Definition 4.2.4 Zolotrev M-parameterization
A random variable Y is $S\,(\alpha, \beta, \gamma, \delta; M)$ if the characteristic function can be defined as follows:

$$E\left(e^{itY}\right) = \begin{cases} \exp(\gamma[it\delta - |t|^\alpha + it(|t|^{\alpha-1} - 1)\beta \tan\frac{\pi\alpha}{2}]) & \alpha \neq 1, \\ \exp(\gamma[it\delta - |t|^\alpha - i\beta\frac{2}{\pi}t\log|\mathrm{t}|), & \alpha = 1. \end{cases} \tag{4.15}$$

Note the similarities between Nolan's S_0-parameterization and Zolotrev M-parameterization where adjustments are made primarily to γ and δ to better align with the traditional concepts of scale and location parameters. Similarly, Nolan S_1-parameterization corresponds to Zolotrev A-parameterization. The cumulative distribution function satisfies $F(y;\gamma) = F(y/\gamma; 1)$, where γ represents the scale parameter in the classical definition for some distributions. Nolan's S_0-parameterization, Nolan's S_1-parameterization and the scale parameter γ fall under this category. Some parameterizations mimic the scale parameter; where, it is observed as a combination of scale parameters and some other parameters, like the Zolotrev A-parameterization.

Definition 4.2.5 Zolotrev B-parameterization
A random variable Y is $S\,(\alpha, \beta, \gamma, \delta; B)$ when the characteristic function is defined as:

$$E\left(e^{itY}\right) = \begin{cases} exp(\gamma[it\delta - |t|^\alpha \exp(-i\frac{\pi}{2}\beta K(\alpha)\mathrm{sign}(t))]) & \alpha \neq 1, \\ \exp(\gamma[it\delta - |t|^\alpha\,(\frac{\pi}{2} + i\beta\log|\mathrm{t}|\mathrm{sign}(t))]), & \alpha = 1, \end{cases} \tag{4.16}$$

where $K = \alpha - 1 + sign(1 - \alpha)$. The parameters in this formulation have the same range of values as in the A-parameterization. Similar to the A-parameterization, the B-parameterization exhibit discontinuities at points where $\alpha = 1$. However, the B-parameterization yields a limit distribution that exists and remains stable as $\alpha^\star \to 1_+, \beta^\star \to \beta, \gamma^\star \to \gamma$ and $\delta^\star \to \delta$. In this expression $\to 1_+$ describes convergence to 1 from above.

Zolotarev (1986) discusses the following parameterizations:

Definition 4.2.6 Zolotrev C-parameterization
A random variable Y is $S(\alpha, \beta, \gamma, \delta; C)$ if the characteristic function is denoted by:

$$E\left(e^{itY}\right) = -\delta|t|^{\alpha}\, exp\Big(-i\left(\frac{\pi}{2}\right)\theta_{\alpha}\text{sign t}\Big), \tag{4.17}$$

where the parameters vary within their respective limits: $0 < \alpha \leq 2,\ \delta > 0,\ |\theta| \leq \theta_{\alpha} = \min(1, 2/\alpha - 1)$.

Definition 4.2.7 Zolotrev E-parameterization
A random variable Y is $S(\alpha, \beta, \gamma, \delta; E)$ when the characteristic function is expressed as:

$$E\left(e^{itY}\right) = -\nu^{\frac{1}{2}}\left(\log|t| + \tau - i\left(\frac{\pi}{2}\right)\theta\text{sign}\, t\right) + \mathbb{C}(\nu^{-\frac{1}{2}} - 1), \tag{4.18}$$

where $\mathbb{C} \approx 0.577$ (Euler constant) and the parameters vary within their limits: $\nu \geq \frac{1}{4}$, $|\theta| \leq (1, 2\sqrt{\nu} - 1)$, $|\tau| < \infty$. It is crucial to establish the appropriate parameterization for stable distributions prior to random variable generation, hypothesis testing, and parameter estimation. Below are some conversions between different parameterizations.

$S_0 \rightarrow S_1$

$$\beta_1 = \beta_0, \gamma_1 = \gamma_0, \delta_1 = \begin{cases} \delta_0 - \beta\gamma\tan\frac{\pi\alpha}{2}, & \alpha = 1, \\ \delta_0 - \beta\frac{2}{\pi}\gamma\ln\gamma, & \alpha \neq 1. \end{cases}$$

$(M) \rightarrow (A)$

$$\begin{aligned} \beta_A = \beta_M, \delta_A = \delta_M - \beta_M\tan\tfrac{\pi\alpha}{2}, \gamma_A = \gamma_M, \quad & \text{if } \alpha \neq 1, \\ \beta_A = \beta_M, \delta_A = \delta_M, \gamma_A = \gamma_M, \quad & \text{if } \alpha = 1. \end{aligned}$$

4.2.4 Distribution and density functions

Yang (2012a) suggests that stable distributions do not have explicit closed-form cumulative distribution functions or probability density functions. Nolan (2001) explores a numerical method for computing densities, and Zolotarev (1986) shows in detail the integral form of density functions for stable distributions.

Zolotarev (1986) states that the integral formula in the M-parameterization, defined by

$$\zeta = \zeta(\alpha, \beta) = \begin{cases} -\beta \tan\frac{\pi\alpha}{2} & \alpha \neq 1, \\ 0 & \alpha = 1. \end{cases} \tag{4.19}$$

$$\theta_0 = \theta_0(\alpha, \beta) = \begin{cases} -\frac{1}{\alpha}\arctan(\beta \tan\frac{\pi\alpha}{2}) & \alpha \neq 1, \\ \frac{\pi}{2} & \alpha = 1. \end{cases} \tag{4.20}$$

$$c_1(\alpha, \beta) = \begin{cases} \dfrac{1}{\pi}\left(\dfrac{\pi}{2} - \theta_0\right) & \alpha < 1, \\ 0 & \alpha = 1; \\ 1 & \alpha > 1. \end{cases} \tag{4.21}$$

$$V(\theta; \alpha, \beta) = \begin{cases} (\cos\alpha\theta_0)^{1/\alpha-1}\left(\frac{\cos\theta}{\sin\alpha(\theta_0+\theta)}\right)^{\alpha/\alpha-1} \frac{\cos(\alpha\theta_0+(\alpha-1)\theta)}{\cos\theta} & \alpha \neq 1, \\ \frac{2}{\pi}\left(\frac{\pi/2+\beta\theta}{\cos\theta}\right)\exp\left(\frac{1}{\beta}(\frac{\pi}{2}+\beta\theta)\tan\theta\right) & \alpha = 1, \beta \neq 0. \end{cases} \tag{4.22}$$

The integral formula is very complex which is the reason why there various setbacks exist for the applications of stable distributions.

Theorem 4.3.1 (Nolan, 2015)
All non-degenerate stable distributions are continuous with an infinitely differentiable densities where, $f(y|\alpha, \beta, \gamma, \delta; k)$ denotes the density function and $F(y|\alpha, \beta, \gamma, \delta; k)$ denotes the distribution function of an $S(\alpha, \beta, \gamma, \delta; k)$ distribution. When the scale parameter γ = 1 and the location parameter $\delta = 0$, the distribution is standardised. The density function and distribution function of the standardised distribution are denoted respectively by $f(y|\alpha, \beta; k)$ and $F(y|\alpha, \beta; k)$. Stable densities are defined over the entire real line or half a line. The half-line scenario occurs when $\alpha < 1$ and $\beta = -1$ or $\beta = 1$. Further details are provided in Lemma 4.3.1 concerning specific limits.

Lemma 4.3.1 (Nolan, 2015)

$$support f(y|\alpha,\beta,\gamma,\delta;0) = \begin{cases} [\delta - \gamma\tan\frac{\pi\alpha}{2}, \infty), & \alpha < 1 \text{ and } \beta = 1, \\ (-\infty, \delta + \gamma\tan\frac{\pi\alpha}{2}], & \alpha < 1 \text{ and } \beta = 1 \\ (-\infty, +\infty), & \text{otherwise.} \end{cases} \tag{4.23}$$

$$support f(y|\alpha,\beta,\gamma,\delta;1) = \begin{cases} [\delta, \infty), & \alpha < 1 \text{ and } \beta = 1, \\ (-\infty, \delta] & \alpha < 1 \text{ and } \beta = 1 \\ (-\infty, +\infty), & \text{otherwise.} \end{cases} \tag{4.24}$$

The term $\tan\frac{\pi\alpha}{2}$ is a constant, as is often seen when working with stable distributions. We observe as $\alpha \uparrow 1$, then $\tan\frac{\pi\alpha}{2} \downarrow -\infty$. There is a discontinuity at $\alpha = 1$. This is troublesome when working with stable distributions. It is also possible that if $|\beta| = 1$, then as $\alpha \uparrow 1$, the support in Lemma 3.3.1 tends to $\mathbb{R}$ naturally.

The reflection property is a basic fact of stable distributions.

Property 4.3.1 Reflection Property (Nolan, 2015)
For any α and β, $P \sim S(\alpha, \beta; k)$ where $k = 0, 1, 2$

$$P(\alpha, -\beta) \stackrel{\mathrm{d}}{=} -P(\alpha, \beta). \tag{4.25}$$

Random variables $P(\alpha, \beta)$ have density and distribution functions that satisfy: $f(y|\alpha, \beta; k) = f(-y|\alpha, -\beta; k)$ and $F(y|\alpha, \beta; k) = 1 - F(-y|\alpha, -\beta; k)$. If $Y \sim S(\alpha, \beta, \gamma, \delta; k)$ then $-Y \sim S(\alpha, -\beta, \gamma, -\delta; k)$. Therefore, $f(y|\alpha, \beta, \gamma, \delta; k) =$ $f(-y|\alpha, -\beta, \gamma, -\delta; k)$ and $F(y|\alpha, \beta, \gamma, \delta; k) = 1 - F(-y|\alpha, -\beta, \gamma, -\delta; k)$.

If $\beta = 0$, the reflection property suggests $f(y|\alpha, 0; k) = f(-y|\alpha, 0; k)$. This implies that the density and distribution functions are symmetric around 0. Graphically, it is observed that as α decreases, the peaks of bell-shaped symmetric stable distributions get higher, the region closest to the peak gets lower and the tails get heavier. The distribution is skewed with the right tail heavier than the left tail $P(Y > y) > P(Y < -y)$ for large $y > 0$ when $\beta > 0$. A stable distribution is classified as entirely right-skewed when $\beta = 1$. $\beta < 0$ is a reflection of $\beta > 0$ by the reflection property. In this case, the left tail is heavier than the right tail. A stable distribution is classified as entirely left-skewed when $\beta = 1$. A non-standardised Normal distribution is obtained when $\alpha = 2$. In this case, $\tan\frac{\pi\alpha}{2}$ in equation (4.9). The distribution is characterised as always symmetric, and the characteristic function is real irrespective of the value of β. It can be represented as, $P(2, -\beta) \stackrel{\mathrm{d}}{=} -P(2, \beta)$. Typically, all stable distributions become symmetric as $\alpha = 2$ and β is challenging to estimate accurately, which diminishes its significance in practical applications.

Stable distributions do not have a known formula for the location of the mode. All stable distributions can be described as unimodal. $m(\alpha, \beta)$ denotes the mode of $Z \sim S(\alpha, \beta; 0)$ distribution. $m(\alpha, -\beta) = -m(\alpha, \beta)$, by the reflection property. It can also be numerically observed that $P(Z > m(\alpha, \beta)) > P(Z < m(\alpha, \beta))$ when $\beta > 0$ (more mass to the right of the mode). By the reflection property, when $\beta < 0$, then $P(Z > m(\alpha, \beta)) < P(Z < m(\alpha, \beta))$ and there is more mass to the left of the mode. If $\beta = 0$, then $P(Z > m(\alpha, \beta)) = P(Z < m(\alpha, \beta)) = 1/2$ (Nolan, 2015).

Property 4.3.2 (Yang, 2012a)
Let $Y \sim S(\alpha, \beta, \gamma, \delta)$ and $f(y)$ and $F(y)$, denote the density and distribution function, respectively. When $\alpha = 2$, the normal distribution exhibits asymptotic tail properties. For $\alpha < 1$, stable distributions have one tail when $\beta = \pm 1$, and two tails otherwise. In certain scenarios, these distributions can exhibit asymptotic power laws with heavy tails.

(i) Paretian tail density
The tail densities and probabilities of non-Normal stable distributions asymptotically follow power laws. If $0 < \alpha < 2$ and $-1 < \beta \leq 1$, then as $x \to \infty$,

$$\frac{1 - F(y)}{\gamma^{\alpha} c_{\alpha}(1 + \beta) x^{-\alpha}} \to 1, \frac{f(y)}{\alpha \gamma^{\alpha} c_{\alpha}(1 + \beta) x^{-(\alpha+1)}} \to 1, \tag{4.26}$$

where $c_{\alpha} = \dfrac{\sin(\frac{\pi\alpha}{2})\Gamma(\alpha)}{\pi}$.
Likewise for the lower tail properties, $-1 \leq \beta < 1$ as $x \to \infty$:

$$\frac{F(-y)}{\gamma^{\alpha} c_{\alpha}(1 - \beta) x^{-\alpha}} \to 1, \frac{f(-y)}{\alpha \gamma^{\alpha} c_{\alpha}(1 - \beta) x^{-(\alpha+1)}} \to 1, \tag{4.27}$$

(ii) Stable distributions are unimodal.

(iii) The laws of stability have densities with uniformly bounded derivatives of every order.

Property 4.3.3 (Yang, 2012a)
For any parameter quadruples $(\alpha, \beta_k, \gamma_k, \delta_k)$ and all real numbers h and $c_k, k = 1, \ldots, n$, uniquely determine a parameter quadruple $(\alpha, \beta, \gamma, \delta)$ where,

$$S(\alpha, \beta, \gamma, \delta) \overset{\mathrm{d}}{=} \sum_k c_k S(\alpha, \beta_k, \gamma_k, \delta_k) + h.$$

In the A-parameterization, the dependence of the quadruple $(\alpha, \beta, \gamma, \delta)$ on the selected parameters and values is:

$$\delta = \sum_k \delta_k \left|c_k\right|^{\alpha},$$

$$\delta\beta = \sum_k \delta_k \beta_k \left|c_k\right|^{\alpha} \mathrm{sign(c_k)},$$

$$\delta\gamma = \sum_k \delta_k \gamma_k c_k + h_0,$$

where $h_0 = h$ if $\alpha \neq 1$ and $h_0 = h - \frac{2}{\pi}\sum_k \delta_k \beta_k c_k \log\left|\mathrm{c_k}\right|$ if $\alpha = 1$.

Property 4.3.4 (Yang, 2012a)
For any two parameter quadruples $(\alpha, \beta, \gamma, \delta)$ and $(\alpha, \beta, \gamma^{'}, \delta^{'})$, we uniquely determine real numbers $a > 0$ and b where,

$$S(\alpha, \beta, \gamma, \delta) \overset{\mathrm{d}}{=} a\, S(\alpha, \beta, \gamma^{'}, \delta^{'}) + \lambda b,$$

with the A-parameterization, the dependence of a and b on the parameters is describes by:

$$a = (\gamma/\gamma')^{1/\alpha}, \tag{4.28}$$

$$b = \begin{cases} \delta - \delta^{'}(\gamma/\gamma')^{\frac{1}{\alpha}-1}, & \alpha \neq 1, \\ \delta - \delta^{'} + \frac{2}{\pi}\beta\log(\gamma/\gamma'), & \alpha = 1. \end{cases} \tag{4.29}$$

This property is used to standardise any stable distribution by letting $\delta = 0$ and $\gamma = 1$.

4.2.5 Properties of stable laws

The basic properties of Nolan's S_1-parameterization are summarised; that is $Y \sim S(\alpha, \beta, \gamma, \delta; 1)$ without proof.

- When $\beta = 0$, it is implied that the stable distribution is symmetric.
- The reflection property is such that: $-Y \sim S(\alpha, -\beta, \gamma, -\delta; 1)$.
- All stable laws have densities $f(y)$ that are smooth and unimodal.
- The support of Y is the whole real line and exceptions occur when $\alpha < 1$ and $\beta = 1$, where the support is $[\delta, +\infty)$ or if $\alpha < 1$ and $\beta = -1$. In this scenario, the support is $(-\infty, \delta]$.
- Tail behaviour: When $\alpha < 2$ and $-1 < \beta \leq 1$, then the density and distribution functions have an asymptotic power law. As $y \to \infty$,

$$1 - F(x) = P(X > x) \sim \gamma^{\alpha} c_{\alpha}(1 + \beta)x^{-\alpha}, \tag{4.30}$$

$$f(x|\alpha, \beta, \gamma, \delta; 0) \sim \alpha\gamma^{\alpha}(1 + \beta)x^{-(\alpha+1)}, \tag{4.31}$$

  where $c_{\alpha} = \dfrac{\sin(\frac{\pi\alpha}{2})\Gamma(\alpha)}{\pi}$. The stable Paretian distribution is employed in the non-Gaussian scenarios due to the similar tail behaviour to the Pareto distribution. For all, $\alpha < 2$ and $-1 < \beta < 1$, both tail probabilities and densities asymptotically follow power laws. However, when $\beta = -1$, the right tail of the distribution does not exhibit an asymptotic power law. Similarly, when $\beta = 1$, the left tail does not exhibit as asymptotic power law.

- The Generalised Central Limit Theorem is fundamental property of stable distributions.
- Fractional moments: When $\alpha < 2$, $E|X|^p$ is finite for $0 < p < \alpha$, and infinite for $p \geq \alpha$. For $\alpha < 2$, the population variance is infinite, and for $\alpha \leq 1$, the population mean is undefined. This is an observed consequence of the power law tail behaviour (Nolan, 2014).

4.2.6 Sum of stable random variables

The fundamental property of stable distributions is that sums of $\alpha-$stable random variables are also $\alpha-$stable. However, the specific results can vary depending on the parameterization used.

Property 4.5.1 (Nolan, 2015)
The properties of $S(\alpha, \beta, \gamma, \delta; 0)$ distributions are as follows:

(a) If $Y \sim S(\alpha, \beta, \gamma, \delta; 0)$, then for any non-zero a and real number b

$$aY + b \sim S(\alpha, (\text{sign}\, a)\beta, |a|\,\gamma, a\delta + b; 0).$$

(b) The characteristic density, distribution and characteristic functions exhibit joint continuity across all four parameters $(\alpha, \beta, \gamma, \delta)$.

(c) If Y_1 follows a $S(\alpha, \beta_1, \gamma_1, \delta_1; 0)$ distribution and Y_2 follows a $S(\alpha, \beta_2, \gamma_2, \delta_2; 0)$; where, Y_1 and Y_2 are independent, then $Y_1 + Y_2$ follows a $S(\alpha, \beta, \gamma, \delta; 0)$ distribution where

$$\beta = \frac{\beta_1\gamma_1^\alpha + \beta_2\gamma_2^\alpha}{\gamma_1^\alpha + \gamma_2^\alpha}, \gamma_\alpha = \gamma_1^\alpha + \gamma_2^\alpha,$$

$$\delta = \begin{cases} \delta_1 + \delta_2 + \tan\frac{\pi\alpha}{2}[\beta\gamma - \beta_1\gamma_1 - \beta_2\gamma_2], & \alpha \neq 1, \\ \delta_1 + \delta_2 + \frac{2}{\pi}[\beta\gamma log\gamma - \beta_1\gamma_1 log\gamma_1 - \beta_2\gamma_2 log\gamma_2], & \alpha = 1. \end{cases}$$

The expression $\gamma_\alpha = \gamma_1^\alpha + \gamma_2^\alpha$ represents the general rule for adding variances of independent random variables and applies for both parameterizations.

Property 4.5.2 (Nolan, 2015)
The properties of $S(\alpha, \beta, \gamma, \delta; 1)$ distributions are as follows:

(a) If $Y \sim S(\alpha, \beta, \gamma, \delta; 1)$ then for any non-zero a and real number b

$$aY + b \sim \begin{cases} S(\alpha, (\text{sign}\, a)\beta, |a|\,\gamma, a\delta + b; 1), & \alpha \neq 1, \\ S(1, (\text{sign}\, a))\beta, |a|\,\gamma, a\delta + b - \frac{2}{\pi}\beta\gamma a log(|a|\,; 1), & \alpha = 1. \end{cases}$$

(b) The characteristic density, distribution, and characteristic functions remain continuous except at $\alpha = 1$, where they exhibit discontinuities.

(c) If Y_1 follows a $S(\alpha, \beta_1, \gamma_1, \delta_1; 1)$ distributions and Y_2 follows a $S(\alpha, \beta_2, \gamma_2, \delta_2; 1)$; where, Y_1 and Y_2 are independent, then $Y_1 + Y_2$ follows a $S(\alpha, \beta, \gamma, \delta; 1)$ where

$$\beta = \frac{\beta_1\gamma_1^\alpha + \beta_2\gamma_2^\alpha}{\gamma_1^\alpha + \gamma_2^\alpha},\ \gamma_\alpha = \gamma_1^\alpha + \gamma_2^\alpha,\ \delta = \delta_1 + \delta_2.$$

Property 4.5.1 demonstrates that γ and δ serve as the standard scale and location parameters in the Nolan's S_0-parameterization but not in the Nolan's S_1-parameterization when $\alpha = 1$. Part (c) in the first parameterization indicates that location parameter, δ is the sum $\delta_1 + \delta_2$.

By induction, formulas can be derived for the sum of *n* stable random variables.
Let $Y_j \sim S(\alpha, \beta_j, \gamma_j, \delta_j; k), j = 1, 2, \ldots n,$ and independent and arbitrary $w_1, ..., w_n$, the sum is denoted as:

$$w_1 Y_1 + w_2 Y_2 + ... + w_n Y_n \sim S(\alpha, \beta, \gamma, \delta; k), \tag{4.32}$$

where

$$\gamma^{\alpha} = \sum_{j=1}^{n} |w_j \gamma_j|^{\alpha},$$

$$\beta = \frac{\sum_{j=1}^{2} \beta_j (\mathrm{sign(w_j)}) \, |w_j \gamma_j|^{\alpha}}{\gamma^{\alpha}},$$

$$\delta = \begin{cases} \sum_j w_j \delta_j + \tan \frac{\pi\alpha}{2} (\beta\gamma - \sum_j \beta_j w_j \gamma_j), & k = 0, \alpha \neq 1, \\ \sum_j w_j \delta_j + \frac{2}{\pi} (\beta\gamma \mathrm{log}\gamma - \sum_j \beta_j w_j \gamma_j \mathrm{log} \, |\mathrm{w_j \gamma_j}|), & k = 0, \alpha = 1, \\ \sum_j w_j \delta_j, & k = 1, \alpha \neq 1, \\ \sum_j w_j \delta_j - \frac{2}{\pi} \sum_j \beta_j w_j \gamma_j \mathrm{log} \, |\mathrm{w_j}|, & k = 1, \alpha = 1. \end{cases}$$

If $\beta_j = 0$, for all j then $\beta = 0$ and $\delta = \sum_j w_j \delta_j$. Additionally, we observe an important property known as the scaling property for random variables.

When the terms are independent and identically distributed,

$$Y_j \sim S(\alpha, \beta, \gamma, \delta; k),$$

then,

$$Y_1 + ... + Y_n \sim S(\alpha, \beta, n^{1/\alpha}\gamma, \delta_n; k), \tag{4.33}$$

where

$$\delta_n = \begin{cases} n\delta + \gamma\beta \tan\frac{\pi\alpha}{2}(n^{1/\alpha} - n), & k = 0, \alpha \neq 1, \\ n\delta + \gamma\beta\frac{\pi\alpha}{2} n\text{logn}, & k = 0, \alpha = 1, \\ n\delta & k = 1. \end{cases}$$

The shape of the sum of *n* terms remains unchanged from the original shape. It should be noted that stable distributions uniquely possess this property, as highlighted by Nolan (2015)).

4.2.7 Stable parameter estimation

In this study, the Nolan's S_0-parameterzation for univariate stable parameter estimation is adopted. The ML estimation method is the most widely used approach in stable parameter estimation. A simulation study by Ojeda (2001) found that the ML method provides the most accurate results followed by the empirical characteristic function method, the quantile method, and lastly, the fractional moment method (Nolan, 2003).

4.2.8 Maximum likelihood estimation

The parameter vector represented by $\vec{\theta} = (\alpha, \beta, \gamma, \delta_0)$, and the density function is denoted by $f(x|\vec{\theta})$. The parameter space is denoted by $\Theta = (0, 2] \times [-1, 1] \times (0, \infty) \times (-\infty, \infty)$. The log-likelihood function for an independent and identically distributed stable sample $Y_1, \ldots, Y_n$ is expressed by:

$$L(\vec{\theta}) = \sum_{i=1}^{n} \log f(Y_i|\vec{\theta}).$$

Due to the absence of closed-form formulas for general stable densities, computing the likelihood function poses challenges. The R program stable by Robust Analysis Inc. (2013) computes reliable stable densities for α greater than 0.1 and any β, γ and δ_0. The McCulloch (1986) quantile method, initially utilised to approximate the parameters and the parameter space can constrain a method to maximise it, called the quasi-Newton method. DuMouchel (1971) and DuMouchel (1973) indicated that if $\vec{\theta_0}$ lies within the interior of the parameter space Θ, the maximum likelihood estimator is consistent and asymptotically Normal with mean $\vec{\theta_0}$ and covariance matrix given by $n^{-1}B$ where $B = (b_{ij})$ is the inverse of the 4×4 Fisher information matrix $\boldsymbol{I}$. Entries in $\boldsymbol{I}$ are expressed by:

$$\boldsymbol{I}_{ij} = \int_{-\infty}^{\infty} \frac{\partial f}{\partial \theta_i} \frac{\partial f}{\partial \theta_j} dy.$$

The behaviour of the estimators is not known when $\vec{\theta}$ is close to the boundary of the parameter space. The distribution of the estimator gets skewed away from the boundary. When $\alpha = 2$ or $\beta = \pm 1, \vec{\theta}$ is on the boundary of the parameter space. At the boundary point, the Normal distribution for the estimators tends to a degenerate distribution. Away from the boundary, large sample confidence intervals for each of the parameters are given by

$$\hat{\theta}_i \pm Z_{\frac{\alpha}{2}} \frac{\sigma_{\hat{\theta}_i}}{\sqrt{n}}.$$

where $\sigma_{\hat{\theta}_1} \cdots \sigma_{\hat{\theta}_n}$ are the diagonal entries of B (Nolan, 2001).

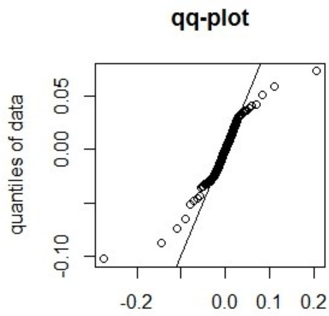
4.3 Empirical results

In this section, the S_0-parameterization stable model by Nolan (2001) is fitted to the daily closing prices of the FTSE/JSE ALSI, FTSE/JSE Banks Index and FTSE/JSE Mining Index and the USD/ZAR exchange rates. Stable parameters are estimated under Nolan's S_0-parameterization using maximum likelihood estimation. Table 4.1 reports the values estimated for each stable parameter on the returns of the financial data under investigation. The results are presented and discussed in the subsequent sections.

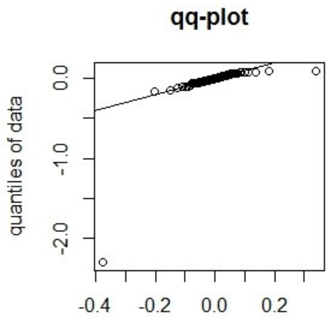
Table 4.1: ML stable parameter estimates for South African financial data returns.

Returns	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\delta}$
All Share Index	1.7316	-0.2464	0.0060	0.0008
Banks Index	1.7522	-0.0894	0.0097	0.0004
Mining Index	1.7436	$<$0.0002	$<$0.0002	$<$0.0002
USD/ZAR exchange rate	1.7729	-0.3250	0.0059	0.0003

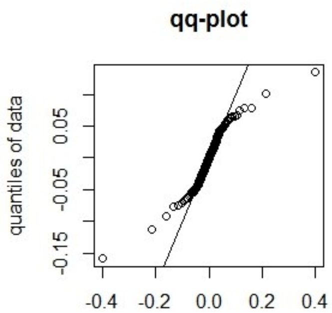
The skewness parameter β is negative for all returns, therefore indicating that each fitted univariate stable distribution is skewed to the left. Since the location parameter $\delta > 0$ for all returns it can be implied that the fitted distributions have a rightward shift.



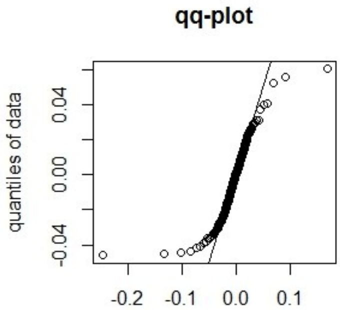
(a) All Share Index



(b) Banks Index



(c) Mining Index

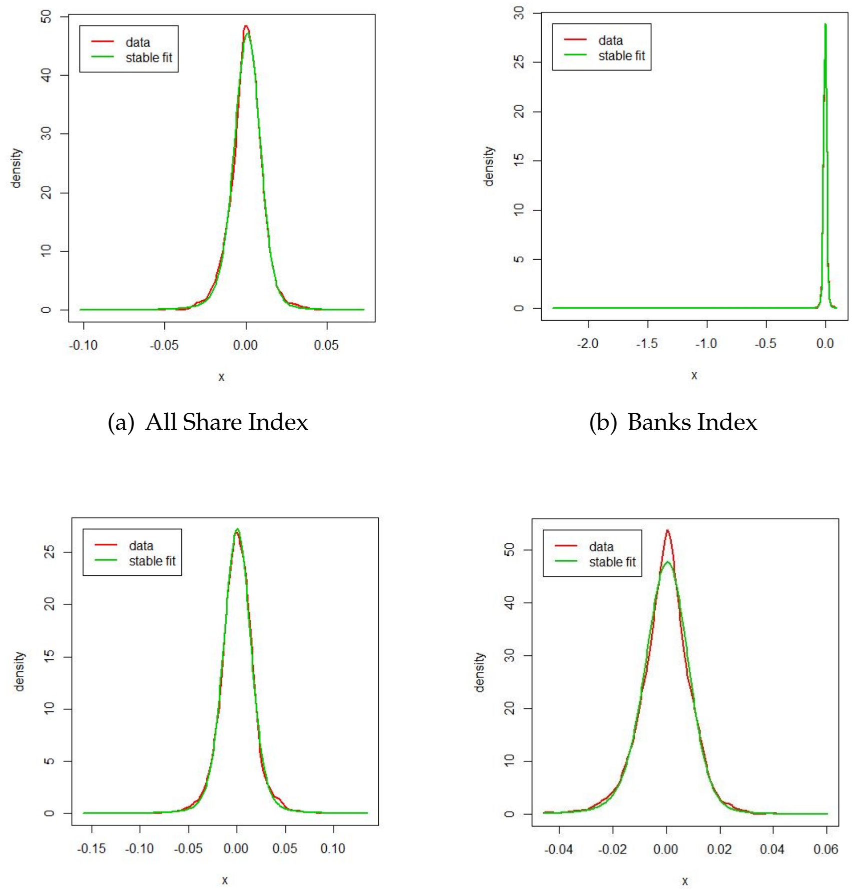


(d) USDZAR

Figure 4.1: Q-Q plots for financial market indices and exchange rate returns.

Figure 4.1 displays Q-Q plots that visually appear compressed and dominated by extreme values. The heavy tails observed in these Q-Q plots suggest significant variability in extreme order statistics. Therefore, deviations from the straight-line Q-Q plot are challenging to assess. This also shows that extreme tails from the data set are lighter than the stable model (Nolan, 2005). The Q-Q plots imply the inadequacy of the fitted stable model at extreme values. Therefore, the discrepancies mentioned about the Q-Q plots highlight the focus on formal model adequacy tests such as the Anderson-Darling goodness-of-fit test.

Empirically, we compare the densities of the daily returns to univariate S_0 stable distribution. Figure 4.2 shows graphically a close fit of the estimated univariate S_0 model to the daily returns of the data, as the fitted stable model does not deviate much from the returns of each financial index and exchange rate. A better fit for the data is provided over most of the range, with extreme tails being overestimated.



(a) All Share Index
(b) Banks Index
(c) Mining Index
(d) USDZAR

Figure 4.2: Stable density plots for financial market indices and exchange rate returns.

The AD test statistic and corresponding p-values are reported in Table 4.2. The fitted stable distributions fits the data well since the p-values are greater than 0.05 for all the returns. We conclude that the fitted $S_0(\alpha, \beta, \gamma, \delta)$ model is adequate for fitting the returns of the financial market indices and exchange rate.

Table 4.2: Anderson-Darling goodness-of-fit test for daily returns.

Returns	Test Statistic	p-value
All Share Index	0.6649	0.5888
Banks Index	0.4813	0.7659
Mining Index	0.4025	0.8462
USD/ZAR exchange rate	1.6934	0.1364

4.3.1 VaR and backtesting

Table 4.3 presents the VaR estimates at both the long and short positions for each of the returns series under various confidence levels.

Table 4.3: VaR estimates of financial market indices and exchange rate price returns using fitting stable model.

	VaR Estimates					
	Short position			**Long position**		
Returns	**1%**	**2.5%**	**5%**	**95%**	**97.5%**	**99%**
All Share Index	-0.0317	-0.0212	-0.0158	0.01535	0.0193	0.0264
Banks Index	-0.0466	-0.0324	-0.0249	0.0246	0.0314	0.04367
Mining Index	-0.0491	-0.0344	-0.0266	0.0271	0.0349	0.0497
USDZAR	-0.0301	-0.0209	-0.0160	0.0142	0.0177	0.0232

VaR estimates in Table 4.3 demonstrate the acceptable limits of risk exposure under different market conditions and allow practitioners to mitigate losses.

In Table 4.4, the p-values of the Kupiec likelihood ratio test are presented, providing insights into model validity.

Table 4.4: Kupiec p-values for financial indices and exchange rate returns.

***p*-value of Kupiec Test**						
	Short position			**Long position**		
Returns	**1%**	**2.5%**	**5%**	**95%**	**97.5%**	**99%**
All Share Index	0.0300	0.9464	0.4644	0.7115	0.2849	0.5527
Banks Index	0.5395	0.4765	0.2758	0.2758	0.6544	0.6900
Mining Index	0.5395	0.5665	0.5212	0.9233	0.3440	0.2986
USD/ZAR	0.2456	0.7009	0.4169	0.9118	0.3855	0.6662

According to Table 4.4, at a 5% level of significance, the Kupiec test suggests that the fitted stable model is a suitable fit at almost all VaR probability levels for each of the returns. This conclusion is drawn as the p-values are greater than 0.05; indicating that, the null hypothesis of model adequacy is not rejected. Indicators for model inadequacy may be observed for the FTSE/JSE All Share Index returns at the 1% VaR probability level based on the observed p-value at a 5% level of significance; however, at a 1% significance level the fitted stable model is a fairly good fit.

4.4 Concluding remarks

To summarise, this chapter provided a comprehensive understanding of stable distributions and focuses on stable model definitions, parameterizations, probability density functions as well as special cases of stable distributions. Stable parameter estimation was also discussed extensively with a core focus on the maximum likelihood method. The shortcomings of stable distributions, specifically focusing on financial industry applications, were discussed. The daily log returns of 3 indices and the Dollar/South African exchange rate were analysed using the fitted univariate Nolan's S_0-parameterization stable distribution. The results validate the work by Nolan (2014) where stable distributions are a flexible class of probability laws that effectively capture the characteristics of South African financial data. The estimation of stable parameters is feasible, and diagnostics show that large sets of financial data with heavy tails and skewness are well described by stable distributions, as evidenced by the Anderson-Darling goodness-of-fit test. Furthermore, VaR estimates and VaR in-sample backtesting using the Kupiec likelihood ratio test emphasise the robustness of the fitted stable models. This study notes some challenges with stable distributions which are highlighted as:

Complexity and computational intensity

Research by Nolan (2003) identifies the difficulty with application of stable laws. The most significant setback is the lack of closed-form expressions for the probability density functions, except for three exclusive cases, that is, the Normal, Cauchy and Lévy distributions. Computing stable densities with this setback is numerically difficult and computationally intensive and poses a problem in analytical calculations for practical applications. Without closed-form densities in stable distributions, computationally intensive numerical methods are applied leading to slower and inaccurate VaR calculations. Inaccurate VaR estimates and added uncertainty in backtesting may overestimate or underestimate risk which have negative implications for financial institutions from an operational and regulatory compliance stand point. For financial analysts, careful consideration is required for stable model applications in financial risk management where accuracy, speed and utmost adherence to financial regulations are of paramount importance.

Parameter estimation challenges

Nolan (2014), Kateregga et al. (2017), Robust Analysis Inc. (2013) and work by several notable researchers emphasize the challenges of estimating stable parameter estimates. Stability, skewness, scale and location parameters are the four parameters characterised by stable distributions that are difficult to estimate simultaneously. Stable parameter estimation requires specialised numerical techniques due to the complexities of the underlying mathematics. Traditional estimation methods like the maximum likelihood estimation procedure are computationally intensive and are prone to convergence issues, thus leading to biased or unreliable parameter estimates. Inaccurate parameter estimates result in inaccurate VaR estimates and impact backtesting results that introduce model risk, where the applied VaR estimation procedure is fundamentally flawed. This risk can translate into other forms of financial risk due to risk profile misjudgment and raise concerns about model validity.

Lack of intuitive interpretability

Stable models that are typically not of traditional Gaussian characteristics are challenging to conceptualise intuitively. Those familiar with the concept of the Normal distribution and the central limit theorem find the ideology of stability counter-intuitive, where linear combinations of i.i.d stable random variables are stable and have the same distribution form. Stable distribution parameters do not correspond to mean and variance in the same way as other distributions are represented thus traditional risk measures like standard deviation and VaR are likely to be misleading. As a result, financial industry professionals may struggle with interpretation and communicating results from stable model analysis with audiences not familiar with the nuances of stable distributions.

Heavy tails and extreme risk estimation

Stable distributions are valued in practice for the ability to model heavy tails which is highly important in capturing extreme events, however, this flexible quality potentially introduces the risk of model overfitting in empirical data applications. Specifically, the ability to model a wide range of data often leads to models that capture noise rather than the underlying data trends. Model overfitting has negative implications for reliability and model predictive power and can provide inadequate results with out-of-sample data. Overestimation could result in conservative VaR measures, where, although the likelihood of unexpected losses is reduced, financial institutions need to account for the capital required for the overestimation of risk which is an inefficient use of capital resources.

Limited adoption and acceptance

The adoption of stable distributions is attractive in theory, however, adoption and acceptance in practical applications within industry, specifically the financial industry has been limited. The challenges and complexities highlighted above, advanced mathematical methods and niche knowledge about stable distributions as a subject matter may deter financial practitioners from industry utilisation as analysts and researchers alike, may prefer to adopt simple and familiar models like the Normal distribution. A major factor contributing to the limited adoption is the lack of software available to apply stable distributions. While there are software libraries available like Robust Analysis Inc. (2013) for working with stable models, these are not commonly available and offer advanced features that require a licensing fee. The lack of integration of stable models imply that practitioners rely on less familiar tools or need to develop custom solutions which is a significant hurdle is stable distribution analyses. The absence of industry-standard support in terms of models, software and regulatory acceptance makes it challenging to implement stable distributions in a professional environment.

While stable distributions offer a theoretically strong framework for modelling data that exhibits heavy tails and skewness, the disadvantages need to be considered. The lack of closed-form densities, challenges with parameter estimation, issues with interpretability and the absence of software availability make stable distributions more challenging to apply that traditional approaches. For practitioners, researchers and statistical analysts, these drawbacks require a thoughtful and deliberate approach in stable distribution application, often emphasing the need for specialised knowledge, mathematical tools and software to address the complexities identified with stable distributions.

Chapter 5 covers the GARCH-Stable hybrid model, in which a practical application of stable distributions is conducted.

CHAPTER 5

GARCH-TYPE MODELS WITH STABLE DISTRIBUTION AND GENERALISED PARETO DISTRIBUTION

5.1 Introduction

Generalised autoregressive conditional heteroskedasticity (GARCH) models, stable models and the Generalised Pareto distribution (GPD) are effective for dealing with volatility and heavy-tailed data. The combination of these concepts brings about the development of the GARCH-stable (GARCH-SD) and GARCH-GPD hybrid models. Atsmegiorgis et al. (2016) modelled the Korean Composite Stock Price Index returns using a GARCH (1,1) framework with normal, t, generalised hyperbolic and generalised Pareto distributed errors. A comparative analysis to estimate model performance was carried out using VaR violations and the Kupiec exceedance test. Ilupeju (2016) compared the relative performance of the GARCH-type model combined with various heavy-tailed distributions including the stable and GPD models for the FTSE/JSE ALSI returns. Accurately forecasting and handling financial risk is highly crucial when dealing with the uncertainty of extreme events. Sometimes, traditional models fail to capture the dynamics of extreme market movements. In this chapter the GARCH-GPD hybrid model is analysed to better understand the volatility clustering and heavy-tailed phenomenon in financial markets, particularly the South African financial market. The fitted GARCH-GPD models will then be compared to the fitted GARCH-SD models.

5.2 GARCH framework

The GARCH (1,1) model is suggested for the volatility clustering phenomenon. The GARCH (1,1) model is known for its simplicity and effectiveness in modelling conditional variance in finance and econometrics.

5.2.1 The ARCH model

The Autoregressive conditional heteroskedasticity (ARCH) model introduced by Engle (1982) has been extensively investigated by many researchers. Consider a log-return series w_t as

$$\begin{aligned} w_t &= \mu_t + a_t, \\ a_t &= \sqrt{\sigma_t}\epsilon_t, \end{aligned} \tag{5.1}$$

where ϵ_t is a white noise, $\epsilon_t \sim N(0,1)$. The ARCH(m) process proposed by Engle (1982),

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2. \tag{5.2}$$

where $\alpha_0 > 0, \beta_j \geq 0$ are considered to ensure strictly positive variance. In general, q is of high order because of the prominent volatility clustering phenomenon in financial markets. The unconditional variance is given by,

$$E[a_t^2] = \frac{\alpha_0}{1 - \sum_{i=1}^{max(m,s)} \alpha_i}. \tag{5.3}$$

The process is covariance stationary if and only if the sum of the autoregressive parameters, $\sum_{i=1}^{m} \alpha_i < 1$ (Poon, 2005).

5.2.2 The GARCH model

An extension of the ARCH model is the generalised ARCH (referred to as GARCH) model. For a high-order ARCH(m) process, it is more parsimonious to model volatility as a GARCH(m,s), as noted by Bollerslev (1986). For a log-return series w_t, we define $a_t = w_t - \mu_t$ as the innovation at time t. Thus, a_t follows a GARCH(m,s) model if

$$\begin{aligned} a_t &= \sigma_t\epsilon_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2, \end{aligned} \tag{5.4}$$

where ϵ_t denote a sequence of iid random variables with mean 0 and variance 1. We observe that $\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0$ and $\sum_{i=1}^{max(m,s)}(\alpha_i + \beta_j) < 1$. It is understood that $\alpha_i = 0$ for $i > m$ and $\beta_j = 0$ for $j > s$. The constraint $\alpha_i + \beta_i$ implies that the unconditional variance of a_t is finite whereas the conditional variance σ_t^2 evolves over time. Equation (5.1) above simplifies to an ARCH(m) model if $s = 0$. The α_i and β_j are referred to as ARCH and GARCH parameters, respectively. For the properties of the GARCH model, the following representation is used. Let $\eta_t = a_t^2 - \sigma_t^2$ so $\sigma_t^2 = a_t - \eta_t$. By substituting $\sigma_{t-i}^2 = a_{t-i}^2 - \eta_{t-i}\,(i = 0, \ldots, s)$ into equation (5.1), ee now formulate the GARCH model as follows:

$$a_t^2 = \alpha_0 + \sum_{i=1}^{max(m,s)} (\alpha_i + \beta_i) a_{t-i}^2 + \eta_t - \sum_{j=1}^{s} \beta_j \eta_{t-j}, \tag{5.5}$$

where $E(\eta_t) = 0$ and $\mathrm{Cov}(\eta_t, \eta_{t-j}) = 0$ for $j \geq 1$. Equation (5.2) denotes an ARMA form for the squared series a_t^2. Thus, the GARCH model resembles an ARMA model but with squared series a_t^2. Using the unconditional mean of the ARMA model, we obtain:

$$E[a_t^2] = \frac{\alpha_0}{1 - \sum_{i=1}^{max(m,s)} (\alpha_i + \beta_i)}, \tag{5.6}$$

where the denominator is non-negative.
We examine the simplest form of GARCH models, specifically the GARCH (1,1) model:

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \qquad 0 \leq \alpha_1, \beta_1 \leq 1 \ (\alpha_1 + \beta_1) < 1. \tag{5.7}$$

A large a_{t-1}^2 or σ_{t-1}^2 leads to a large σ_t^2. This indicates that a large a_{t-1}^2 tends to be followed by another large a_t^2 generating volatility clustering. It can be demonstrated that if
$1 - 2\alpha_1^2 - (\alpha_1 + \beta_1)^2 > 0$, then

$$\frac{E(a_t^4)}{\left[E(a_t^2)\right]^2} = \frac{3\left[1 - (\alpha_1 + \beta_1)^2\right]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3. \tag{5.8}$$

Like the ARCH models, the GARCH (1,1) model exhibits heavier tails than that of the normal distribution. For the GARCH (1,1) model, we consider that the forecast origin is h. The one-step ahead forecast is:

$$\sigma_{h+1}^2 = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2,$$

where a_h and σ_h^2 are known at time index h. The one-step ahead forecast is:

$$\sigma_h^2(1) = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2,$$

$a_t^2 = \sigma_t^2 \epsilon_t^2$ is used for multistep forecasts. Therefore, the volatility equation is formulated as:

$$\sigma_{t+1}^2 = \alpha_0 + (\alpha_1 + \beta_1)\sigma_t^2 + \alpha_1 \sigma_t^2 (\epsilon_t^2 - 1),$$

when $t = h + 1$, we have

$$\sigma_{h+1}^2 = \alpha_0 + (\alpha_1 + \beta_1)\sigma_{h+1}^2 + \alpha_1 \sigma_{h+1}^2 (\epsilon_{h+1}^2 - 1).$$

Since $E\left(\epsilon_{h+1}^2 - 1|\mathfrak{F}_h\right) = 0$, the two-step ahead volatility forecast at the origin satisfies the equation

$$\sigma_h^2(2) = \alpha_0 + (\alpha_1 + \beta_1)\sigma_h^2(1).$$

Generally,

$$\sigma_h^2(\ell) = \alpha_0 + (\alpha_1 + \beta_1)\sigma_h^2(\ell - 1), \qquad \ell > 1. \tag{5.9}$$

From equation (5.6), repeated substitutions show that the ℓ-step ahead forecast is:

$$\sigma_h^2(\ell) = \frac{\alpha_0\left[1 - (\alpha_1 + \beta_1)^2\right]}{1 - \alpha_1 - \beta_1} + (\alpha_1 + \beta_1)^{\ell-1}\sigma_h^2(1).$$

Hence,

$$\sigma_h^2(\ell) \to \frac{\alpha_0}{1 - \alpha_1 - \beta_1},\ \ell \to \infty,$$

only if $\alpha_1 + \beta_1 < 1$. In the GARCH (1,1) model, as the forecast horizon tends to infinity, the multistep ahead volatility forecasts converge to the unconditional variance of a_t, assuming $\text{Var}(a_t)$ exists (Tsay, 2005).

Parameter estimation

The GARCH (m, s) is discussed by Yang (2012b) as:

$$\sigma_t = \alpha_0 + \sum_{i=1}^{m} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{s} \beta_i \sigma_{t-i},$$
$$\epsilon_t = v_t\sqrt{\sigma_t},$$

where σ_t defines the conditional variance and v_t is the white noise term.

To estimate the parameters of GARCH models with given k, m and s, we have

$$y_t = C + \sum_{i=1}^{k} a_i y_{t-i} + \epsilon_t, \tag{5.10}$$

$$\epsilon_t = v_t\sqrt{\sigma_t}, \tag{5.11}$$

$$\sigma_t = \alpha_0 + \sum_{i=1}^{m} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{s} \beta_i \sigma_{t-i}, \tag{5.12}$$

where v_t represents the white noise term. ϵ_t follows a normal distribution with mean zero and conditional variance σ_t, i.e.

$$p(\epsilon_t|\epsilon_{t-1}, \ldots, \epsilon_0) = \frac{1}{\sqrt{2\pi\sigma_t}}\exp\left(\frac{-\epsilon_t^2}{2\sigma_t}\right). \tag{5.13}$$

The log-likelihood function of the parameter vector is expressed as:

$$\boldsymbol{\theta} = (\alpha_0, \alpha_1, \cdots, \alpha_s, \beta_1, \cdots, \beta_m)^T \text{ is}$$

$$L(\boldsymbol{\theta}) = \sum_{t=s+1}^{n} l_t(\boldsymbol{\theta}) = \sum_{t=s+1}^{n} \Bigg(-\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\sigma_t - \frac{\epsilon_t^2}{2\sigma_t} \Bigg). \tag{5.14}$$

Therefore,

$$\frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \left(\frac{\epsilon_t^2}{2\sigma_t^2} - \frac{1}{2\sigma_t} \right) \frac{\partial \sigma_t}{\partial \boldsymbol{\theta}}, \tag{5.15}$$

$$\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} = \left(\frac{\epsilon_t^2}{2\sigma_t^2} - \frac{1}{2\sigma_t} \right) \frac{\partial^2 \sigma_t}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} + \left(\frac{1}{2\sigma_t^2} - \frac{\epsilon_t^2}{\sigma_t^2} \right) \frac{\partial \sigma_t}{\partial \boldsymbol{\theta}} \frac{\partial \sigma_t}{\partial \boldsymbol{\theta}^T}, \tag{5.16}$$

where

$$\frac{\partial \sigma_t}{\partial \boldsymbol{\theta}} = \left(1, \epsilon_{t-1}^2, \cdots, \epsilon_{t-s}^2, \sigma_{t-1}, \cdots, \sigma_{t-m}\right)^T + \sum_{i=1}^{m} \beta_i \frac{\partial \sigma_{t-i}}{\partial \boldsymbol{\theta}}. \tag{5.17}$$

The gradient is

$$\nabla L(\boldsymbol{\theta}) = \frac{1}{2} \sum_{t=s+1}^{n} \left(\frac{\epsilon_t^2}{\sigma^2} - \frac{1}{\sigma_t} \right) \frac{\partial \sigma_t}{\partial \boldsymbol{\theta}}, \tag{5.18}$$

The Fisher information matrix is denoted as:

$$\begin{aligned} J &= \sum_{t=s+1}^{n} E\Bigg[\left(\frac{\epsilon_t^2}{2\sigma_t^2} - \frac{1}{2\sigma_t} \right) \frac{\partial^2 \sigma_t}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} + \left(\frac{1}{2\sigma_t^2} - \frac{\epsilon_t^2}{\sigma_t^3} \right) \frac{\partial \sigma_t}{\partial \boldsymbol{\theta}} \frac{\partial \sigma_t}{\partial \boldsymbol{\theta}^T} \Bigg] \\ &= -\frac{1}{2} \sum_{t=s+1}^{n} E\left(\frac{1}{h_t^2} \frac{\partial h_t}{\partial \boldsymbol{\theta}} \frac{\partial h_t}{\partial \boldsymbol{\theta}^T} \right). \end{aligned} \tag{5.19}$$

Consider the GARCH (1,1) model

$$\epsilon_t = v_t \sqrt{\sigma_t}, \tag{5.20}$$

$$\sigma_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}, \tag{5.21}$$

to estimate the coefficients $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1)^T$, where

$$\nabla L(\boldsymbol{\theta}) = \frac{1}{2} \sum_{t=2}^{n} \left(\frac{\epsilon_t^2}{\sigma_t^2} - \frac{1}{\sigma_t} \right) \frac{\partial \sigma_t}{\partial \boldsymbol{\theta}}, \tag{5.22}$$

and

$$J = -\frac{1}{2} \sum_{t=2}^{n} E\left(\frac{1}{\sigma^2} \frac{\partial \sigma_t}{\partial \boldsymbol{\theta}} \frac{\partial \sigma_t}{\partial \boldsymbol{\theta}^T} \right), \tag{5.23}$$

with

$$\frac{\partial \sigma_t}{\partial \boldsymbol{\theta}} = \left(1, \epsilon_{t-1}^2, \sigma_{t-1}\right)^T + \beta_1 \frac{\partial \sigma_{t-1}}{\partial \boldsymbol{\theta}}. \tag{5.24}$$

5.2.3 Exponential generalised autoregressive conditional heteroskedasticity (EGARCH)

The Exponential generalized autoregressive conditional heteroskedasticity (EGARCH) model is a time series model where the conditional volatility of financial returns is modelled. This model allows for asymmetries between positive and negative shocks on the conditional variance, where the volatility clustering phenomenon observed in financial data is captured. The EGARCH model is specified as:

$$\begin{aligned} y_t &= \mu + a_t, \\ \ln \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i \left(|\epsilon_{t-i}| + \gamma_i \epsilon_{t-i}\right) + \sum_{j=1}^{q} \beta_j lin \sigma_{t-j}^2, \end{aligned} \tag{5.25}$$

where y_t is the time series at time t, the mean of the GARCH model is depicted by μ, a_t is the residual at time t, σ_t notes the volatility at time t, the series α_i are parameters of the ARCH component of the model and β_i are the GARCH parameters of the model. The standardized residuals are ϵ_t.

5.2.4 Threshold GARCH (TGARCH)

The threshold GARCH (TGARCH) model was created by Zakoian (1994) and Glosten et al. (1993) to capture the effect of asymmetric volatility for positive and negative shocks. A threshold parameter is used to investigate the different reactions to positive and negative shocks where typically volatility increases more with negative shocks than positive ones.
The TGARCH (p,q) model is given by:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \left(\alpha_i + \gamma_i N_{t-i}\right) \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \tag{5.26}$$

where σ_t^2 is the variance, α_0 is a constant, α_i is the ARCH effect of the model, $\gamma_i N_{t-i}$ represents the asymmetric effect that is multiplied by the effect of a positive or negative shock. When $\epsilon_{t-i} > 0$ then it represents a positive shock, conversely when $\epsilon_{t-i} < 0$ then a negative shock is represented. $\beta_j \sigma_{t-j}^2$ represents the GARCH component of the model.

5.2.5 Asymmetric Power ARCH (APARCH)

The Asymmetric power ARCH (APARCH) model was introduced by Ding et al. (1993) as an extension of GARCH models. Tsay (2014) describes the APARCH model as:

$$\sigma_t^\delta = \alpha_0 + \sum_{i=1}^{p} \alpha_i \left(|\epsilon_{t-i}| \gamma_i \epsilon_{t-i}\right)^d + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^\delta, \tag{5.27}$$

where $\alpha_0, \alpha_i, \beta_j$ is positive. α_i and β_j are the ARCH and GARCH coefficients respectively. γ_i is the leverage coefficient. δ is a symmetric power transformation of σ_t.

This study considers the scenario where $\delta = 1$ and $p = q = 1$; therefore, the volatility equation for the APARCH (1,1) model is represented as:

$$\sigma_t = \alpha_0 + \alpha_i \left(|\epsilon_{t-1}| \gamma_i \epsilon_{t-1}\right) + \beta_1 \sigma_{t-1}.$$

5.3 Generalised Pareto distribution (GPD)

The Generalised Pareto distribution (GPD) is a classic probability model used in EVT where the tails beyond a certain threshold value are modelled.

Coles et al. (2001) assumes $Y_1, \ldots Y_n$ be a sequence of independent and identical random variables, over a suitable threshold u the excess $Y - u$ can be approximated by the GPD model $F(y|u, \sigma_u, \xi)$.
The GPD is characterised by the probability distribution function defined as:

$$F\left(y|u, \sigma_u, \xi\right) = \begin{cases} 1 - \left[1 + \xi \left(\dfrac{y-u}{\sigma_u}\right)\right]_+^{-1/\xi} & \xi \neq 1, \\ 1 - exp\left[-\left(\dfrac{y-u}{\sigma_u}\right)\right]_+ & \xi = 1, \end{cases} \tag{5.28}$$

where $y > u$, σ_u and $\left[1 + \xi \left(\dfrac{y-u}{\sigma_u}\right)\right]$ are greater than 0. σ_u is the scale parameter with threshold u. The shape parameter is determined by the tail distribution of the GPD model and is represented by ξ. $\xi = 0$ refers to an exponential tail, $\xi > 0$ indicates a heavy tail, and $\xi < 0$ is a short tail with a finite end upper point $u - \dfrac{\sigma_u}{\xi}$.

5.3.1 Maximum likelihood estimation of the GPD model

Assume $y_1, \ldots, y_n$ be a sequence of iid random variables which are exceedances above u, the threshold parameter. Let nu be the number of observations above u:

$$logL\left(\sigma_u, \xi, Y\right) = -nu\ log\ \sigma_u - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{nu} log\left[1 + \xi\left(\frac{y_i - u}{\sigma_u}\right)\right], \qquad (5.29)$$

where $\xi \neq 0$ and $\left[1 + \xi\left(\frac{y_i - u}{\sigma_u}\right)\right] > 0$. The expression $\left[1 + \xi\left(\frac{y_i - u}{\sigma_u}\right)\right]$ suggests $y < \mu - \frac{\sigma_u}{\xi}$.

When $\xi = 0$, equation 5.29 is:

$$logL\left(\sigma_u, Y\right) = -nu\ log\ \sigma_u - \sum_{i=1}^{nu}\left(\frac{y_i - u}{\sigma_u}\right). \qquad (5.30)$$

5.3.2 Threshold selection

Hu (2013) notes that threshold selection remains an area of ongoing research in the literature. The tradeoff between variance and bias is imminent in the threshold selection process. Coles et al. (2001) states if the chosen threshold is too low, then there is a violation of the underlying asymptotic arguments of the derived GPD model. Contrastingly, high variance is experienced when the chosen threshold is too high. Therefore, threshold selection becomes an important choice where the model provides an apt approximation as compared to parameter estimate variance. Commonly, various diagnostic plots investigate the fit of the model as well as selecting the threshold. This study mentions three approaches for threshold choice described by Coles et al. (2001). These include the evaluation of the parameter stability plot, mean life residual plot and the model fit diagnostic plots.

5.3.2.1 Parameter stability plot

Assume the excesses over a high threshold u follow the GPD model with parameters ξ and σ_u. If $b > u$, for any higher threshold, then the excesses still follow a GPD model with parameter ξ and a scale parameter of:

$$\sigma_b = \sigma_u + \xi\left(b - u\right).$$

If the scale parameter σ_b is re-parameterized:

$$\sigma^* = \sigma b - \xi b.$$

Given that u is a high threshold, σ^* is independent of b. Work by Hu (2013) suggests that u should ideally be selected where the shape and re-parameterized scale parameter remains constant after sampling variability is considered. When a suitable threshold is reached, the excess over the threshold u follows GPD.

5.3.2.2 Mean residual life plot

Hu (2013) notes if the excess $Y - u$ is approximated by the GPD model, then for a threshold u, the mean excess is described as:

$$E\left(Y - u|Y > u\right) = \frac{\sigma_u}{1-\xi}, \qquad \text{for } \xi < 1.$$

For a higher threshold $b > u$, the mean excess function is defined as:

$$E\left(Y - b|Y > b\right) = \frac{\sigma_b}{1-\xi} = \frac{\sigma_u + \xi\left(b-u\right)}{1-\xi}, \qquad \text{for } \xi < 1.$$

Once a suitable high threshold u is chosen, the mean excesses $E\left(Y - u|Y > u\right)$ is a linear function of u.
The plot points of the sample mean residual life plot is drawn using:

$$\left(u, \frac{1}{n_u}\sum_{i=1}^{n_u}\left(y_i - u\right)\right),$$

where y_i is the observation above the chosen threshold u and n_u is the number of observations above u. The excesses $b > u$ in the mean residual plot changes in a linear fashion with u for a high enough threshold u. When sampling variability is included, the threshold to be selected should be chosen where the relationship of mean excesses of all higher thresholds are linear. Coles et al. (2001) acknowledge the interpretation of the mean residual plot to be a difficult undertaking. The parameter stability plot and the mean residual life plot are frequently used and popular choices for graphically choosing the threshold.

5.3.2.3 Model fit diagnostic plot

Standard statistical model diagnostic plots like the quantile plot, probability plot, return level plot and the empirical versus the fitted density plot can be utilised for not only checking the model fit but also the suitability of the chosen threshold.

5.4 Empirical results

5.4.1 Fitting the GARCH type models

This study examines the optimal GARCH-type model capable of effectively capturing the volatility clustering phenomenon. The ML approach is used to get the GARCH (1,1) model with Normal innovations to the return series data. The ML parameter estimates and associated p-values for the fitted GARCH (1,1) model are shown in Panel A of Table 5.1 and Panel B presents the results of the Ljung-Box and ARCH LM test statistics.

Table 5.1: ML estimates of parameters for the GARCH (1,1) model using normal innovations with corresponding goodness-of-fit statistics for financial stock and exchange rate returns.

Panel A				
Financial stock returns				
Parameter estimates	FTSE/JSE ALSI	FTSE/JSE Banks Index	FTSE/JSE Mining Index	USD/ZAR exchange rate
$\hat{\alpha_0}$	0.0000	0.0000	0.0000	0.0000
$\hat{\alpha_1}$	0.0884	0.1479	0.0583	0.0442
$\hat{\beta}_1$	0.8891	0.7691	0.9289	0.9447
Panel B				
p-values of the corresponding to the parameter estimates or test statistics				
$\hat{\alpha_0}$	0.0466	< 0.0000	0.0549	0.1417
$\hat{\alpha_1}$	< 0.0000	< 0.0000	< 0.0000	< 0.0000
$\hat{\beta}_1$	< 0.0000	< 0.0000	< 0.0000	< 0.0000
Ljung-Box	0.1308	0.8106	0.1455	0.2262
ARCH-LM	0.5060	0.9833	0.1892	0.1726

From Table 5.1, the coefficient β_1 for each stock returns under investigation is significant as the p-value is greater than 0.05. This is related to the measurement of volatility persistence, which suggests that periods of high volatility is followed by periods of low volatility, and vice versa. The p-values of the Ljung-Box test exceed 0.05; therefore, at a 5% level of significance, there is evidence to reject the null hypothesis of serial correlation. This suggests that the GARCH (1,1) model with Normal innovations indicates no serial correlation in the returns of each financial stock index and exchange rate. According to the ARCH LM test confirms the ARCH effects of each return series.

Table 5.2 shows the test statistics and associated p-values for the sign bias test.

Table 5.2: Sign bias test of return series.

Panel A				
Financial stock returns				
Sign-bias estimates	FTSE/JSE ALSI	FTSE/JSE Banks Index	FTSE/JSE Mining Index	USD/ZAR exchange rate
Sign-bias	1.5437	1.1174	0.7800	0.5062
Negative sign-bias	0.1009	0.0020	1.3182	0.6277
Positive sign-bias	2.5654	0.5379	1.1302	1.6774
Joint effect	24.3508	1.2914	4.0531	4.1492
Panel B				
***p*-values of the corresponding to sign-bias estimates**				
Sign-bias	<0.0001	0.2639	0.4355	0.6127
Negative sign-bias	<0.0001	0.9984	0.1876	0.5303
Positive sign-bias	< 0.0001	0.5907	0.2583	0.0936
Joint effect	< 0.0001	0.7312	0.2558	0.2458

The sign bias test indicates p-values greater than 0.05, except for the FTSE/JSE ALSI. The negative and positive sign bias p-values greater than 0.05 suggests that there were significant positive and negative reaction shocks to the returns, except for FTSE/JSE ALSI. The joint effect was notably significant for ALSI, indicating that returns are asymmetrical. Table 5.2 shows that ALSI returns exhibit asymmetry and best captured by asymmetric GARCH-type models.

Table 5.3 shows the parameter estimates for the asymmetric GARCH-type models for the FTSE/JSE ALSI returns.

Table 5.3: ML parameter estimates for asymmetric GARCH-type models with Normal innovations on FTSE/JSE ALSI returns.

Asymmetric GARCH-type model			
Parameter estimates	**EGARCH (1,1)**	**APARCH (1,1)**	**TGARCH (1,1)**
Panel A			
FTSE/JSE ALSI returns			
$\hat{\alpha_0}$	–0.2015	0.0002	0.0000
$\hat{\alpha_1}$	–0.1342	0.0682	0.0000
$\hat{\beta_1}$	0.9783	0.9244	0.9070
$\hat{\gamma_1}$	0.0863	1.0000	–
$\hat{\delta_1}$	–	1.0000	–
$\hat{\eta_{11}}$	–	–	0.1397
Panel B			
***p*-values of the corresponding to the parameter estimates or AIC/BIC**			
$\hat{\alpha_0}$	< 0.0000	< 0.0000	0.0025
$\hat{\alpha_1}$	< 0.0000	< 0.0000	< 0.0000
$\hat{\beta_1}$	< 0.0000	< 0.0000	< 0.0000
$\hat{\gamma_1}$	< 0.0000	< 0.0000	–
$\hat{\delta_1}$	–	< 0.0000	–
$\hat{\eta_{11}}$	–	–	< 0.0000
AIC	–6.5637	–6.5603	–6.5594
BIC	–6.5520	–6.5486	–6.5478

Table 5.3 presents the results of fitting the EGARCH (1,1), APRACH (1,1) and TGARCH (1,1) models to the FTSE/JSE ALSI returns. According to the AIC and BIC values, a GARCH-type model is the best asymmetric model for a set of returns, while the EGARCH (1,1) model is for the FTSE/JSE ALSI returns.

5.4.2 Fitting the hybrid GARCH-type-SD models

Standardised residuals of the best GARCH-type model are extracted, and stable parameters are estimated using ML estimation with Nolan's S_0– parameterization. This approach yields a hybrid GARCH-type-SD model, where the volatility of innovations is governed by stable distributions.

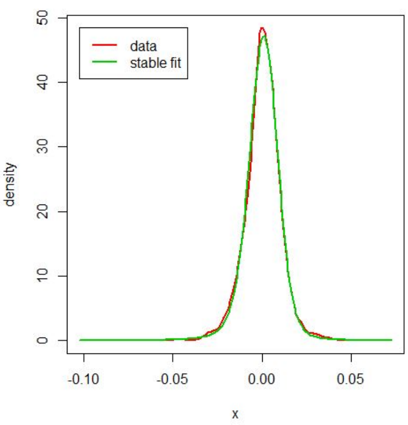
Table 5.4 shows the ML parameter estimates of SD fitted to the standardised residuals derived from the suitable GARCH-type model along with the Anderson Darling goodness-of-fit test statistics and corresonding p-value.

Table 5.4: ML parameter estimates of the hybrid GARCH-type-SD model.

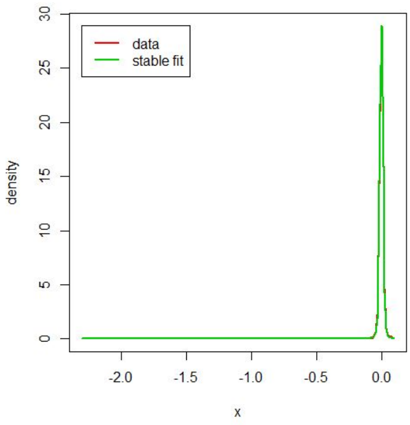
Parameter Estimates	**Financial stock returns**			
	FTSE/JSE ALSI	FTSE/JSE Banks Index	FTSE/JSE Mining Index	USD/ZAR exchange rate
$\hat{\alpha}$	1.9012	1.9052	1.9149	1.8474
$\hat{\beta}$	–0.8719	–0.1299	–0.1567	–0.5224
$\hat{\gamma}$	0.6689	0.6586	0.6714	0.6396
$\hat{\delta}$	0.0834	–0.0102	–0.0028	0.0634
AD test (p-value)	0.8937 (0.4181)	$-^{*}$ (< 0.0001)	0.5147 (0.7319)	1.5819 0.1581)
Note. * No AD statistic available for FTSE/JSE Banks Index				

Stable parameter estimates are presented in Table 5.4, and the Anderson-Darling goodness-of-fit test provides a good fit of the stable model to FTSE/JSE ALSI, FTSE/JSE Mining Index and USD/ZAR exchange rate GARCH residuals. Inconclusive results are obtained for the FTSE/JSE Banks Index, thus leading to an evaluation of the fitted stable density plots.

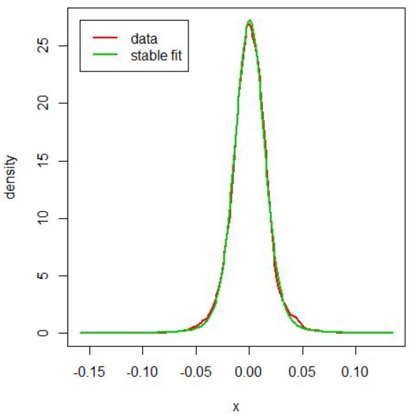
Figure 5.1 shows the stable density plots of fitted GARCH residuals for each return series.



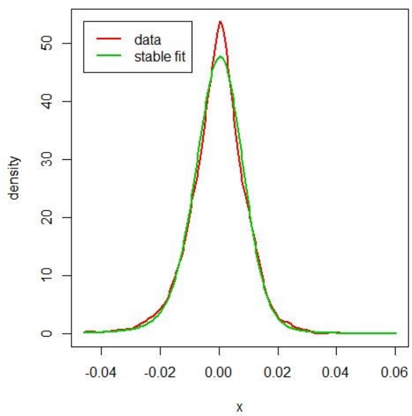
(a) FTSE/JSE ALSI



(b) FTSE/JSE Banks Index



(c) FTSE/JSE Mining Index



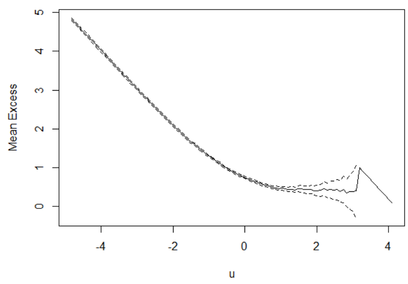
(d) USD/ZAR

Figure 5.1: Stable density plots of fitted GARCH residuals for each stock market indices and exchange rate return series.

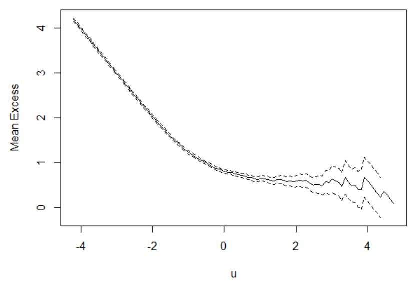
Stable density plots in Figure 5.1 demonstrate that the estimated univariate $S_0(\alpha, \beta, \gamma, \delta)$ model is effective in describing the residuals obtained through the GARCH-type model. This is evidenced by the good fit observed between the stable fitted model and the return series data.

5.4.3 Fitting the hybrid-GARCH-type GPD models

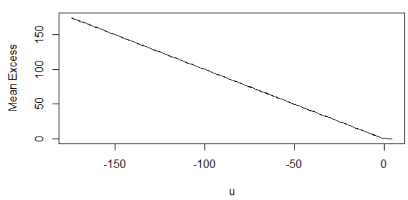
In the previous section, a GARCH-type-SD model was fitted to the return series. In this section, the standardised residuals was extracted from the GARCH-type model, and the GPD was fitted by the ML estimation approach. The new model is combines a GARCH-type model and GPD, which is referred to as a hybrid GARCH-type-GPD model. The GPD model is applied to the upper (gains) and lower tails (losses). Thresholds for the model are determined using the mean residual life plot and the parameter stability plot. Additionally, the Pareto quantile plot will confirm the threshold. Figure 5.2 presents the mean residual life plots, and Figure 5.3 shows the Pareto quantile plots for the return series data evaluated in this study.



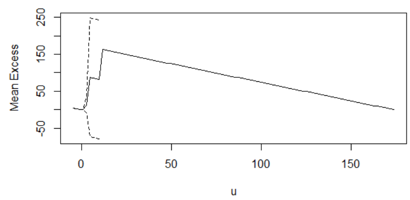
(a) FTSE/JSE ALSI positive standardised residuals



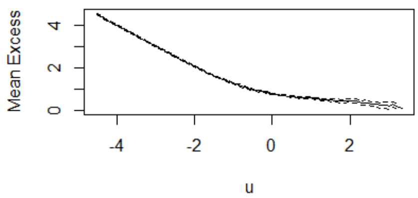
(b) FTSE/JSE ALSI negative standardised residuals



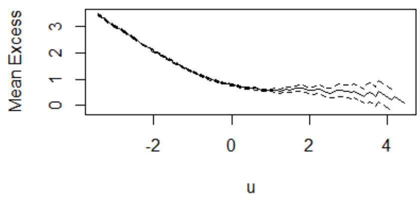
(c) FTSE/JSE Banks Index positive standardised residuals



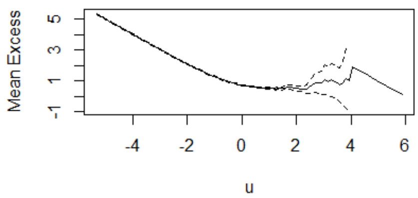
(d) FTSE/JSE Banks Index negative standardised residuals



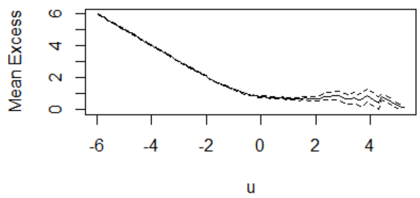
(e) FTSE/JSE Mining Index positive standardised residuals



(f) FTSE/JSE Mining Index negative standardised residuals

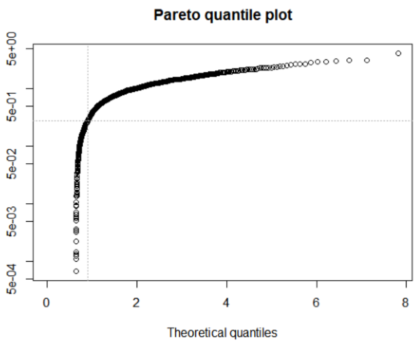


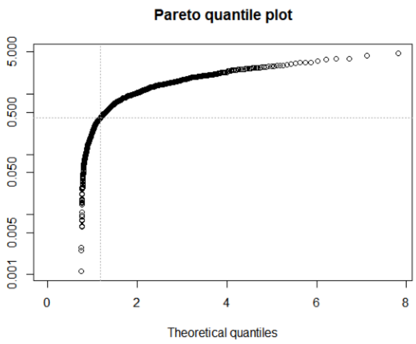
(g) USD/ZAR exchange rate positive standardised residuals



(h) USD/ZAR exchange rate negative standardised residuals

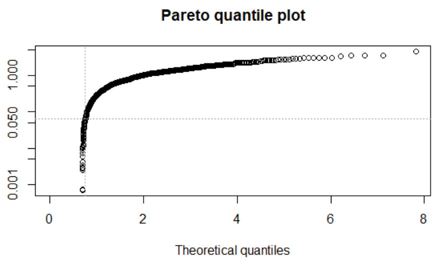
Figure 5.2: Mean residual life plots of fitted GARCH residuals for each stock market indices and exchange rate return series.

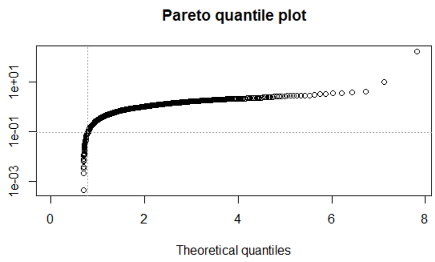




(a) FTSE/JSE ALSI positive standardised residuals

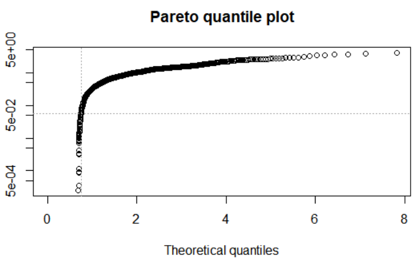
(b) FTSE/JSE ALSI negative standardised residuals

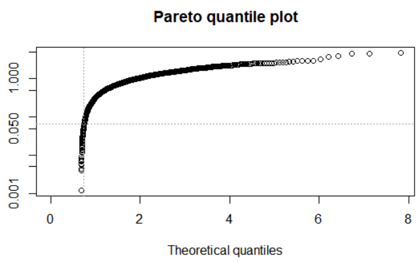




(c) FTSE/JSE Bank Index positive standardised residuals

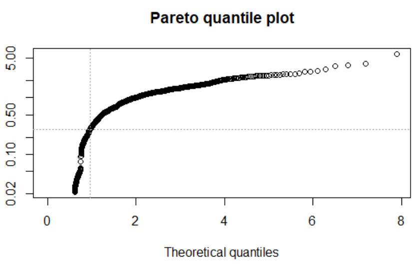
(d) FTSE/JSE Bank Index negative standardised residuals

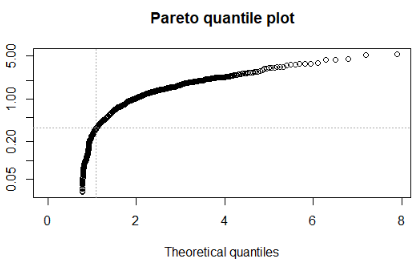




(e) FTSE/JSE Mining Index negative standardised residuals

(f) FTSE/JSE Mining Index negative standardised residuals





t

(g) USD/ZAR exchange rate positive standardised residuals

(h) USD/ZAR exchange rate negative standardised residuals

Figure 5.3: Pareto quantile plots of fitted GARCH residuals for each stock market indices and exchange rate return series.

Table 5.5 shows the GPD residual parameters of the GARCH-type GPD model.

Table 5.5: ML parameter estimates of hybrid GARCH-type-GPD model.

	Financial stock and exchange rate returns							
	FTSE/JSE ALSI		FTSE/JSE Banks Index		FTSE/JSE Mining Index		USD/ZAR exchange rate	
Tail	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower
Threshold (u)	1.3200	1.500	1.0600	1.0900	1.0600	1.0600	1.3100	1.4000
Number of exceedances	188	172	313	313	306	334	200	211
$\hat{\epsilon}$	0.4499	0.6640	0.5282	0.5019	0.5752	0.6278	0.4558	0.6064
SE($\hat{\epsilon}$)	0.0456	0.0699	0.0445	0.0371	0.0469	0.0448	0.0469	0.0630
$\hat{\sigma}$	–0.0172	–0.0687	0.0007	0.2584	–0.0370	–0.0491	0.0905	0.0649
SE($\hat{\sigma}$)	0.0706	0.0728	0.0624	0.0497	0.0582	0.0460	0.0750	0.0779

In the next section, Value-at-Risk estimates and backtesting procedures are used to compare the fitted GARCH-type hybrid model in both SD and GPD cases.

5.4.4 VaR estimation and backtesting

VaR values are calculated at various levels. Table 5.6 shows VaR estimates at different VaR probability levels of the suitable hybrid GARCH-GPD and GARCH-SD models for the financial market and USD/ZAR exchange rate returns. Table 5.6 shows that the VaR models resulting from VaR estimates are within the same ranges for both the short and long positions. The VaR estimates from the fitted hybrid GARCH-type models are backtested using the Kupiec likelihood ratio test. The corresponding p-values of the Kupiec likelihood ratio test at different VaR levels are summarized in Table 5.7. The fitted GARCH-GPD and GARCH-SD models are highly suitable for the observed return series in this study across nearly all VaR levels.

Table 5.6: VaR estimates for financial market indices and exchange rate returns using the fitted hybrid GARCH-GPD and GARCH-SD models.

Financial stock returns	Hybrid Model	VaR estimates at different levels			
		Short position		Long position	
		2.5%	**5%**	**95%**	**97.5%**
FTSE/JSE ALSI	EGARCH (1,1)–GPD	1.8110	1.5032	1.7099	2.1495
	EGARCH (1,1)–SD	–2.0928	–1.6655	1.5604	1.8478
FTSE/JSE Banks Index	GARCH (1,1)–GPD	1.9116	1.5451	1.61010	2.0931
	GARCH (1,1)–SD	–1.9652	–1.6100	1.5513	1.8868
FTSE/JSE Mining Index	GARCH (1,1)–GPD	1.9476	1.5668	1.6627	2.0703
	GARCH (1,1)–SD	–1.9859	–1.6295	1.5818	1.9170
USD/ZAR exchange rate	GARCH (1,1)–GPD	1.8350	1.4968	1.6806	2.1234
	GARCH (1,1)–SD	–2.0848	–1.6355	1.5124	1.8219

Table 5.7: p-values of the Kupiec likelihood ratio test for financial indices and exchange rate returns.

Financial stock returns	Hybrid Model	VaR estimates at different levels			
		Short position		Long position	
		2.5%	5%	95%	97.5%
FTSE/JSE ALSI	EGARCH (1,1)–GPD	0.7479	0.6851	0.9304	0.5618
	EGARCH (1,1)–SD	0.9464	0.5212	0.0480	0.4766
FTSE/JSE Banks Index	GARCH (1,1)–GPD	0.9514	0.5527	0.3170	0.3985
	GARCH (1,1)–SD	0.2849	0.3170	0.7156	0.5665
FTSE/JSE Mining Index	GARCH (1,1)–GPD	0.7496	0.3288	0.9963	0.5665
	GARCH (1,1)–SD	0.2849	0.5212	0.0600	0.9514
USD/ZAR exchange rate	GARCH (1,1)–GPD	0.8156	0.5356	0.9823	0.8895
	GARCH (1,1)–SD	0.4547	0.4685	0.6717	0.8156

Since the observed p-values exceed 0.05, we accept the null hypothesis of model adequacy. The GARCH-GPD model indicates model inadequacy for the 99% VaR level for the FTSE/JSE Banks Index returns, whereas the fitted GARCH-SD models show model inadequacy for the FTSE/JSE Mining Index returns at the 99% VaR level, considering a 5% level of significance. However, at a 1% significance level, the robustness of the fitted GARCH-SD models is highlighted across all VaR levels, whereas the fitted GARCH-GPD still shows model inadequacy at the 99% VaR level. The most robust VaR model is summarised at each VaR level. Table 5.8 presents the most appropriate hybrid GARCH-type model selected at different VaR levels for the returns of the financial indices and USD/ZAR exchange rate returns.

Table 5.8: Most suitable hybrid GARCH type model chosen for financial indices and USD/ZAR exchange rate returns across different VaR levels.

Financial stock returns	Selected hybrid GARCH type models at different VaR levels			
	Short position		Long position	
	2.5%	5%	95%	97.5%
FTSE/JSE ALSI	EGARCH (1,1)-SD	EGARCH (1,1)-GPD	EGARCH (1,1)-GPD	EGARCH (1,1)-GPD
FTSE/JSE Banks Index	GARCH (1,1)-GPD	GARCH (1,1)-GPD	GARCH (1,1)-SD	GARCH (1,1)-SD
FTSE/JSE Mining Index	GARCH (1,1)-GPD	GARCH (1,1)-SD	GARCH (1,1)-GPD	GARCH (1,1)-SD
USD/ZAR exchange rate	GARCH (1,1)-GPD	GARCH (1,1)-GPD	GARCH (1,1)-GPD	GARCH (1,1)-GPD

Table 5.8 shows that there is no hybrid model that outperforms another across all VaR levels for all financial stock returns. For FTSE/JSE ALSI, the EGARCH (1,1)-GPD model is the most robust VaR model at the 5% short position and at all long positions. For the FTSE/JSE Banks Index, GARCH (1,1)-SD is the robust model at all levels of the long position, and the GARCH (1,1)-GPD model is the robust model at all levels of the short position. For the FTSE/JSE Mining Index, the GARCH (1,1)-GPD is the best VaR model at the 2.5% VaR level at short position and 5% VaR level at long position. At the 5% VaR level short position and 2.5% VaR level at long position, the GARCH (1,1)-SD is the most appropriate VaR model for the FTSE/JSE Mining

index. The GARCH (1,1)-GPD model performs well at all VaR levels for USD/ZAR exchange rate returns.

Preliminary tests were undertaken to examine the characteristics of each return series and to highlight the stylised properties of financial data such as volatility clustering and heteroskedasticity. In this study, each FTSE/JSE indices and USD/ZAR exchange rate exhibited leptokurtic behaviour and stationarity; therefore, the hypothesis of normality is rejected. Serial correlation is observed in the FTSE/JSE ALSI Index and FTSE/JSE Mining Index, contradicting the assumption of no serial correlation and heteroskedasticity when fitting an appropriate statistical distribution. However, according to McNeil et al. (2015), some financial return series exhibit serial correlation. Empirical evidence of financial data properties, including ARCH effects in each return series, suggests employing a GARCH-type framework with heavy-tailed distributions such as the GPD and Nolan's S_0-parameterization stable distribution (SD). The suitability of the GARCH (1,1) model is confirmed for each return series, except for FTSE/JSE ALSI, where asymmetric effects are evident. The EGARCH (1,1) model was found to best capture the volatility present in FTSE/JSE ALSI returns. The GDP model applied to the tail distribution, has shown to capture tail asymmetry consistent with findings from McNeil and Frey (2000), which uses a conditional method for estimating tail innovations of a GARCH-type model using a heavy-tailed distributions recommended by extreme value theory (EVT). The stable density plots and model diagnostics indicate a good fit of the residuals for each financial return series, aligning with Nolan (2003). This suggests a robust approach to modelling the tail behavior of fitted GARCH-type models. The fitted GARCH-type SD model demonstrates superior performance compared to the fitted GARCH-type GPD model at various VaR levels using the Kupiec likelihood backtesting procedure, particularly in the upper tail (gains) for the FTSE/JSE indices. This highlights the robustness of the fitted hybrid GARCH-type SD model when assessing tail innovations in South African financial data.

5.5 Concluding remarks

Each return series was fitted with a GARCH-type model, with the FTSE/JSE ALSI returns being best described by the asymmetric EGARCH (1,1) model. The standardised residuals are drawn out and fitted with the Nolan's S_0-parameterization model. A graphical evaluation of the stable density plots shows a close fit between the fitted model and the actual values of the data return series for each stock market indices and exchange rate. The GARCH-type-SD model combines the strength of GARCH models and the stable model's favourable characteristics of dealing with data that deviates largely or experiences extremes, providing a nuanced approach to understanding market-driven fluctuations. The hybrid GARCH-SD model has shown to be an advantageous model for capturing the volatility clustering phenomenon in South African financial markets. This chapter also describes the GPD model where the threshold selection and parameter estimation are evaluated. The hybrid GARCH type-SD model and the hybrid GARCH type-GPD were fitted to the return series where model diagnostics show that both the GARCH (1,1)-GPD and GARCH (1,1)-SD models provide a parsimonious fit to the South African financial data. A VaR analysis was conducted where VaR estimates were calculated, and the Kupiec likelihood backtesting procedure was applied to both fitted models to assess model performance. VaR estimates and VaR in-sample backtesting validate the robustness of the fitted GARCH (1,1)-SD and GARCH (1,1)-GPD model in both long and short positions. However, from both models, neither outperforms one another. This study highlights some challenges with the GARCH-type hybrid models as:

Model complexity and estimation challenges

Hybrid GARCH-type SD models add a stable distribution component to the innovations of a fitted GARCH model. These hybrid models are criticised for the increased complexity. The main issue is the lack of closed-form densities for stable distributions which complicates the parameter estimation process. Computational inefficiencies can arise due to the estimation of additional parameters. This complexity can render models less practical for real-time financial applications, as the demand for more advanced estimation techniques is required and longer processing times. Similarly with GARCH-type GPD hybrid models, computational issues arise as there is a need to estimate additional GPD parameters to the fitted GARCH-type model.

Assumptions and model limitations

Model assumptions of the GARCH-type SD models often are not in alignment with real world data. The characteristics of heavy tails and skewness may not be appropriate for all financial time series. The heavy-tailed nature of stable distributions can lead to overestimation of risk or misrepresentation of the underlying volatility process. Additionally, assuming a constant tail index over time may fail to account for the dynamic changes in financial markets. The choice of threshold for the GPD can greatly impact the model performance, and selecting an inappropriate threshold may result in misleading conclusions about tail risks. Moreover, the GPD assumption of independence among extreme values might not accurately capture the reality observed in financial markets, where extreme events can be clustered or influenced by shared factors.

Risk of overfitting

The GARCH-type hybrid SD and GPD models at likely to introduce the risk of overfitting. Model overfitting can result in poor predictive performance and unreliable risk forecasts when applied to new data. The overfitting risk highlights the need for rigorous validation and testing to ensure that the model provides reliable forecasts and risk assessments.

Data sensitivity and practical implementation

The practical implementation of GARCH-type SD models can be challenging due to the sensitivity of stable parameters and assumptions. Small variations in these parameters can lead to significant changes in model outputs, making the models less robust in practical applications. Additionally, the need for specialised parameter estimation techniques and statistical software can limit usability in real-world financial settings. Lastly, although GARCH models with GPD models are intended to effectively capture tail risks, there is excessive emphasis on extreme events, leading to overly conservative risk estimates. This focus on rare occurrences can result in a risk management approach that is overly cautious and may not accurately reflect the actual risk profile of financial portfolios.

In financial modelling, while GARCH-type SD and GARCH-type GPD hybrid models provide advanced techniques for capturing volatility and extreme value behaviour, such models are not without criticisms. The challenges related to model complexity, assumptions, overfitting, and practical implementation must be meticulously addressed to ensure that these models deliver valuable insights and precise risk assessments. Tackling these issues effectively requires a balanced approach that combines robust estimation techniques, thorough validation, and a comprehensive understanding of market dynamics. In conclusion, the GARCH model paired with the stable distribution or GPD as a hybrid model suggests a powerful approach in financial econometrics to attain data-driven insights for the South African financial market and in Chapter 6 various mixture models are explored including models with a stable distribution component.

CHAPTER 6

EXTREME MIXTURE MODELS

6.1 Introduction

In this chapter, extreme value theory is integrated with mixture modelling as an alternative approach to understanding data complexities. Extreme mixture modelling contributes to understanding not only extreme events but also the modelling of data between those extremes, that is, the bulk model. Noteworthy work by Zhao (2010) and Zhao et al. (2010) proposed an extreme value mixture model that fits the GPD model at the upper and lower tails, and the Normal distribution is fitted as the bulk model between the two tails. The mixture model proposed by Zhao (2010) and Zhao et al. (2010) is an extension of literature by McNeil and Frey (2000) where a two-stage model is developed where, at the initial stage, a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model is fitted to capture volatility clustering and the latter stage fits a Generalized Pareto Distribution (GPD) to the tails. This chapter builds on the work by Zhao (2010) by considering stable distributions to extreme mixture modelling in South African financial data.

6.2 Mixture models

With statistical model development, there is often the challenge of depicting data that may have a varied combination of underlying descriptors. A powerful way to tackle such complexities is to harness mixture models by making an assumption that any given data set may be described by a mixture of probability distributions. This chapter presents sophisticated frameworks, namely the GPD-Normal-GPD, GPD-KDE-GPD, GPD-Stable-GPD, Stable-Normal-Stable and the Stable-KDE-Stable model for dealing with skewness and the heavy-tailed phenomenon in data.

6.2.1 GPD-Normal-GPD (GNG)

The work of De Melo Mendes and Lopes (2004) inspired Zhao (2010) to consider the GPD-Normal-GPD mixture model, with the lower and upper tail modelled by the GPD model to capture both low and high extreme values. The center of the model distribution or bulk model is a fitted Normal distribution.

Zhao (2010) shows the GNG distribution function as:

$$G\left(f|\boldsymbol{\theta}\right)=\begin{cases}\phi\left(u_l|u,\beta\right)\left[1-G\left(-y|\eta_l,\sigma_l,-u_l\right)\right], & y\leq u_l,\\ \phi\left(u_l|u,\beta\right), & u_l<y<u_r,\\ \phi\left(u_r|u,\beta\right)+1-\phi\left(u_r|u,\beta\right)G\left(-y|\eta_r,\sigma_r,-u_r\right), & y\geq u_r,\end{cases}\tag{6.1}$$

where, $\boldsymbol{\theta}=\left(u_l,\sigma_{u_l},\eta_l,\mu,\beta,u_r,\sigma_{u_r},\eta_r\right), G\left(\cdot|\eta,\sigma,u\right)$ are the GPD distribution function for the upper (shown by subscript r) and lower tail (shown by subscript l), where η denotes the shape parameter and σ is the scale parameter and the threshold expressed as u. $\phi\left(\cdot|\mu,\beta\right)$ is the Normal distribution with mean μ and β as standard deviation.

This study considers the parameterized tail fraction approach specified by Hu (2013).

Parameterised Tail Fraction Approach

Define $\phi_{u_l}=P\left(Y<\ u_l\right)$ and $\phi_{u_r}=P\left(Y>\ u_r\right)$ where the distribution function is defined as:

$$G\left(f|\boldsymbol{\theta}\right)=\begin{cases}\phi_{u_l}\left[1-G\left(-y|\eta_l,\sigma_l,-u_l\right)\right], & y\leq u_l,\\ \phi_{u_l}\ +\left(1-\phi_{u_l}-\phi_{u_r}\right)\frac{\phi(y|\mu,\beta)-\phi(u_l|\mu,\beta)}{\phi(u_r|\mu,\beta)-\phi(u_l|\mu,\beta)}, & u_l<y<u_r,\\ \left(1-\phi_{u_r}\right)+\phi_{u_r}G\left(y|\eta_r,\sigma_r,-u_r\right), & y\geq u_r,\end{cases}\tag{6.2}$$

where, $\boldsymbol{\theta}=\left(u_l,\sigma_{u_l},\varepsilon_l,\mu,\beta,u_r,\sigma_{u_r},\eta_r\right)$.

GNG parameter estimation

This work makes use of the evmix R package by Hu and Scarrott (2018) along with the evir and ismev packages. In this study the evmix package estimates GNG and GKG mixed model parameters where the threshold is treated as a parameter and is also estimated.

The GNG parameter vector is described by $\boldsymbol{\theta}\ =\ \left(\mu,\beta,u_l,\sigma_{u_l},\eta_l,u_r,\sigma_{u_r},\eta_r\right)$ where:

- μ represents the normal mean,
- β represents the normal standard deviation,
- u_l and u_r are the lower and upper tail thresholds respectively,

- σ_{u_l} and σ_{u_r} denote the lower tail and upper tail GPD scale parameter respectively,
- η_l and η_r specify the lower and upper tail GDP shape parameter.

6.2.2 Univariate kernel density estimator

Hu (2013) describes the traditional kernel density estimator as:

$$\hat{f}\,(y; y, \lambda) = \frac{1}{n\lambda} \sum_{i=1}^{n} K((y - y_i)/\lambda), \tag{6.3}$$

where $K(\boldsymbol{.})$ is the kernel density function which is usually symmetric and λ is the bandwidth parameter. $K(y) \geq 0$ and $\int K(y)\ dy = 1$ are two conditions that are satisfied by the kernel function.

Scale notation $K_\lambda(y) = \lambda^{-1} K\left(\frac{y}{\lambda}\right)$ is used to denote another formula for the kernel function:

$$\hat{f}\,(y; y, \lambda) = \frac{1}{n} \sum_{i=1}^{n} K_\lambda\,(y - y_i)\,. \tag{6.4}$$

An interpretation of the kernel estimate at a point y by Wand and Jones (1995) is the average of n kernel ordinates at that specific point. Uniform, normal and bi-weight among other functions are suggested by literature for kernel functions. The Gaussian probability density function is a popular choice for the kernel function, where the bandwidth λ takes the role of standard deviation and controls the spread of the kernel. Wand and Jones (1995) also note that the choice of the kernel function is not as critical as the choice of the bandwidth.

6.2.3 Selecting the bandwidth

Various methods for choosing the bandwidth λ are discussed in the literature. The simplest way to select the bandwidth parameter λ is to plot various values and choose the parameter value that best suits the underlying density. The other possible method to select the bandwidth parameter is to minimise an error criterion, for example, the Mean Square Error (MSE).

6.2.4 GPD-KDE-GPD

This study also assesses the adequacy of the GPD-KDE-GPD extreme value mixture model to South African financial data. The KDE is the kernel density estimate for the central bulk model between a lower and upper threshold with the GPD model applied beyond these thresholds.

The evmix package in R by Hu and Scarrott (2018) describes the cumulative distribution function (CDF) with three components. The lower tail model with trail fraction ϕ_{ul} is described by the KDE model up and till the lower threshold $y < u_l$:

$$F(y) = H(u_l)\left[1 - G_l(y)\right], \tag{6.5}$$

where, $H(y)$ represents the kernel density estimator of the CDF and $G_l(Y)$ is the conditional GPD with negated threshold and y value.

Between the thresholds $u_l \leq y \leq u_r$, the KDE bulk model is defined as:

$$F(y) = H(y). \tag{6.6}$$

Above the upper threshold u_r, that is, where $y > u_r$ is the conditional GPD denoted as:

$$F(y) = H(u_r) + \left[1 - H(u_r)\right] G_r(y), \tag{6.7}$$

where $G_r(y)$ is the CDF of the GPD tail model.

6.2.5 GPD-Stable-GPD (GSG)

In this study, the distribution function of the GSG model is derived in a similar approach to that of the GNG model by Zhao (2010) using Nolan's S_0-parameterization for the bulk stable model component. The GSG distribution function is described as:

$$G(f|\boldsymbol{\theta}) = \begin{cases} \phi(u_l|u,\beta)\left[1 - G(-y|\eta_l,\sigma_l,-u_l)\right], & y \leq u_l, \\ \phi(u_l|\alpha,\beta,\delta,\gamma), & u_l < y < u_r, \\ \phi(u_r|u,\beta) + 1 - \phi(u_r|u,\beta)\, G(-y|\eta_r,\sigma_r,-u_r), & y \geq u_r, \end{cases} \tag{6.8}$$

where, $\boldsymbol{\theta} = (u_l, \sigma_{u_l}, \eta_l, \alpha, \beta, \gamma, \delta, u_r, \sigma_{u_r}, \eta_r), G(\cdot|\eta,\sigma,u)$ are the GPD distribution function for the upper (shown by subscript r) and lower tail (shown by subscript l) with η the shape parameter and σ the scale parameter and threshold u. $\phi(\cdot|\mu,\beta)$ is the stable distribution function of S_0-parameterization with parameters α, β, γ and δ.

GSG parameter estimation

This study fits the GSG model in a piecewise manner. Each of the return data series is spliced where the GPD distribution is fitted at the tails and the S_0-parameterization stable distribution is fitted as the bulk model. The GSG parameter vector is described by $\boldsymbol{\theta}$ where,

- α, β, γ and δ denote stability, skewness, scale and location parameters, respectively, for the stable bulk distribution,
- u_l and u_r are the lower and upper tail thresholds respectively,
- σ_{u_l} and σ_{u_r} denote the lower tail and upper tail GPD scale parameter respectively,
- η_l and η_r specify the lower and upper tail GDP shape parameter.

6.2.6 Stable Normal Stable (SNS)

This study also explores the fit of a newly proposed Stable-Normal-Stable (SNS) model using Nolan's S_0-parameterization on the South African currency and stocks. The SNS mixture model is a two-tailed model where the Normal distribution is fitted as the bulk model, and the S_0-parameterization stable distribution is fitted to the tails. The SNS model is described as:

$$F(y|\boldsymbol{\theta}) = \begin{cases} \phi(u_l|\mu,\sigma)\left[1 - G(-y|\alpha_l,\beta_l,\gamma_l,\delta_l)\right], & y \le u_l; \\ \phi(u_l|\mu,\sigma), & u_l < y < u_r; \\ \phi(u_r|\mu,\sigma) + \left[1 - \phi(u_r|\mu,\sigma)\, G(-y|\alpha_r,\beta_r,\gamma_r,\delta_r)\right], & y \ge u_r, \end{cases} \quad (6.9)$$

where, $\boldsymbol{\theta} = (\alpha_l,\beta_l,\gamma_l,\delta_l,\mu,\sigma,\alpha_r,\beta_r,\gamma_r), G(\text{.}|\alpha,\beta,\gamma,\delta,u)$ are the fitted stable distribution function for the upper (shown by subscript r) and lower tail (shown by subscript l) with ϵ, the four stable parameters and threshold $u.\phi(\text{.}|\mu,\sigma)$ is the normal distribution with mean μ and σ as standard deviation.

6.2.7 SNS parameter estimation

The SNS parameter vector is decribed by $\boldsymbol{\theta}$ where,

- μ represents the Normal mean,
- σ represents the Normal standard deviation,
- $\alpha_l, \beta_l, \gamma_l$ and δ_l denote stability, skewness, scale and location parameters, respectively, for the lower tail model,
- $\alpha_r, \beta_r, \gamma_r$ and δ_r denote stability, skewness, scale and location parameters, respectively, for the upper tail model.

In this study, a piecewise method of estimation is used to create the SNS model. The basic principle behind the piecewise method is the assumption that data follows different distributions over its range; therefore, the mixture model should be modelled in "pieces". Huang et al. (2017) implemented a piecewise mixture model instead of a single parametric approach to improve the accelerated evaluation of automated vehicles with results that confirmed the accuracy and efficiency of the piecewise mixture distribution. Also, to the best of our knowledge, analytic software that makes use of stable models at the tails of extreme value mixture models has not been explored before. In this context, the thresholds need to be pre-determined before proceeding to fit the respective stable tail and normal bulk model.

6.2.8 Stable-KDE-Stable (SKS) mixture model

A two-tailed Kernel Stable model, namely, the Stable-KDE-Stable model, is evaluated in this study. MacDonald et al. (2011) proposed a two-tailed model by joining a standard kernel density estimator between two extreme value GPD tail models. Hu (2013) describes the two-tailed model introduced by MacDonald et al. (2011) as:

$$F\left(y|\boldsymbol{\theta}\right)=\begin{cases}\phi_{u_l}\left[1-G\left(-y|\epsilon_l,\sigma_{u_l},-u_{u_l}\right)\right], & y\leq u_l;\\ \phi_{u_l}+\left(1-\phi_{u_l}-\phi_{u_r}\right)\dfrac{H\left(y|Y,\lambda\right)-H\left(u_l|Y,\lambda\right)}{H\left(u_r|Y,\lambda\right)-H\left(u_l|Y,\lambda\right)}, & u_l<y<u_r;\\ \left(1-\phi_{u_r}\right)+\phi_{u_r}G\left(y|\epsilon_r,\sigma_{u_r},u_r\right), & y\geq u_r,\end{cases} \tag{6.10}$$

where $\boldsymbol{\theta}=(Y,\lambda,u_l,\sigma_{u_l},\epsilon_l,\phi_{u_l},u_2,\sigma_{u_r},\epsilon_r,\sigma_{u_r})$ is the parameter vector. ϕ_{u_l} and σ_{u_r} are estimated as sample proportions less than the lower threshold u_l and above the upper threshold u_r. $G\left(-y|\epsilon_l,\sigma_{u_l},-u_{u_l}\right)$ and $G\left(y|\epsilon_r,\sigma_{u_r},u_r\right)$ represents the unconditional GPD function for $y<u_l$ and $y>u_r$ respectively. $H(\boldsymbol{.}|y,\lambda)$ is the distribution function of the kernel density estimator where λ is the bandwidth.

6.2.9 SKS parameter estimation

In this research, a similar modelling approach as the SNS mixture model is applied to the SKS mixture model. The notable difference is in the central bulk distribution which is the fitted KDE, with bandwidth parameter λ. The parameter vector is expressed as $\boldsymbol{\theta}=(\alpha_l,\beta_l,\gamma_l,\delta_l,\lambda,\alpha_r,\beta_r,\gamma_r,\delta_r)$.

The rationale behind choosing the SNS and SKS mixture models is primarily because of the flexible characteristics of the S_0-parameterization stable model. Stable distributions are known to have a robust fit to the underlying data and, therefore, an apt choice when considering the tail component of EVT mixture models. With this, the prediction of extreme events is likely to be more accurate than traditional EVT modelling approaches.

6.3 Empirical results

Each stock market indices and exchange rate return series will be fitted with the mentioned mixture models to investigate and evaluate the performance of models that are a combination of several probability distributions. This will allow for an in-depth look at both the bulk and tail components of each mixture model. It is noted that mixture models provide flexibility when commencing model building in data analysis. The separate parameters of each component that make up the mixture distribution can shed light on some valuable insights within the data return series.

6.3.1 Fitting the GNG and GKG models

This work explores the evmix package available in R by Hu (2013) as well as the evir and ismev packages. This section uses the evmix package to estimate the GNG and GKG mixed model parameters, where the threshold is considered a parameter and is also estimated.

Table 6.1 provides the estimates of the GNG model using the evmix package in R for the FTSE ALSI, FTSE Banks Index, FTSE Mining Index and United States American Dollar to South African Rand exchange rate.

Table 6.1: ML parameter estimates with corresponding standard errors in prentice of the GNG mixture model using the evmix package in R.

	Financial Stock return			
GNG parameter estimates	FTSE/JSE ALSI	FTSE/JSE Banks Index	FTSE/JSE Mining Index	USD/ZAR
$\hat{\mu}$	0.0007 (0.0002)	0,0003 (0.0004)	0,0003 (0.0003)	-0,0001 (NULL)
$\hat{\beta}$	0.0084 (0.0002)	0,0138 (0.0005)	0,0149 (0.0003)	0,0086 (NULL)
$\hat{u}_l$	-0,0033 (0.0022)	-0,0151 (0.0035)	0,0110 (0.0008)	0,0086 (NULL)
$\hat{\sigma}_l$	0.0070 (0.0003)	0,0087 (0.0014)	0,0096 (0.0006)	0,0101 (NULL)
$\hat{\eta}_l$	0,0600 (0.0315)	0,3304 (0.0598)	0,1829 (0.0571)	0,0069 (NULL)
$\hat{u}_l$	0.0106 (0.0012)	0,0040 (0.0026)	0,0212 (0.0009)	0,0170 (NULL)
$\hat{\sigma}_r$	0.0054 (0.0005)	0,0106 (0.0005)	0,0116 (0.0012)	0,0115 (NULL)
$\hat{\eta}_r$	0.1473 (0.0649)	0,0336 (0.0321)	0,1224 (0.0831)	0,1463 (NULL)

GPD-KDE-GPD parameter estimates are shown in Table 6.2 using the evmix package in R for the three stock market indices and the United States American Dollar to South African Rand exchange rate.

Table 6.2: ML estimates of GKG mixture model using the evmix package in R.

GPD-KDE-GPD Parameter Estimates	Financial Stock return			
	FTSE/JSE ALSI	FTSE/JSE Banks Index	FTSE/JSE Mining Index	USD/ZAR
$\hat{\lambda}$	0.0015	0.0023	0.0022	-0.0007
$\widehat{u_l}$	-0.0116	-0.0175	-0.0200	-0.0118
$\widehat{\sigma_{u_l}}$	0.0064	0.0092	0.0111	0.0068
$\widehat{\eta_l}$	0.1744	0.1790	0.1224	0.0170
$\widehat{u_r}$	0.0115	0.0197	0.0203	0.0109
$\widehat{u_r}$	0.0106	0.0040	0.0212	0.0170
$\widehat{\sigma_{u_r}}$	0.0057	0.0080	0.0098	0.0048
$\widehat{\eta_r}$	0.1426	0.3349	0.1829	0.1463

In Table 6.2, $\hat{\lambda}$ denotes the estimated bandwidth of the kernel. $\widehat{u_l}, \widehat{\sigma_{u_l}}, \widehat{\eta_l}$ represents the respective threshold, scale and shape parameters for the lower GPD tail estimates. Likewise, $\widehat{u_r}, \widehat{\sigma_{u_r}}, \widehat{\eta_r}$ describes the upper GPD tail estimated parameters for respective threshold, scale and shape parameters.

Figures 6.1 to Figure 6.4 shows the density plots for the fitted GNG. GKG and normal models for the FTSE/JSE ALSI, FTSE/JSE Banks Index, FTSE/JSE Mining Index and United States American Dollar to South African Rand exchange rate.

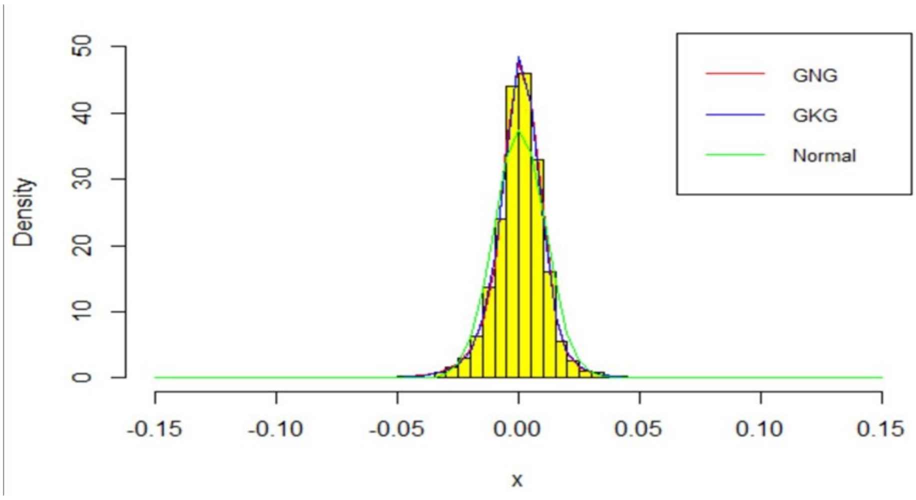


Figure 6.1: Fitted GNG and GKG mixture model FTSE/ALSI.

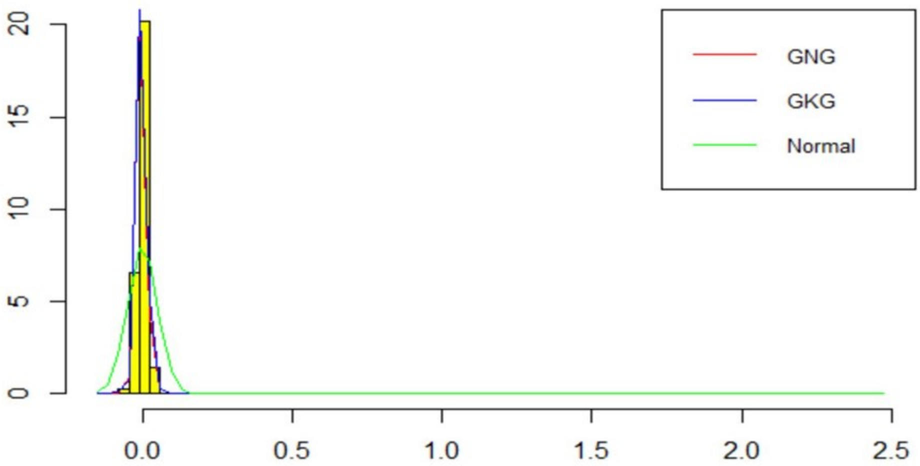


Figure 6.2: Fitted GNG and GKG mixture model FTSE/Banks Index.

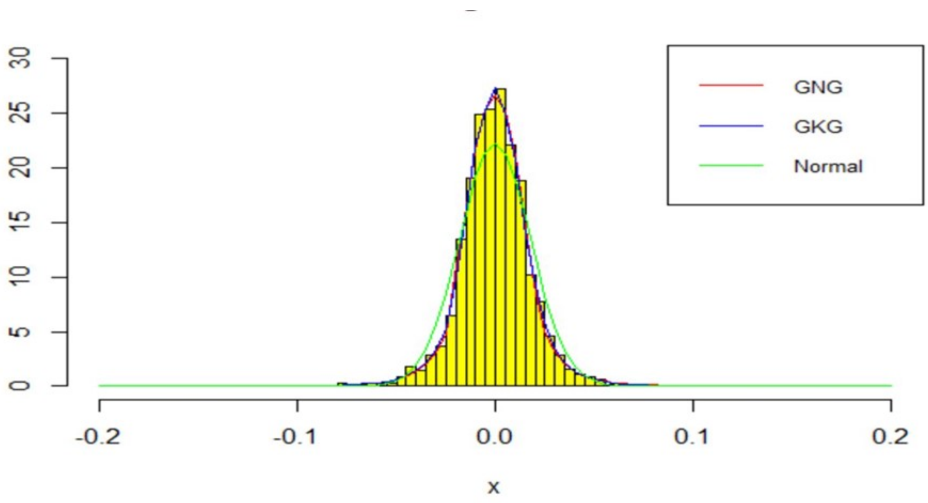


Figure 6.3: Fitted GNG and GKG mixture model of the FTSE/Mining Index.

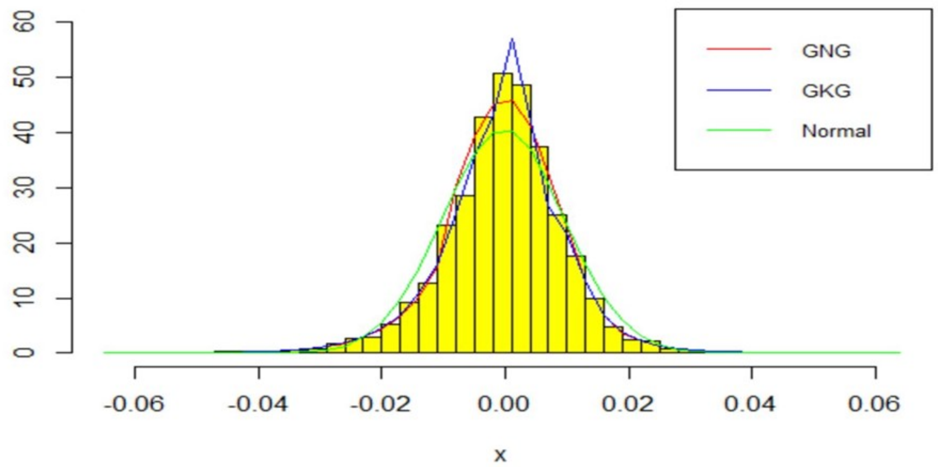


Figure 6.4: Fitted GNG and GKG model USD/ZAR.

The fitted GNG and GKG model distribution plots are shown in Figures 6.1 to 6.4 for each stock index and American Dollar to South African exchange rate. Notably, the two-tailed GNG and GKG model is a good fit across the entire range of data as compared to the commonly used normal distribution. For the FTSE/JSE Banks Index the plot is visually compressed due to the return series having extreme values; nevertheless, the GNG and GKG model outperforms the Normal model. In the case of the USD/ZAR exchange rate the GKG model outperforms the GNG and Normal model over the entire return series. Observations from these plots suggest the need for more formal model adequacy testing. Application of VaR estimation and Kupiec likelihood test could offer a more formal method of evaluating for model robustness.

Table 6.3 and Table 6.4 show the VaR estimates and p-values of the Kupiec likelihood backtesting procedure, respectively, to the South African financial data sets evaluated in this study.

Table 6.3: VaR estimates of financial market indices and exchange rate price returns using fitted GNG and GKG model.

Fitted Model	Returns	VaR level					
		Short position			Long position		
		1%	2.5%	5%	95%	97.5%	99%
GNG	FTSE/JSE ALSI	2.1880	2.7394	1.0394	0.1328	0.0360	0.1591
	FTSE/JSE Banks Index	2.1880	2.3034	1.0011	2.5317	3.2157	< 0.0002
	FTSE/JSE Mining Index	1.5811	0.3367	0.0076	0.2122	0.2004	0.1645
	USD/ZAR	0.0214	0.0696	0.0599	50.1711	66.2705	34.5848
GKG	FTSE/JSE ALSI	-0.4199	-0.2601	-0.1392	0.0159	0.0205	0.0274
	FTSE/JSE Banks Index	-0.0454	-0.0334	-0.0255	0.0251	0.0328	0.0461
	FTSE/JSE Mining Index	-0.0491	-0.0364	-0.0276	0.0273	0.0355	0.0480
	USD/ZAR	-0.0278	-0.0214	-0.0166	0.0145	0.0184	0.0242

Table 6.4: p-values of the Kupiec likelihood ratio test for financial indices and exchange rate returns.

Fitted Model	Returns	VaR level					
		Short position			Long position		
		1%	2.5%	5%	95%	97.5%	99%
GNG	FTSE/JSE ALSI	0.1391	0.0980	0.3080	0.7156	0.8495	0.6900
	FTSE/JSE Banks Index	0.1391	0.1290	0.3170	0.1116	0.0729	0.9984
	FTSE/JSE Mining Index	0.2086	0.5617	0.9304	0.6451	0.6544	0.6851
	USD/ZAR	0.8836	0.7935	0.8066	< 0.0002	< 0.0002	< 0.0002
GKG	FTSE/JSE ALSI	-	-	-	0.3080	0.7496	0.6900
	FTSE/JSE Banks Index	0.8401	0.2665	0.7858	0.3170	0.3985	0.6851
	FTSE/JSE Mining Index	0.5395	0.8457	0.9963	0.9233	0.9514	0.6851
	USD/ZAR	0.8836	0.8895	0.4877	0.4877	0.5411	0.4553

VaR estimates for the FTSE/JSE indices and USD/ZAR and corresponding Kupiec p-values are presented in Table 6.3 and Table 6.4, respectively. From Table 6.4, it is observed that at a 5% level of significance, the Kupiec test indicates the fitted GNG model is most suitable at almost all VaR levels for each of the returns since the p-values are greater than 0.05. Therefore, there is no basis to reject the null hypothesis. Conversely, the observed p-values for the long position on the USD/ZAR exchange rate returns are below 0.05, indicating a model misfit; however, a substantial fit is observed in the short position for the return series. Inconclusive results are seen for the short position of the fitted GKG model; however, for all return series, p-values are greater than 0.05. Thus, it can be inferred that at 5% level of significance the null hypothesis of model adequacy is accepted, therefore highlighting the robustness of the fitted GKG model.

6.3.2 Fitting the GNG and GSG using the piecewise modelling technique

This section investigates the fitting of EVT mixture models using a technique that separates the return series into various quantile ranges or segments and then modelling the tail and bulk components at the quantile cut-off points. This allows for ease of computationally fitting varied probability distributions over the same data range. Table 6.5 shows the estimated parameters for the GNG and GSG at each of the defined quantile thresholds.

Table 6.5: ML parameter estimates of the GNG (top) and GSG (bottom) mixture models using the piecewise modelling technique.

GNG parameter	**Financial Stock Indices and exchange rate returns**															
	FTSE/JSE ALSI				FTSE/JSE Banks Index				FTSE/JSE Mining Index				USD/ZAR			
Quantiles																
	$\hat{q}_{0.01}$	$\hat{q}_{0.025}$	$\hat{q}_{0.05}$	$\hat{q}_{0.1}$	$\hat{q}_{0.01}$	$\hat{q}_{0.025}$	$\hat{q}_{0.05}$	$\hat{q}_{0.1}$	$\hat{q}_{0.01}$	$\hat{q}_{0.025}$	$\hat{q}_{0.05}$	$\hat{q}_{0.1}$	$\hat{q}_{0.01}$	$\hat{q}_{0.025}$	$\hat{q}_{0.05}$	$\hat{q}_{0.1}$
$\hat{u}_l$	-0.1230	-0.1200	-0.1150	-0.1120	-2.1000	-2.1000	-1.7000	-1.8000	-0.1585	-0.1555	-0.1560	-0.1550	-0.1200	-0.1050	-0.1090	-0.0008
$\hat{\sigma}_l$	0.2513	0.3282	0.3388	0.3104	9.7170	12.5033	9.6455	10.0075	0.3368	0.3314	0.3883	0.4993	0.3641	0.2657	0.2769	0.2192
$\hat{\xi}_l$	-2.6545	-3.3214	-3.4394	-3.1298	-4.7297	-6.0451	-5.7614	-5.6181	-3.0326	-2.7508	-3.0199	-3.7030	-3.9406	-3.1871	-2.9850	-3.2219
Bulk Model																
$\hat{\mu}$	0.0004	0.0004	0.0004	0.0003	-0.0001	-0.0002	-0.0004	-0.0004	-0.0003	< 0.0001	-0.0002	-0.0007	-0.0003	-0.0002	-0.0002	-0.0001
$\hat{\beta}$	0.0089	0.0080	0.0070	0.0059	0.0140	0.0127	0.0113	0.0092	0.0153	< 0.0001	0.0120	0.0099	0.0086	0.0078	0.0069	0.0056
Quantiles																
	$\hat{q}_{0.99}$	$\hat{q}_{0.975}$	$\hat{q}_{0.95}$	$\hat{q}_{0.90}$	$\hat{q}_{0.99}$	$\hat{q}_{0.975}$	$\hat{q}_{0.95}$	$\hat{q}_{0.90}$	$\hat{q}_{0.99}$	$\hat{q}_{0.975}$	$\hat{q}_{0.95}$	$\hat{q}_{0.90}$	$\hat{q}_{0.99}$	$\hat{q}_{0.975}$	$\hat{q}_{0.95}$	$\hat{q}_{0.90}$
$\hat{u}_r$	-0.0150	-0.0210	-0.0100	-0.0250	0.0020	0.0040	0.0050	0.0060	0.0370	0.0348	0.0295	0.0220	-0.0350	-0.0380	-0.0280	-0.0350
$\hat{\sigma}_r$	0.1021	0.0768	0.0444	0.0591	0.1272	0.0668	0.0448	0.0138	0.0319	0.0150	0.0124	0.0126	0.1272	0.1142	0.0752	0.0772
$\hat{\eta}_r$	-1.1657	-0.8168	-0.5284	-0.6019	-1.3098	-0.6929	-0.4588	-0.1685	-0.2163	0.0341	0.0848	0.0601	-1.3342	-1.1613	-0.8501	-0.8088

GSG parameter	**Financial Stock Indices and exchange rate returns**															
	FTSE/JSE ALSI				FTSE/JSE Banks Index				FTSE/JSE Mining Index				USD/ZAR			
Quantiles																
	$\hat{q}_{0.01}$	$\hat{q}_{0.025}$	$\hat{q}_{0.05}$	$\hat{q}_{0.1}$	$\hat{q}_{0.01}$	$\hat{q}_{0.025}$	$\hat{q}_{0.05}$	$\hat{q}_{0.1}$	$\hat{q}_{0.01}$	$\hat{q}_{0.025}$	$\hat{q}_{0.05}$	$\hat{q}_{0.1}$	$\hat{q}_{0.01}$	$\hat{q}_{0.025}$	$\hat{q}_{0.05}$	$\hat{q}_{0.1}$
$\hat{u}_l$	-0.1230	-0.1200	-0.1150	-0.1120	-2.1000	-2.1000	-1.7000	-1.8000	-0.1585	-0.1555	-0.1560	-0.1550	-0.1200	-0.1050	-0.1090	-0.0008
$\hat{\sigma}_l$	0.2513	0.3282	0.3388	0.3104	9.7170	12.5033	9.6455	10.0075	0.3368	0.3314	0.3883	0.4993	0.3641	0.2657	0.2769	0.2192
$\hat{\xi}_l$	-2.6545	-3.3214	-3.4394	-3.1298	-4.7297	-6.0451	-5.7614	-5.6181	-3.0326	-2.7508	-3.0199	-3.7030	-3.9406	-3.1871	-2.9850	-3.2219
Bulk Model																
$\hat{\alpha}$	1.9999	1.9999	2.0000	2.0000	1.9999	1.9999	1.9999	1.9999	1.9999	1.9999	1.9999	1.9999	1.9999	1.9999	1.9999	1.9999
$\hat{\beta}$	0.1069	-0.2649	0.4857	0.4112	0.2477	-0.1741	0.3221	-0.4663	-0.4371	0.6793	0.7088	-0.6409	0.2473	0.2680	0.4544	0.3474
$\hat{\gamma}$	0.0063	0.0056	0.0050	0.0042	0.0100	0.0090	0.0080	0.0065	0.0109	0.0097	0.0085	0.0070	0.0061	0.0055	0.0049	0.0039
$\hat{\delta}$	0.0004	0.0004	0.0004	0.0003	0.0002	0.0002	0.0002	0.0003	0.0003	0.0003	0.0002	0.0003	-0.0003	-0.0002	-0.0002	-0.0001
Quantiles																
	$\hat{q}_{0.99}$	$\hat{q}_{0.975}$	$\hat{q}_{0.95}$	$\hat{q}_{0.90}$	$\hat{q}_{0.99}$	$\hat{q}_{0.975}$	$\hat{q}_{0.95}$	$\hat{q}_{0.90}$	$\hat{q}_{0.99}$	$\hat{q}_{0.975}$	$\hat{q}_{0.95}$	$\hat{q}_{0.90}$	$\hat{q}_{0.99}$	$\hat{q}_{0.975}$	$\hat{q}_{0.95}$	$\hat{q}_{0.90}$
$\hat{u}_r$	-0.0150	-0.0210	-0.0100	-0.0250	0.0020	0.0040	0.0050	0.0060	0.0370	0.0348	0.0295	0.0220	-0.0350	-0.0380	-0.0280	-0.0350
$\hat{\sigma}_r$	0.1021	0.0768	0.0444	0.0591	0.1272	0.0668	0.0448	0.0138	0.0319	0.0150	0.0124	0.0126	0.1272	0.1142	0.0752	0.0772
$\hat{\eta}_r$	-1.1657	-0.8168	-0.5284	-0.6019	-1.3098	-0.6929	-0.4588	-0.1685	-0.2163	0.0341	0.0848	0.0601	-1.3342	-1.1613	-0.8501	-0.8088

Figure 6.5 to Figure 6.8 show the density plots for the fitted model on the bulk model segment of the return data series.

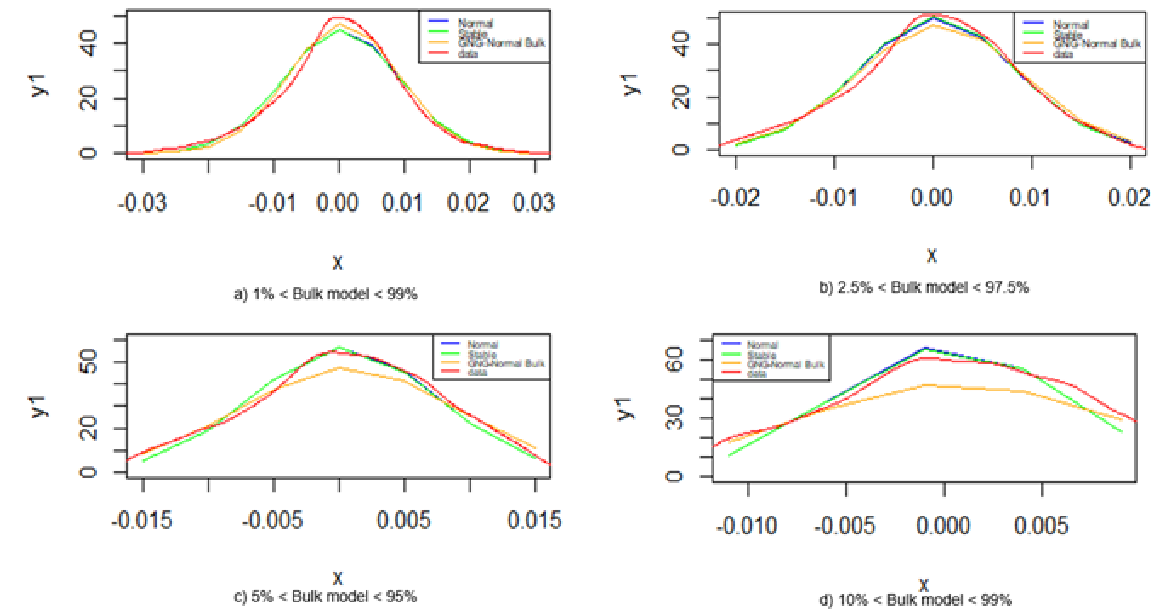


Figure 6.5: FTSE/JSE ALSI returns Bulk density plots for fitted models.

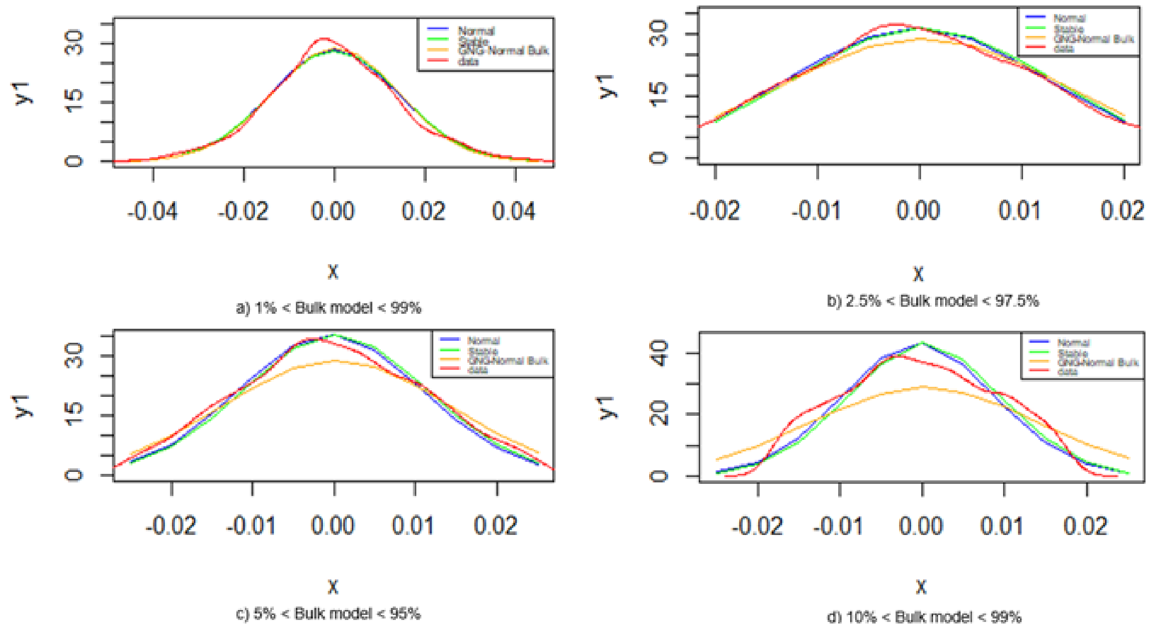


Figure 6.6: FTSE/JSE Banks Index returns Bulk density plots for fitted models.

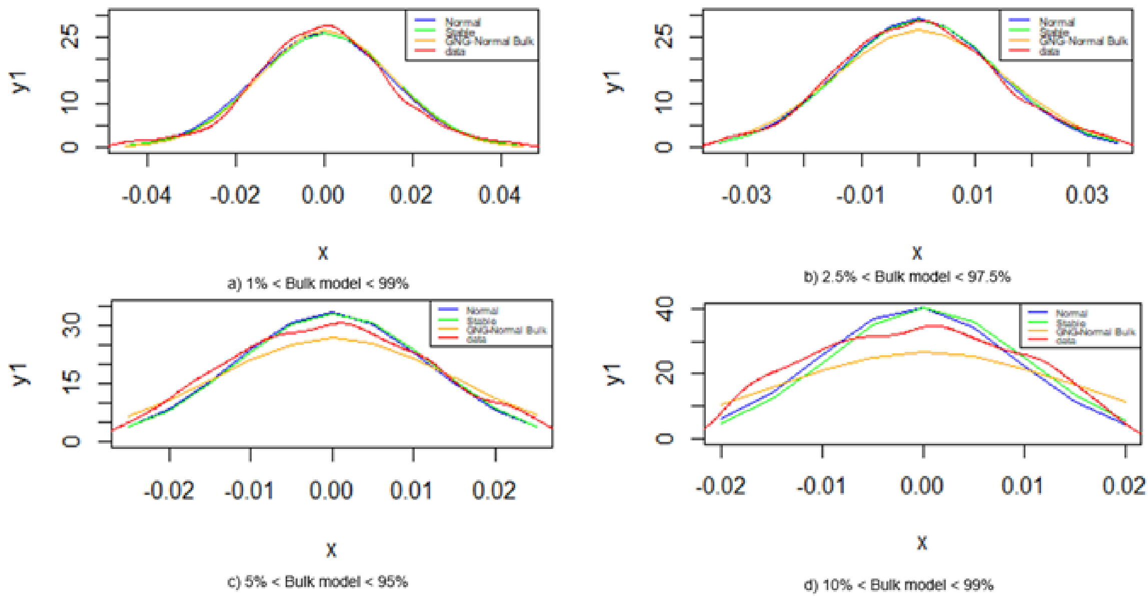


Figure 6.7: FTSE/JSE Mining Index returns Bulk density plots for fitted models.

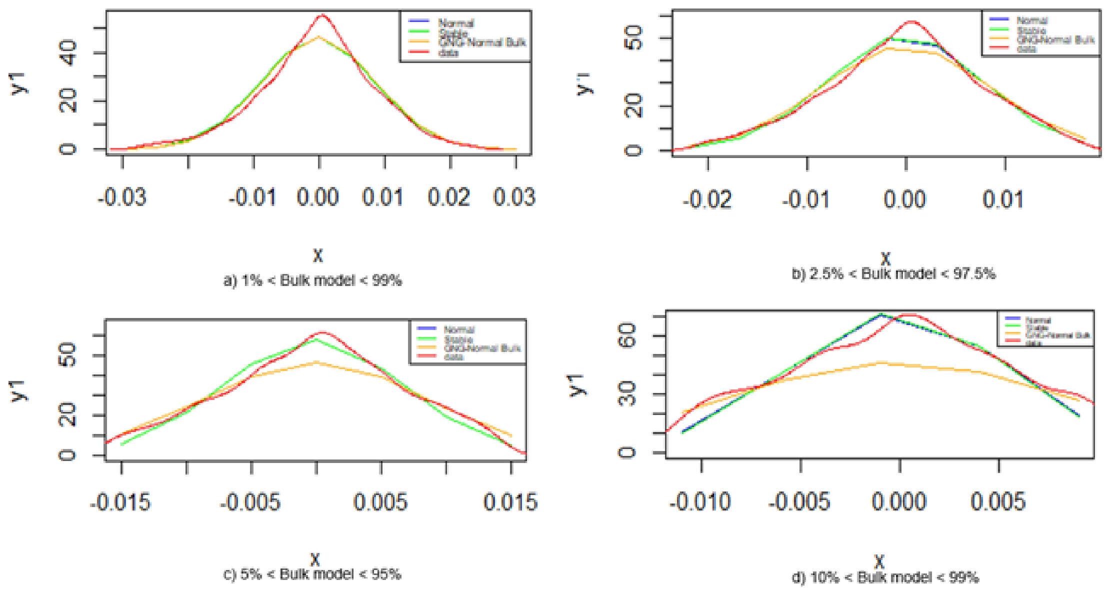


Figure 6.8: USD/ZAR returns Bulk density plots for fitted models.

Figures 6.5, 6.6, 6.7 and 6.8 depict the fitted models over each quantile range. The Normal and stable bulk models were fitted using the piecewise method, and the GNG model was fitted using the evmix package, where only the bulk model is compared. The stable density plots show that for each return series, over the specified quantile range, the fitted stable and normal models are almost equivalent and suggest a better fit than the normal bulk model estimates than that of the GNG model fitted using the evmix package. Graphically, these density plots show that the Stable model is a fitting alternative model when considering the bulk portion of extreme value mixed models.

Table 6.6 shows the RMSE test statistics to evaluate bulk model fitting for each return series.

Table 6.6: Calculated Root Mean Square Error (RMSE) for the fitted bulk model.

FTSE/JSE ALSI returns				
	$\hat{q}_{0.01} < GNG^* < \hat{q}_{0.99}$	$\hat{q}_{0.025} < GNG^* < \hat{q}_{0.975}$	$\hat{q}_{0.05} < GNG^* < \hat{q}_{0.95}$	$\hat{q}_{0.10} < GNG^* < \hat{q}_{0.90}$
RMSE	22.0654	28.0044	33.9154	43.6685
	$\hat{q}_{0.01} <$ GSG $< \hat{q}_{0.99}$	$\hat{q}_{0.025} <$ GSG $< \hat{q}_{0.975}$	$\hat{q}_{0.05} <$ GSG $< \hat{q}_{0.95}$	$\hat{q}_{0.10} <$ GSG $< \hat{q}_{0.90}$
RMSE	22.0534*	28.1462	33.7439	43.5186
	$\hat{q}_{0.01} <$ GNG $< \hat{q}_{0.99}$	$\hat{q}_{0.025} <$ GNG $< \hat{q}_{0.975}$	$\hat{q}_{0.05} <$ GNG $< \hat{q}_{0.95}$	$\hat{q}_{0.10} <$ GNG $< \hat{q}_{0.90}$
RMSE	22.7114	27.3275*	30.9252*	35.9534*
FTSE/JSE Banks Index				
	$\hat{q}_{0.01} < GNG^* < \hat{q}_{0.99}$	$\hat{q}_{0.025} < GNG^* < \hat{q}_{0.975}$	$\hat{q}_{0.05} < GNG^* < \hat{q}_{0.95}$	$\hat{q}_{0.10} < GNG^* < \hat{q}_{0.90}$
RMSE	14.5759	22.0992	21.3052	23.6056
	$\hat{q}_{0.01} <$ GSG $< \hat{q}_{0.99}$	$\hat{q}_{0.025} <$ GSG $< \hat{q}_{0.975}$	$\hat{q}_{0.05} <$ GSG $< \hat{q}_{0.95}$	$\hat{q}_{0.10} <$ GSG $< \hat{q}_{0.90}$
RMSE	14.5023*	22.0722*	21.2906	23.6152
	$\hat{q}_{0.01} <$ GNG $< \hat{q}_{0.99}$	$\hat{q}_{0.025} <$ GNG $< \hat{q}_{0.975}$	$\hat{q}_{0.05} <$ GNG $< \hat{q}_{0.95}$	$\hat{q}_{0.10} <$ GNG $< \hat{q}_{0.90}$
RMSE	14.6812	21.1061	19.2387*	19.2320*
FTSE/JSE Mining Index				
	$\hat{q}_{0.01} < GNG^* < \hat{q}_{0.99}$	$\hat{q}_{0.025} < GNG^* < \hat{q}_{0.975}$	$\hat{q}_{0.05} < GNG^* < \hat{q}_{0.95}$	$\hat{q}_{0.10} < GNG^* < \hat{q}_{0.90}$
RMSE	13.9428	16.5807	20.6607	25.1563*
	$\hat{q}_{0.01} <$ GSG $< \hat{q}_{0.99}$	$\hat{q}_{0.025} <$ GSG $< \hat{q}_{0.975}$	$\hat{q}_{0.05} <$ GSG $< \hat{q}_{0.95}$	$\hat{q}_{0.10} <$ GSG $< \hat{q}_{0.90}$
RMSE	13.8906*	16.5575*	20.6415	25.1573
	$\hat{q}_{0.01} <$ GNG $< \hat{q}_{0.99}$	$\hat{q}_{0.025} <$ GNG $< \hat{q}_{0.975}$	$\hat{q}_{0.05} <$ GNG $< \hat{q}_{0.95}$	$\hat{q}_{0.10} <$ GNG $< \hat{q}_{0.90}$
RMSE	14.1288	15.8854	18.4725*	20.1934
USD/ZAR				
	$\hat{q}_{0.01} < GNG^* < \hat{q}_{0.99}$	$\hat{q}_{0.025} < GNG^* < \hat{q}_{0.975}$	$\hat{q}_{0.05} < GNG^* < \hat{q}_{0.95}$	$\hat{q}_{0.10} < GNG^* < \hat{q}_{0.90}$
RMSE	22.4805	28.3340	34.1813	44.8735
	$\hat{q}_{0.01} <$ GSG $< \hat{q}_{0.99}$	$\hat{q}_{0.025} <$ GSG $< \hat{q}_{0.975}$	$\hat{q}_{0.05} <$ GSG $< \hat{q}_{0.95}$	$\hat{q}_{0.10} <$ GSG $< \hat{q}_{0.90}$
RMSE	22.4455*	28.3735	34.1076	45.2201
	$\hat{q}_{0.01} <$ GNG $< \hat{q}_{0.99}$	$\hat{q}_{0.025} <$ GNG $< \hat{q}_{0.975}$	$\hat{q}_{0.05} <$ GNG $< \hat{q}_{0.95}$	$\hat{q}_{0.10} <$ GNG $< \hat{q}_{0.90}$
RMSE	22.4802	26.9819*	30.5722*	35.5594*

Note.
* represents the minimum RMSE within the quantile range
GNG^* represents the GDP-Normal-GPD piecewise mixture model
GNG represents the GPD-Normal-GPD model with Bayesian inference by Zhao (2010) and Zhao et al. (2010)

Results from Table 6.6 show that for the FTSE/JSE ALSI returns, the GNG model performs well at the 2.5%, 5% and 10% lower quantile thresholds and 97.5%, 95% and 90% upper quantiles thresholds; however, for the 1% (lower) and 99% (upper) quantile threshold, the GSG model outperforms both the piecewise GNG as well as the GNG model by Zhao (2010) and Zhao et al. (2010). For the FTSE/JSE Banks Index returns, the GSG provides a better fit at the 1% and 2.5% (lower quantiles) as well as the 99% and 97.5% upper quantile thresholds, whereas the GNG model with Bayesian inference performs better at the 5%, 10% (lower) and 95% and 90% (upper) threshold. The FTSE/JSE Mining Index returns have similar results to that of the FTSE/JSE Banks Index return series data. However, it is observed that the piecewise GNG model is a good fit for the 10% and 90% threshold since the RMSE test statistic yielded the lowest value when compared to the other two models. The GNG model by Zhao (2010) and Zhao et al. (2010) is observed to be a good fit for the USD/ZAR return series and is outperformed by the GSG at the 1% and 99% quantile thresholds. Generally, results from Table 6.6 favours the Bayesian GNG model at certain threshold quantile points, although in some instances, the GSG model is shown to be a better-fitted model. We check goodness-of-fit using the Anderson-Darling goodness-of-fit tests to evaluate the GSG model performance. The AD-test and corresponding p-values are shown in Table 6.7

Table 6.7: Anderson Darling goodness of fit test of the GSG mixture model.

FTSE/JSE ALSI returns				
	$\hat{q}_{0.01} < \text{GSG} < \hat{q}_{0.99}$	$\hat{q}_{0.025} < \text{GSG} < \hat{q}_{0.975}$	$\hat{q}_{0.05} < \text{GSG} < \hat{q}_{0.95}$	$\hat{q}_{0.10} < \text{GSG} < \hat{q}_{0.90}$
AD Test Statistic	4.1462	2.2196	2.9447	6.0759
p-value	0.0074	0.06979*	0.0292***	0.00089
FTSE/JSE Banks Index				
	$\hat{q}_{0.01} < \text{GSG} < \hat{q}_{0.99}$	$\hat{q}_{0.0025} < \text{GSG} < \hat{q}_{0.975}$	$\hat{q}_{0.05} < \text{GSG} < \hat{q}_{0.95}$	$\hat{q}_{0.10} < \text{GSG} < \hat{q}_{0.90}$
AD Test Statistic	1.8054	0.82216	2.2493	7.1648
p-value	0.1178*	0.4653*	0.0673*	0.0003
FTSE/JSE Mining Index				
	$\hat{q}_{0.01} < \text{GSG} < \hat{q}_{0.99}$	$\hat{q}_{0.0025} < \text{GSG} < \hat{q}_{0.975}$	$\hat{q}_{0.05} < \text{GSG} < \hat{q}_{0.95}$	$\hat{q}_{0.10} < \text{GSG} < \hat{q}_{0.90}$
AD Test Statistic	2.2836	0.4490	2.8948	8.0201
p-value	0.0645*	0.7991*	0.0310***	0.0001
USD/ZAR				
	$\hat{q}_{0.01} < \text{GSG} < \hat{q}_{0.99}$	$\hat{q}_{0.0025}$ GSG $< \hat{q}_{0.975}$	$\hat{q}_{0.05} < \text{GSG} < \hat{q}_{0.95}$	$\hat{q}_{0.10} < \text{GSG} < \hat{q}_{0.90}$
AD Test Statistic	4.7302	2.9236	3.4248	5.7556
p-value	0.0039	0.0299***	0.0168***	0.0013
Note. * and *** indicates significance at 5% and 1% respectively				

The results from the Anderson-Darling goodness-of-fit tests confirm the results seen in the stable density plots where the suggested GSG model serves as a good fit at the 2,5% and 97,5% as well as the 5% and 95% quantile thresholds. For the FTSE/JSE ALSI return index series, the fitted stable bulk model favours the null hypothesis at a 5% level of significance for the 2,5% and 97,5% quantiles. At a 1% level of significance, the 5% and 95% quantiles indicates that the fitted stable model is significant. Favourable results are seen for the Banks Index and Mining Index returns where the suggested stable bulk model is considered fitting, at a 1% level of significance, for the 1% and 99%, 2,5% and 97,5% in addition to 5% and 95% quantile thresholds. The recommended fitted bulk stable model is significant at a 1% level of significance for the 2,5% and 97,5% along with the 5% and 95% quantile thresholds.

6.3.3 Fitting the SNS and SKS models

SNS parameter estimates are shown in Table 6.8, where $\hat{\mu}$ denotes the estimated normal mean, $\hat{\sigma}$ is the estimated normal standard deviation of the normal bulk model. $\hat{\alpha}_l, \hat{\beta}_l, \hat{\gamma}_l, \hat{\delta}_l$ represents the lower stable tail estimated parameters. Similarly,

$\hat{\alpha}_r, \hat{\beta}_r, \hat{\gamma}_r, \hat{\delta}_r$ describes the upper stable tail estimated parameters. Table 6.8 also shows the bulk parameter estimates for the SKS model. Since a piecewise estimation approach was adopted, the fitted stable tail model is the same as Table 6.8 for each quantile threshold point; therefore, only the bandwidth parameter λ for the KDE bulk model is estimated.

Table 6.8: ML parameter estimates of the SNS (top) and SKS (bottom) mixture models.

SNS parameter	Financial Stock Indices and exchange rate returns															
	FTSE/JSE ALSI				FTSE/JSE Banks Index				FTSE/JSE Mining Index				USD/ZAR			
	Quantiles															
	$\hat{q}_{0.01}$	$\hat{q}_{0.025}$	$\hat{q}_{0.05}$	$\hat{q}_{0.1}$	$\hat{q}_{0.01}$	$\hat{q}_{0.025}$	$\hat{q}_{0.05}$	$\hat{q}_{0.1}$	$\hat{q}_{0.01}$	$\hat{q}_{0.025}$	$\hat{q}_{0.05}$	$\hat{q}_{0.1}$	$\hat{q}_{0.01}$	$\hat{q}_{0.025}$	$\hat{q}_{0.05}$	$\hat{q}_{0.1}$
$\hat{\alpha}_l$	0.7046	0.8555	0.9106	0.9757	0.6116	0.7299	0.7153	0.9576	0.0091	0.9772	0.9267	0.8950	1.9999	1.9999	1.0880	0.9889
$\hat{\beta}_l$	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.000	-1.0000	-1.0000	-1.0000	-1.0000	-0.4954	0.0683	-1.0000	-1.0000
$\hat{\gamma}_l$	0.0022	0.0024	0.0022	0.0022	0.0050	0.0040	0.0026	0.0032	0.0045	0.0042	0.0037	0.0032	0.0039	0.0045	0.0025	0.0021
$\hat{\delta}_l$	-0.0309	-0.0241	-0.0193	-0.0158	-0.0505	-0.0363	-0.0288	-0.0232	-0.0532	-0.0414	-0.0324	-0.0244	-0.0350	-0.0285	-0.0199	-0.0149
	Bulk Model															
$\hat{\mu}$	0.0004	0.0004	0.0004	0.0003	0.0002	0.0002	0.0002	0.0003	0.0003	0.0003	0.0002	0.0003	-0.0003	-0.0002	-0.0002	-0.0001
$\hat{\sigma}$	0.0089	0.0080	0.0070	0.0059	0.0141	0.0127	0.0113	0.0091	0.0153	0.0137	0.0120	0.0098	0.0086	0.0078	0.0069	0.0056
	Quantiles															
	$\hat{q}_{0.99}$	$\hat{q}_{0.975}$	$\hat{q}_{0.95}$	$\hat{q}_{0.90}$	$\hat{q}_{0.99}$	$\hat{q}_{0.975}$	$\hat{q}_{0.95}$	$\hat{q}_{0.90}$	$\hat{q}_{0.99}$	$\hat{q}_{0.975}$	$\hat{q}_{0.95}$	$\hat{q}_{0.90}$	$\hat{q}_{0.99}$	$\hat{q}_{0.975}$	$\hat{q}_{0.95}$	$\hat{q}_{0.90}$
$\hat{\alpha}_r$	0.9738	0.8633	0.8532	0.9300	1.1938	0.7726	0.8786	0.9580	0.9670	1.0251	1.0707	0.9118	0.7644	0.9562	0.9539	0.9769
$\hat{\beta}_r$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$\hat{\gamma}_r$	0.0023	0.0022	0.0020	0.0018	0.0071	0.0037	0.0032	0.0033	0.0050	0.0040	0.0044	0.0037	0.0017	0.0020	0.0018	0.0016
$\hat{\delta}_l$	0.0311	0.0230	0.0180	0.0140	0.0559	0.0363	0.0294	0.0226	0.0535	0.0426	0.0340	0.0241	0.0255	0.0206	0.0166	0.0132

SKS parameter	Financial Stock Indices and exchange rate returns															
	FTSE/JSE ALSI				FTSE/JSE Banks Index				FTSE/JSE Mining Index				USD/ZAR			
	Quantiles															
	$\hat{q}_{0.01}$	$\hat{q}_{0.025}$	$\hat{q}_{0.05}$	$\hat{q}_{0.1}$	$\hat{q}_{0.01}$	$\hat{q}_{0.025}$	$\hat{q}_{0.05}$	$\hat{q}_{0.1}$	$\hat{q}_{0.01}$	$\hat{q}_{0.025}$	$\hat{q}_{0.05}$	$\hat{q}_{0.1}$	$\hat{q}_{0.01}$	$\hat{q}_{0.025}$	$\hat{q}_{0.05}$	$\hat{q}_{0.1}$
$\hat{\alpha}_l$	0.7046	0.8555	0.9106	0.9757	0.6116	0.7299	0.7153	0.9576	0.0091	0.9772	0.9267	0.8950	1.9999	1.9999	1.0880	0.9889
$\hat{\beta}_l$	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.000	-1.0000	-1.0000	-1.0000	-1.0000	-0.4954	0.0683	-1.0000	-1.0000
$\hat{\gamma}_l$	0.0022	0.0024	0.0022	0.0022	0.0050	0.0040	0.0026	0.0032	0.0045	0.0042	0.0037	0.0032	0.0039	0.0045	0.0025	0.0021
$\hat{\delta}_l$	-0.0309	-0.0241	-0.0193	-0.0158	-0.0505	-0.0363	-0.0288	-0.0232	-0.0532	-0.0414	-0.0324	-0.0244	-0.0350	-0.0285	-0.0199	-0.0149
	Bulk Model															
$\hat{\lambda}_l$	0.0016	0.0011	0.0010	0.0004	0.0025	0.0014	0.0010	0.0010	0.0023	0.0016	0.0008	0.0005	0.0004	0.0007	0.0009	0.0009
	Quantiles															
	$\hat{q}_{0.99}$	$\hat{q}_{0.975}$	$\hat{q}_{0.95}$	$\hat{q}_{0.90}$	$\hat{q}_{0.99}$	$\hat{q}_{0.975}$	$\hat{q}_{0.95}$	$\hat{q}_{0.90}$	$\hat{q}_{0.99}$	$\hat{q}_{0.975}$	$\hat{q}_{0.95}$	$\hat{q}_{0.90}$	$\hat{q}_{0.99}$	$\hat{q}_{0.975}$	$\hat{q}_{0.95}$	$\hat{q}_{0.90}$
$\hat{\alpha}_r$	0.9738	0.8633	0.8532	0.9300	1.1938	0.7726	0.8786	0.9580	0.9670	1.0251	1.0707	0.9118	0.7644	0.9562	0.9539	0.9769
$\hat{\beta}_r$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$\hat{\gamma}_r$	0.0023	0.0022	0.0020	0.0018	0.0071	0.0037	0.0032	0.0033	0.0050	0.0040	0.0044	0.0037	0.0017	0.0020	0.0018	0.0016
$\hat{\delta}_l$	0.0311	0.0230	0.0180	0.0140	0.0559	0.0363	0.0294	0.0226	0.0535	0.0426	0.0340	0.0241	0.0255	0.0206	0.0166	0.0132

Figure 6.9 shows the upper and lower tail density plots at various threshold quantiles for the FTSE/JSE ALSI returns.

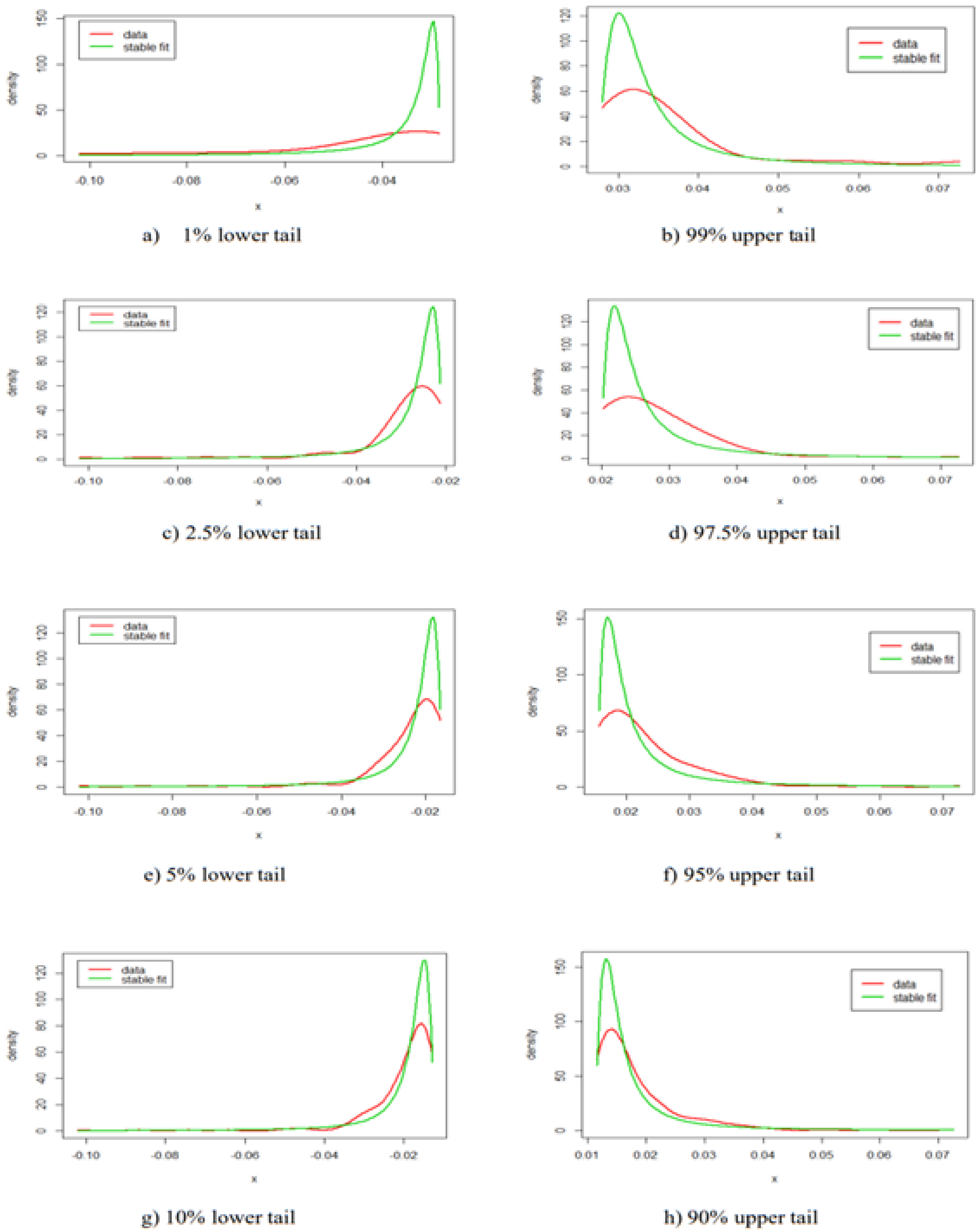
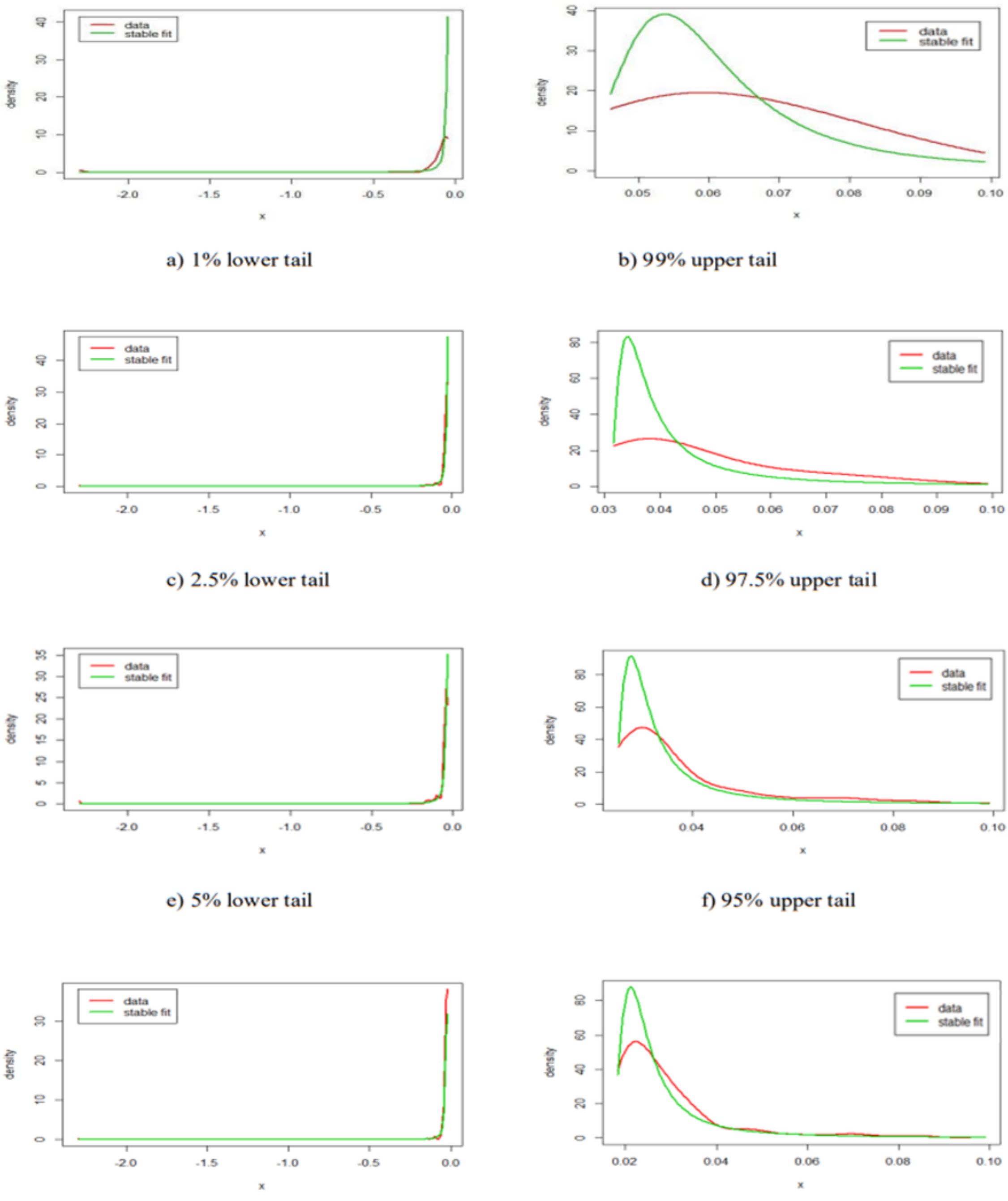


Figure 6.9: FTSE/JSE ALSI stable lower tail model density plots (left) and upper tail density plots (right).

Figure 6.10 shows the upper and lower tail density plots at various threshold quantiles for the FTSE/JSE Banks Index returns.



a) 1% lower tail

b) 99% upper tail

c) 2.5% lower tail

d) 97.5% upper tail

e) 5% lower tail

f) 95% upper tail

g) 10% lower tail

h) 90% upper tail

Figure 6.10: FTSE/JSE Banks Index stable lower tail model density plots (left) and upper tail density plots (right).

Figure 6.11 shows the upper and lower tail density plots at various threshold quantiles for the FTSE/JSE Mining Index returns.

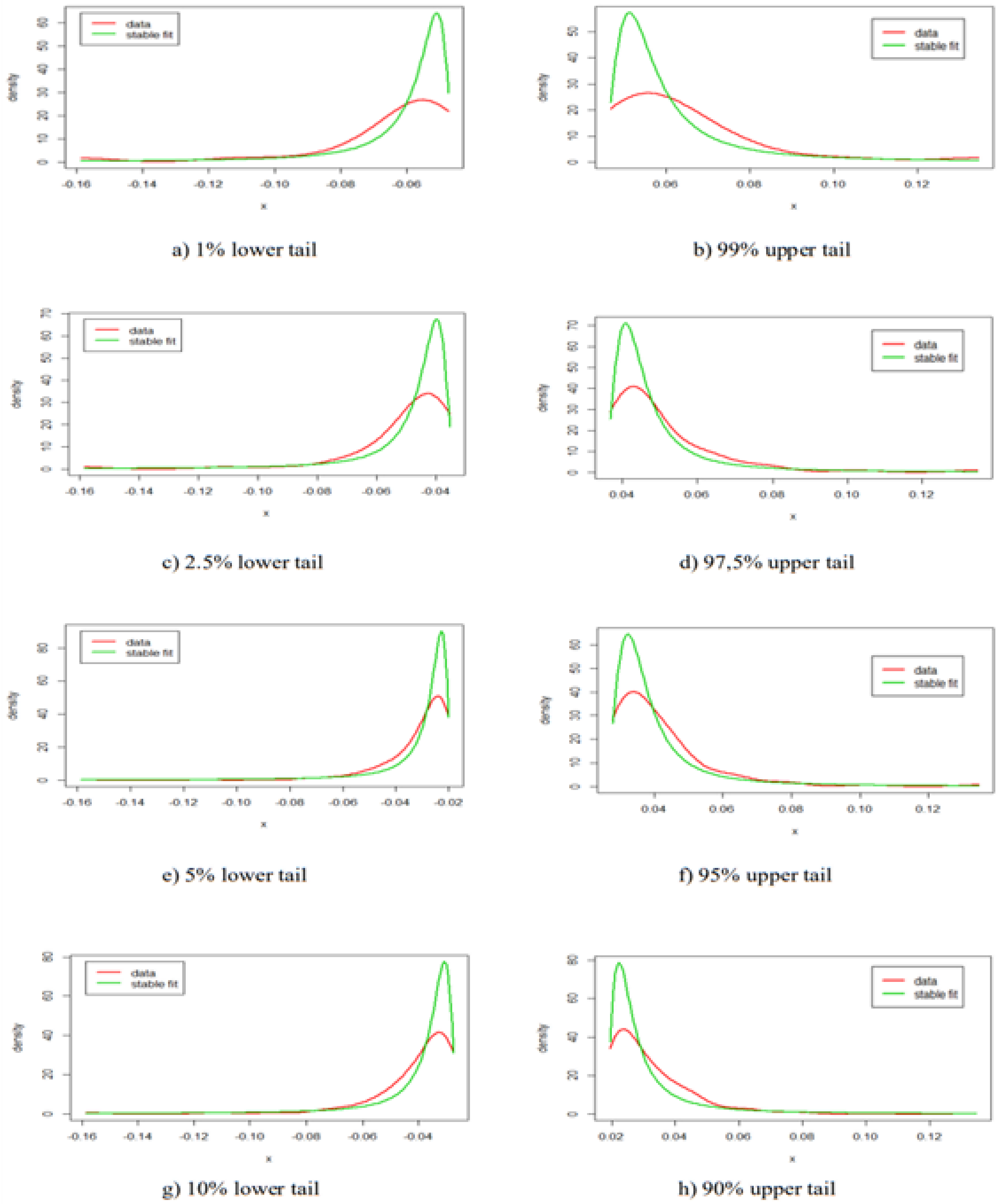
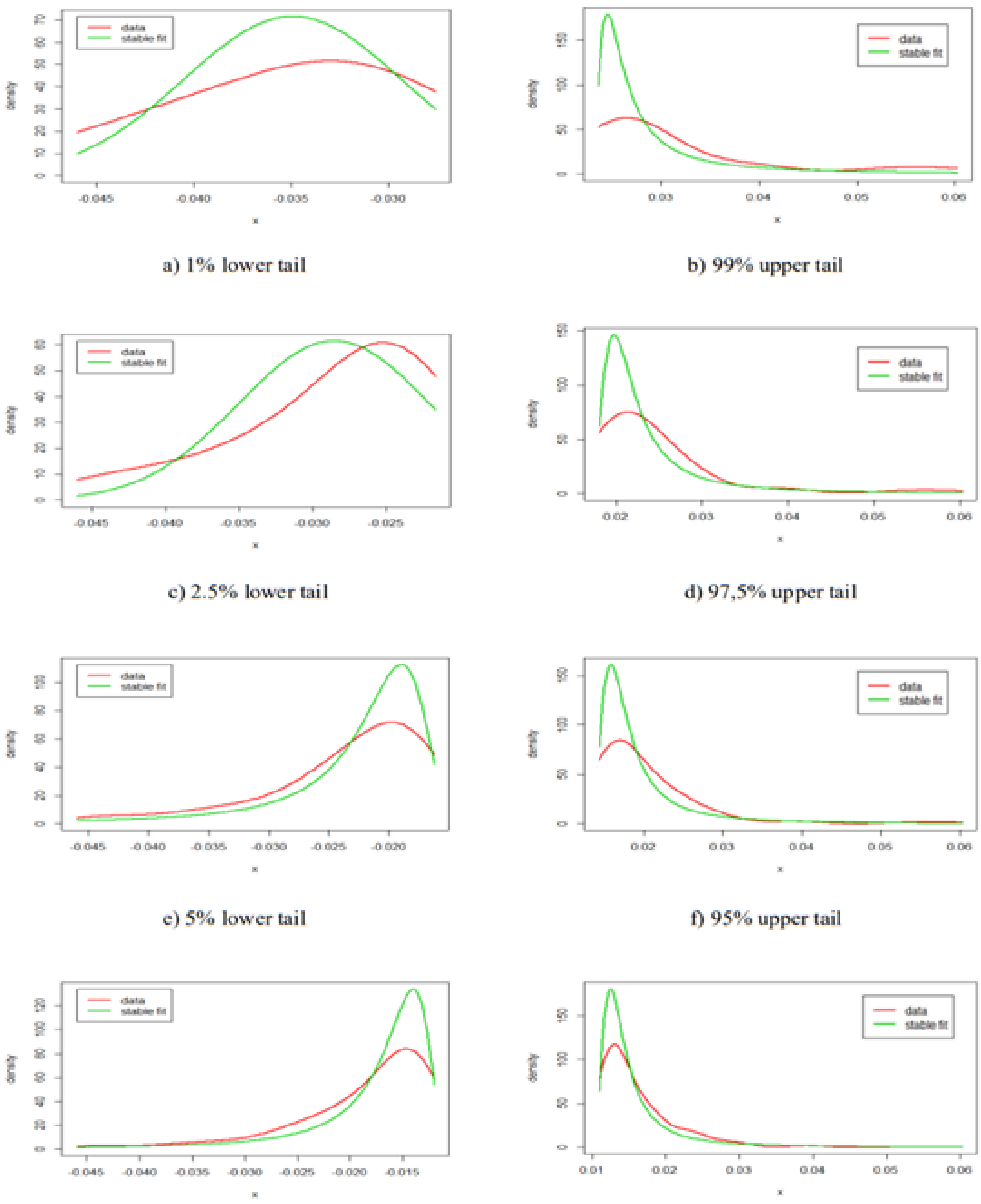


Figure 6.11: FTSE/JSE Mining Index stable lower tail model density plots (left) and upper tail density plots (right).

Figure 6.12 shows the upper and lower tail density plots at various threshold quantiles for the USD/ZAR returns.



a) 1% lower tail

b) 99% upper tail

c) 2.5% lower tail

d) 97,5% upper tail

e) 5% lower tail

f) 95% upper tail

g) 10% lower tail

h) 90% upper tail

Figure 6.12: USD/ZAR stable lower tail model density plots (left) and upper tail density plots (right).

Figures 6.9, 6.10, 6.11 and 6.12 shows that the fitted stable tail segment of the SNS and SKS model fits best for all returns at all tail quantiles specifically the extreme data values. Overfitting is noticed as the tail section reaches the bulk component, which is expected behaviour of stable models where variability is noticed. Mixed results are seen for the 99% upper tail of the Banks Index, where descriptive statistics have previously shown extreme kurtosis and heavy tails. Generally, this suggests the stable model for the tails as a worthy alternative to the tail model for risk-averse practitioners. Table 6.9 shows the results of the Anderson Darling Goodness-of-fit test on the lower tail component of the fitted SNS and SKS models.

Table 6.9: Anderson Darling goodness-of-fit test for the lower tail of the fitted SNS and SKS piecewise mixture models.

	Lower Tail Quantiles			
	1% lower tail	**2.5% lower tail**	**5% lower tail**	**10% lower tail**
		FTSE/JSE ALSI		
AD Test Statistic	0.2363	0.5988	0.8688	1.0311
p-value	0.9771	0.6485	0.4337	0.3413
		FTSE/JSE Banks Index		
AD test Statistic	0.2298	0.4069	0.7775	1.3615
p-value	0.9800	0.8414	0.4973	0.2133
		FTSE/JSE Mining Index		
AD Test Statistic	0.2286	0.4062	0.9290	2.0496
p-value	0.9805	0.8421	0.3966	0.0863
		USD/ZAR		
AD Test Statistic	0.6671	3.2959	0.7694	1.7949
p-value	0.5855	0.0195	0.5034	0.1195

The results from the Anderson-Darling goodness-of-fit tests in Table 6.9 confirm the results seen in the stable density lower tail plots where the suggested stable model serves as a good fit at all quantile thresholds for the lower tail. For each of the four return index series, the fitted stable tail model favours the null hypothesis of model adequacy at a 5% level of significance for all quantiles, thus confirming formally that the stable models ought to be considered for the tail portion of extreme value mixture models. Table 6.10 shows the results of the Anderson Darling goodness-of-fit test for the upper tail innovations for the SNS and SKS models.

Table 6.10: Anderson Darling goodness-of-fit test for the upper tail of the fitted SNS and SKS piecewise mixture models.

Upper Tail Quantiles				
	99% lower tail	**97.5% lower tail**	**95% lower tail**	**90% lower tail**
FTSE/JSE ALSI				
AD Test Statistic	0.1315	0.8604	1.2868	1.3250
p-value	0.9996	0.4390	0.2366	0.2244
FTSE/JSE Banks Index				
AD Test Statistic	0.5916	0.8516	0.5783	1.5101
p-value	0.6543	0.85164	0.6685	0.1741
FTSE/JSE Mining Index				
AD Test Statistic	0.2703	0.3339	0.7417	2.729
p-value	0.9584	0.9104	0.5247	0.03771
USD/ZAR				
AD Test Statistic	0.2674	0.6262	0.9962	1.1420
p-value	0.9602	0.6230	0.3591	0.2906

Table 6.10 shows the results of the Anderson Darling goodness-of-fit test on the upper tail component of the fitted SNS and SKS models. Similar to Table 6.9, results from the Anderson-Darling goodness-of-fit in Table 6.10 confirm that the suggested stable upper tail model serves as a good fit at all quantile thresholds except for the 90% upper tail quantile threshold for the FTSE/JSE Mining Index return series. The fitted stable tail model fails to reject the null hypothesis of model adequacy at a 5% level of significance. At a 1% level of significance, the upper tail fitted stable model is deemed to be a good fit at all upper quantile threshold points, once again confirming the flexibility and robustness of fitted stable distributions for the tail portion of extreme value mixture models.

Figure 6.13 shows the density plots of the fitted bulk model at various threshold quantiles for the FTSE/JSE ALSI return series.

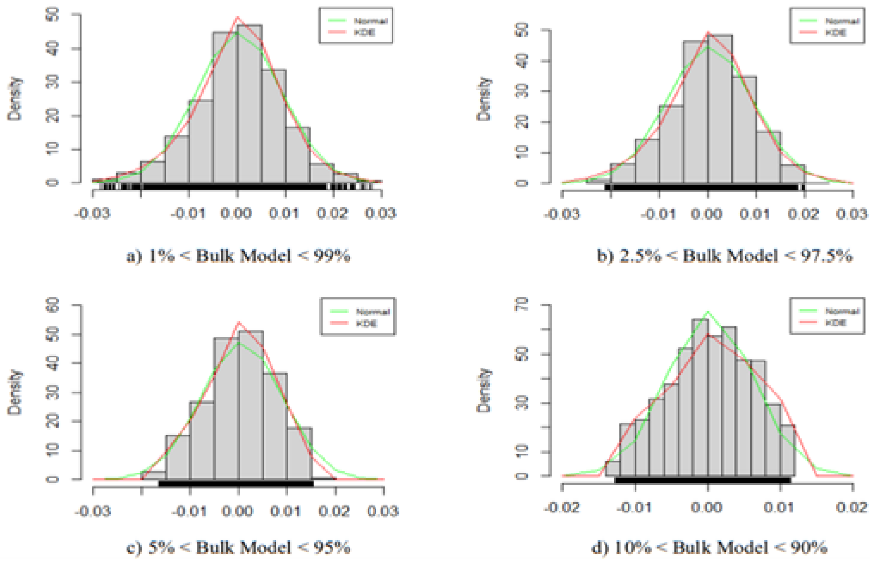


Figure 6.13: FTSE/JSE ALSI returns Bulk density plots for fitted models.

Figure 6.14 shows the density plots of the fitted bulk model at various threshold quantiles for the FTSE/JSE Banks Index return series.

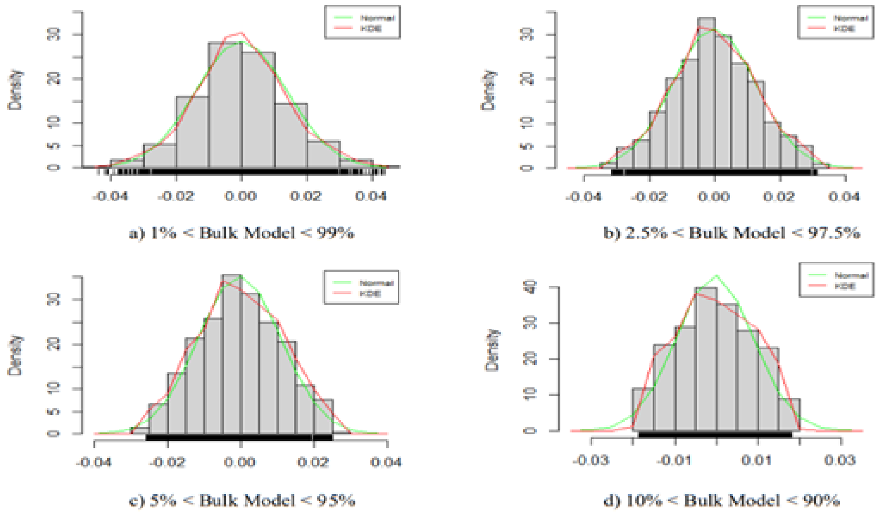


Figure 6.14: FTSE/JSE Banks Index returns Bulk density plots for fitted models.

Figure 6.15 shows the density plots of the fitted bulk model at various threshold quantiles for the FTSE/JSE Mining Index return series.

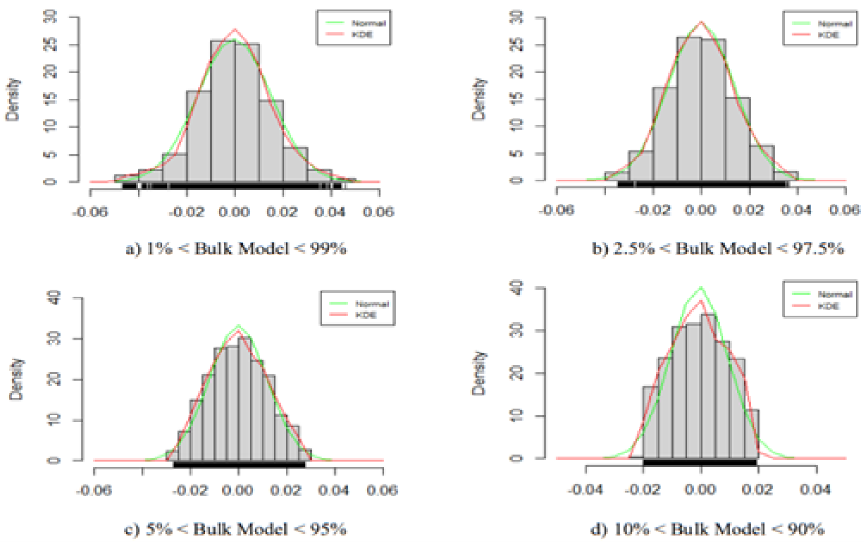


Figure 6.15: FTSE/JSE Mining Index returns Bulk density plots for fitted models.

Figure 6.15 shows the density plots of the fitted bulk model at various threshold quantiles for the USD/ZAR return series.

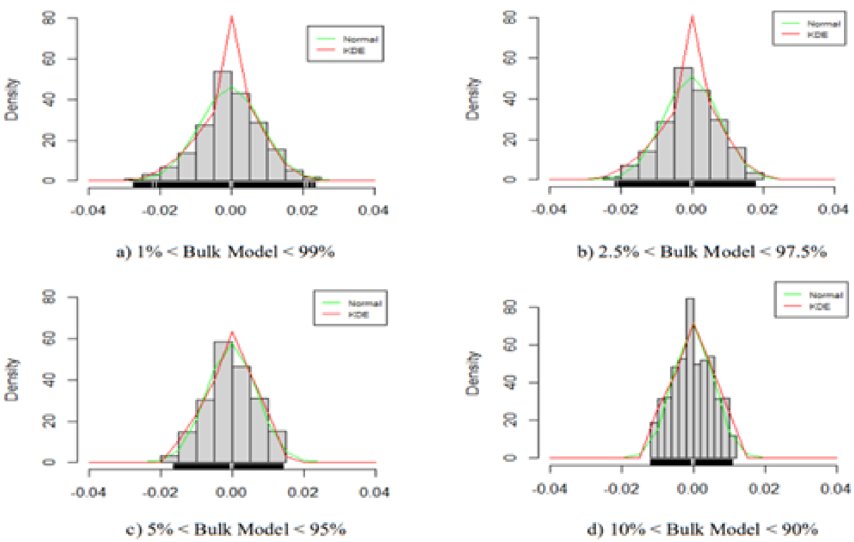


Figure 6.16: USD/ZAR returns Bulk density plots for fitted models.

The fitted normal and KDE model density plots for the SNS and SKS models are shown in Figures 6.13, 6.14, 6.15 and 6.16 above for each stock indices and South African exchange rate. Evidently, the KDE model is a good fit across the entire range of the bulk data as compared to the commonly used normal distribution. However, the Normal model is a close second alternative to model bulk data over the bulk return series range. This seems to be the opposite for the FTSE/JSE Banks Index where the fitted Normal model outperforms the KDE model. This study places more focus on the stable tail component where model robustness is of great interest, especially from an extreme value mixture modelling perspective.

6.4 Concluding remarks

In this chapter, the theoretical framework of the GNG, GSG, SNS and SKS piecewise models was explored with defined probability density functions and parameter vector descriptions. The rationale behind implementing an empirical analysis on these mixture models was to account for the various behaviours of South African financial data over multiple probability distributions, providing a tailored approach to represent the data and account for rare events or market shocks.

The evmix package in R by Hu and Scarrott (2018) provided a computationally sound method of applying the GNG and GKG models. Parameter estimates for both the GNG and GKG were estimated and the density plots were investigated. VaR estimates and the VaR in-sample backtesting procedure, the Kupiec likelihood test indicated that the fitted GNG and GKG models are well-suited at most VaR levels except for the USD/ZAR long position and the GKG short position. This insinuated the limitations of the fitted models and prompted the application of a more suited heavy tailed model in the mixture model gains and losses position.

This chapter also used a piecewise modelling technique to fit the GNG and GSG. The main focus in the piecewise fitted models, in this case, was to compare the stable bulk model component to the Normal fitted models. The fitted stable bulk model performed well when observing the density plots of the fitted bulk models over the specified quantile range, thus suggesting a fitting alternative for considering the bulk component model in EVT mixture modelling. RMSE test statistics depicted mixed results in terms of the GSG bulk model and favoured the fitted GNG bulk model. However, the AD test results for the GSG model confirmed a good fit. Overall, inconclusive results were observed when considering a stable bulk model; nevertheless, this probed an investigation to inspect stable model performance in the tail component of EVT mixture models. A possible solution for dealing with inconclusive results from the stable bulk model is to consider model re-specification or alternative robust estimation techniques.

Likewise, the SNS and SKS models were fitted using the piecewise method to achieve a more nuanced estimate of the stable tail model component. A graphical analysis of the stable tail density plots implied a relatively good fit; however, overfitting was seen as the tail models tended towards the bulk model threshold. This is typical behaviour of stable models to overfit when variablity is noticed in the data series. This suggested suitability of stable models from an EVT modelling point of view to capture and predict rare or unusual events. The AD test confirmed the appropriateness of the lower and upper stable tail models.

Extreme mixture models, when applied to financial data, present several challenges. One major issue is the need to carefully select components that accurately represent the tail behaviour of financial returns, which can lead to overfitting, particularly with smaller datasets. To address this, it is essential to employ regularisation techniques and cross-validation methods. Regularisation helps mitigate overfitting by penalising excessively complex models, while cross-validation assesses the model performance on new data to ensure it fits effectively. Additionally, leveraging domain knowledge and expertise to guide the bulk or tail model component selection and using Bayesian methods for parameter estimation can help balance model complexity with predictive accuracy.

CHAPTER 7

CONCLUSION AND RECOMMENDATIONS

In this concluding chapter, the findings of this study are summarised and the aims and contributions are reflected upon. In addition, this chapter discusses the limitations of the study and provides recommendations for future research.

7.1 Conclusion

In the financial industry, there has been a longstanding tradition of assuming that financial returns conform to a normal distribution. The Gaussian paradigm comprises of many favourable analytical properties align with those found in the stable distribution family (Yang, 2012a).

The Normal distribution is favoured in financial modeling for several reasons:

- It is a simple and practical distribution that allows for the implementation of numerical methods.
- Normally distributed random variables cluster around a central mean; as one moves away from the mean, the odds of deviation exponentially decrease.
- The Central Limit Theorem and the Law of Large Numbers are fundamental properties in statistics that simplify complexities by approximating distributions to be normal. Stoyanov et al. (2011) mentions the Black-Scholes option pricing model, Capital asset pricing model and Markowitz's modern portfolio theory as notable financial modelling frameworks which utilise the Normal distribution as the analytical foundation.

Empirical analysis by Pele and Mazurencu-Marinescu (2012) on the Bucharest Stock Exchange demonstrates that the normal distribution underestimates the likelihood of extreme events, whereas the application of stable models significantly enhances predictions of extreme events. Jama (2009) rejects the Gaussian model with evidence from the Johannesburg Stock Exchange (FTSE/JSE) and notes more reasonable results for VaR and Conditional VaR for stable models as compared to estimates derived from the Gaussian distribution. Numerous studies with empirical evidence suggest the normal distribution fails to adequately capture extreme returns often observed in financial markets. Thus, to overcome the shortcomings and inadequacies of the Gaussian approach, modelling financial asset returns using the family of stable models is suggested as a better alternative.

Initial descriptive statistics tests were conducted to ascertain the underlying nature of each return series. Each of the daily log returns of the three FTSE/JSE indices and the Dollar/South African exchange rate has shown to be non-stationary, with evidence of heteroskedasticity and volatility clustering. The Jarque-Bera test for normality rejects the assumption for normality, thus reinforcing the criticisms of the Gaussian approach highlighted by Pele and Mazurencu-Marinescu (2012). This has prompted the exploration of stable distributions as an alternative for financial modelling. The fitting of a statistical distribution usually presumes no serial correlation and heteroskedasticity. However, the empirical properties of financial returns, as recognized by McNeil et al. (2015) describes that some returns in financial data show serial correlation. This is the case with FTSE/JSE All Share Index and FTSE/JSE Mining Index where serial correlation is noticed. The maximum likelihood method was utilised to estimate stable parameters under Nolan's S_0-parameterization. Results from the Anderson-Darling test validated the suitable fit of stable distributions for all return series. VaR estimates were computed and the Kupiec likelihood test confirms the robustness of the stable model at most VaR levels. These findings emphasize the significance of fitting stable models when describing South African financial data where fat-tails and asymmetry are prevalent.

In this study, evidence for empirical properties of financial data with ARCH effects in reach return series implies the use of the GARCH-type framework with heavy-tailed distributions such as the GPD and Nolan's S_0-parameterization stable distribution. The suitability of the GARCH (1,1) model was confirmed for each return series, except for FTSE/JSE ALSI, where asymmetric effects were observed. The EGARCH (1,1) model was identified to best account for the volatility present in FTSE/JSE ALSI returns. The GDP model fitted to the tail distribution has been shown to capture asymmetry in the tail model in research by McNeil and Frey (2000). Tail innovations are estimated using a conditional method of GARCH-type models using heavy-tailed distributions recommended by extreme value theory.

Extreme behaviour, rare phenomena and disasters have long been of interest to researchers and analysts alike. The ability to model and predict these events is a powerful statistical tool. This study centers around the principles of Extreme Value

Theory and mixture models with a core focal point around stable distributions. The theory of extremes allows for the understanding of data complexities, especially in the tail model, and when combined with mixture models, the underlying data series is modelled in a nuanced manner that is more descriptive, leading to precise predictions. Stable distributions are largely justified in finance, due to the accuracy of capturing large fluctuations and various empirical properties such as heavy tails, asymmetry, and volatility, which are frequently found in financial data.

The over-arching aim of this thesis was to propose novel EVT stable mixture modelling methodologies to unpack insights within South African stock market indices and exchange rates. This was studied in detail by exploratory data analysis, sound stable modelling using the ML estimation procedure and applying the theory of extreme values practically. The stable package by Robust Analysis Inc. (2013) and evmix by Hu (2013) packages were used to explore mixture models in an EVT context.

Several conclusions can be obtained from the research outputs of this study. Referring to the aims outlined in Chapter 1, the following inferences can be made:

(i) Each data return series were analysed using the fitted univariate Nolan's S_0-parameterization stable distribution. This study substantiates the results of Nolan (2014) that stable distributions are a flexible class of probability laws that can sufficiently capture the characteristics of South African financial data.

(ii) The GARCH (1,1) model was applied to each return series, and the sign bias test showed asymmetry for FTSE/JSE ALSI returns only. Subsequently, the GPD and Nolan's S_0 -parameterization stable models were fitted to the best GARCH-type model residuals. Model diagnostics indicate that both the GARCH (1,1)-GPD and GARCH (1,1)-SD models provide a parsimonious fit to the South African financial data. VaR estimates and VaR in-sample backtesting using the Kupiec likelihood ratio test highlight the robustness of the fitted GARCH (1,1)-SD model in both the long and short positions.

(iii) The GNG and GKG, as well as the GSG models have shown to be an adequate fit of the underlying data series through the graphical fit evaluations either by observing the density plots, the AD test or by evaluation using the Kupiec likelihood ratio backtesting procedure.

(iv) The proposed EVT, SNS and SKS models were introduced, and model diagnostics show promising models that are flexible in nature in capturing the primary characteristics of heavy tails, skewness and volatility clustering in South African financial data.

Extreme value models and stable distributions have been shown, through previous research, to be invaluable in the contribution to financial analyses where the properties of heavy tails and skewness are prevalent. Designing and implementing innovative extreme value stable mixture models that combine the strengths of stable distributions and mixture models provides an improved representation of tail behaviour for extreme events. The developed models were applied to various empirical South African financial data sets to provide insights into tail risk analysis and demonstrate the practical utility of the proposed extreme value mixture models. Individuals concerned with extreme losses, such as risk analysts, insurers, policy makers and risk-averse investors, may find stable mixture models beneficial. This research notes that the versatility of stable distributions paves the way for future work in extreme value stable distribution mixture modelling.

7.2 Limitations of the study

The limitations and inadequacies of the traditional Gaussian model fitting approach can be discussed at length and often suggests the exploration of newer models as done by the work of Zhao (2010) and MacDonald et al. (2011), where the GNG and GKG models are investigated respectively. Bayesian inference is considered throughout the analysis, where the threshold is considered to be a parameter needed to be estimated.

Research by Mukhodobwane et al. (2022), Chinhamu and Chifurira (2019) and Makatjane and Molefe (2020) illustrate the satisfactory performance of the GPD model in the context of the South African financial industry. Hu (2013) notes that the GPD is commonly fitted at the tails of extreme value mixture models. While the GNG and GKG may be great model alternatives, Hu (2013) highlights a possible drawback where model mis-specification is probable. The Normal bulk distribution influences the tail fit if mis-specified. The shortcomings of the Gaussian framework are frequently emphasised in this study and thus leads to the research of applying a stable bulk model. Stable distributions have proven to be a flexible class of models that may remedy the issue of model mis-specification in the bulk component of EVT mixture models. In this regard, the GSG model is proposed as a possible alternative.

MacDonald et al. (2011) also highlights a drawback with proposed mixture models in industry, which is the mis-specification of the bulk model, which has a large impact of the fit of the tails. The non-parametric KDE model is applied to the bulk model to deal with computational and inference challenges. With this, new EVT mixture models, namely the SNS and SKS models, are proposed as possible alternative models to be explored since Nolan's S_0 stable distribution is a flexible class of distributions that account for asymmetry and heavy tails provides a reasonable model for tail behaviour. It is also important to note stable distributions have an absence of an analytical PDF with infinite and undefined moments making empirical analysis challenging. This study looks into novel models that prove to be sound alternatives to exploring Extreme Value theory in a South African financial context; however, with novelty comes limitations. The SNS and SKS models in the current state can only be considered in a piecewise manner and not in its entirety due to computational challenges as stable distributions describe a family of models; this includes the Gaussian model among other parametric distributions such as the Cauchy and Lévy model. Model validation also becomes a challenge due to computational complexities. The data sets used in this study have minimal influence on the methodologies of these models as the ideology behind the SNS and SKS models were derived from the extensive research and empirical analysis of univariate stable distributions that lead to a natural progression to explore mixture model applications. In essence, both the SNS and SKS models can be probed further with other datasets in various industries that exhibit heavy tails and skewness.

7.3 Recommendations

Based on the research findings, exploratory analysis, financial extreme value mixture modelling, the following recommendations can be made:

- The application of extreme value mixture models in achieving a better understanding of the South African financial industry where financial modelling risk strategies are applied and validated.
- The application of the SNS and SKS models across the African continent with similar economies or in other industries where the novel mixed models that depict real-world occurrences are required and valued.
- This study identifies that there is no statistical software available to estimate and evaluate the goodness-of-fit of the SNS and SKS mixture models; thus, as future work, the development of a software package is recommended.
- VaR and backtesting procedures for the proposed stable mixture models.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEET Transactions on Automatic Control*, *19*(6), 716–723. doi: 10.1109/tac.1974.1100705 .

Anderson, T. W., and Darling, D. A. (1954). A test of goodness of fit. *Journal of the American Statistical Association*, *49*(268), 765–769.

Atsmegiorgis, C., Kim, J., and Yoon, S. (2016). The GARCH-GPD in market risks modeling: An empirical exposition on KOSPI. *Journal of the Korean Data Information Science Society*, *27*(6), 1661–1671. http://dx.doi.org/10.7465/jkdi.2016.27.6.1661 .

Behrens, C. N., Lopes, H. F., and Gamerman, D. (2004). Bayesian analysis of extreme events with threshold estimation. *Statistical Modelling*, *4*(3), 227–244. https://doi.org/10.1191/1471082x04st075oa.

BIS (2016). Large exposures. [Accessed 05-07-2024].
URL `https://www.bis.org/basel_framework/standard/LEX.htm?tldate=20240705/`

Bobbitt, Z. (2024). The complete guide: How to interpret q-q plots. [Accessed 24-08-2021].
URL `https://www.statology.org/qq-plot-interpretation/`

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1.

Burnham, K. P., and Anderson, D. R. (2004). Multimodel inference: Understanding aic and bic in model selection. *Sociological Methods and Research*, *33*(2), 261–304. https://psycnet.apa.org/doi/10.1177/0049124104268644.

Campbell, S. D. (2006). A review of backtesting and backtesting procedures. *The Journal of Risk*, *9*(2), 1. http://dx.doi.org/10.21314/jor.2007.146.

Chambers, J. M. (2008). *Software for data analysis: Programming with R*, vol. 2. Springer.

Chan, K.-S., Ripley, B., Chan, M. K.-S., and Chan, S. (2022). Package 'tsa'. *R package version*, *1*.

Changole, A. (2019). *Bad news about economy proves too much for SA's bank stocks*. Finance 24.[Accessed 24-08-2021].

URL `https://www.news24.com/fin24/companies/bad-news-about-economy-proves-too-much-for-sas-bank-stocks-20190806`

Chinhamu, K., and Chifurira, R. (2019). Evaluating South Africa's market risk using asymmetric power auto-regressive conditional heteroscedastic model under heavy-tailed distributions. *Journal of Economic and Financial Sciences*, *12*(1), 1–11. https://doi.org/10.4102/jef.v12i1.475 .

Chinhamu, K., Huang, C.-K., and Chikobvu, D. (2015). Evaluating risk in gold prices with generalized hyperbolic and stable distributions. In *Annual Proceedings of the South African Statistical Association Conference*, 1. 17-24. South African Statistical Association (SASA).

Coles, S., Bawa, J., Trenner, L., and Dorazio, P. (2001). *An introduction to statistical modeling of extreme values*, vol. 208. Springer.

Coulibaly, B. D., Chaibi, G., and Mohammed, E. (2024). Parameters estimation of the alpha-stable distribution and applications to financial data. *Chilean Journal of Statistics*, *15*, 60–80.

De Melo Mendes, B. V., and Lopes, H. F. (2004). Data driven estimates for mixtures. *Computational Statistics and Data Analysis*, *47*(3), 583–598. http://dx.doi.org/10.1016/j.csda.2003.12.006.

De Wet, R., Smit, E., and Krige, J. (2007). An investigation into the return distributions of the South African general equity universe. *Management Dynamics: Journal of the Southern African Institute for Management Scientists*, *16*(3), 2–17.

Ding, Z., Granger, C. W., and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, *1*(1), 83–106.

Do Nascimento, F. F., Gamerman, D., and Lopes, H. F. (2012). A semiparametric bayesian approach to extreme value estimation. *Statistics and Computing*, *22*, 661–675. https://doi.org/10.1007/s11222-011-9270-z.

DuMouchel, W. (1971). *Stable distributions in statistical inference, unpublished PhD*. Ph.D. thesis, Yale University.

DuMouchel, W. H. (1973). On the asymptotic normality of the maximum-likelihood estimate when sampling from a stable distribution. *The Annals of Statistics*, *1(5)*, 948–957.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, *50(4)*. 987-1007.https://doi.org/10.2307/1912773.

Engle, R. F., and Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The Journal of Finance*, *48*(5), 1749–1778.

Escanciano, J. C., and Olmo, J. (2010). Backtesting parametric value-at-risk with estimation risk. *Journal of Business and Economic Statistics*, *28*(1), 36–51. https://doi.org/10.1198/jbes.2009.07063.

Evans, D. L., Drew, J. H., and Leemis, L. M. (2008). The distribution of the Kolmogorov–Smirnov, Cramer–von Mises, and Anderson–Darling test statistics for exponential populations with estimated parameters. *Communications in Statistics—Simulation and Computation®*, *37*(7), 1396–1421.

Frigessi, A., Haug, O., and Rue, H. (2002). A dynamic mixture model for unsupervised tail estimation without threshold selection. *Extremes*, *5*, 219–235.

Ghalanos, A., Ghalanos, M. A., and Rcpp, L. (2018). Package rugarch. *R Team Cooperation,43*.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*(5), 1779–1801.

Gujarati, D. N., and Porter, D. C. (2009). *Basic Econometrics*. McGraw-Hill.

Hu, Y. (2013). *Extreme value mixture modelling with simulation study and applications in finance and insurance*. Master's thesis, University of Canterbury.

Hu, Y., and Scarrott, C. (2018). evmix: An r package for extreme value mixture modeling, threshold estimation and boundary corrected kernel density estimation. *Journal of Statistical Software*, *84*(5), 1–27. https://doi.org/10.18637/jss.v084.i05.

Huang, Z., Lam, H., LeBlanc, D. J., and Zhao, D. (2017). Accelerated evaluation of automated vehicles using piecewise mixture models. *IEEE Transactions on Intelligent Transportation Systems*, *19*(9), 2845–2855.

Ilupeju, Y. E. (2016). *Modelling South Africa's market risk using the APARCH model and heavy-tailed distributions*. Master's thesis, University of KwaZulu-Natal.

Jama, S. (2009). *An alternative model for multivariate stable distributions*. Master's thesis, University of Cape Town. http://hdl.handle.net/11427/8959.

Kallah-Dagadu, G. (2013). *Modelling Ghana stock exchange Indicies and exchange rates with stable distributions*. Master's thesis, University of Ghana.

Kateregga, M. (2017). *Stable processes: theory and applications in finance*. Ph.D. thesis, University of Cape Town. http://hdl.handle.net/11427/27069.

Kateregga, M., Mataramvura, S., and Taylor, D. (2017). Parameter estimation for stable distributions with application to commodity futures log-returns. *Cogent Economics and Finance*, *5*(1), 1318813.

Katsenga, G. Z. (2013). *Value at Risk (VaR) Backtesting:'evidence from a South African Market Portfolio*. Master's thesis, University of Witwatersrand.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, *3*, 73–84.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, *54*(1-3), 159–178.

Lee, D., Li, W. K., and Wong, T. S. T. (2012). Modeling insurance claims via a mixture exponential model combined with peaks-over-threshold approach. *Insurance: Mathematics and Economics*, *51*(3), 538–550.

Liu, T., and Shi, Y. (2022). Innovation of the component garch model: simulation evidence and application on the chinese stock market. *Mathematics*, *10*(11), 1903.

MacDonald, A., Scarrott, C. J., Lee, D., Darlow, B., Reale, M., and Russell, G. (2011). A flexible extreme value mixture model. *Computational Statistics and Data Analysis*, *55*(6), 2137–2157.

Makatjane, K. D., and Molefe, E. K. (2020). Predicting Regime Shifts in Johannesburg Stock Exchange All-Share Index (JSE-ALSI): A Markov-Switching Approach. *Eurasian Journal of Economics and Finance*, *8*(2), 95–103.

McCulloch, J. H. (1986). Simple consistent estimators of stable distribution parameters. *Communications in Statistics-Simulation and Computation*, *15*(4), 1109–1136.

McCulloch, J. H. (1997). Measuring tail thickness to estimate the stable index α: A critique. *Journal of Business and Economic Statistics*, *15*(1), 74–81.

McNeil, A. J., and Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, *7*(3-4), 271–300.

McNeil, A. J., Frey, R., and Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques and tools-revised edition*. Princeton University Press.

Mukhodobwane, R. M., Sigauke, C., Chagwiza, W., and Garira, W. (2022). Stochastic modelling of the brics equity markets' risks. *Journal of Statistics Applications Probability*, *11*(1), 215–239.

Naradh, K. (2016). *Multivariate elliptically contoured stable distributions with applications to BRICS financial data*. Master's thesis, University of KwaZulu-Natal. http://hdl.handle.net/10413/15527.

Nguyen, T. T., and Sampson, A. R. (1991). A note on characterizations of multivariate stable distributions. *Annals of the Institute of Statistical Mathematics*, *43*(4), 793–801.

Nolan, J. P. (2001). Maximum likelihood estimation and diagnostics for stable distributions. In *Lévy processes*. Springer, 379-400.

Nolan, J. P. (2003). Modeling financial data with stable distributions. *Handbook of Heavy Tailed Distributions in Finance, Handbooks in Finance: Book*, *1*, 105–130.

Nolan, J. P. (2005). Multivariate stable densities and distribution functions: General and elliptical case. In *Deutsche Bundesbank s 2005 annual fall conference*.

Nolan, J. P. (2014). Financial modeling with heavy-tailed stable distributions. *WIREs Computational Statistics*, *6*(1), 45–55.

Nolan, J. P. (2015). *Stable distributions - Models for Heavy Tailed Data*. Birkhauser.

Ojeda, D. (2001). *Comparison of stable estimators*. Ph.D. thesis, American University.

Pele, D. T., and Mazurencu-Marinescu, M. (2012). Modelling stock market crashes: The case of Bucharest Stock Exchange. *Procedia-Social and Behavioral Sciences*, *58*, 533–542.

Poon, S.-H. (2005). *A practical guide to forecasting financial market volatility*. John Wiley and Sons.

Qiu, Y. (2024). Estimation of tail risk measures in finance: Approaches to extreme value mixture modeling. *arXiv preprint arXiv:2407.05933*.

R Core Team (2016). *graphics: The R Graphics Package*. R package version 3.3.2. URL `https://cran.r-project.org/web/packages/graphics/index.html`

Robust Analysis Inc. (2013). *User Manual for STABLE 5.3*. R package version 5.3.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6(2)*. 461-464.https://www.jstor.org/stable/2958889.

StataCorp (2015). user's guide-stata. Release 18. stata Press. [Accessed 23-06-2024]. URL `http://www.stata.com/manuals14/u.pdfAccessed(22June2024)`

Stoyanov, S. V., Rachev, S. T., Racheva-Yotova, B., and Fabozzi, F. J. (2011). Fat-tailed models for risk estimation. *Journal of Portfolio Management*, *37*(2), 107.

Syczewska, E. M. (2010). Empirical power of the Kwiatkowski-Phillips-Schmidt-Shin test. [Working Papers] 45, Department of Applied Econometrics, Warsaw School of Economics.
URL `https://ideas.repec.org/p/wse/wpaper/45.html`

Team, R. C., Team, M. R. C., Suggests, M., and Matrix, S. (2018). Package stats. *The R Stats Package*.

Teimouri, M. (2020). Maximum likelihood estimators for -stable distribution. *Journal of Statistical Sciences*, *14*, 75–96.

Trapletti, A., Hornik, K., LeBaron, B., and Hornik, M. K. (2015). Package 'tseries'. *R project*.

Tsay, R. S. (2005). *Analysis of financial time series*, vol. 543. John Wiley and Sons.

Tsay, R. S. (2014). *An introduction to analysis of financial data with R*. John Wiley and Sons.

Wand, M. P., and Jones, M. C. (1995). *Kernel smoothing*. Taylor Francis.

Wei, W. (2006). *Time Series Analysis: univariate and multivariate methods*. Pearson Addison Wesley.

Wuertz, D., Setz, T., Chalabi, Y., Maechler, M., and Setz, M. T. (2017). Package 'fbasics'. *Rmetrics-markets and basic statistics. R foundation for statistical computing*, *100*.

Xu, W., Wu, C., Dong, Y., and Xiao, W. (2011). Modeling Chinese stock returns with stable distribution. *Mathematical and Computer Modelling*, *54*(1-2), 610–617. https://doi.org/10.1016/j.mcm.2011.03.004.

Yang, G. (2012a). *The energy goodness-of-fit test for univariate stable distributions*. Ph.D. thesis, Bowling Green State University.

Yang, Y. (2012b). Parameter Estimation of GARCH Model. *Optimization*, *1*, 1.

Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, *18*(5), 931–955. https://doi.org/10.1016/0165-1889(94)90039-6.

Zhao, X. (2010). *Extreme value modelling with application in finance and neonatal research*. Ph.D. thesis, University of Canterbury.

Zhao, X., Scarrott, C., Oxley, L., and Reale, M. (2010). Extreme value modelling for forecasting market crisis impacts. *Applied Financial Economics*, *20*(1-2), 63–72. http://dx.doi.org/10.1080/09603100903262947.

Zolotarev, V. M. (1986). *One-dimensional stable distributions*, vol. 65. [Translations of Mathematical Monographs Vol. 65] American Mathematical Society.