Estimation of the value at risk using a long-memory GARCH application to JSE Indices

UNIVERSITY OF
KWAZULU - NATAL

INYUVESI
YAKWAZULU-NATALI

Moses Bhekinhlahla Khumalo: 200276813

June, 2020

Estimation of the value at risk using a long-memory GARCH application to JSE Indices

by

Moses Bhekinhlahla Khumalo: 200276813

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

Thesis Supervisors: Dr K. Chinhamu
Dr R. Chifurira

UNIVERSITY OF KWAZULU-NATAL
INYUVESI
YAKWAZULU-NATALI

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

Declaration - Plagiarism

I, Moses Bhekinhlahla Khumalo: 200276813, 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 acknowlegded 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 re-written 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.

Moses Bhekinhlahla Khumalo: 200276813 (Student) Date

Dr K. Chinhamu (Supervisor) Date

Dr R. Chifurira (Co-Supervisor) Date

Disclaimer

This document describes work undertaken as a Masters 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

Financial data are characterized by stylized facts; this makes it difficult to model financial assets if these stylized facts are not taken into account. Therefore, the implementation of accurate risk management tools such as value at risk (VaR), which is crucial in the management of market risk, becomes a futile exercise. This study aims to compare the performance of the long-memory GARCH-type models with heavy-tailed innovations in estimating the value at risk of the All Share Index, the Mining Index, and the Banking Index. This was achieved by investigating the empirical properties of the JSE Indices, fitting the FIGARCH, HYGARCH, and FIAPARCH with the Student's t-distribution (STD), skewed Student's t-distribution (SSTD), and generalized error distribution (GED). The study further estimates VaR for the short and long-trading positions on the 95th, 99th, and 99,7th quantiles, as well as backtests the results. The main findings indicate that the JSE All Share index returns is best captured by the FIGARCH-SSTD model, whereas the JSE Mining Index retuns most robust model is the FIAPARCH-STD model. For the JSE Banking Index returns, the FIAPARCH-STD model is predominantly appropriate at most of different VaR levels. The findings of the study provide a solution to both risk practitioners and asset managers for better understanding the behaviour of the financial indices' returns. Finally, this can assist the role players in fastidiously managing risks and assets' returns.

Acknowledgements

This has been an electrifying journey to me. I express my sincere gratitude to God (the Father, the Son Jesus, and the Holy Spirit) and will forever remain grateful to God, who stood by me. My supervisor, Dr. Knowledge Chinhamu, and co-supervisor, Dr. Retius Chifurira, deserve special mention for their support. Although you have a very tight agenda, you always took ample time to give my dissertation a ”thorough bleeding”. I want to thank these people who have provided my life with a vehicle to forge ahead. In no precise order in this regard are Khaya Madela, Petros Gumede, Dr. Ntokozo Mthembu, Mthoko Nzama, Sivu Ngwane, Jim Matsemela, and Nkosana Kgasi. You gave me comfort and support, helped me and were company in all good and bad moments of my life. Thank you for opening your homes, virtually and physically, to me and allowing me in sharing so many special occasions with you and your families.

Dedication

I cannot utter enough words to express my gratitude for having such a wonderful wife, Thandazile, and her efforts remain fresh in my memory. Your constant love and motivations were like a tonic that provided strength to steady progress in my study. My special thanks go to my girls, Simphiwe, Nokulunga, and Zamantungwa, for their understanding throughout my absences, even when I was in their midst; my girls have grown accustomed to the reality that I might be in their presence, but far away in my thoughts. I dedicate this study to my Dad, Vukani, whose shortened life prevents me from sharing this and so much more with him, and my mom, Jeslina, who always encouraged me to pursue things that I found fascinating. My siblings, Skhu and Jabulile, thank you for supporting me. I know I hardly say this; however, I love you. I pay tribute to my grandparents, Albert and Lydia, for their tenacity and vision. They were not educated, but they valued schooling and understood the equalizing power it grants. I thank you, Gogo, and I can confirm that you made a significant contribution to my success: I am who I am today because of you.

Contents

List of Figures

List of Tables

Abbreviations

ACF:	Autocorrelation function
ADF:	Augmented Dickey-Fuller
AIC:	Akaike information criteria
APARCH:	Asymmetric power autoregressive conditional heteroskedasticity
AR:	Autoregressive
ARCH:	Autoregressive conditional heteroskedasticity
ARCH LM:	Autoregressive conditional heteroskedasticity Lagrange multiplier
ARFIMA:	Autoregressive fractionally integrated moving average
ARMA:	Autoregressive moving average
BBM:	Ballie, Bollerslev and Mikkelsen
BIC:	Bayesian information criterion
CDF:	Cumulative distribution function
CLT:	Central limit theorem
df:	Degrees of freedom
EGARCH:	Exponential generalized autoregressive conditional heteroskedasticity
ES:	Expected shortfall
EVT:	Extreme value theory
EWMA:	Exponentially weighted moving average
fBm:	Fractional Brownian motion
fGn:	Fractional Gaussian noise
FIAPARCH:	Fractionally integrated asymmetric power autoregressive conditional heteroskedasticity
FIEGARCH:	Fractionally integrated exponential generalized autoregressive conditional heteroskedasticity
FIGARCH:	Fractionally integrated generalized autoregressive conditional heteroskedasticity
FTSE:	Financial Times Stock Exchange
GARCH:	Generalized autoregressive conditional heteroskedasticity
GJRGARCH:	Glosten-Jagannathan-Runkle generalized autoregressive conditional heteroskedasticity
GPD:	Generalized Pareto distribution
GPH:	Geweke and Porter-Hudak
H:	Hurst exponent
HYGARCH:	Hyperbolic generalized autoregressive conditional heteroskedasticity
HGARCH:	New generalized autoregressive conditional heteroskedasticity
IGARCH:	Integrated generalized autoregressive conditional heteroskedasticity
i.i.d:	Independent and identically distributed

JALSH:	Johannesburg Stock Exchange Africa All Share Index
JB:	Jarque-Bera
JSE:	Johannesburg Stock Exchange
JSEALSI:	Johannesburg Stock Exchange All Share Index
JSEBNKS:	Johannesburg Stock Exchange Banking Index
JSEJMNNG:IND:	Johannesburg Stock Exchange Africa Mining Index
KPSS:	Kwiatkowski- Phillips- Schmidt- Shin
LM:	Lagrange multiplier
LM:	Long-memory
MAPE:	Mean absolute percentage error
MAE:	Mean absolute error
ML:	Maximum likelihood
MLE:	Maximum likelihood estimation
MSE:	Mean square error
MZ:	Mincer-Zarnowitz
PACF:	Partial autocorrelation function
PDF:	Probability density function
PIVD	Pearson type-IV distribution
PP:	Phillips-Perron
QMLE:	Quasi-Maximum likelihood Estimator
Q-Q:	Quantile-quantile
RMSE:	Root mean square error
RS:	Range scale
SDF:	Spectral density function
SSR:	Residuals sum of squares
T.Dist:	T-distribution
TGARCH:	Threshold generalized autoregressive conditional heteroskedasticity
VaR:	Value at risk

Chapter 1

Introduction

In this chapter, the background of this study, a review of the relevant literature, the stylized facts of the financial asset returns, the research problem, the aims and objectives, and the importance of the study are discussed.

1.1 Background of the study

The empirical studies in finance emphasize some stylized facts such as long memory process, excess volatility, volatility clustering, heavy-tails, and asymmetry in the asset returns. Volatility in the financial assets' returns has been viewed as an indicator of vulnerability of the financial markets, where volatility measure has been essential for value at risk modeling. In many cases, the value at risk (VaR) assumes the normality, and the crucial setback of the assumption is that model disregards the presence of heavy-tailed and skewed characteristics in the return distributions.

According to Jorion, (2007), VaR represents the maximum expected loss for a given time horizon and a pre-specified confidence interval under normal market conditions. Furthermore, the unconditional volatility models assume that the variance is constant over time, which poses challenges in VaR estimation. The researchers have established that volatility cannot assume homoskedasticity as it evolves over time and shocks persist for a long time. In the early 80's an introduction of Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models was due to the need to adequately capture some stylized facts observed on financial markets returns, more specifically the time-dependent volatility property.

According to Klar et al., (2012), the basic GARCH model is symmetrical and does not capture the asymmetric impact called the leverage effects that is inherent within the most market stock returns. In the financial time series context, the asymmetry impact on the asset price refers to when bad news tends to increase volatility more than good news (Black, 1976). Furthermore, the asymmetric GARCH models show significant evidence of symmetry in stock returns, confirming the presence of a leverage effect in the return series. In the recent past, GARCH-type models combined with heavy-tailed distributions, to model the VaR, have become an integral part of theoretical and time series data analysis to capture the stylized facts, in particular, the findings of heavy-tailed property observed in asset returns (Chifurira and Chinhamu, 2019).

The availability of high-frequency data for financial markets, another common behaviour, has posed another challenge in estimating VaR. High-frequency data usually show the presence of long memory patterns, which is the squared returns (Caporin, 2003). Long memory in volatility refers to a slow hyperbolic decay in autocorrelation functions of the squared returns i.e. autocorrelation function of squared returns slowly convergent toward zero. The squared returns are employed as a proxy of the volatility of the returns in this study. The autocorrelation of the squared returns seems to decay at a slower rate, slow hyperbolic rate of decay in the autocorrelation of squared returns. This structure is not compatible with the basic GARCH model, whose implied theoretical autocorrelations exponentially converge to zero. Most of the GARCH models employed in many types of research do not account for long memory in volatility (Wojtowicz and Gurgul, 2009).

The Integrated Generalized Autoregressive Conditional Heteroskedasticity (IGARCH) model was proposed by Engle and Bollerslev (1986) to address the deficiencies of the GARCH models, hence it is persistent in variance given that the current information remains essential for forecasts on all horizons. The Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity (FIGARCH) model shows some flexibility processes for the conditional variance, that are accomplished in explaining and representing the observed temporal dependencies in financial market volatility. It permits only a slow hyperbolic rate of decay for the lagged squared returns in the conditional variance function, and over and above accommodate the time dependence of the variance and a leptokurtic unconditional distribution for the returns with a long memory behaviour for the conditional variance (Tayeti and Rammanathan, 2012).

The Fractionally Integrated Asymmetric Power Autoregressive Conditional Heteroskedasticity (FIAPARCH) model was an improvement of the FIGARCH model, even though the FIGARCH model is capable to capture the long memory process, the FIAPARCH captures asymmetric response of volatility to positive and negative shocks which FI-

GARCH is not capable of, and also permitting the determination of the power of the returns for which the predictable structure in the volatility is the strongest. The Hyperbolic Generalized Autoregressive Conditional Heteroskedasticity (HYGARCH) model improves the FIGARCH as proposed by Davidson (2004), by introducing the hyperbolically decaying response coefficients which address the weakness of the FIGARCH. Furthermore, the presence of the long memory structure will (theoretically) allow the computation of consistent long term forecasts of the volatility. News impact is also affected as in such a model shocks will produce their effects far in the future, even if not indefinitely as in the IGARCH case. This particular behaviour is due to the long memory pattern of the conditional variances. A central issue will, therefore, be choice of the correct model for the series of interest, distinguishing among the possible short memory (GARCH), and long memory (FIGARCH) specifications for the variances.

The evidence of long memory in volatility within the South African markets is analyzed using three prominent JSE indices, in highlighting the presence of the long memory properties in the volatility of the returns. In South Africa, it seems to have little previous research establishing the existence of its long memory properties. In order to ascertain the presence of the long memory of the indices, this study employs the long memory tests, which is measured by the d parameter. The presence of the long memory in these indices will suggest the use of long memory volatility GARCH models. The crucial area of the study is the computation of the value at risk, and this approach is used by financial and regulatory agencies (i.e. SARB) to measure, monitor, and manage market risk factors. Several recent types of research ignore the long memory which leads to an underestimation of risk posed by volatility in the returns, and only the heavy-tailed distributions can capture the tail risk.

The South African indices used are JSE All Share index (JSEALSI), JSE Mining index (JSEJMMNG), and JSE Banking index (JSEBNKS), and are highly volatile. This research study therefore uses three major indices listed in the JSE. The overall performance of South Africa's stock market is measured by the JSEALSI, JSEJMMNG, and JSEBNKS. According to the JSE Bulletin (2006), the indices offer an easy way to determine the overall performance of the stock market over some time. The JSEALSI is an index that reflects the movement on the equity market and accounts for 99% of the full market capital value (i.e., before the application of any investability weighting) of all ordinary securities listed on the Main Board of the JSE that qualify under the regulations of eligibility. The JSE All Share Index contains the main securities, measured by way of market capitalization, and includes 166 listed companies across 41 sectors.

The Banking Index is a market capitalization-weighted index whose mother index is JSE All Share index. The index assumed ICB classification (effective since 3 January 2006); historical values before 2006 were related to the FTSE GCS classification. Capitec, Barclays Africa Group, Standard Bank Group, First Rand, RMB Holdings, and the Nedbank Group are examples of large companies on this market capitalization-weighted index. The JSE Mining Index is a market capitalization weighted index whose mother index is JALSH. It involves platinum, palladium, gold, coal, and iron ore. A short background of these indices, as well as investigating the mechanisms in which data returns from the time series data are generated, became part of the research, given the significance to investors for establishing a way to look ahead.

1.2 Stylized facts of financial asset returns

According to Cont (2001), stylized facts concern taking a common view of the properties ascertained in studies of various markets and instruments. The stylized facts of the financial returns are the following:

- **Volatility clustering** is a measure of volatility showing that a positive autocorrelation quantifies the fact that high-volatility events tend to cluster in time over several days. This implies that a period of low volatility follows other periods of low volatility.
- **Long-memory** is the statistical dependence that decays more slowly as oppose to an exponential decay.
- **Leverage effect** displays a negative correlation between the volatility of an index and the returns of the indices; in simpler terms, this indicates that volatility rises when index prices go down and decrease when index prices go up. The consequence of bad news and index market volatility is greater than the consequence of good news
- **Slow decay of autocorrelation** is a process during which the autocorrelation function of absolute returns decays slowly as a function of the time lag, roughly as a power law with an exponent $\beta \varepsilon [0.2, 0.4]$. This is, on occasion, interpreted as a signal of long-range dependence
- **Heavy-tails** are characterized by the distribution's returns that appear to show a power-law or Pareto-like tail with a finite tail index greater than two, but much less than five for most data sets studied. In particular, this excludes stable laws with infinite variance and normal distribution. The kurtosis tends

to be extremely sensitive to outliers, which implies that the distribution with such features is non-normal

- **Conditional heavy-tails** occur as a result of the correction of the returns exhibiting volatility clustering, but the residuals time series nonetheless show heavy tails. However, the tails are much less heavy than in the unconditional distribution of returns; and

- **Absence of autocorrelation** takes place when the autocorrelation of asset returns is regularly insignificant, barring very small intraday time scales in which microstructure effects come into play.

1.3 Relevant literature review

In this section, related literature on the JSE indices used in the study will be discussed, namely the All Share Index, Mining Index, and Banking Index.

Mabrouk, (2016) attempts to conclude the comparisons of the capabilities of the two integrated GARCH models with a FIAPARCH, employing skew Students' t-distribution, and the results are primarily based entirely on MZ regression approach (Mincer and Zarnowitz, 1969). The results confirm that the FIAPARCH model has an extremely better property by increasing the flexibility of the conditional variance. Chkili et al. (2014) investigated the consequences of asymmetric behaviour and long-memory process models in forecasting the conditional variance of the market risk factors.

In their study, they used four widely listed commodities, namely gold, gas, crude oil, and silver, and the study confirms that the FIAPARCH model is superior. Li et al., (2015) used the daily Hang Seng Index and Koren Won to compare the performances of the three innovations, namely HYGARCH, New Generalized Autoregressive Conditional Heteroskedasticity (HGARCH), and FIAPARCH to estimate the value at risk (according to the internet). Their conclusion was that the FIGARCH innovation with normal distribution is a robust model to estimate the value at risk forecasts. Sethaparamote et al. (2014) concluded that value-at-risk estimations using the FIGARCH model with normal distribution are accurate, as opposed to those generated by the short-memory GARCH model that appear in the analysis of the accuracy of the value-at-risk estimation in the study of the Stock Exchange of Thailand.

Yaya (2013) attempted to establish the most effective GARCH model for the All Share Index of the Nigerian Stock Exchange by using HYGARCH, IGARCH, and Fractionally Integrated Asymmetric Power Autoregressive Conditional Heteroskedasticity (FIEGARCH). Supported by the normality assumptions, HYGARCH parameters were not considerably exclusive; the HYGARCH was unsuccessful and reverted to FIGARCH. However, the HYGARCH model exhibits extraordinary stability across pre-crises and post-crises. In their study of the stock exchange of the United Arab Emirates, Maghyereh and Awartani (2012) confirmed that value at risk accuracy improves once they used FIAPARCH applied with a skewed Student's t-distribution. Mighri et al., (2010) argues against the proposition given by Kasman (2009); the author proves his assertion by studying the consequences of asymmetric long-memory volatility models on estimating value at risk by using stock index returns. The findings are that the VaR model produces better results in both in-sample and out-sample procedures. Reddy et al.(2017) explore the behavior of the daily JSEALSI returns using GARCH, IGARCH, GJR-GARCH, and FIGARCH, they conclude that out-sample results consequences exhibit that structural break model outperforms all static GARCH models over the forecasting horizon, and GJR-GARCH model stands to be the robust model. The literature on long-memory GARCH models is presented in the Table 1.1:

Table 1.1: Literature on long-memory GARCH-type models

Author	Data	Methodology	Robust Models
Mabrouk(2016)	Daily crude oil,gas	FIGARCH, HYGARCH, and FIAPARCH model	FIAPARCH
Reddy et al.,(2017)	Daily closing JSE All Share Index(03/01/00-31/12/12)	IGARCH,FIGARCH, GJR-GARCH	GJR- GARCH
Chkili et al,(2014)	Daily gas,oil,crude oil,silver(07/01/97-31/12/09)	FIAPARCH,IGARCH,and FIGARCH	FIAPARCH
Li et al.,(2015)	Daily Heng Seng Index,Koren Won(31/12/86-7/01/10)	HYGARCH, HGARCH,and FIGARCH	HYGARCH
Sethaparamote et al.,(2014)	Daily Thailand SE	FIAPARCH,FIGARCH,and HYGARCH	FIGARCH/Normal distribution
Yaya(2013)	Daily Nigerian All Share Index(01/2007-12/2011)	HYGARCH,APARCH,and FIGARCH	HYGARCH
Maghyereh and Awartami(2012)	Daily United Arab Emirates Stock Exchange(31/12/03-30/06/09)	FIAPARCH	FIAPARCH/skewed student's t-distribution.
Mighri et al.,(2010)	Thailand returns daily(02/01/97-25/08/08)	FIAPARCH,FIGARCH,and HYGARCH	FIAPARCH

Chifurira and Chinhamu (2017) conducted a study employing the daily JSE Mining Index by using generalized Pareto and Pearson type-IV distributions to measure value at risk. Their findings concluded that the models provided better results in comparison with generalized hyperbolic distributions. Reddy et al. (2017) explored the behaviour of daily JSE All Share Index returns by investigating the empirical evidence of structural breaks in stock return volatility. Their conclusion was that leverage effects have to take a quadratic shape when utilized for South Africa's Equity Market, given the most suitable overall performance of Glosten, Jagannathan, Runkle-GARCH models.

Elenjical et al. (2016) revealed that the daily JSE All Share Index indicated a need for the implementation of model-switching policies that can provide significant improvements in forecasting and minimizing chances of value at risk estimates falling short of actual losses. Katzke and Garbers (2015) used the GARCH (1,1) model to investigate the existence of asymmetric volatility in the daily JSE All Share Index. The results ensured that controlling asymmetries and long memory in volatility models improved risk management calculation. Makhwiting et al. (2014) applied the daily JSE All Share Index with generalized Pareto distribution; empirical results highlighted that the ARMA-GARCH generalized Pareto distribution model produced more accurate estimates of extreme returns than the ARMA-GARCH model.

Huang et al. (2014a) also used the daily JSE All Share Index for subsequent probability distributions, particularly hyperbolic distribution, generalized extreme value distribution, and generalized Pareto distribution. Their findings concluded that the most effective model choice was not variant; hence, neither the extreme value theory nor the generalized hyperbolic distribution always produced the effective fit. Toerien et al. (2014) studied the reverse probability distribution model using the daily JSE All Share Index. The results showed that the matched-pair target, ranging from ± two to ± eight, and the maximum probability of the negative return on a timeline laid to the left of the maximum probability of the equivalent positive.
Using the IGARCH, Exponentially Generalized Autoregressive Conditional Heteroskedasticity (EGARCH), Threshold Generalized Autoregressive Conditional Heteroskedasticity (TGARCH), FIGARCH and FIEGARCH models and the VaR measure on the daily JSE All Share Index, McMillan and Thupayagale (2010) established that results generated by FIEGARCH proved to incorporate both asymmetric and long-memory attributes, thus generally outperforming all other methods in estimating value at risk across three percentiles considered 95%, 97.5%, and 99%. This study differs in several respects, as the FIGARCH, HYGARCH and FIAPARCH models, combined with Student's t-, skewed Student's t-, and generalized error distribution

have been employed. Bonga-Bonga and Mutema (2009) analyzed the daily JSE All Share Index returns by using GARCH and EGARCH models. The results showed that assumptions of conditional heteroskedasticity ought to have been taken into consideration when estimating the value at risk in emerging markets.
The Table 1.2 presents a literature on the JSE financial data used in this research.

Table 1.2: JSE financial data used in this research

Author	Data	Models Fitted	Robust Models
Chifurira and Chinhamu(2017)	Daily JSE Mining(20/02/01-30/12/18)	GARCH-PIVD/GARCH-GPD	GARCH-GPD/GARCH-Pearson type-IV
Reddy et al.,(2017)	Daily JSE All Share Index(03/01/00-31/12/12)	IGARCH,GJRGARCH,FIGARCH	GJRGARCH
Elenjical et al.,(2016)	Daily JSE All Share Index(03/01/03-17/08/12)	IGARCH,EGARCH,and FIEGARCH	EGARCH
Katzke and Garbers(2015)	Daily JSE All Share Index	APARCH,FIGARCH,and FIAPARCH	FIGARCH
Makhwiting et al.,(2014)	Daily JSE All Share Index(2002-2011)	Generalized Pareto, Generalized Extreme Value	Weibull distribution
Huang et al.,(2014a)	Daily JSE All Share Index(17/12/03-17/12/13)	Generalized Hyperbolic, Normal Inverse Gaussian, Generalized Pareto, and Generalized Extreme Value	Mixture distributions/Model-switching procedures.
Huang et al.,(2014)	Daily JSE Mining Index(02/01/01-22/08/13)	Hyperbolic, Normal Inverse Gaussian, and Generalized Skew-*t*	GHSSTD
Toerin et al.,(2014)	Daily JSE All Share Index(1995-2012)	Reverse Probability distribution	Reverse statistics probability distributions.
McMillan and Thupayagale(2010)	Daily JSE All Share Index	FIGARCH,EGARCH,and FIEGARCH	EGARCH.
Bonga Bonga and Mutema(2009)	Daily JSEALSI	GARCH,EGARCH	EGRACH

From literature reviewed, there seems to be limited literature on, modeling the JSE All Share Index, JSE Mining Index and JSE Banking Index using FIGARCH, HYGARCH and FIAPARCH in the South African context. The literature further reveals that the JSE Banking Index has never been used in any study in South Africa's financial market. Reddy et al., 2017 focused on the GARCH(1,1), Glosten, Jagannathan, and Runkle-GARCH (1,1), and FIGARCH (1,1) combined with Student's t- distribution to model the value at risk. They concluded that Glosten, Jagannathan, and Runkle-GARCH (1,1) (GJR-GARCH) combined Student's t- distribution was the robust model. In this study, we aim to improve on the work of Reddy et al., (2017) by using FIGARCH, HYGARCH and FIAPARCH with heavy-tailed distribution on the All-Share Index, Mining Index and Banking Index combined with heavy tailed distributions.

1.4 Research problem

The financial data returns are characterized by volatility clustering, leptokurtosis, leverage effects, and long memory. The research in time series has attracted enormous interest in recent years. This has led to intensive studies in models that have the ability to capture the properties of the stylized facts. In an attempt to solve and contribute to the body of knowledge, this study presents a long memory in volatility to model and forecast the volatilities of JSE indices mentioned in the background of the study.

1.5 Aim and objectives of study

The study aims to examine the overall performance of the long-memory GARCH-type models with heavy-tailed innovations in estimating the value at risk of the All Share Index, Mining Index and Banking Index. This is achieved by:

- investigating the empirical parameters of the JSE Indices
- fitting the FIGARCH, HYGARCH, and FIAPARCH combined with Student's t-, skewed Student's t-, and generalized error distributions
- estimating the value at risk of short and long trading positions on 95^{th}, 99^{th}, and 99.7^{th} quantiles; and
- backtesting the results as well as comparing the relative performance of the fitted models at different VaR levels.

1.6 Significance of the study

The financial turbulence in South Africa's stock market resulted in the designing of extra sophisticated risk measurement tools in order to aid risk management. The active participation of foreign investors in South Africa's financial markets provided important initiatives to improve risk measurement tools. These tools play a significant role in understanding various financial assets, including indices' returns behaviour. These indices are used as a gauge of the financial markets, which are characterized by high levels of volatilities. The volatility is directly associated with risk, and value at risk is widely considered to be effective to measure the risks. The financial indices used in this study account for specific behaviour of the returns series, which revealed heavy-tailed distributions to model the value at risk. The results of this study will be of interest to risk managers, researchers and financial statisticians.

1.7 Research study structure

This research thesis is structured as follows. **Chapter 1** presents the introduction, background of study , related literature review stylized facts of financial asset returns, research problem, aims and objectives and significance of the study including the research study structure. **Chapter 2** delves into econometrics models related to the research topic, and giving deeper perception into a theoretical review of GARCH-type models, heavy-tailed distributions, parameter estimation methods.

Chapter 3 discusses the methods used to check for empirical properties of the data, model adequacy and backtesting VaR. **Chapter 4** gives insight to the empirical results based on the chosen time series data. **Chapter 5** concludes the research and presents recommendations on some ideas.

Chapter 2

Long process volatility models

2.1 Introduction

In this chapter, the theory of long-memory GARCH-type models used in the study is provided. The heavy-tailed distributions used in the study are also discussed.

2.2 Long-memory process

Long-memory behaviour is important in the way time-series returns modeling can be approached, given that it has a significant effect on the financial world. To differentiate between short- and long-memory processes, the following definitions are given below, time series data have a short memory if:

$$\sum_{i=1}^{n} |\rho(h)| < \infty, \tag{2.1}$$

then the following exists:

$$\sum_{i=1}^{n} |\rho(h)| = \infty, \tag{2.2}$$

then the sequence is said to exhibit long-range dependence, where n is the total number of time series realizations.
Let, $y_1, y_2, \ldots, y_n$ be a stochastically generated observations, the mean is given by

$$\bar{y}_n = \frac{y_1 + y_2 + \cdots + y_n}{n}, \tag{2.3}$$

then the sample variance of the series is given by:

$$var(\bar{y}_n) = \frac{1}{n}\left(1 - \frac{|h|}{n}\right)\rho(h) = \frac{1}{n}\sum_{h=-\infty}^{\infty}\left(1 - \frac{|h|}{n}\right)_{+}\rho(h), \tag{2.4}$$

where $(a)_+ = max(a,0)$ is a if $a \geq 0$ and 0 if $a < 0$. If $\sum |\rho(h)| < \infty$, then:

$$var(\bar{y}_n) = \frac{1}{n}\left[1 - \frac{|h|}{n}\right]\rho(h) + 0.\left(\frac{1}{n}\right), \tag{2.5}$$

as $n \longrightarrow \infty$, then $\frac{|h|}{n} \longrightarrow 0$.

So, $nvar(\bar{y}_n) \longrightarrow \sum_{h=-\infty}^{\infty} \rho(h)$.

$$var(\bar{y}_n) = \frac{1}{n}\left[1 - \frac{|h|}{n}\right]\rho(h) + 0.\left(\frac{1}{n}\right). \tag{2.6}$$

Thus, for a short memory time series, $var(\bar{y}_n)$ goes to 0 as the size of the sample increases at the rate, $\frac{\sigma^2}{n}$ but with different multiplier, the integrated fractionally white noise process is given as:

$$(1-B)^d y_t = w_t, \; 0 < d < 0.5, \tag{2.7}$$

where the fractional difference operator $(1-B)^d$ is defined by a binomial series:

$$(1-B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k = 1 - dB - \frac{d(1-d)B^2}{2} - \frac{d(1-d)(2-d)B^3}{6} - \ldots. \tag{2.8}$$

$$(1-B)^d = \sum_{k=0}^{\infty} \frac{\Gamma(k-d)z^k}{\Gamma(-d)\Gamma(k+1)}. \tag{2.9}$$

The autocorrelation of the fractionally integrated series is defined as:

$$\rho(h) = \frac{\Gamma(h+d)\Gamma(1-d)}{\Gamma(h-d+1)\Gamma(d)} \sim h^{2d-1}. \tag{2.10}$$

So,that $0 < d < 0.5$, $\sum_{h=-\infty}^{\infty} |\rho(h)| = \infty$. The $var(\bar{y}_n)$ decays like n^{2d-1}, so $\left(\frac{1+\alpha}{2}\right)$, where α is the slope of variance-time graph, gives a rough empirical estimate of d. Traditional financial modeling assumes that returns of the series are independent of each other and have short memories. The existence of long-range persistence is tested and modeled using long-memory in the mean and in volatility.

2.3 Volatility models

The ARCH models were conceived by Engle, (1982) and were developed specifically to capture variance that is non-constant over time. These models were enhanced by Bollerslev, (1986). The deficiency of GARCH-type models is due to the fact that they do not account for leverage effects. The interesting feature of the volatility models is that GARCH models have parameter $d = 0$, IGARCH has $d = 1$, and volatility GARCH type models have $0 < d < 1$ which captures long-memory behaviour.

2.3.1 ARCH(p) model

The ARCH model is characterized by symmetric features, that assume, that negative and positive shocks have the symmetric impact on the conditional volatility. It is given by:

$$h_t^2 = \omega + \sum_{k=1}^{p} \alpha_k \varepsilon_{t-k}^2, \tag{2.11}$$

where $\alpha_1, \cdots, > \alpha_p$, $\omega > 0$, $\alpha_k \geq 0$, and $0 \leq \sum_{k=1}^{p} \alpha_k < 1$.

2.3.2 GARCH(p,q) model

The general GARCH (p,q) model conceived by Bollerslev (1986), is defined by the following:

$$h_t^2 = \omega + \sum_{k=1}^{p} \alpha_k \varepsilon_{t-k}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}^2, \tag{2.12}$$

where p is the lag on the disturbance term, ε_t^2, and q is the lag on the conditional variance, h_t^2. The GARCH specification allows the conditional variance to follow an autoregressive and moving average process, which is a parsimonious specification to capture the time series properties in volatility. The lag orders of the AR and MA components are denoted by p and q, respectively.

2.4 Variants volatility GARCH models

There are many variants of volatility models but for the purpose of the study we briefly review the IGARCH and APARCH for symmetry.

2.4.1 The IGARCH(p,q) model

The advantage of IGARCH and GARCH type models is the ability to capture non-linear dynamics presented by financial returns. However, the IGARCH process which considers infinite memory is not appropriate given that the situation of long-memory process is very unlikely to happen in the real world.

$$h_t^2 = \frac{\omega}{[1-\beta(L)]} + 1 - \varphi(L)(1-L)[1-\beta(L)]^{-1}\varepsilon_t^2, \tag{2.13}$$

and $\sum_{k=1}^{p}\alpha_k + \sum_{j=1}^{q}\beta_j = 1$, with p is the distribution term lag, ε_t^2, and q related to the conditional variance lag, h_t^2. The GARCH (p,q) model researched by Bollerslev (1986) is an extension of the basic ARCH model and it includes lags h_t^2 to address the long lag lengths on ε_t^2. In the study by Bollerslev and Mikkelsen (1986), they developed an IGARCH model, that accounted for volatility persistence. If $\sum_{k=1}^{p}\alpha_k + \sum_{j=1}^{q}\beta_j = 1$, then the shocks of the conditional variance are always present and exhibit the characteristics of persistence, in contrast to where they are dying out when $\sum_{k=1}^{p}\alpha_k + \sum_{j=1}^{q}\beta_j < 1$. According to Poon and Granger (2003), the IGARCH model conditional variance is a hyperbolic function representing a gradual decay in the effects of shocks.

2.4.2 APARCH(p,q) model

Ding et al., (1993) introduced an asymmetric power ARCH power ARCH (APARCH). The APARCH (p,q) model can be defined as:

$$\sigma_t^\delta = \alpha_0 + \sum_{i=1}^{q}\left(\alpha_i|\varepsilon_{t-1}| - \gamma_i\varepsilon_{t-i}\right)^\delta + \sum_{j=1}^{p}\beta_j\sigma_{t-j}^\delta, \tag{2.14}$$

where $\delta > 0$ and $-1 < \gamma_i < 1$, $i = 1, \cdots, q$. The leverage effect in the model lead to higher volatility, vice versa. An APARCH (p,q) model of asymmetric differs from other GARCH-type volatility models with the introduction of the power term, δ, which is to be estimated. The estimation of the power term is an attempt to account for the true distribution underlying volatility. The idea behind the introduction of a power term arose from the fact that in modeling financial data, the assumption of normality, which restricts δ to either 1 or 2 is often unrealistic due to significant skewness and kurtosis.

2.5 Long-memory volatility GARCH models

In this section we discuss the long memory volatility models and the estimation methods of the FIGARCH, HYGARCH, and FIAPARCH models.

2.5.1 The FIGARCH model

Barkoules and Baum (1996) advocated that the existence of long-range persistence in the time series returns pointed out evidence against a weak form of the financial market efficiency. A further study conducted by Fama (1965) demonstrated that financial markets are called efficient only when asset prices always reflect the available information fully. Hence, the FIGARCH model was developed by the pioneers Bailey et al., (1965) to reflect the full available information of the time series data. The FIGARCH model, is given by the equation:

$$\sigma_t^2 = \omega\big[1 - \beta(L) - \phi(L)(1-L)^d\big]\varepsilon_t^2 + \beta(L)\sigma_t^2, \tag{2.15}$$

with backshift or lag operator represented by L, and parameters of the model are ω, β, ϕ and d, where d lies within $0 \leq d \leq 1$. When $d = 1$, the FIGARCH (p,d,q) model is reduced to IGARCH(p,q) and GARCH(p,q) model provides the parameter $d = 0$.

According to Conrad and Haag (2006), they derived "necessary and sufficient conditions for the non-negativity of the conditional variance in the FIGARCH (p,d,q) model and sufficient conditions for the general model. Theses constraints of Conrad and Haag use an ARCH(∞) representation of the FIGARCH (1,d,1), namely the coefficients g_i and f_i are the functions of the fractional differencing parameter d such that $g_j = f_i g_{j-1} = \Pi_{i=1}^{j} f_i$ with $f_j = j - 1 - d$ for $j = 1, 2, \cdots$ and $g_0 = 1$.
The conditions are that:

$$\psi_i = \beta_1^2 + \psi_{i-2} + [\beta_1(f_{i-1} - \phi_1)(-g_{i-1}), i \geq 2. \tag{2.16}$$

or else

$$\psi_i = \beta_1^2 + \psi_{i-2} + [\beta_1(f_{i-1} - \phi_1)(-g_{i-1}), i \geq 3. \tag{2.17}$$

Corollary 1:The conditional variance of the FIGARCH(1,d,1) is non-negative if and only if.

Case 1: $0 < \beta_1 < 1$.

Either: $\psi_1 \geq 0$ and $\phi_1 \leq f_2$ or $k > 2$ with $f_{k-1} < \phi_1 < f_k$ it holds that $\psi_{k-1} \geq 0$.

Case 2: $-1 < \beta_1 < 0$.

Either: $\psi_1 \geq 0$, $\psi_2 \geq 0$ and $\phi_1 \leq \frac{f_2(\beta_1+f_2)}{(\beta_1+f_2)}$ or $k > 3$, with $f_{k-2}\frac{(\beta_1+f_{k-1})}{(\beta_1+f_{k-2})} < \phi_1 \leq f_{k-1}\frac{(\beta_1+f_k)}{(\beta_1+f_{k-1})}$, then $\psi_{k-1} \geq 0$ and $\psi_{k-2} \geq 0$. This corollary can be derived from the recursions as given below:

$-g_i > 0$ for $i \geq 1$,in which it applies when:

$F_i = f_i - \phi_1$ and $F_i^{(1)} = \beta_1(f_{i-1} - \phi_1) + (f_i - \phi_1)f_{i-1}$ are increasing and when there exists a k such that $F_{k-1} < 0 \leq F_k$ and
$F_{k-1}^{(1)} < 0 \leq F_k^{(1)}$. The FIGARCH was developed, which takes into account an intermediate level of long-memory captured by the parameter d. Therefore, the advantage of the FIGARCH is that it nests both GARCH ($d = 0$) and IGARCH ($d = 1$) models, as special cases. In contrary, the FIGARCH model does not outline a covariance stationary process.

2.5.2 Estimating FIGARCH model

The maximum likelihood estimation is used to compute the parameters of the FIGARCH (p,d,q) model with the normality assumption of y_t. The likelihood of a FIGARCH (p,d,q) model is then derived on the sample $\varepsilon_1, \cdots, \varepsilon_T$ and can be written as follows (Baillie, 1996):

$$L_\Omega(\theta, \varepsilon_1, \cdots, \varepsilon_T) = -\frac{1}{2} ln(2\pi) - \frac{1}{2} \sum_{t=1}^{T} \left[ln(h_t) - \frac{\varepsilon_t^2}{h_t} \right], \tag{2.18}$$

where $\theta = (\alpha_0, d, \beta_1, \cdots, \beta_q, \varphi_1, \cdots, \varphi_p)$. Therefore, the likelihood function is maximized conditionally on the set of initial values. The maximum likelihood method is given by:

$$\frac{\partial L_\Omega(\theta, \varepsilon_1, \cdots, \varepsilon_T)}{\partial \theta} = -\frac{1}{2} \frac{ln(2\pi)}{\partial \theta} - \frac{1}{2} \frac{\sum_{t=1}^{T} \left[ln(h_t) + \frac{\varepsilon_t^2}{h_t} \right]}{\partial \theta}, \tag{2.19}$$

where θ is the estimated set of parameter, L_Ω is the likelihood function, and T is the number of observations.

2.5.3 The HYGARCH model

Davidson (2004) advocated the use of the HYGARCH model to capture the long-range dependence measured by the geometric of hyperbolic decay of the coefficients in the $ARCH(\infty)$ model. The conditional variance of the FIGARCH the model was extended by incorporating weights to its difference operator. The HYGARCH model can be defined as:

$$\sigma_t^2 = \frac{\omega}{(1-\beta(L))} + \left\{1 - \frac{\phi(L)(1-L)^d}{(1-\beta(L))}\right\}\varepsilon_t^2, \tag{2.20}$$

where $0 \leq d \leq 1$. The HYGARCH model conditional variance can be derived from FIGARCH (p,d,q) by introducing weights to HYGARCH difference operator. Then $1-[1-\beta(L)]^{-1}\phi(L)(1-L)^d$, is replaced by $1-\left[1-\beta(L)\right]^{-1}\phi(L)\left[1+\alpha\left\{(1-L)^d\right\}\right]$.

$$\sigma_t^2 = \omega\left[1-\beta(L)\right]^{-1} + \left\{1-\left[1-\beta(L)\right]^{-1}\phi(L)\left[1+d\left\{(1-L)^d\right\}\right]\right\}\varepsilon_t^2, \tag{2.21}$$

where L is the lag operator, $\omega > 0$, $\beta < 1$, $0 < d < 1$.

The model above yields the properties of volatility clustering, leptokurtosis, and long memory; however, it disregards asymmetry and the fact that the conditional volatility is best represented by non-integer powers of the absolute value of the observations. The HYGARCH reduces to a FIGARCH whenever $\alpha = 1$.

Inequality constraints of HYGARCH (p,d,q) model (Conrad 2010) states that the conditional variance of the HYGARCH(1,d,1) is nonnegative if and only if.

Case 1: $0 < \beta < 1$.

Either: $\psi_i^{HY} = \iota\psi_i^{F_1} + \left(1-\iota\right)\psi_i^{GA}\left(L\right)$ and $\phi_1 \leq f_2$ for $k > 2$ with $f_{k-1} < \phi_1 \leq f_k$ it holds that $\psi_{k-1}^{HY} \geq 0$.

Case 1: $-1 < \beta < 0$.

Either: $\psi_1^{HY} \geq 0$, $\psi_2^{HY} \geq 0$ and $\phi_1 \leq \frac{f_2(\beta_1+f_3)}{(\beta_1+f_2)}$ or for $k = 3$ with $\frac{f_{k-2}(\beta_1+f_{k-1})}{(\beta_1+f_{k-2})} < \phi_1 \leq \frac{f_{k-1}(\beta_1+f_k)}{(\beta_1+f_{k-1})}$ it holds that $\phi_{k-1}^{HY} \geq 0$ and $\phi_{k-2}^{HY} \geq 0$.
These are the sufficient conditions for the HYGARCH(1,d,1); more detailed conditions for higher-order HYGARCH can be sourced from Conrad(2010).

2.5.4 Estimating HYGARCH model

The HYGARCH model is given by Davidson(2004):

$$y_t = \varepsilon_t\sqrt{h_t}, \tag{2.22}$$

where $h_t = \frac{\gamma}{\beta(1)} + \omega\left\{1 - \frac{\delta(B)}{\beta(B)}(1-B)^d\right\}y_t$, where $0 < d < 1$, $\gamma > 0$, $\alpha > 0$, $\beta(x) = 1 - \sum_{j=1}^{p}\beta_j x^j$, and $\delta(x) = 1 - \sum_{i=1}^{p}\delta_i x^i$. Denote the parameter vector by $\theta =$

$(\alpha, \acute{\delta}, \acute{\beta}, \omega, d)' \in R^{p+q+s}$, where $\delta = (\delta_1, \cdots, \delta_q)'$ and $\beta = (\beta_1, \cdots, \beta_p)'$. The Gaussian log-likelihood function is given by:

$$L_n(\theta) = \sum_{t=1}^{n} l_t(\theta), \tag{2.23}$$

where $l_t(\theta) = \frac{y_t^2}{h_t(\theta)} + ln[h_t(\theta)]$,and $h_t(\theta) = \frac{\gamma}{\beta(1)} + \sum_{j=1}^{\infty} b_j(\theta) y_{t-j}^2$ with $b_j(\theta)$s being functions of θ. The Gaussian Quasi Maximum likelihood estimation is given by:

$$\hat{\theta_n} = argmin\tilde{L_n}(\theta). \tag{2.24}$$

The HYGARCH excels with the GARCH model to outline the required property of covariance stationarity, while accounting for the decaying impulse response coefficients as the FIGARCH. HYGARCH further offers a framework for testing geometric versus hyperbolic decay, and capable of modelling the long-run dynamics in the second conditional moments of several financial time series returns. Conrad (2010) studied the HYGARCH in detail and offered specific necessary and sufficient conditions for the non-negativity of the conditional variance of the HYGARCH model. However, the conditions for the HYGARCH proposed by Conrad (2010) are complex and can be probably violated in applications. Another important limitation of the HYGARCH model (and other long-memory) exists when structural breaks occur in time series data when structure breaks are present, d tends to be overestimated.

2.5.5 The FIAPARCH model

The FIAPARCH model is the extension of the FIGARCH model with the APARCH model of Ding et al., (1993). The APARCH (p,q) model can be defined as:

$$\begin{aligned} \sigma_{t+h|t}^{\delta} &= E\big(\sigma_{t+h}^{\delta}|\Omega_t\big) \\ &= E\big[\hat{\omega} + \textstyle\sum_{i=1}^{\delta} \sigma_i \big(|\varepsilon_{t+h-i}| - \hat{\gamma}_i \varepsilon_{t+h-i}\big)^{\delta} + \sum_{j=1}^{p} \hat{\beta}_j \sigma_{t+h-j}^{\delta} |\Omega_t\big] \\ &= \hat{\omega} + \textstyle\sum_{i=1}^{q} \sigma_i E\big[\big(\varepsilon_{t+h-i} - \hat{\gamma}_i \varepsilon_{t+h-i}\big)^{\delta} |\Omega_t + \sum_{j=1}^{p} \hat{\beta}_j \sigma_{t+h-j}^{\hat{\delta}}, \end{aligned}$$

where $E\big[\big(\varepsilon_{t+k} - \gamma_i \varepsilon_{t+k}\big)^{\delta} |\Omega_t = k\sigma_{t+k|t}^{\delta}$, for $k > 1$, and $k_i = E\big(|\varepsilon_{t+h-i}| - \gamma_i \varepsilon_{t+h-i}|\big)^{\delta}$. If $d = 0$, the FIAPARCH (p,d,q), specification can be reduced to APARCH (p,q) specification. The FIAPARCH model(p,d,q) can be written as follows:

$$\sigma^{\delta} = \omega + \big[1 - \big(1 - \beta(L)\big)^{-1}\big(1 - \phi(L)\big)\big(1 - L\big)^{d}\big]\big(|\varepsilon_t| - \gamma \varepsilon_t\big)^{\delta}, \tag{2.25}$$

where $\omega > 0$, $\delta > 0$, $-1 < \gamma < 1$ and $0 < d < 1$. When $\gamma > 0$, then negative shock increases volatility than positive shock and vice versa. The FIAPARCH model becomes a FIGARCH model when $\delta = 2$ and $\gamma = 0$. The $(1 - L)^d$ is the differencing operator in terms of a hypergeometric function(Bentes, 2015). When $d = 0$, the HYGARCH process reduces to the APARCH(1,1). The model can capture both long memory and asymmetry in the conditional variance. Conrad and Haag (2006) outlined the advantages of FIAPARCH model. The FIAPARCH model allows for an asymmetric response of volatility to positive and negative shocks, so being able to traduce the leverage effect. However, the statistical properties of the general FIAPARCH process remain unestablished. The stationarity is not a certainty as well as the source of long-memory on volatility or even its existence are controversial. However, in constrast, the statistical properties of the general FIAPARCH process remain unestablished. The stationarity is not a certainty as well as the source of long-memory on volatility or even its existence are controversial.

2.5.6 Estimating FIAPARCH model

The FIAPARCH (p,d,q) log-likelihood (L) can be defined as follows (Tse, 2002):

$$L(\varepsilon|\theta) = \Pi_{t=1}^{n} \frac{1}{\sqrt{\pi\sigma_t^2}} e^{\frac{-\varepsilon_t^2}{2\sigma_t^2}}, \tag{2.26}$$

where $\theta = (\omega, \alpha, \gamma, \beta, \delta, d)$, n is the number of observations and $\varepsilon = \{\varepsilon_1, \cdots, \varepsilon_n\}$. The MLE is given by:

$$\frac{\partial L(\varepsilon|\theta)}{\partial\theta} = \frac{\Pi_{t=1}^{n} \frac{1}{\sqrt{\pi\sigma_t^2}} e^{\frac{-\varepsilon_t^2}{2\sigma_t^2}}}{\partial\theta}. \tag{2.27}$$

2.6 Heavy tailed distributions

In this section, we provide the heavy-tailed distributions used in the study and the maximum likelihood estimation method of deriving the parameters.

2.6.1 Student's t- distribution

The Student's t- distribution has characteristics of the normal distribution, as it is symmetrical and bell-shaped, but differs from the normal distribution as a result of heavier tails. The probability density function of the univariate Student t- distribution is given (Arfken, 2013):

$$f(y) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sigma\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{y^2}{\nu\sigma^2}\right)^{\frac{-\nu+1}{2}}, \tag{2.28}$$

for $-\infty < y < \infty$, with μ being the location parameter, $\sigma > 0$ is understood as the scale parameter, and $\nu > 0$ degrees of freedom.

2.6.2 Maximum likelihood estimation of the Student's t- distribution

Given that error terms which assume a Student's t- distribution, it follows that Z_t-distribution also tends to follows symmetric Student's t- distribution with n degrees of freedom, and μ=0 with $\sigma_t^2 = \frac{\nu}{\nu-2}$ for $n > 2$ (Green, 2005). The log likelihood function is given by:

$$lnL(\mu,\sigma,\nu) = n\left(ln\Gamma\left(\frac{\nu+1}{2}\right) - ln\Gamma\left(\frac{\nu}{2}\right) - ln(\sigma\sqrt{\nu\pi})\right) - \frac{\nu+1}{2}\sum_{i=1}^{n} ln\left(1+\frac{(y_i-\mu)^2}{\nu\sigma^2}\right). \tag{2.29}$$

The partial derivatives of the log-likelihood function are:
then

$$\partial\frac{ln(\mu,\sigma,\nu)}{\partial\mu} = (1+\nu)\sum_{i=1}^{n}\frac{y_i-\mu}{\nu\sigma^2+(y_i-\mu)^2}, \tag{2.30}$$

and

$$\partial\frac{ln(\mu,\sigma,\nu)}{\partial\sigma} = \frac{1}{\sigma}\left((1+\nu)\sum_{i=1}^{n}\frac{(y_i-\mu)^2}{\nu\sigma^2+(y_i-\mu)^2} - n\right), \tag{2.31}$$

then

$$\partial\frac{ln(\mu,\sigma,\nu)}{\partial\nu} = \frac{n}{2}\left(\psi\left(\frac{\nu+1}{2}\right) - \psi\left(\frac{\nu}{2}\right) - \frac{1}{\nu}\right) + \frac{1}{2}\sum_{i=1}^{n}\left[\frac{(1+\nu)(y_i-\mu)^2}{\nu^2\sigma^2+\nu(y_i-\mu)^2}\right], \tag{2.32}$$

where $\psi(y)$ is the digamma function, defined by $\psi(y) = \frac{d}{dy}ln\Gamma(y) = \frac{\Gamma(y)'}{\Gamma(y)}$. The maximum likelihood is obtained through the application of the numerical optimization methods.

2.6.3 Skewed Student's t- distribution

The Student's t- distribution is capable of capturing the heavier tails, but as a result of the symmetric nature of this distribution, it is still unable to deal with time series data that are asymmetric. Hansen (1994) conceived the skewed Student's t-distribution in an endeavour to deal with an asymmetry conferred by long-tailed data series. This distribution includes a random variable with μ=0, and σ^2=1, and then tends towards the Student's t- distribution if and given that λ=0; through the generalization of the parent Student's t- distribution, one will acquire the skewed Student's t- distribution, defined by Hansen as follows:

$$f(y,\lambda,z) = \frac{b\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi(\nu-2)\Gamma(\frac{\nu}{2})}}.\left(1+\frac{\zeta^2}{\nu-2}\right)^{(\frac{-\nu+1}{2})}, \tag{2.33}$$

where $\zeta = \frac{(by+a)}{1-\lambda}$ if $y < -\frac{a}{b}$, and $\frac{(by+a)}{1+\lambda}$ if $y \geq -\frac{a}{b}$. The constants terms a and b are defined by:

$a = 4\lambda c\frac{z-2}{z-1}$, and $b = 1+3\lambda^2 - a^2$, then $c = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi(\nu-2)\Gamma(\frac{\nu}{2})}}$
In this probability density function, $2 < \nu < \infty$ defines ν (degrees of freedom parameter), and $-1 < \lambda < 1$ is the asymmetry parameter.

2.6.4 Maximum likelihood estimation of the skewed Student's t- distribution

The log-likehood function of the skewed Student's t- distribution is defined by (Peters, 2001):

$$\partial\frac{log(Y_t)}{\partial L} = \frac{\kappa^2-1}{\kappa(\kappa^2+1)} + \frac{\sigma'}{\sigma} + \left(\frac{(\nu+1)\omega_t(sgn(\omega_t)\nu_t - \kappa\omega_t')}{\kappa\big(\omega_t^2+(\nu-2)\kappa^{2sgn(\omega_t)}\big)}\right), \tag{2.34}$$

then the maximum likelihood is derived as follows:

$\partial\frac{log(Y_t)}{\partial\nu} = \frac{1}{2}\left(\frac{2\sigma'}{\sigma} - \frac{1}{\nu-2} + \psi\left(\frac{\nu+1}{2}\right) - \psi\left(\frac{\nu}{2}\right) - log\left[\frac{1+\kappa^{-2sgn(\omega_t)}}{\nu-2}\omega_t\right] - \frac{(\nu+1)\omega_t(2(\nu-2))\omega_t'' - \omega_t}{\nu-2(\omega_t^2+(\nu-2)\kappa^{2sgn(\omega_t)}}\right)$
where $\omega_t' = \mu' + \sigma' Z_t$, and $\omega_t'' = \mu'' + \sigma'' Z_t$,with $\mu' = \partial\frac{\mu}{\partial\kappa} = \frac{(\kappa^2+1)\mu}{\kappa(\kappa^2-1)}$,
$\mu'' = \partial\frac{\mu}{\partial\nu} = \frac{\mu}{2}\left(\frac{1}{\nu-2} + \psi\left(\frac{\nu-1}{2}\right) - \psi\left(\frac{\nu}{2}\right)\right)$, $\sigma'' = \partial\frac{\sigma}{\partial\nu} = -\frac{\mu^2}{2}\left(\frac{1}{\nu-2} + \psi\left(\frac{\nu-1}{2}\right) - \psi\left(\frac{\nu}{2}\right)\right)$
, where $\psi(.)$ indicates the psi function obtained by the logarithmic derivative of the gamma function.

$$\mu_{sst} = \Gamma\left(\frac{\nu-1}{2}\right)\sqrt{\nu-2}\left(\kappa - \frac{1}{\kappa}\right), \tag{2.35}$$

and

$$\sigma_{sst} = \left(\kappa^2 + \frac{1}{\kappa^2}\right) - \mu_{sst}^2, \tag{2.36}$$

then $g_{sst}(Y_t) = \left(\frac{2}{\kappa+\frac{1}{\kappa}}\right)\frac{\sigma_{sst}\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi(\nu-2)}\Gamma\left(\frac{\nu}{2}\right)}\left[1+\frac{\kappa^{-2sgn(\omega_t)}}{\nu-2}\omega_t^2\right]^{-\frac{(\nu+1)}{2}}$, where $\omega_t = \mu_{sst} + \sigma_{sst}Y_t$.

2.6.5 Generalized Error distribution

The generalized error distribution is a family of distributions which assumes a range of specific types relying on the value parameter ν, and which consists of the normal distribution as an exceptional case. The GED is a more flexible generalization of the normal distribution and consequently defined by the way of three parameters:

- Mean (μ), which determines the peak of the distribution. In the standard normal distribution, the median and mode are equal to the mean (μ)
- Standard deviation (σ), which determines the dispersion; and
- Shape parameter (β), which refers to as kurtosis, and reveals how much data is in the tails.

Given the above background, the definition of the generalized error distribution as follows (Giller, 2005):

$$f(y) = \frac{\lambda.z}{2\Gamma(12)} exp\Bigg(- \lambda^z |y-\mu|^2 \Bigg), for - \infty < y < \infty, \tag{2.37}$$

where $\Gamma(1k)$ is a Euler function, z a shape parameter, λ a scale parameter, and μ is a location parameter.

2.6.6 Maximum likelihood estimation of the generalized error distribution

The new approach to determine the estimates of the generalized error distribution through MLE estimation was conceived by Bednarz (2012) and is given by Purczynski and Bednarz-Okrzynska(2014):

$$f(y) = \frac{\lambda.z}{2\Gamma(12)} exp\Bigg(- |\lambda.y|^z \Bigg), for - \infty < y < \infty, \tag{2.38}$$

then MLE is given by the following:

$$ln\Bigg(L(\lambda, z)\Bigg) = N.ln(\lambda) + N.\Bigg(\frac{z}{2\Gamma(\frac{1}{2})}\Bigg) - \sum_{i=1}^{N} |\lambda.y|^z, \tag{2.39}$$

then $\frac{\partial ln\big((\lambda,z)\big)}{\partial \lambda}$=0 and $\frac{\partial ln\big((\lambda,z)\big)}{\partial z}$=0 gives

$$\lambda = \Bigg(\frac{N}{z.\sum_{i=1}^{N}|y_i|^z}\Bigg)^{\frac{1}{z}}, \tag{2.40}$$

and $z+\psi(\frac{1}{z})+ln\left(\frac{z}{N}\sum_{i=1}^{N}|y_i|^z\right)-z.\frac{\sum_{i=1}^{N}|y_i|^z ln|y_i|}{\sum_{i=1}^{N}|y_i|^z}$=0 where $\psi(p)=\frac{d}{dp}\big[ln\Gamma(p)\big]$.

2.7 Summary

In this chapter we discussed the theory of long-memory GARCH type models used in this study namel, HYGARCH, FIGARCH, and FIAPARCH models. We also discussed the three heavy-tailed distributions STD, SSTD, and GED. In the following chapter we present the methods used to investigate the empirical properties of the data used in the study.

Chapter 3

Methodology

3.1 Introduction

In this chapter, the methods used to check for empirical properties of the data and model adequacy are outlined. These methods include tests for normality, autocorrelation, ARCH effect, unit roots, and long memory. Value at risk and Kupiec's LR tests are also discussed.

3.2 Test for normality

3.2.1 Quantile-quantile (Q-Q) normality plot

The quantile-quantile plot is a graphical tool that assesses if a set of data is plausibly generated by theoretical distribution such as a normal. It is simply a visible display, no longer an air-tight proof, so it is incredibly subjective. It allows us, though, to see at a glance whether our assumption is plausible and if not, how the assumption is violated and what data points make contributions to the violation. A Q-Q plot is a scatter plot, created by way of plotting two sets of quantiles against one another. If both sets of quantiles came from the same distribution, we ought to see the points forming a roughly straight line. However, if the data points deviate from the straight line, the conclusion is that the data are non-normally distributed.

3.2.2 Jarque-Bera test

The time series data are only normally distributed when the coefficients of the skewness and kurtosis parameters are observed to be zero and three, respectively. The leptokurtic is represented by the coefficient of kurtosis greater than three and has a fat tail. The skewness of a random variable Y is defined as:

$$Skew(Y) = E\left(\frac{y-\mu}{\sigma}\right)^3, \tag{3.1}$$

and kurtosis of the random variable Y is defined as:

$$Kurt(Y) = E\left(\frac{y-\mu}{\sigma}\right)^4. \tag{3.2}$$

The skewness of zero implies that the distribution of Y is symmetrical around its mean μ, while negative skewness shows that values of Y larger than μ are more probable. Kurtosis is a measures of how the variance is affected by means of severe departure from the mean μ. The test ordinarily used to capture normality by using the third and fourth moments is Jarque-Bera Test (1987). This test measures the distinction in kurtosis and skewness of a variable in contrast to those of the normal distribution. Therefore, Jarque Bera test examines whether or not a specific distribution is normal or not. The Jarque Bera (JB) value is calculated as follows:

$$JB = \frac{N}{6}\left(S^2 + \frac{(K-3)^2}{4}\right), \tag{3.3}$$

where N is the sample size, S the skewness function, and K the kurtosis function. The intuition behind this test is that the larger the Jarque Bera value is, the lower the probability that the given series is drawn from a normal distribution. For large sample size, the test statistic of the Jarque Bera test is $\chi^2_{(2)}$ distribution with two degrees of freedom under the null hypothesis that the return series is normally distributed.

3.2.3 Shapiro- Wilk test

The Shapiro- Wilk, (1965) test ascertain whether a sample $Y_1, Y_2, \cdots, Y_n$ emanates from a normally distributed population. The null hypothesis is to ascertain whether the series $\{r_t\}_{t=1}^{T}$ is generated by normally distribution; in simple terms, the null hypothesis states that the population is normally distributed is rejected when the p-value of the test is below a defined significance value (0.05). The Shapiro- Wilk test can be calculated as follows:

$$W = \frac{\left(\sum_{t=1}^{T} w_t r_t\right)^2}{\sum_{t=1}^{T}\left(r_t - \bar{r}\right)^2}, \tag{3.4}$$

where r_t is the $t-th$ order statistic, $\bar{r}$ is the sample mean, $\left(w_1, w_2, \cdots, w_T\right)$ are weights.

3.3 Test for autocorrelation

3.3.1 Autocorrelation function plot

An autocorrelation plot shows the value of the autocorrelation function (acf) on the vertical axis, and the height of each spike shows the value of the autocorrelation function for the lag.

3.3.2 Ljung-Box test

Ljung and Box, (1978) advocated the test that ascertain the assumption that the residuals contain no autocorrelation up to any order k. Hence, Ljung-Box test is performed to test jointly whether several autocorrelations of data series are significant or not. The Ljung-Box value is calculated by:

$$Q_{LB} = T(T+2)\sum_{j=1}^{k}\frac{r_j^2}{T-j}, \tag{3.5}$$

where T is sample size, k is number of lags, and r_j is the j^{th} autocorrelation. If Q_{LB} is larger than the probability, then the process has an uncorrelated time series data decline. The null hypothesis for the test is that there exists no correlation and under that hypothesis, Q_{LB} is an $\chi^2_{(k)}$ distribution with k degrees of freedom.

3.4 Test for Arch effect

3.4.1 ARCH-LM test

The heteroskedasticity emanates from a series of a random variable with trending variances. That indicates that the series has a non-constant variance. Therefore, heteroskedasticity is tested using Engle's ARCH-LM test (Engle,1982). The test is defined as:
H_0: There is no heteroskedasticity

H_1: There is heteroskedasticity

Test statistic:

$$LM = nR^2, \tag{3.6}$$

where n is the number of observations, and R^2 is the coefficient of determination of the augmented residuals. The rejection of the null hypothesis is when p-value$\leq$ level of significance and the conclusion is then that there is heteroskedasticity.

3.5 Unit root and stationarity tests

The stationarity of a returns series will have powerful bearing on its behaviour and properties, with persistence of the shocks likely to be infinite for non-stationary series.

Definition 3: *Stochastic process* $\{y_t : t = 1, 2, \cdots\}$ *is stationary if, and only if, a set of indices,* $1 \leq t_1, \leq t_2 \leq \cdots$*; if the joint distribution of a draw* $\{y_{t1}, y_{t2}, \cdots\}$ *is the same than* $\{y_{t1+h}, y_{t2+h}, \cdots\}$ *for* $h \geq 1$*, the sequence is identically distributed*. Then, time series y_t, $t = 1, \cdots, T$ is known as covariance stationary if the following holds:

- $E[y_t] = \mu$, for all t
- $Var[y_t] = \gamma_0$ $(< \infty)$, for all t
- $Cov(y_t, y_{t-k}) = \gamma_k$, for all t.

Stationarity suggests that the covariance and correlation of the observations collected at k periods are only a characteristic of the lag k, but not depend on the time point t. The autocorrelation is given by:

$$\rho_k = \frac{\gamma_k}{\gamma_0}. \tag{3.7}$$

The covariance stationarity process and autocorrelations, as well as the autocovariances, are symmetrical. This implies that $\gamma_k = \gamma_{-k}$ and $\rho_k = \rho_{-k}$.. The a_t is a white noise process if, and solely if, following assumptions suffice:

- $E[a_t] = \mu$, for all t
- $var[a_t] = \sigma_u^2 < \infty$, for all t
- $Cov(a_t, a_s) = 0$, for all $t \neq s$, and is denoted by $a_t \sim (\mu, \sigma_u^2)$. The white noise process assumes that $\mu = 0$ and is stationary process.

3.5.1 Augmented Dickey-Fuller test

An augmented Dickey-Fuller test is solely valid if a_t is white noise, and a_t will be autocorrelated if there is autocorrelation. This test was conceived by Dickey and Fuller (1979). The test ascertains the existence of the unit root and stationarity of the time series returns over time. The essential difference between the Dickey-Fuller test and the augmented Dickey-Fuller test is that the latter is employed for an outsized and, additionally, complicated set of time series models. The augmented Dickey-Fuller statistic could be a negative number and the more negative it is, the more likely the rejection of the speculation that there is a unit root. The test is based on the preferred Dickey-Fuller test:

$$P_t = \theta P_{t-1} + a_t, \tag{3.8}$$

then

$$\begin{aligned} P_t - P_{t-1} &= \theta P_{t-1} - P_{t-1} + a_t \\ P_t - P_{t-1} &= (\theta - 1)P_{t-1} + a_t, \\ \Delta P_t &= \delta P_{t-1} + a_t, \end{aligned}$$

where $\delta = \theta - 1$, P_t is the level of the index at time t, $\Delta P_t = P_t - P_{t-1}$ and $a_t \sim N(0, \sigma^2)$. This equation represents the type 1. Type 2 can then be presented as:

$$\Delta P_t = \alpha_0 + \delta P_{t-1} + a_t, \tag{3.9}$$

this equation tests for a random walk with a drift term, where α_0 is the drift term, and the equation below gives the illustration of type 3:

$$\Delta P_t = \alpha_0 + \delta P_{t-1} + \alpha_1 + a_t, \tag{3.10}$$

test for a random walk with both drift and linear trend. The Augmented Dickey-Fuller test is given by:

$$\Delta y_t = \alpha + \beta y_{t-1} + \delta_t + \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_k \Delta y_{t-k}, \tag{3.11}$$

where k is the number of lags specified. The non-constant removes the constant term α from the above equation, and the trend includes the time trend δ_t, which, by default is not included. The Augmented Dickey-Fuller test incorporates three types:

Type 1: *is a linear model with no drift and linear trend for time.*

Type 2: *is a linear model with drift but no linear trend.*

Type 3: *is a linear model with both drift and linear trends.*

Testing of the Augmented Dickey-Fuller is given as follows:

$H_0 : \gamma = 0$
$H_1 : \gamma < 0$

Test statistic is:

$$F_t = \frac{\hat{\gamma}}{SE(\hat{\gamma})}, \tag{3.12}$$

with $SE(\hat{\gamma})$ being the standard error of γ. The computed value, F_t, is compared with the critical value from augmented Dickey-Fuller to reject or accept the null hypothesis, and is a lower-tailed test. Thus, if F_t is less than the critical value, the null hypothesis of unit root is rejected; therefore the conclusion is that the variable of the series will contain a unit root and is non-stationary.

3.5.2 Phillips- Perron (PP) test

Phillips and Perron, (1988) developed a plausible theory of unit root non-stationarity. The tests are comparable to augmented Dickey- Fuller tests. One can differentiate Phillips- Perron unit root test from the augmented Dickey-Fuller test based on how best to understand serial correlation and heteroskedasticity in the errors. The Phillips- Perron unit root test generally provides similar conclusions to the augmented Dickey-Fuller test, and thus the computation of the test statistic of the Phillips-Perron unit root test is complicated and given by the following model:

$$y_t = \theta_0 + \phi y_{t-1} + a_t, \tag{3.13}$$

then ADF: $a_t \sim i.i.d$ and PP: $a_t \sim$ serially correlated. The PP test equation:

$$\Delta y_t = \theta_0 + \zeta y_{t-1} + a_t, \tag{3.14}$$

add a correction factor to the Dickey-Fuller test statistics. To add lagged Δy_t to the augmented Dickey-Fuller is to "whiten" the serially correlated residuals and the hypothesis is given:
$H_0 : \delta = 0$

$H_1 : \delta < 0$

The modified statistics denoted Z_t and Z_δ are given by:

$$Z_t = \sqrt{\frac{\hat{\sigma}^2}{\hat{\lambda}^2}} t_{\hat{\delta}} - \frac{1}{2}\left(\frac{\hat{\lambda}^2 - \hat{\sigma}^2}{\hat{\lambda}^2}\right)\left(\frac{n(s.e(\hat{\delta}))}{\hat{\sigma}^2}\right). \tag{3.15}$$

$$Z_\delta = n\hat{\delta} - \frac{1}{2}n^2 \frac{s.e(\hat{\delta}))}{\hat{\sigma}^2}\left(\hat{\lambda}^2 - \hat{\sigma}^2\right). \tag{3.16}$$

The terms $\hat{\sigma}^2$ and $\hat{\lambda}^2$ are consistent estimates of the variance parameters.

$$\sigma^2 = \lim_{n\to\infty} n^{-1} \sum_{t=1}^{n} E(a_t^2), \tag{3.17}$$

and

$$\lambda^2 = \lim_{n\to\infty} \sum_{t=1}^{n} E\left[\frac{1}{n}\sum_{t=1}^{n} a_t^2\right]. \tag{3.18}$$

Under the null hypothesis that $\delta = 0$, the Phillips- Perron Z_t and Z_δ statistics have equivalent asymptotic distributions to the augmented Dickey- Fuller t-statistic and normalized bias statistics. One advantage of the Phillips- Perron tests over the augmented Dickey-Fuller tests is that the Phillips-Perron tests are robust to general forms of heteroskedasticity within the error term a_t.

3.5.3 Kwiatkowski-Phillips-Schmidts-Shim (KPSS) test

Kwiatkowski, Phillips, Schmidt and Shim (KPSS), (1992) introduced a stationary test that assumes that the series has a short memory in the null hypothesis; in other words, its partial sum satisfies an invariance principle. The KPSS test assesses the null hypothesis that a univariate time series is trend stationary against the alternative that it is a nonstationary unit root process. The test makes use of the following structural model:

$$y_t = c_t + \delta_t + u_{1t}, \tag{3.19}$$

and $c_t = c_{t-1} + u_{2t}$, where δ is the trend coefficient, u_{1t} is a stationary process, and u_{2t} is an independent and identically distributed process with mean zero and variance σ^2. The null hypothesis is that $\sigma^2 = 0$, which implies that the random walk term (c_t) is constant and acts as the model intercept. The alternative hypothesis is that $\sigma^2 > 0$, which introduces the unit root in the random walk. The KPSS test can be calculated:

$$\eta = \frac{1}{T^2\sigma_T^2(q)} \sum_{t=1}^{T} S_t^2, \tag{3.20}$$

where S_t^2 is a consistent estimator of the long run variance of α_t. The $e_t = r_t - \left(\hat{\beta}_t + \hat{\alpha}_t\right)$ are residuals of the regression of r_t on an intercept, and time $S_t = \sum_{i=1}^{t} e_t$, $t = 1, 2, \cdots, T$ as the partial sum process of the residuals.

3.6 Long-memory tests

In this section, we discuss we discuss the long-memory tests used to detect the long-memory behaviour in returns. The tests used in the study are Geweke and Porter-Hudak test (GPH), Whittle estimation, and Hurst exponent:R/S test.

3.6.1 Geweke and Porter-Hudak test

The Geweke and Porter-Hudak, (1983) estimator is the widely used to distinguish between long-memory and short-memory effects and is called spectral regression method. The spectral density of the fractionally integrated process Y_t is:

$$f(\omega) = \left[4sin^2\left(\frac{\omega}{2}\right)\right]^{-d} f_u(\omega), \tag{3.21}$$

where ω is the Fourier frequency, and $f_u(\omega)$ is the spectral density corresponding to u_t. The difference parameter d can be estimated as:

$$ln(f(\omega_j)) = \beta - dln\left[4sin^2\left(\frac{\omega}{2}\right)\right] + e_j, \tag{3.22}$$

for $j = 1, 2, \cdots\cdots, n_f T$, Geweke and Porter-Hudak showed that the least squares estimate $\hat{d}$ using regression is normally distributed in large samples if $n_f(T) = T^\alpha$, with $0 < \alpha < 1$:

$$\hat{d} \sim N\left[d, \frac{\pi^2}{6\sum_{j=1}^{n_f}(U_j - \bar{U})^2}\right], \tag{3.23}$$

where $U_j = \left[4sin^2\left(\frac{\omega}{2}\right)\right]$ and $\bar{U}$ is the sample mean of U_j. Under the null hypothesis of no memory $d = 0$, the test statistic is:

$$t_{d=0} = \hat{d}\left[d, \frac{\pi^2}{6\sum_{j=1}^{n_f}(U_j - \bar{U})^2}\right]^{-0.5}. \tag{3.24}$$

3.6.2 Whittle estimation

The long-memory process is also characterized in the frequency domain with the aid of a spectral density function proportional to λ^{-2d} as the frequency λ approaches zero at a rate dictated by the memory parameter d. The Whittle estimator is akin to the maximum likelihood estimator (MLE) in the frequency domain. The Whittle estimation method has becom famous due to its likelihood interpretation, asymptotic characteristics (smaller asymptotic variance in contrast to log-periodogram estimators), and mild assumption (no need for normality assumption). It is described as follows:

$$Q(G_0, d) = \frac{1}{m}\sum_{j=1}^{m}\left[log\big(G_0\lambda_j^{-2d}\big) + \frac{I_z(\lambda_j)}{G_0, \lambda_j^{-2d}}\right], \tag{3.25}$$

where $G_0 = G(0)$, $m = m(T)$ is the bandwidth which goes to infinity as $T \to \infty$, but at a slower rate than T, $\lambda_j = \frac{2\pi_j}{T}$ are the Fourrier frequencies. The estimator $\hat{d}$ of d is obtained by minimizing the equation (3.26) to derived the following:

$$G.\hat{d}_{LW} = argmin_d\left[log\hat{G}_0(d) - 2dm^{-1}\sum_{m}^{j=1} log\lambda_j\right], \tag{3.26}$$

where $\hat{G}_0(d) = m^{-1}\sum_{m}^{j=1}\lambda_j^{2d} I_z(\lambda_j)$.

3.6.3 Hurst exponent: R/S test

The Hurst, (1951) method estimates H via spectral regression in the exploitation of the relationship between β and the Hurst coefficient. The short-memory process has $H = 0.5$, and the autocorrelation function also decays faster, but when it is completely related to long-memory process, it is characterized by the Hurst exponent in the interval $0.5 < d < 1$. The Hurst exponent will be modeled by the subsequent equation:

$$log(\frac{R}{S}) = logk + Hlogm, \tag{3.27}$$

thus it can be interpreted as follows:

- If $H = 0.5$, time series assumes a random walk and independent
- If $0 < H < 0.5$, time series are anti-persistent and the process occupies only a small distance in contrast to a random walk; and
- If $0.5 < H < 1$, time series is persistent and the process covers larger distance than a random walk (long-memory process). The Hurst parameter is understood as self-similarity parameter and is defined by $H = d + 0.5$.

3.7 Model selection criteria

3.7.1 Akaike information criteria

The Akaike information criteria (AIC) apply the diagnostic test on residuals to deduce which model is most preferable. They propose the measure of the model's goodness of fit by balancing the error of fit against the number of parameters in the model (Tsay, 2005). The criteria offer the measure of information lost when a given model is used to represent reality. If the entire model fits poorly, AIC will not give any warning of the poor fit. The AIC criteria are defined as:

$$AIC(q) = T\, ln\left(\frac{SSR}{T}\right) * 2q, \tag{3.28}$$

where T is the sample size to which the model is fitted, SSR is the sum of squared residuals and q is the number of parameters equal to $n + 2$.

3.7.2 Bayesian information criterion

The other crucial measure of accuracy that has similarly interpretation as the Akaike information criteria is the Bayes information criterion, which (BIC) defined as follows:

$$BIC(p) = ln\left(\frac{SSR_{(p)}}{T}\right) + (p+1)\frac{ln(T)}{T}. \tag{3.29}$$

It resembles the AIC, where the best model is that with the lowest value of the BIC, by nature it considers the smallest models, hence BIC is good for the parsimonious models as oppose to AIC.

3.8 Value at risk

The Value at Risk for given probability p is defined as:

$$VaR_p = infu : F(u) \geq p. \tag{3.30}$$

That is, VaR is the quantile of F, exceeded with probability $1 - p$.

3.8.1 Value at risk of the skewed Student's t- distribution

The skewed Student's t- distribution value at risk of α quantile for long and short positions are:

$$VaR_L = \hat{\mu}_t + sstd_{\alpha,\nu,\xi}\hat{\sigma}_t. \tag{3.31}$$

$$VaR_S = \hat{\mu}_t + sstd_{1-\alpha,\nu,\xi}\hat{\sigma}_t, \tag{3.32}$$

where $sstd_{\alpha,\nu,\xi}\hat{\sigma}_t$ is the left quantile at $\alpha\%$ for the skewed Student's t-distribution and $sstd_{1-\alpha,\nu,\xi}\hat{\sigma}_t$ is the long position, with ν being the degrees of freedom and ξ an asymmetry coefficient, and μ_t is the conditional mean process.

3.8.2 Value at risk of the generalized error distribution

Once $\nu = 2$, the normal distribution could be a special case of the generalized error distribution. If $\nu < 2$, then the generalized has fatter tails than ordinary normal distribution. Then, the value at risk is then:

$$VaR_L = \hat{\mu}_t + \phi_p^{-1}(\varepsilon_t)\hat{\sigma}_t, \tag{3.33}$$

where $\phi_p^{-1}(\varepsilon_t)$ is the left quantile of the generalized error distribution at p level.

3.8.3 Value at risk of the Student's t- distribution

The Student's t- distribution is typical as an example of a heavy-tailed distribution. The value at risk:

$$VaR_t(1-\alpha) = \left(\hat{\mu}_t + t_\nu^{-1}(\alpha)\sqrt{\frac{\nu-2}{\nu}}\sigma\right), \tag{3.34}$$

where ν denotes degrees of freedom and t_ν^{-1} is the α-quantile of the standard Student's t- distribution with ν degrees of freedom.

3.8.4 Backtesting VaR using Kupiec's LR test

What is the backtesting Backtesting is a technique for simulating a model or strategy on past data to gauge its accuracy and effectiveness. In this study, we use Kupiec's (1995) to back test value at risk. The failure rate is the number of times return series exceeds the forecasted value at risk. The assumption is that the range of exceedances over time follows the binomial distribution and the purpose of the Kupiec's test is to establish the consistency of these violations with given confidence level. If the range of exceedances differs drastically from what is expected, the risk model's adequacy is questionable. To perform the test, the wide variety of actual violations E, number of observations N, and the VaR probability level (p) are needed. Assuming E is distributed as Bin(N, p) the hypothesis is given as below:

$H_0 : p = p_0$,

$H_1 : p \neq p_0$,

and p is estimated by $\frac{E}{N}$. As Kupiec (1998) proposed this test is based on the likehood ratio test.

Test statistic is given:

$$LR = 2log\left[\frac{\left(1 - \frac{E}{N}\right)^{N-E} \frac{E^E}{N}}{\left(1 - p_0\right)^{N-E} p_0^E}\right] \sim \chi^2_{(1)}, \tag{3.35}$$

where N is the number of observations used to forecast VaR values and E is the observed number of actual exceedances.

3.9 Summary

In this chapter, we explored the diagnostics tests used in the study, model selection criteria were also discussed. We discussed the theory of value at risk and backtesting VaR. In the subsequent chapter we outline the empirical results of the study.

Chapter 4

Empirical Results

4.1 Introduction

In this chapter, data and their sources are discussed. The results of data exploration are also presented. Finally, the results of fitting the long-memory GARCH-type model are presented.

4.2 Data and data sources

The data used composed of daily closing indices' prices of the All-Share Index (JSEALSI), Mining Index (JSEJMNNG), and Banking Index (JSEBNKS) as defined according the Global Industry Classification Standard. The data was obtained from INET Bridge and Bloomberg for period 7 June 2008 to 7 June 2018 with a total of 2 500 day by day observations for each index. For the cause of this research study, each day's returns were calculated, comprising a total of 2 449 return observations for each index.

Let Y_t, $t = 1, 2, \cdots\cdots, T$, where the return series is denoted by R_t, and the log returns of the indices are defined as:

$$R_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \Big(\ln\left(P_t\right) - \ln\left(P_{t-1}\right)\Big), \tag{4.1}$$

with P_t and P_{t-1} as the current and one lagged of each index on day t and $t-1$, respectively.

4.3 Data exploration

The plots of the time series and return series for the three indices are displayed in Figure 4.1-4.3. The returns of all the indices exhibit periods of high and low volatilities, which confirm volatility clustering.

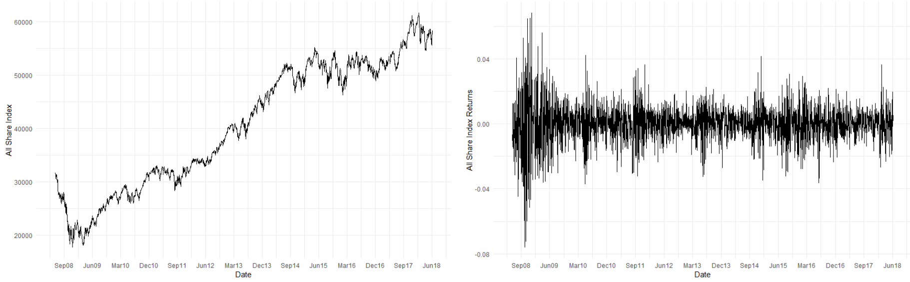

Figure 4.1: Graphical representation of the All Share Index series (left) and Returns (right)

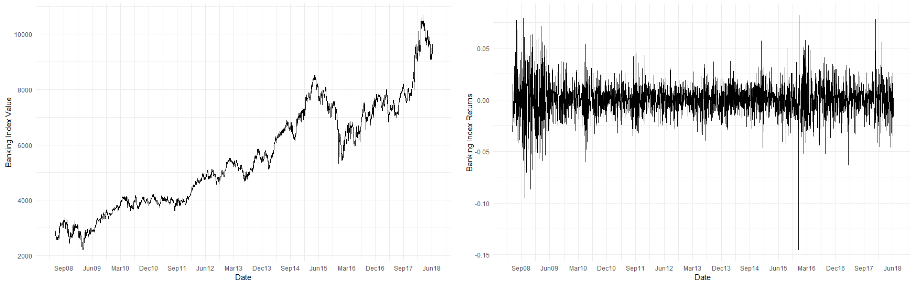

Figure 4.2: Graphical representation of the Banking Index series (left) and Returns (right)

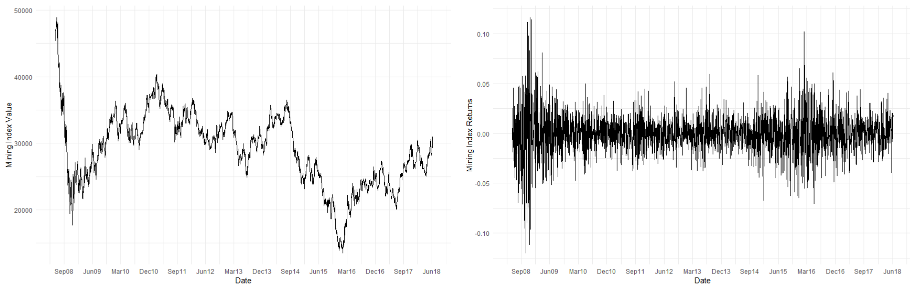

Figure 4.3: Graphical representation of the Mining Index series (left) and Returns (right)

However, for the mining index the time series plot generally decreases in trends. This can be attributed to the slowdown in domestic economic conditions includes factors related to labour unrest, electricity constraints, and the political landscape in South Africa, which contributed to a downwards shock that took place in the March 2015 period. These difficult conditions were mirrored by the non-stationary of the returns, which were trending downwards. Generally, the All Share and Banking Indices are increasing overtime.

Descriptive statistics, autocorrelation, normality, ARCH-LM, unit roots and stationarity tests are reported in Table 4.1. Panel A reveals that the JSEBNKS constitutes 2 449 observations, ranging from -0.14521 to 0.08189, with the highest mean returns of 0.00046, followed by the JSEALSI, ranging from -0.07581 to 0.06834, with the mean of 0.00025, and lastly the JSEJMNNG, ranging from -0.11966 to 0.11616, with the mean of -0.00017. The skewness values of the JSEALSI and JSEBNKS are negative for the returns; negatively skewed returns suggest a greater probability of large declines in these stock returns during the sample period. The kurtosis is greater than three for all the indices, thus revealing heavy-tailed distributions. The kurtosis greater than three suggests that extreme changes in the prices occurred a lot throughout the sampling period. Furthermore, the return series have standard deviations higher than the returns of their means, revealing the possibility of volatilities.

From Panel B, the Jarque-Bera and Shapiro-Wilk tests for normality have p-values less than 0.0001 for all three indices returns, thereby rejecting normality assumptions for all levels of significance. The ARCH-LM test shows the presence of conditional heteroskedasticity in all the three daily returns. Further anaalysis was done using Ljung Box statistics, based on the Ljung Box statistics we reject null hypothesis of zero autocorrelation in returns and squared returns.

Finally, Panel C reports the ADF and PP unit roots and KPSS as stationarity tests. The ADF and PP test reject null hypothesis of unit root test for all the three daily returns. We can conclude that the three daily returns are stationary. The KPSS test also show that we cannot reject the stationary null hypothesis for daily returns.

4.4 Q-Q Plots for normality

Figures 4.4-4.6 show the normal Q-Q plots for the All Share Index, Banking Index, and Mining Index, respectively. The normal QQ plots show that the tails of all the three daily indices' returns are heavier than the tails of normal distribution.

Table 4.1: Descriptive statistics and unit roots tests of the log returns of the three indices

		JSEBNKS		JSEJMNNG		JSEALSI	
Panel A *Descriptive Statistics*							
Minimum		-0.1452		-0.1197		-0.0758	
Maximum		0.0819		0.1162		0.0683	
Mean		0.0005		-0.0002		0.0002	
Std.Dev		0.0169		0.0203		0.0119	
Skewness		-0.1757		0.0469		-0.1379	
Kurtosis		7.6664		6.8978		7.2814	
Panel B *Testing for correlation, normality, and heteroskedasticity*							
		Statistic	*p-value*	*Statistic*	*p-value*	*Statistic*	*p-value*
$Q(5)$		44.87	<0.0001	22.34	<0.0001	23.35	<0.0001
$Q(10)$		55.18	<0.0001	44.85	<0.0001	34.67	<0.0001
Jarque-Bera		2277.2	<0.0001	1582.9	<0.0001	1917.2	<0.0001
Shapiro-Wilk		0.9586	<0.0001	0.9590	<0.0001	0.9509	<0.0001
$Q^2(5)$		344.15	<0.0001	894.55	<0.0001	1085.0	<0.0001
$Q^2(10)$		543.92	<0.0001	1986.8	<0.0001	2417.4	<0.0001
Arch LM Test		48.292	<0.0001	115.49	<0.0001	142.87	<0.0001
Panel C *Unit root and stationary tests*							
		Statistic	*p-value*	*Statistic*	*p-value*	*Statistic*	*p-value*
ADF	No constant	-48.9	<0.0001	-47.5	<0.0001	-48.5	<0.0001
	No Trend	-48.9	<0.0001	-47.5	<0.0001	-48.6	<0.0001
	With Trend	-48.9	<0.0001	-47.5	<0.0001	-48.6	<0.0001
PP	No constant	-2079	<0.0001	-2145	<0.0001	-2151	<0.0001
	No Trend	-2074	<0.0001	-2145	<0.0001	-2149	<0.0001
	With Trend	-2074	<0.0001	-2145	<0.0001	-2148	<0.0001
KPSS	No constant	0.9970	>0.1	0.3340	>0.1	0.5130	>0.1
	No Trend	0.0197	>0.1	0.1340	>0.1	0.1230	>0.1
	With Trend	0.0815	>0.1	0.0674	>0.1	0.1100	>0.1

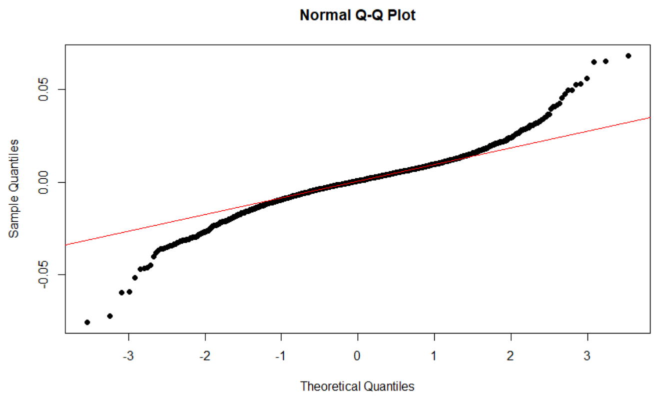

Figure 4.4: Q-Q plot of the All Share Index returns

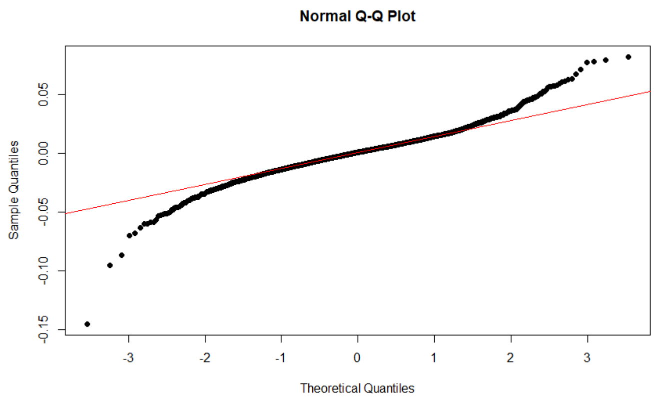

Figure 4.5: Q-Q plot of the Banking Index returns

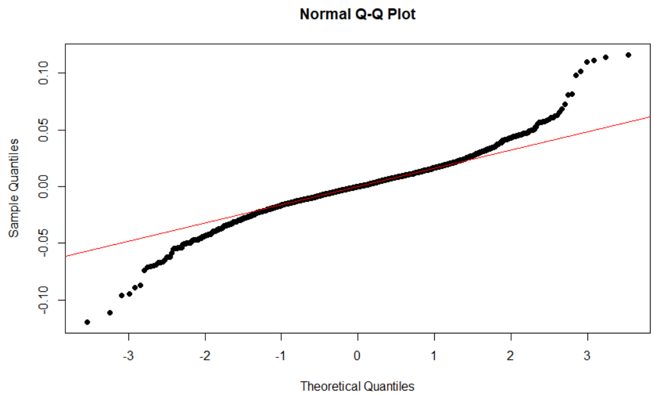

Figure 4.6: Q-Q plot of the Mining Index returns

The Q-Q plots show the presence of heavy-tailed distributions and asymmetric dispersion of all the indices' returns. The evidence advocates the use of volatility models such as asymmetric models or heavy-tailed distribution to account for leptokurtic and asymmetric factors. The reason for the series to depart from normality is volatility clustering that is present in the series data. This was also confirmed by excess kurtosis in Table 4.1 which indicate the heavy tailness of the three daily returns.

4.5 ACF and PACF plots of returns

Further analysis was done using sample ACF and sample PACF daily All Share Index returns, Banking Index returns, and Mining Index returns. Figures 4.7-4.12 show sample ACF and sample PACF of daily returns.

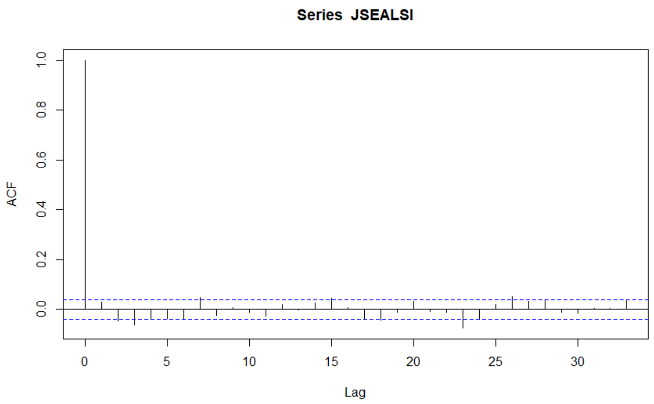

Figure 4.7: ACF plot for All Share Index returns

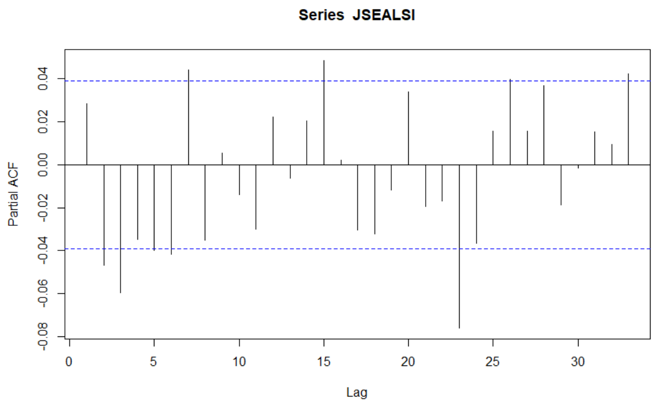

Figure 4.8: PACF plot for All Share Index returns

The ACF and PACF plots suggest that there is no autocorrelation in the All Share and Banking Indices, this is rather surprising since Ljung Box suggests that there is autocorrelation in the daily returns.

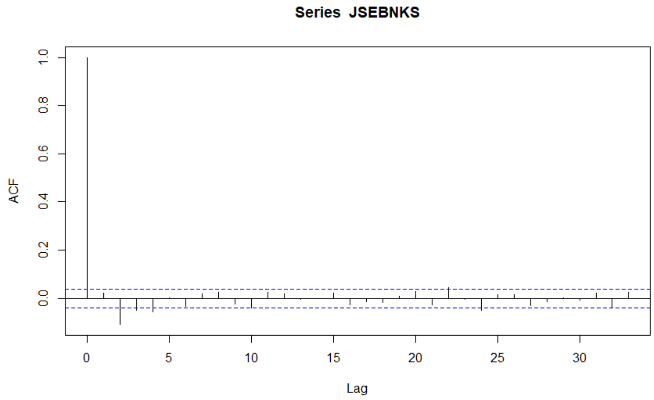

Figure 4.9: ACF plot for Banking Index returns

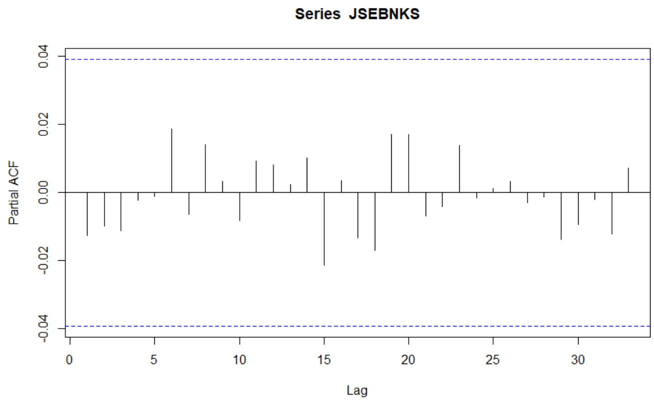

Figure 4.10: PACF plot for Banking Index returns

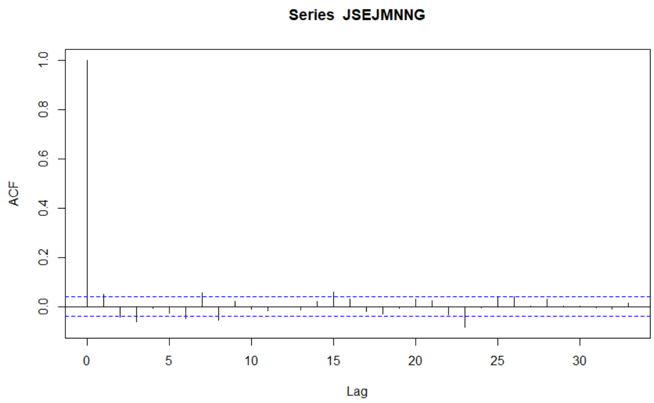

Figure 4.11: ACF plot for Mining Index returns

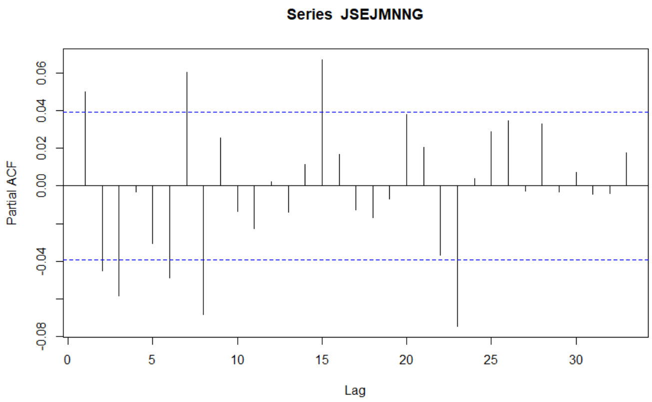

Figure 4.12: PACF plot for Mining Index returns

4.6 Testing for long-memory in returns and squared returns

Long-memory tests are carried out using Whittle estimator, GHP, and Hurst R/S methods. The results for the Whittle Estimator are presented in Tables 4.2 and 4.3:

Table 4.2: Whittle estimator (fG_n) long-memory test of the returns

Returns	Hurst	Standard Error	t-value	p-value
JSEBNKS	0.4677	0.0123	38.0386	<0.0001
JSEJMNNG	0.5076	0.0125	40.6302	<0.0001
JSEALSI	0.4899	0.0124	39.4828	<0.0001

Table 4.3: Whittle estimator (fG_n) long-memory test of the squared returns

Returns	Hurst	Standard Error	t-value	p-value
JSEBNKS	0.6373	0.0129	49.1413	<0.0001
JSEJMNNG	0.6564	0.0130	50.3995	<0.0001
JSEALSI	0.6615	0.0130	50.7236	<0.0001

In Tables 4.2 and 4.3, the Whittle estimator is used on both log-returns and squared log-returns to ascertain the existence of the long-memory process. For the log-returns, the Hurst index values for the three indices are close to 0.5 and the p-values are less than a significance level of 0.05, implying the absence of long memory features. For the squared log-returns the Hurst indices are significantly greater than 0.5, with p-values at less than significance level of 0.0001, revealing the presence of long-memory.

Table 4.4: GPH's long-memory test for returns

Returns	Bandwidths	d	Standard Error	t-value	p-value
JSEBNKS	$m = T^{0.5}$	-0.1923	0.1008	-11.8332	<0.0001
	$m = T^{0.6}$	-0.1357	0.0608	-18.6839	<0.0001
	$m = T^{0.7}$	-0.1455	0.0403	-28.4297	<0.0001
JSEJMNNG	$m = T^{0.5}$	-0.0409	0.0681	-15.2742	<0.0001
	$m = T^{0.6}$	-0.0063	0.0492	-20.3630	<0.0001
	$m = T^{0.7}$	-0.0153	0.0414	-24.4951	<0.0001
JSEALSI	$m = T^{0.5}$	0.0307	0.0945	-10.2155	<0.0001
	$m = T^{0.6}$	0.0232	0.0731	-13.3656	<0.0001
	$m = T^{0.7}$	-0.0682	0.0469	-22.7351	<0.0001

Table 4.5: GPH's long-memory test for long squared returns

Returns	Bandwidths	d	Standard Error	t-value	p-value
JSEBNKS	$m = T^{0.5}$	0.4150	0.0927	-6.3128	<0.0001
	$m = T^{0.6}$	0.3293	0.0616	-10.8941	<0.0001
	$m = T^{0.7}$	0.2210	0.0372	-20.9142	<0.0001
JSEJMNNG	$m = T^{0.5}$	0.7674	0.0731	-3.1839	0.0007
	$m = T^{0.6}$	0.6913	0.0533	-5.7867	<0.0001
	$m = T^{0.7}$	0.5605	0.0419	-10.4649	<0.0001
JSEALSI	$m = T^{0.5}$	0.8497	0.0796	-1.8860	0.0297
	$m = T^{0.6}$	0.7379	0.0511	-5.1375	<0.0001
	$m = T^{0.7}$	0.6058	0.0412	-9.5639	<0.0001

Tables 4.4-4.6, show the results for the GPH and Hurst tests for the returns and squared returns. The following bandwidths are employed: $m = T^{0.5}$; $m = T^{0.6}$; and $m = T^{0.7}$. The squared log-returns show the presence of long-memory process; this indicates that long-memory possesses mean reverting, but it is not covariance stationary. The p-values of the squared log-returns are less than 0.0001, indicating the presence of long-memory in the squared returns.

Table 4.6 shows the results of long memory test, for the returns Hurst values are close to 0.5 and insignificant, implying the absence of long-memory features in some of the indices. For squared log-returns the Hurst values are significantly greater than 0.5 ie, within $0.5 < H < 1$ revealing the presence of long-memory. For all three tests results indicate the presence of long-memory in the squared returns.

Table 4.6: Hurst exponent's long-memory test

	Returns	Squared Log Returns
	Simple R/S Hurst Estimation	
JSEBNKS	0.46249	0.71008
JSEJMMNG	0.49765	0.72270
JSEALSI	0.52956	0.73689

4.7 Summary of empirical properties of daily JSEALSI, Banking and Mining returns

From data exploration, it can be concluded that daily JSEALSI, JSEJMMNG, and JSEBNKS returns exhibit the following empirical properties:

- Heavy tails
- Volatility clustering
- Long-memory (squared returns); and
- Stationary

Hence, suggested models are long-memory GARCH-type models with heavy-tailed innovations. These suggested models are:

- FIGARCH
- HYGARCH; and
- FIAPARCH

The suggested heavy tailed innovations are:

- Student's t-distribution
- skewed Student's t-distribution; and
- generalized error distribution.

4.7.1 Estimating long models with heavy tailed innovations

Tables 4.7-4.14 presents the parameter estimation results of the FIGARCH, HYGARCH, and FIAPARCH models with normal, STD, SSTD, and GED distributions for all three daily indices' returns. For all the models, the long range dependence parameter of the GARCH-type models are within $0.5 < d_v < 1$. This suggests strong evidence of long memory. It further implies that even historical shocks seem to influence the present shocks (Arouri et al., 2012). It is important to note that FIAPARCH results for JSEALSI Index could not been obtained, further study must be conducted to ascertain the reason. Bold values in the Tables 4.7-4.14 indicate that parameters are insignificant at 5% significance.

Table 4.7: FIGARCH parameter estimation with different innovations(All-Share-Index)

	Normal		STD		SSTD		GED	
Parameters	statistic	$p-$value	statistic	$p-$value	statistic	$p-$value	statistic	$p-$value
Cst(M)	0.0006	0.0013	0.0006	0.0002	0.0005	0.0042	0.0006	0.0002
Cst(V)	0.0174	0.0110	0.0165	0.0164	0.0150	0.0218	0.0171	0.0118
d	0.6937	0.0002	0.6537	0.0001	0.6637	0.0001	0.6662	0.0001
ϕ_1	0.1405	**0.1250**	0.1397	0.0329	0.1352	**0.0676**	0.1448	0.0486
β_1	0.7652	0.0001	0.7396	0.0001	0.7447	0.0001	0.7477	0.0001
ν	-	-	11.9096	0.0001	-	-	1.6415	0.0001
ξ	-	-	-	-	-0.1313	0.0001	-	-
$Q(10)$	17.5015	0.0639	17.2867	0.0683	17.3618	0.0667	17.3932	0.0661
$Q(20)$	27.8493	0.1130	27.5149	0.1214	27.5212	0.1212	27.7095	0.1165
$Q(50)$	71.0899	0.0266	70.6646	0.0287	70.4997	0.0299	70.9636	0.0272
$Q^2(10)$	8.7385	0.3648	8.8576	0.3545	8.9313	0.3481	8.6626	0.3716
$Q^2(20)$	16.1965	0.5788	16.2393	0.5759	16.5890	0.5515	16.0069	0.5921
$Q^2(50)$	45.7637	0.5649	46.2716	0.5439	46.7294	0.5249	45.8468	0.5615
ARCH(5)	0.8732	0.4982	0.9119	0.4721	0.9535	0.4452	0.8576	0.5089
ARCH(10)	0.8446	0.5855	0.8632	0.5675	0.8714	0.5596	0.8391	0.5908
LL	7975.57	-	7988.61	-	7998.35	-	7986.79	-
AIC	-6.3790	-	-6.3886	-	-6.3956	-	-6.3872	-
SBI	-6.3673	-	-6.3747	-	-6.3793	-	-6.3732	-

The Ljung Box test, ARCH-LM test, AIC, and SBI are presented in Tables 4.7-4.14. The diagnostic tests are used to capture the accuracy of the model specifications. The Ljung Box (Box-Pierce) test is given, and Q statistic on squared standardized residuals for all models are over 1% level of significance. This implies that one cannot reject null hypothesis of independently and identically distributed standardized residuals. Therefore, Ljung Box test on squared standardized residuals shows that the volatility equations are adequate.

Table 4.8: HYGARCH parameter estimation with different innovations(All-Share-Index)

	Normal		STD		SSTD		GED	
Parameters	statistic	$p-$value	statistics	$p-$value	statistic	$p-$value	statistic	$p-$value
Cst(M)	0.0005	0.0029	0.0006	0.0003	0.0005	0.0090	0.00062	0.0005
Cst(V)	0.0222	**0.1293**	0.0251	0.0151	0.0213	**0.3879**	0.0248	**0.1199**
d	0.9121	0.0001	0.7553	0.0005	0.8286	**0.2225**	0.8403	0.0060
ϕ_1	0.0191	**0.8852**	0.0936	**0.3880**	0.0573	**0.8718**	0.0565	**0.7344**
β_1	0.8607	0.0001	0.7855	0.0001	0.8242	0.0180	0.8267	0.0001
Log AlphaHY	-0.0191	**0.2349**	-0.0206	**0.1540**	-0.0177	**0.4994**	-0.0215	**0.2478**
ν	-	-	12.9624	0.0001	-	-	1.6616	0.0001
ξ	-	-	-	-	-0.1298	0.0001	-	-
$Q(10)$	17.2888	0.0682	17.1919	0.0702	17.3444	0.0671	17.2767	0.0685
$Q(15)$	28.1752	0.1053	27.6679	0.1175	27.9076	0.1116	28.0265	0.1088
$Q(20)$	71.7410	0.0236	70.9505	0.0273	71.0117	0.0269	71.4896	0.0247
$Q^2(10)$	9.1853	0.3269	9.3400	0.3144	8.9769	0.3443	9.0850	0.3352
$Q^2(20)$	15.7575	0.6095	16.2284	0.5766	16.3818	0.5660	15.7948	0.6069
$Q^2(50)$	42.8594	0.6829	44.3404	0.6236	44.2695	0.6265	43.2604	0.6671
ARCH(5)	0.6999	0.6235	0.9314	0.4593	0.8879	0.4882	0.8051	0.5458
ARCH(10)	0.8905	0.5413	0.9096	0.5232	0.8779	0.5533	0.8814	0.5500
LL	7978.35	-	7989.43	-	7999.27	-	7988.38	-
AIC	-6.3804	-	-6.3885	-	-6.3956	-	-6.3877	-
SBI	-6.3665	-	-6.3722	-	-6.3769	-	-6.3714	-

The ARCH-LM tests confirm the absence of heteroskedasticity in the residuals. Finally, we use Akaike information criterion (AIC) and Schwarz information criterion (SBI) to identify the best possible model for the conditional dependence in volatility process. The suitable model is given by the smallest value of the AIC and SBI. The suitable model for the JSEALSI Index returns that capture the dependence in the conditional variance is FIGARCH-SSTD confirmed by both AIC and SBI. For Banking Index returns, both AIC and SBI confirms the FIAPARCH-STD as the adequate model. In the case of Mining Index returns, both AIC and SBI give the FIAPARCH-STD as a preferred model.

Table 4.9: FIGARCH parameter estimation with different innovations(Mining Index)

	Normal		STD		SSTD		GED	
Parameters	statistic	$p-$value	statistic	$p-$value	statistic	$p-$value	statistic	$p-$value
Cst(M)	0.0030	**0.8220**	0.0001	**0.9143**	-0.0003	**0.9796**	0.0017	**0.8957**
Cst(V)	0.0026	0.0096	0.0024	**0.0848**	0.0024	**0.0906**	0.0025	**0.0890**
d	0.9506	0.0001	0.9574	0.0001	0.9562	0.0001	0.9530	0.0001
ϕ_1	0.0116	**0.8671**	0.0023	**0.9713**	0.0039	**0.9525**	0.0084	**0.8993**
β_1	0.9329	0.0001	0.9363	0.0001	0.9364	0.0001	0.9349	0.0001
ν	-	-	12.9119	0.0001	-	-	1.6609	0.0001
ξ	-	-	-	-	-0.0196	**0.5008**	-	-
$Q(10)$	18.0049	0.0549	18.0809	0.0536	18.0797	0.0536	18.0494	0.0541
$Q(20)$	24.7436	0.2114	24.7892	0.2096	24.7893	0.2096	24.7689	0.2104
$Q(50)$	54.4706	0.3083	54.5530	0.3056	54.5761	0.3048	54.5301	0.3063
$Q^2(10)$	5.3575	0.7188	5.5879	0.6933	5.5753	0.6947	5.4335	0.7104
$Q^2(20)$	14.7725	0.6775	14.6385	0.6866	14.5739	0.6910	14.5996	0.6893
$Q^2(50)$	62.6929	0.0755	62.9759	0.0722	62.8840	0.0732	62.6927	0.0755
ARCH(5)	0.6923	0.6292	0.7150	0.6121	0.7096	0.6162	0.6910	0.6303
ARCH(10)	0.5475	0.8571	0.5726	0.8376	0.5719	0.8381	0.5568	0.8899
LL	-2862.04	-	-2852.18	-	-2851.96	-	-2852.45	-
AIC	2.2936	-	2.2865	-	2.2872	-	2.2868	-
SBI	2.3053	-	2.3005	-	2.3035	-	2.3007	-

Table 4.10: HYGARCH parameter estimation with different innovations (Mining Index)

	Normal		STD		SSTD		GED	
Parameters	statistic	$p-$value	statistic	$p-$value	statistic	$p-$value	statistic	$p-$value
Cst(M)	0.0032	**0.8138**	0.0016	**0.9068**	0.0002	**0.9903**	0.0019	**0.8844**
Cst(V)	0.0060	0.0365	0.0053	0.0385	0.0052	0.0415	0.0056	0.0354
d	0.9701	0.0001	0.9737	0.0001	0.9728	0.0001	0.9711	0.0001
ϕ_1	-0.0041	**0.9399**	-0.0101	**0.8540**	-0.0087	**0.8734**	-0.0059	**0.9127**
β_1	0.9324	0.0001	0.9359	0.0001	0.9361	0.0001	0.9345	0.0001
Log(AlphaHY)	-0.0096	**0.0771**	-0.0082	**0.0899**	-0.0081	**0.0931**	-0.0089	**0.0794**
ν	-	-	14.4532	0.0001	-	-	1.6800	0.0001
ξ	-	-	-	-	-0.1298	0.0001	-	-
$Q(10)$	18.3018	0.0501	18.3313	0.0496	18.3295	0.0497	18.3240	0.0497
$Q(20)$	25.3856	0.1871	25.3365	0.1889	25.3319	0.1890	25.3650	0.1878
$Q(50)$	55.1982	0.2847	55.1584	0.2859	55.1688	0.2856	55.1929	0.2849
$Q^2(10)$	5.7018	0.3176	5.8923	0.6593	5.8752	0.6612	5.7729	0.6727
$Q^2(20)$	13.4053	0.7669	13.5545	0.7576	13.5157	0.7601	13.4128	0.7665
$Q^2(50)$	60.8550	0.1008	61.2166	0.0954	61.1626	0.0962	60.9230	0.0997
ARCH(5)	0.7149	0.6122	0.7373	0.5955	0.8879	0.4882	0.7162	0.6112
ARCH(10)	0.5814	0.8304	0.6010	0.8142	0.8779	0.5533	0.5894	0.8239
LL	-2858.96	-	-2850.22	-	-2850.03	-	-2850.14	-
AIC	2.2919	-	2.2858	-	2.2864	-	2.2857	-
SBI	2.3059	-	2.3021	-	2.3051	-	2.3020	-

Table 4.11: FIAPARCH parameter estimation with different innovations (Mining Index)

Parameters	Normal		STD		SSTD		GED	
	statistic	$p-$value	statistic	$p-$value	statistic	$p-$value	statistic	$p-$value
Cst(M)	-0.0086	**0.5448**	-0.0097	**0.4835**	-0.0115	**0.4077**	-0.0084	**0.5440**
Cst(V)	0.0043	0.0436	0.0036	0.0522	0.0034	0.0506	0.0039	0.0468
d	1.0023	0.0001	1.0052	0.0001	1.0065	0.0001	1.0047	0.0001
ϕ_1	-0.0208	**0.5701**	-0.0234	**0.5366**	-0.0247	**0.5070**	-0.0226	**0.5409**
β_1	0.9619	0.0001	0.9624	0.0001	0.9634	0.0001	0.9624	0.0001
Γ_1	0.5467	0.0241	0.5264	0.0165	0.5413	0.0166	0.5331	0.0202
Δ	1.2602	0.0001	1.3379	0.0001	1.3317	0.0001	1.2999	0.0001
ν	-	-	15.6893	0.0002	-	-	1.70476	0.0001
ξ	-	-	-	-	-0.0250	**0.3823**	-	-
$Q(10)$	20.0740	0.0286	19.9651	0.0296	19.9979	0.0293	20.0240	0.0290
$Q(20)$	27.6801	0.1172	27.4292	0.1236	27.4630	0.1227	27.5541	0.1204
$Q(50)$	54.9654	0.2921	54.6952	0.3009	54.7195	0.3001	54.8414	0.2961
$Q^2(10)$	5.3585	0.7187	5.2164	0.7342	5.3468	0.7199	5.3011	0.7249
$Q^2(20)$	11.5124	0.8714	11.5902	0.8677	11.5403	0.8700	11.5319	0.8704
$Q^2(50)$	52.8450	0.2924	53.3513	0.2761	53.2602	0.2790	53.0626	0.2853
ARCH(5)	0.1703	0.9736	0.7373	0.5955	0.1860	0.9680	0.1818	0.9695
ARCH(10)	0.5769	0.8340	0.6010	0.8142	0.5736	0.8367	0.5701	0.8395
LL	-2847.011	-	-2839.21	-	-2838.839	-	-2839.672	-
AIC	2.2832	-	2.2778	-	2.2783	-	2.2781	-
SBI	2.2995	-	2.2964	-	2.2992	-	2.2968	-

Table 4.12: FIGARCH parameter estimation with different innovations (Banking Index)

Parameters	Normal		STD		SSTD		GED	
	statistic	$p-$value	statistic	$p-$value	statistic	$p-$value	statistic	$p-$value
Cst(M)	0.0008	0.0007	0.0113	**0.9143**	0.0006	0.0217	0.0008	0.0058
Cst(V)	0.0606	0.0035	0.0649	0.0222	0.0639	0.0227	0.0615	0.0097
d	0.6500	0.0001	0.5061	0.0001	0.5063	0.0001	0.5766	0.0001
ϕ_1	0.1429	**0.0974**	0.2311	0.0001	0.2336	0.0001	0.1898	0.0051
β_1	0.7376	0.0001	0.6650	0.0001	0.6666	0.0001	0.7015	0.0001
ν	-	-	8.2202	0.0001	-	-	1.4665	0.0001
ξ	-	-	-	-	-0.0202	**0.4940**	-	-
$Q(10)$	25.5177	0.0045	25.7203	0.0041	25.7284	0.0041	25.5940	0.0043
$Q(20)$	34.9507	0.0204	35.2186	0.01897	35.2245	0.0189	35.0811	0.0197
$Q(50)$	73.0295	0.0185	73.7149	0.01619	73.7199	0.0162	73.3436	0.0174
$Q^2(10)$	4.2368	0.8351	3.5032	0.8989	3.5178	0.8978	3.7992	0.8748
$Q^2(20)$	7.2819	0.9875	6.3572	0.9945	6.4061	0.9943	6.7009	0.9924
$Q^2(50)$	16.6990	0.9999	15.8556	0.9999	15.8919	0.9999	16.0954	0.9999
ARCH(5)	0.6982	0.6248	0.5897	0.7090	0.5914	0.7066	0.6304	0.6766
ARCH(10)	0.4197	0.9379	0.3447	0.9688	0.3458	0.9684	0.3748	0.9579
LL	6935.403	-	6988.518	-	6988.76	-	6974.52	-
AIC	-5.5443	-	-5.5860	-	-5.5854	-	-5.5748	-
SBI	-5.5327	-	-5.5720	-	-5.5691	-	-5.5608	-

Table 4.13: HYGARCH parameter estimation with different innovations (Banking Index)

Parameters	Normal		STD		SSTD		GED	
	statistic	$p-$value	statistic	$p-$value	statistic	$p-$value	statistic	$p-$value
Cst(M)	0.0008	0.0065	0.0007	0.0110	0.0006	0.0190	0.0008	0.0059
Cst(V)	0.0917	0.0077	0.1026	0.0208	0.1014	0.0192	0.0929	**0.0741**
d	0.7683	0.0001	0.6533	0.0098	0.6492	0.0085	0.7430	0.0086
ϕ_1	0.0904	**0.1656**	0.1729	**0.1117**	0.1766	**0.0975**	0.1181	**0.3669**
β_1	0.7833	0.0001	0.7309	0.0001	0.7298	0.0001	0.7752	0.0001
Log(AlphaHY)	-0.0315	**0.1615**	-0.0434	**0.1044**	-0.0429	**0.1048**	-0.0354	**0.1966**
ν	-	-	8.6468	0.0001	-	-	1.4777	0.0001
ξ	-	-	-	-	-0.0177	**0.5434**	-	-
$Q(10)$	25.9290	0.0038	26.1443	0.0036	26.1383	0.0036	26.0541	0.0037
$Q(20)$	34.9825	0.0202	35.3091	0.0185	35.3111	0.0185	35.1203	0.0195
$Q(50)$	73.2314	0.0178	73.7946	0.0159	73.7939	0.0159	73.4298	0.0171
$Q^2(10)$	4.1655	0.8419	4.1370	0.8422	4.1387	0.8444	4.1115	0.8469
$Q^2(20)$	7.1202	0.9891	7.0464	0.9897	7.0746	0.9895	7.0759	0.9895
$Q^2(50)$	17.4329	0.9999	17.2976	0.9999	17.2933	0.9999	17.3534	0.9999
ARCH(5)	0.7294	0.6013	0.7218	0.6070	0.7207	0.6078	0.7149	0.6122
ARCH(10)	0.4162	0.9396	0.4072	0.9439	0.4069	0.9440	0.4073	0.9438
LL	6937.42	-	6989.67	-	6989.86	-	6975.95	-
AIC	-5.5451	-	-5.5861	-	-5.5855	-	-5.5752	-
SBI	-5.5312	-	-5.5698	-	-5.5668	-	-5.5588	-

Table 4.14: FIAPARCH parameter estimation with different innovations (Banking Index)

	Normal		STD		SSTD		GED	
Parameters	statistic	$p-$value	statistic	$p-$value	statistic	$p-$value	statistic	$p-$value
Cst(M)	0.0004	**0.1782**	0.0005	**0.0898**	0.0004	**0.1777**	0.0005	**0.0586**
Cst(V)	0.7198	**0.6817**	0.2004	**0.4272**	0.2107	**0.4468**	0.3126	0.4900
d	0.6168	0.0001	0.4964	0.0001	0.4954	0.0001	0.5607	0.0001
ϕ_1	0.1521	0.0087	0.2219	0.0001	0.2233	0.0001	0.1862	0.0004
β_1	0.7207	0.0001	0.6516	0.0001	0.6510	0.0001	0.6885	0.0001
Γ_1	0.3685	**0.2138**	0.2346	0.0303	0.2507	0.0307	0.2749	**0.0604**
Δ	1.5279	0.0017	1.7930	0.0001	1.7846	0.0001	1.6951	0.0001
ν	-	-	8.6015	0.0001	-	-	1.4972	0.0001
ξ	-	-	-	-	-0.0343	**0.2556**	-	-
$Q(10)$	25.1990	0.0049	25.0822	0.0052	24.9674	0.0054	25.1400	0.0051
$Q(20)$	34.1524	0.0251	34.3193	0.0240	34.1892	0.0249	34.2874	0.0242
$Q(50)$	71.2649	0.0257	71.8164	0.0232	71.5974	0.0242	71.5694	0.0243
$Q^2(10)$	4.4738	0.8120	2.6978	0.9519	2.6339	0.9552	3.2099	0.9205
$Q^2(20)$	7.0588	0.9896	5.2102	0.9985	5.1752	0.9986	5.7610	0.9917
$Q^2(50)$	21.9195	0.9999	18.0808	0.9999	18.2853	0.9999	19.1373	0.9999
ARCH(5)	0.8185	0.5363	0.4413	0.8198	0.1860	0.4279	0.5515	0.7373
ARCH(10)	0.4503	0.9216	0.2677	0.9880	0.2612	0.9891	0.3208	0.9760
LL	6948.28	-	6993.15	-	6993.84	-	6980.79	-
AIC	-5.5530	-	-5.5881	-	-5.5879	-	-5.5782	-
SBI	-5.5367	-	-5.5695	-	-5.5669	-	-5.5596	-

4.8 VaR estimation and backtesting models

VaR is calculated at 0.3%, 0.1%, 0.5%, 0.95%, 0.99%, and 0.99.7%. The VaR estimates are then backtested using the Kupiec LR test. The model with the highest possible p-value at a given level is selected as the robust model. The p-values of Kupiec LR test for in sample VaR backtesting are summarized in Tables 4.15-4.17.

Table 4.15: In-sample VaR backtesting: FIGARCH with different innovations for the three indices

Returns	Model	p-values of Kupiec LR test					
		Long positions			Short positions		
		0.3%	1%	5%	99.7%	99%	95%
JSEALSI	FIGARCH-N	0.0011	0.0003	0.1050	0.9205	0.0529	0.0001
	FIGARCH-STD	0.0411	0.0242	0.0077	0.3349	0.0159	0.0003
	FIGARCH-SSTD	0.7675	0.8401	0.7803	0.9205	0.1391	0.8915
	FIGARCH-GED	0.0183	0.0378	0.0724	0.3349	0.0300	0.0001
JSEJMNNG	FIGARCH-N	0.7682	0.0580	0.2066	0.0861	0.0093	0.0295
	FIGARCH-STD	0.0469	1.0000	0.5847	0.9197	0.6836	0.0134
	FIGARCH-SSTD	0.0469	0.2079	0.7839	0.7682	0.6915	0.0176
	FIGARCH-GED	0.1473	1.0000	0.7839	0.5019	0.8396	0.0056
JSEBNKS	FIGARCH-N	0.0077	0.0580	0.5169	0.1674	0.2450	0.0229
	FIGARCH-STD	0.6039	0.6836	0.9269	0.1473	0.6836	0.4034
	FIGARCH-SSTD	0.6039	0.5382	0.6442	0.2961	0.8396	0.5789
	FIGARCH-GED	0.9197	0.8396	0.4582	0.3444	0.5382	0.0737

4.8.1 FIGARCH models

The FIGARCH-N for the JSEALSI returns is not adequate at 0.3%, 1%, and 95% VaR levels. While the FIGARCH-SSTD is acceptable at all levels of the long and short positions. The VaR estimates for the FIGARCH-STD produced the lowest p-values at 0.3%, 1%, 5%, 95%, and 99% VaR levels, however adequate at 99.7% VaR level. The FIGARCH-GED produced the similar results to the FIGARCH-STD with similar VaR levels of lowest p-values and adequate at 99.7% VaR level.

For the JSEJMMNG returns, the FIGARCH-STD and FIGARCH-GED are adequate at 0.3%, 1%, 5%, 99%, and 99.7% VaR levels, respectively. Meanwhile, FIGARCH-SSTD gives the largest p-values at 0.3%, 1%, 5%, 99%, and 99.7% VaR levels, however not adequate at 95% VaR level. The FIGARCH-N is adequate at 0.3%, 1%, 5%, 99.7%

VaR levels, but not acceptable at 99% and 95% VaR levels.

For the JSEBNKS returns, FIGARCH-SSTD is adequate at all VaR levels, furthermore, FIGARCH-GED also produce the same results which indicate that is adequate at all levels. The FIGARCH-STD is also noticeable that it is adequate at all the VaR levels, meanwhile FIGARCH-N is not adequate at at 0.3% and 95% VaR levels. The JSEBNKS returns conclude that all the three models namely: FIGARCH-SSTD, FIGARCH-GED, and FIGARCH-STD are adequate at all VaR levels.

4.8.2 HYGARCH models

Table 4.16 presents the Kupiec LR test results, for JSEALSI returns HYGARCH-SSTD has the highest p-values at all VaR levels. The HYGARCH-STD has the lowest p-values at all VaR levels implying that HYGARCH-STD is not an adequate model. The JSEJMMNG has the largest p-values at all VaR levels, while HYGARCH-STD has the highest p-values at 0.3%, 1%, 5%, 99%, and 99.7%, however not adequate at 95% VaR level. For the JSEBNKS returns, both HYGARCH-SSTD and HYGARCH-GED have the highest p-values at all VaR levels. The HYGARCH-STD presents the highest p-values at all VaR levels.

Table 4.16: In-sample VaR backtesting: HYGARCH with different innovations for the three indices

Returns	Model	***p*-values of Kupiec LR test**					
		Long positions			**Short positions**		
		0.3%	**1%**	**5%**	**99.7%**	**99%**	**95%**
JSEALSI	HYGARCH-N	0.0004	0.0001	0.0077	0.5012	0.0300	0.0003
	HYGARCH-STD	0.0411	0.0151	0.0127	0.3349	0.0159	0.0007
	HYGARCH-SSTD	0.3012	0.5527	0.2758	0.5012	0.2986	0.2267
	HYGARCH-GED	0.0183	0.0242	0.0162	0.6045	0.0159	0.0002
JSEJMNNG	HYGARCH-N	0.7683	0.0580	0.2066	0.0861	0.0093	0.0295
	HYGARCH-STD	0.1473	0.3299	0.1060	0.9197	0.6915	0.0376
	HYGARCH-SSTD	0.1473	0.5541	0.2403	0.5019	0.5514	0.1106
	HYGARCH-GED	0.1473	0.3299	0.5242	0.3017	0.6915	0.0229
JSEBNKS	HYGARCH-N	0.0184	0.0282	0.7122	0.1674	0.0580	0.0907
	HYGARCH-STD	0.3344	0.8417	0.7839	0.3344	0.8396	0.5169
	HYGARCH-SSTD	0.3344	0.6836	0.7269	0.3344	0.8396	0.4442
	HYGARCH-GED	0.9197	0.6836	0.6442	0.4039	0.8417	0.1106

4.8.3 FIAPARCH models

As in the Table 4.17, for JSEJMMNG returns, FIAPARCH-N offers the largest p-values at all VaR levels except at 95% and 99.7% VaR levels. The FIAPARCH-STD and FIAPARCH-SSTD give the highest p-values at all VaR levels. The FIAPARCH-GED affords the largest p-values at all VaR levels except 95% VaR level.

For the JSEBNKS returns, the FIAPARCH-STD, FIAPARCH-SSTD, and FIAPARCH-GED are adequate at all VaR levels, meanwhile, the FIAPARCH-N was not adequate at 5% and 95% VaR levels. Overall, the long-memory GARCH models combined with heavy-tailed distributions used in this study were adequate in estimating VaR of JSEALSI, JSEJMMNG, and JSEBNKS returns at 0.3%, 1%, 5%, 95%, 99%, and 99.7% levels.

Table 4.17: In-sample VaR backtesting: FIAPARCH with different innovations for the three indices

Returns	**Model**	***p*-values of Kupiec LR test**					
		Long positions			**Short positions**		
		0.3%	**1%**	**5%**	**99.7%**	**99%**	**95%**
JSEJMNNG	FIAPARCH-N	0.6039	0.5541	0.4140	0.0412	0.2450	0.0176
	FIAPARCH-STD	0.8473	0.8977	0.3193	0.7019	0.8417	0.9029
	FIAPARCH-SSTD	0.1473	0.2079	0.5243	0.5019	0.8417	0.0737
	FIAPARCH-GED	0.1473	0.0942	0.5243	0.5019	0.8417	0.0134
JSEBNKS	FIAPARCH-N	0.0030	0.0093	0.5169	0.0030	0.0580	0.1603
	FIAPARCH-STD	0.9197	0.8915	0.7839	0.8473	0.8417	0.3058
	FIAPARCH-SSTD	0.6039	0.8396	0.7823	0.1473	0.6915	0.3442
	FIAPARCH-GED	0.5019	0.6915	0.6442	0.1473	0.8322	0.1106

4.9 Summary

This chapter presented the empirical results of fitting FIGARCH, HYGARCH, and FIAPARCH combined with SSTD, STD, and GED to JSEALSI, JSEJMMNG and JSEBNKS returns. We further calculated VaR at different levels and backtested using Kupiec LR test. The chapter gives a detailed conclusion, further study, and limitations of the study.

Chapter 5

Conclusion

From a risk management perspective, it is essential to hedge against the losses rather than the returns, consequently, this study is based on the loss distributions. The crucial issue regarding risk is that financial risk is not directly observable. We analyzed the statistical properties of the JSEALSI, JSEBNKS, and JSEJMMNG returns. The three returns exhibited the heavy tails, asymmetry, volatility clustering, and long memory. The long memory GARCH type models with heavy tailed innovations were used to capture the silent features of the three returns. VaR estimates were calculated using the long memory GARCH type models with heavy tailed innovations at 0.3%, 1%, 5%, 95%, 99%, and 99.7% levels. In-sample backtesting was employed to assess the adequacy of the heavy-tailed distributions and GARCH-type models used in this study by using the Kupiec LR test. For All Share Index returns, the FIGARCH-SSTD model seems to be overall adequate model at all the three long and short positions, in simpler terms provides highest p-values at 0.3%, 1%, 5%, 95%, 99%, and 99.7% VaR levels. This implies that the FIGARCH-SSTD model produces adequate VaR estimations at both long and short positions. For the Mining Index returns, the FIAPARCH-STD offers the highest p-values at 0.3%, 1%, 5%, 95%, 99% and 99.7% VaR levels. Finally, for the Banking Index returns, the FIAPARCH-STD affords the largest p-values at 0.3%, 1%, 5%, 95%, 99%, and 99.7% one can conclude that FIAPARCH-STD is the adequate model at all three long and short positions. The table 5.1 summarized most appropriate model at each VaR levels for each JSE indices.

Table 5.1: Most appropriate model selected for JSE Indices

VaR Level	JSEALSI	JSEJMMNG	JSEBNKS
0.3%	FIGARCH-SSTD	FIAPARCH-STD	FIAPARCH-STD
1%	FIGARCH-SSTD	FIAPARCH-STD	FIAPARCH-STD
5%	FIGARCH-SSTD	FIAPARCH-STD	FIAPARCH-STD
99.7%	FIGARCH-SSTD	FIAPARCH-STD	FIAPARCH-STD
95%	FIGARCH-SSTD	FIAPARCH-STD	FIAPARCH-STD
99%	FIGARCH-SSTD	FIAPARCH-STD	FIAPARCH-STD

Our findings confirm the importance of taking into account volatility clustering, heavy tails, asymmetry, and long memory in the behaviour of JSE indices. The empirical results show that FIGARCH, FIAPARCH, and HYGARCH combined with STD, SSTD, and GED innovations are suitable for depicting JSEALSI, JSEJMMNG, and JSEBNKS returns and can be used for VaR estimation.

The results are consistent with the study conducted by Mabrouk (2016), who used daily crude oil and gas applying the FIGARCH, HYGARCH, and FIAPARCH to accomplish the comparisons of the capabilities and concluded that FIAPARCH was the most suitable model. Similarly, Chkili et al. 2014 used crude oil, gas, and silver applying IGARCH, FIGARCH, and FIAPARCH, their conclusion was that FIAPARCH is the robust model. Sethaparamote (2014) used Thailand Stock Exchange time series data to FIGARCH, HYGARCH, and FIAPARCH and concluded that FIGARCH was the robust model. In the South African context, Reddy et al., 2017 came to a contrary conclusion, when applying JSEALSI to IGARCH, FIGARCH, and GJR-GARCH, and their conclusion indicated that GJR-GARCH was the robust to model the behaviour of the index.

In a future undertaking, the author intends to explore the multivariate time series approach using the three indices and applying the bivariate extreme value theory distributions and copulas.

5.1 Limitations of the study

The main objective of the study was to estimate the VaR using long-memory GARCH-type models incorporating heavy-tailed distributions to JSE Indices. The scope of the study was well defined and had no visible limitations in achieving the main objective of the study. We could have investigated the relative performance of our models on other financial indices.

References

[1] Acerbi, C. and Szekely, B., 2015. "Backtesting Expected Shortfall." *Available: http://workspace. Imperial.ac.uk/mathfin/Public/Seminars*[Accessed 30^{th} September 2017]

[2] Akaike, H., 1974. "A new look at the statistical model identification." *IEEE Transactions on Automatic Control Vol.19(6):716-723*

[3] Arfken, G., and Harris, F., 2013. "Mathematical Methods for Physicists." *Seventh Edition,2013*

[4] Artzner, P., Dealbean, F., Eber, J. and Heath, D., 1999. "Coherent measures of risk. Mathematical Finance, Vol.9(3):208-228." *Available: http://www.math.ethz.ch/del-bean/ftp/preprints/Coherent MF.pdf.*[Accessed 13^{th} October 2017]

[5] Ballie, R.T., 1996. "Long memory processes and Fractional Integration in Econometrics." *Journal of Econometrics, Vol.73(1):3-30.*

[6] Bentes, S.R., 2015. "Forecasting volatility in gold returns under the GARCH, IGARCH, and FIGARCH frameworks: New evidence." *Physica, Vol.438:355-364.*

[7] Black, F., 1976. "Studies of stock price volatility changes." *Proceedings of the Business and Economics Section of the American Statistical Association.*

[8] Bonga-Bonga, L. and Mutema, G., 2009. "Volatility forecasting and value at risk estimation in emerging markets: the case of the stock market index portfolio in South Africa" *South African Journal of Economic and Management Sciences, Vol.12(1):401-411*

[9] Caporin, 2003. "Identification of Long Memory in GARCH Models. Statistical Methods and Applications, Vol.12(2):133-151.

[10] Chifurira, R., and Chinhamu, K., 2019. "Evaluating South Africa's market risk using asymmetry power autoregressive conditional heteroskedasticity model

under heavy-tailed distributions." *Journal of Economics and Financial services, Vol.12(1):1-11.*

[11] Chifurira, R., and Chinhamu, K., 2017. "Using the Generalized Pareto and Pearson Type-IV Distributions to measure Value at Risk for the South African Mining Index." *Studies in Economics and Econometrics, Vol.51(1):159-182.*

[12] Chkili, W, Hammoudek, S.M., and Nguyen, D.K., 2014. "Volatility forecasting and risk management for commodity markets in the presence of asymmetry and long memory." *Jpag Business School Working Paper Series.*

[13] Cont, R., 2001. "Empirical properties of assets returns: Stylized facts and Statistical Issues." *Quantitative Finance, Vol.1:223-236*

[14] Conrad, C. and Haag, B., 2006. "Inequality constraints in the Fractionally Integrated GARCH Model." *Journal of Financial Econometrics, Vol.4(3):413-449*

[15] Conrad, C., 2010. "Non-negativity conditions for the Hyperbolic GARCH model." *Journal of Econometrics, Vol.157(2):441-457*

[16] Davidson, J., 2004. "Moment and memory properties of linear conditional heteroskedasticity models, and new model." *Journal of Business and Economics Statistics, Vol.22(1):16-29*

[17] Diaz, J., 2016. "Do scarce precious metals equate to safe harbour investment, the case of platinum and palladium." *Economic Research International,pages:1279-1313*

[18] Dickey, D. and Fuller, W., 1979. "Distribution of the estimators for Autoregressive time series with unit root." *Journal of American Statistical Association, No.74:427-431*

[19] Ding, Z. and Granger, C. and Engle, R., 1993. "A long memory property of stock market returns and a new model." *Journal of Empirical Finance, Vol.1:83-106.*

[20] Ding, Z. and Granger, C., 1996. "Modeling volatility persistence of speculative returns. A new approach." *Journal of Econometrics, Vol.73(1):185-215.*

[21] Elenjical, T., Mwangi, P., Panulo, B., and Haung, C.S., 2016. "A Comparative cross-regime analysis on the performance based value at risk models: Evidence of JSEALSI." *Risk Management International Journal, Vol.18.*

[22] Engle, F. and Manganelli, S., 2004. "CAViaR: Conditional Autoregressive Value at Risk by Regression quantiles." *Journal of Business Economics Statistics, Vol.22(4):367-381.*

[23] Fama, F., 1965. "The behaviour of stock market prices." *The Journal of Business, Vol.38:34-105.*

[24] Geweke, J. and Porter-Hudak, S., 1983. "The estimation and application of long memory time series models." *Journal Time Series Analysis Vol.4:221-238.*

[25] Giller, LG., 2005. "A generalized error distribution" *Viewed 20 October 2017:http://dx.doi.org/cggreen/10.2139/ssrn.22650.*

[26] Green, C., 2005. "Heavy-Tailed Distributions in Finance: An Empirical Study. Department of Statistics, University of Washington." *Viewed 18 September 2018:* `http://students.washington.edu/cggreen/uwstat/papers/computing_prelim_2005_green.pdf`

[27] Jarque, C. and Bera, A., 1987. "A test for normality of observations and regression residuals." *International Statistical Review, Vol.55:163-172.*

[28] JSE Bulletin., 2006. " JSE SA: Access to information Manual."

[29] Jorion, P., 2007. "Value at Risk: The new benchmark for managing financial risk. 2^{nd} Edition" *Chicago: Irwin Professional Publishing.*

[30] Huang, CS., and Huang, CK., 2014a. "Assessing the Relative Performance of Heavy-tailed Distributions: Empirical Evidence from the Johannesburg Stock Exchange." *Journal of Applied Business Research Vol.30(4):1263-1286.*

[31] Huang, CS., Huang, CK., Chinhamu, K., and Hammujuddy, J., 2014b."Generalized Hyperbolic Distributions and Value at Risk Estimation for the South African Mining Index." *International Business and Economics Research Journal, Vol. No.13(2):319-328.*

[32] Hansen, B., 1994. "Autoregressive Conditional Density Estimation." *International Economic Review, Vol.35:127-161*

[33] Hull, J., 2012. "Risk Management and Financial Institution, 3^{rd} Ed. New Jersey" *John Wiley and Sons Inc.*

[34] Hurst, E., 1951. "Long term storage capacity of reservoirs." *Transactions of the American Society of Engineers.*

[35] Kasman, A., 2009. "Estimating Value at Risk for the Turkish Stock Index Futures in the Presence of Long Memory Volatility." *Central Bank Review, Vol.1:1-14.*

[36] Katzke, N. and Garbers, C., 2015. "Do Long Memory and Asymmetries Matter When Assessing Downside Return Risk?." *Stellenbosch Economic Working Papers.*

[37] Klar, B., Lindner, F., and Meintanis, S., 2012. "Specification tests for the Error Distribution in GARCH-models." *Journal of Computational Statistics and Data Analysis, No.56(2012):3587-3598.*

[38] Kunsch, H., 1987. "Statistical aspects of self-similar processes, In: Prohorov, Y., Sazarov, V. (Eds.)." *Proceedings of the First World Congress of the Bernoulli Society. Vol.1.VNU Science Press, Utrecht.*

[39] Kupiec, P., 1995. "Techniques for verifying the accuracy of risk management models." *J. Deriv, 3(2):78-84.*

[40] Kwiatkowski, D., Phillips, P., Schmidt, P., and Shin, Y., 1992. "Testing the null Hypothesis of Stationarity against the Alternative of a unit root." *Journal of Econometrics, Vol.54(13):159-178.*

[41] Ljung, G. and Box, G., 1978. "On a measure of a lack fit in time series models." *Biometrika, Vol.65(2):297-303.*

[42] Li, M., Li, W.K., and Li, G., 2015. "A new hyperbolic GARCH model." *Hong Kong Research Grants Council.*

[43] Lo, A., 1991. "Long memory in stock market prices." *Econometrica, Vol.59:1279-1313*

[44] Mabrouk, S., 2016. "Forecasting financial assets volatility using Integrated GARCH type models: International evidence." *Journal of Finance and Economics, Vol.4(2):54-62. Available: http://pubs.sciepub.com/;fe/4/2/3.*[Accessed 21st October 2017]

[45] Maghyereh, A., and Awartani, B., 2012. "Modelling and forecasting Value at Risk in the UAE Equity markets: the role of long memory, fat tails and asymmetrics in return innovations." *Review of Middle East Economics and Finance, Vol.8(1):64-79*

[46] Makhwiting, R., Sigauke, C., and Lesaona, M., 2014. "Modelling tail behaviour of returns using the generalized extreme value distribution." *Masters of Science, University of Limpopo*[Accessed 24st October 2017]

[47] McMillan, D. and Thupayagale, P. 2010. ”Evaluating Stock Index Return Value at Risk Estimates in South Africa: Comparative Evidence for Symmetric, Asymmetric and Long Memory GARCH Models.” *Journal of Emerging Market Finance, 20,Vol.9(3):325-345*

[48] Mighri, Z.B., Mokni, K., and Mansouri, F., 2010. ”Empirical analysis of asymmetric long memory volatility models in value at risk estimation.” *Journal of Risk, Vol.13(1):55-128.*

[49] Mincer, J. and Zarnowitz, V., 1969. ”The evaluation of economic forecasts and expectations.” *Columbia University Press, New York.*

[50] Morgan, J. and Reuters, 1969. ”Riskmetrics Technical document.” *Available: http://pascal.iscg.utl.pt/aafonso/eif/rm/TD4ePt.4.pdf* [Accessed 19^{th} October 2017]

[51] Newey, W. and West, K., 1987. ”A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix.” *Econometrica, Vol.55:335-346*

[52] Perron, P. and Qu, Z., 2008. ”Long memory and level shifts in the volatility of stock market return indices.” *Unpublished Manuscript, Department of Economics, Boston University.*

[53] Peters, J. P, 2001. ”Estimating and forecasting volatility of stock indices using asymmetric GARCH models and (Skewed) Student’ t-densities. University of Liege” *available at: https://pdfs.semanticsholar.org/f445/493c46b364d52e958f69d502ccf8f03d51c3.pdf* [Accessed 17^{th} September 2017]

[54] Poon, S. and Granger, C., 2003. ”Forecasting volatility in financial markets.” *A Review,Journal.*

[55] Purczynski, J., 2014. ”Estimation of the shape parameter of GED distribution for a small sample size.” *Folia Oeconomica Stetinensia. DOI:* `10.2478/foli_2014_0103`

[56] Reddy, L., Pillai, D., Haung, C.S., and Haung, C.K., 2017. ”Forecasting Stock market volatility in the presence of structural breaks: An application to Value at Risk Estimation in South Africa.” *SASA Proceedings of the* 59^{th} *Annual Conference, 2017.*

[57] Robinson, P., 1995a. ”Gaussian semi-parametric estimation of long range dependence.” *Annals of Statistics, Vol.23:515-539*

[58] Sethaparamote, Y., Prukumpai, S., and Kanyamee, T., 2014. ”Evaluation of value at risk estimation using long memory volatility: evidence from stock exchange of Thailand.” *SSRN Paper, available at: https://papers.ssrn.com/sol3/papers.cfm? abstract-id=2396531*[Accessed 15^{th} September 2017]

[59] Schwarz, G., 1978. ”Estimating the dimension of a model.*Änn Statist., Vol.6:461-464.*

[60] Shapiro, S. and Wilk, M., 1965. ”An analysis of variances test for normality (complete samples).” *Biometrika, Vol.52:3-4.*

[61] Tayeti, M. and Rammanathan, T.V., 2012. ”An overview of FIGARCH and related Time Series Models.” *Austrian Journal of Statistics ,Vol.41:44-54.*

[62] Toerin, F., Rosenberg, D. and Kruger, R., 2014. ”The asymmetry of gain loss time horizon on the JSEALSI.” *Journal OF Studies in Economics and Econometrics , No.38(1):65-74.*

[63] Tsay, R. S., 2005. ”Analysis of Financial Time Series. 2^{nd} Edition” *John Wiley & Sons.*

[64] Whittle, P., 1951. ”Hypothesis testing in time series analysis.” *Uppsala: Almqvist and Wikssells Boktrychen AB.*

[65] Wojtowicz, T. and Gurgul, H., 2009. ”Long memory of volatility measures in time series.” *Available at: https://www.researchgate.net/publication/227653938*[Accessed 18^{th} September 2017

[66] Yaya, S., 2013. ”Nigerian Stock Index: A search of optimal GARCH model using high frequency data.” *Journal of Applied Statistics, No.4(2):69-85.*

Appendix A

Proof 1:*The conditional variance of FIGARCH model is given by:*

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2,$$

where L is the operator, $\alpha(L) = \sum_{i=1}^{q} \alpha_i L^i$, $\beta(L) = \sum_{j=1}^{p} \beta_j L^j$, then $\sigma_t^2 - \beta(L)\sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2$, but $\nu_t = \varepsilon_t^2 - \sigma_t^2 \Longrightarrow \varepsilon_t^2 = \nu_t + \sigma_t^2$, then $\sigma_t^2 = \varepsilon_t^2 - \nu_t$

$$\begin{aligned}
[1-\beta(L)]\sigma_t^2 &= \omega + \alpha(L)\varepsilon_t^2 \\
[1-\beta(L)](\varepsilon_t^2 - \nu_t) &= \omega + \alpha(L)\varepsilon_t^2 \\
[\varepsilon_t^2 - \beta(L)\varepsilon_t^2 + \nu_t + \beta(L)\nu_t] &= \omega + \alpha(L)\varepsilon_t^2 \\
\varepsilon_t^2 - \beta(L)\varepsilon_t^2 - \alpha(L)\varepsilon_t^2 &= \omega - \beta(L)\nu_t - \nu_t \\
[1-\beta(L)-\alpha(L)]\varepsilon_t^2 &= \omega + [1-\beta(L)]\nu_t \\
[1-\beta(L)-\alpha(L)]\varepsilon_t^2 &= \omega + [1-\beta(L)](\varepsilon_t^2 - \sigma_t^2),
\end{aligned}$$

and the differencing operator $(1-L)^d$. The Taylor- expansion formula below is:

$$\begin{aligned}
(1-L)^d &= \sum_{k=0}^{\infty} \frac{\Gamma(d+1)L^k}{\Gamma(k+1)\Gamma(d-k+1)} \\
&= 1 - dL - \frac{d(1-d)L^2}{2!} - \frac{d(1-d)(2-d)L^3}{3!} - \cdots \\
(1-L)^d &= 1 - \sum_{k=0}^{\infty} C_k(d)L^k \\
\phi(L)(1-L)^d\varepsilon_t^2 &= \omega + [1-\beta(L)\varepsilon_t^2 - \phi(L)(1-L)^d\varepsilon_t^2] \\
[1-\beta(L)]\sigma_t^2 &= \omega + [(1-\beta(L)\varepsilon_t^2 - \phi(L)(1-L)^d\varepsilon_t^2] \\
\sigma_t^2 &= \omega[1-\beta(L)]^{-1}[1-(1-\beta(L))^{-1}\phi(L)(1-L)^d]\varepsilon_t^2
\end{aligned}$$

Appendix B

Proof 2:*The HYGARCH is reduced into a GARCH and a FIGARCH by the following equation:*

$$\begin{aligned}
\Theta(L) &= 1 - \tfrac{\varphi(L)}{1-\beta(L)} \\
\Psi(L) &= 1 - \tfrac{\varphi(L)(1-L)^d}{1-\beta(L)} \\
\Xi(L) &= 1 - \tfrac{\varphi(L)(1+\eta(1-L)^d-1)}{1-\beta(L)}
\end{aligned}$$

, this denotes the $ARCH(\infty)$ lag polynomials for GARCH, FIGARCH and HYGARCH, respectively, where for $\Psi(L)d = 0$ holds. Then, it easily follows for $\Xi(L)$ by adding an absolute zero.

$$\begin{aligned}
\Xi(L) &= \eta - \eta\tfrac{\phi(L)(1-L)^d}{1-\beta(L)} + (1-\eta) - (1-\eta)\tfrac{\phi(L)}{1-\beta(L)} \\
&= \eta\left(1 - \tfrac{\phi(L)(1-L)^d}{1-\beta(L)}\right) + (1-\eta)\left(\tfrac{\phi(L)}{1-\beta(L)}\right)
\end{aligned}$$

The bigger the value for η in this linear combination, the higher the influence of the long memory FIGARCH part and the less the short memory GARCH part. Secondly, restrictions must be derived for which the process assures weak stationarity.

$$E(\varepsilon_t) = 0, \forall_t$$

, and

$$Cov(\varepsilon_t, \varepsilon_{t-j}) = 0, \ \forall_t \forall_j \in N \tag{4.13}$$

GARCH polynomials offer weak stationary if $\Theta(1) < 1$ is fulfilled, which is an alternative definition of the additional common condition $\varphi(1) = 1 - \alpha(1) - \beta(1) > 0$ from the ARIMA representation of GARCH). Since, FIGARCH is not able to provide weak stationarity $\psi(1) = 1$ for $d \in (0, 1)$ must hold. Thus, $\eta + (1-\eta)\Theta(1) < 1$, is fulfilled if and only if

$$\Theta(1) = 1 - \frac{1 - \alpha(1) - \beta(1)}{1 - \beta(1)} = \frac{\alpha(1)}{1 - \beta(1)} < 1 \tag{4.16}$$

, and $\eta \in (0, 1)$ constitutes a linear combination as mean between GARCH and FIGARCH polynomial. This is true for $\eta = 0$ and the parameter restrictions for the HYGARCH is considered to be weak stationary result, given in the above equation.

$$1 - \frac{\alpha(1)}{1 - \beta(1)} > 0 \tag{4.17}$$

Conrad (2010) discovered that a weak stationary HYGARCH under minor modifications is feasible to be obtained, even for $\eta \geq 1$.

Appendix C

Derivation of the Student's t- Distribution

The characteristic function of x is derived as follows:

$$\begin{aligned}
\phi(t) &= E\Big[exp\Big(\tfrac{-t^2x^2}{2}\Big)\Big] \\
&= \frac{\Gamma\Big(\frac{\nu+1}{2}\Big)}{\sigma\sqrt{\nu\pi}\Gamma\Big(\frac{1}{2}\Big)}\int_{-\infty}^{\infty}\Bigg(1+\frac{x^2}{\nu\sigma^2}\Bigg)^{\frac{-\nu+1}{2}} exp\Big(\frac{-t^2x^2}{2}\Big)dx \\
&= 2\frac{\Gamma\Big(\frac{\nu+1}{2}\Big)}{\sigma\sqrt{\nu\pi}}\int_{0}^{\infty}\Bigg(1+\frac{x^2}{\nu\sigma^2}\Bigg)^{\frac{-\nu+1}{2}} exp\Big(\frac{-\nu\sigma^2t^2x^2}{2\nu\sigma^2}\Big)dx
\end{aligned}$$

$$Let \quad y = \tfrac{x^2}{\nu\sigma^2}, x \in (0,\infty), \text{then } x = \sqrt{\nu\sigma^2}y^{\frac{1}{2}}, \text{and } dx = \tfrac{\sqrt{\nu\sigma^2}}{2}y^{\frac{-1}{2}}dy$$

$$\begin{aligned}
\phi(t) &= \frac{\Gamma\Big(\frac{\nu+1}{2}\Big)}{\sigma\sqrt{\nu\pi}\Gamma\Big(\frac{1}{2}\Big)}\frac{\sqrt{\nu\sigma^2}}{2}\int_{0}^{\infty} y^{\frac{-1}{2}}\big(1+y\big)^{\frac{-\nu+1}{2}} exp\Big(\frac{-\nu\sigma^2t^2y}{2}\Big)dy \\
&= \frac{\Gamma\Big(\frac{\nu+1}{2}\Big)}{\sqrt{\pi}\Gamma\Big(\frac{\nu}{2}\Big)}\int_{0}^{\infty} y^{\frac{-1}{2}}\big(1+y\big)^{\frac{-\nu+1}{2}} exp\Big(\frac{-\nu\sigma^2t^2y}{2}\Big)dy \\
&= \frac{\Gamma\Big(\frac{\nu+1}{2}\Big)}{\sqrt{\pi}\Gamma\Big(\frac{\nu}{2}\Big)}\Gamma\Big(\tfrac{1}{2}\Big)\psi\Big(\tfrac{1}{2},\tfrac{3}{2}-\tfrac{\nu+1}{2},\tfrac{\nu\sigma^2t^2}{2}\Big) \\
&= \frac{\Gamma\Big(\frac{\nu+1}{2}\Big)}{\Gamma\Big(\frac{\nu}{2}\Big)}\psi\Big(\tfrac{1}{2},\tfrac{3}{2}-\tfrac{\nu+1}{2},\tfrac{\nu\sigma^2t^2}{2}\Big)
\end{aligned}$$

, where $\int_0^{\infty} y^{\alpha-1}\big(y+m\big)^{-q}exp^{-ny}dy = \Gamma\big(\alpha\big)m^{\alpha-q}\psi\big(\alpha,\alpha+1-q,nm\big)$, and $\alpha=\frac{1}{2}$, $m=1$, $q=\frac{\nu+1}{2}$ and $n=\frac{\nu\sigma^2t^2}{2}$.