University of Venda

MODELLING EQUITY RISK AND EXTREMAL DEPENDENCE: A SURVEY OF FOUR AFRICAN STOCK MARKETS

By

Richard Abayomi Samuel

(Student no: 14010224)

SUBMITTED IN FULFILLMENT OF THE
REQUIREMENTS FOR
MASTER OF SCIENCE DEGREE IN STATISTICS
AT
UNIVERSITY OF VENDA
THOHOYANDOU, SOUTH AFRICA
AUGUST 2018

UNIVERSITY OF VENDA
DEPARTMENT OF
STATISTICS

The undersigned hereby certify that they have read and recommend to the School of Mathematical and Natural Sciences for acceptance a project entitled "**Modelling equity risk and extremal dependence: A survey of four African stock markets**" by **Richard Abayomi Samuel(Student no: 14010224)** in fulfillment of the requirements for the degree of **Master of Science Degree in Statistics**.

Dated: August 2018

Supervisor: _____
Dr. Caston Sigauke

Co-Supervisor: _____
Dr. Alphonce Bere

_____

UNIVERSITY OF VENDA

Date: **August 2018**

Author: **Richard Abayomi Samuel**
(Student no: 14010224)

Title: **Modelling equity risk and extremal dependence: A survey of four African stock markets**

Department: **Statistics**

Degree: **M.Sc.** Convocation: **May** Year: **2019**

Signature of Author

Abstract

The ripple effect of a stock market crash due to extremal dependence is a global issue with key attention and it is at the core of all modelling efforts in risk management. Two methods of extreme value theory (EVT) were used in this study to model equity risk and extremal dependence in the tails of stock market indices from four African emerging markets: South Africa, Nigeria, Kenya and Egypt. The first is the "bivariate-threshold-excess model" and the second is the "point process approach".

With regards to the univariate analysis, the first finding in the study shows in descending hierarchy that volatility with persistence is highest in the South African market, followed by Egyptian market, then Nigerian market and lastly, the Kenyan equity market. In terms of risk hierarchy, the Egyptian EGX 30 market is the most risk-prone, followed by the South African JSE-ALSI market, then the Nigerian NIGALSH market and the least risky is the Kenyan NSE 20 market. It is therefore concluded that risk is not a brainchild of volatility in these markets.

For the bivariate modelling, the extremal dependence findings indicate that the African continent regional equity markets present a huge investment platform for investors and traders, and offer tremendous opportunity for portfolio diversification and investment synergies between markets. These synergistic opportunities are due to the markets being asymptotic (extremal) independent or (very) weak asymptotic dependent and negatively dependent. This outcome is consistent with the findings of Alagidede (2008) who analysed these same markets using co-integration analysis. The bivariate-threshold-excess and point process models are appropriate for modelling the markets' risks. For modelling the extremal dependence however, given the same marginal threshold quantile, the point process has more access to the extreme observations due to its wider sphere of coverage than the bivariate-threshold-excess model.

Keywords: Bivariate-threshold-excess model, Extreme value theory, Generalized Pareto distribution, Poisson point process, Tail dependence, Volatility.

To God Almighty.

Table of Contents

List of Tables

List of Figures

Acknowledgements

My first and foremost appreciation goes to the Lord and Saviour Jesus Christ who has helped me enormously to complete this project and has put another feather in the cap of my career. He is the Professor in my career and to Him be the glory. I want to thank and give my warmest gratitude to my supervisor, Dr. Caston Sigauke for his unreserved generosity, kindness, patience, guidance, various financial support from him during the research and for his high professional skills throughout the entire period amidst his busy schedules. He has groomed my research skill far beyond my expectation. I am very grateful sir. I also extend this thanks to my co-supervisor, Dr. Alphonce Bere, for his thorough supervision and valuable feedback and he made me to always cross my t's and dot my i's in the write-up with attention to details.

My heart-felt regards and gratefulness goes to Dr. Kyei Kwabena, for his fatherly care, thoughtful advice, relentless efforts and always being there to render help. I also want to thank the two anonymous external reviewers for their thorough comments and suggestions during the reviews. Finally, I am extremely grateful to my family members, especially my sisters, and my brother James Idowu Samuel, whose love and support were crucial to the successful completion of this project. Thank you all.

Chapter 1

Introduction

1.1 Overview

Investment and trading are major sources of income to investors and traders alike. Hence, day to day activities and any unusual movement in the markets where these two activities take place is of paramount concern to these market participants. Accurate modelling and probabilistic forecasting of a rare event whose magnitude could be worse than historical data outliers has become a niche to actuaries and econometrists.

After series of global financial crises to date, empirical researchers and theorists are coming up with continuous modelling of tail dependence underpinned with the aim of accurate probabilistic prediction of the return levels of rare events in the financial field. Such forecasts should adequately use previous extreme values to proffer cautious information on the magnitude and probability of potential values that can pose more extreme threat than those previously observed. This is in line with a potential "black swans" episode described by Taleb (2010).

Two methods of extreme value theory (EVT) were used in this study to model equity risk and extremal dependence in the tails of stock market indices from four African emerging markets: South Africa, Nigeria, Kenya and Egypt. The first is the "bivariate-threshold-excess model" and the second is the "point process approach".

The first fits a generalized Pareto distribution (GPD) model to excesses above a sufficiently high threshold, while the second has a non-homogeneous Poisson process as the limiting distribution for excesses above the same selected threshold as the first.

Risks are adverse market movements that are of major concern to risk managers and investors. The practice of detecting potential risks in advance, analyzing them and taking precautionary steps to mitigate the risk is referred to as risk management. The behaviour of a portfolio during a financial crisis is an important element of risk management because the general perception about investing is that the future could either yield gratifying rises or worrisome drops when the price swing is overvalued or undervalued.

A stock market crash is a quick and usually an unexpected drop in stock prices which may occur as a result of key disastrous events, economic watershed, or the failure of a long-term projected bubble where securities and other investments are trading at prices far beyond their basic values.

Traders, investors and fund managers practice risk management so as to control and minimize their investment risks locally and internationally. Inadequate plans for risk management could be disastrous on investment during financial turmoil in an economy, and could result in a doom and gloomy end for an investor. As a result of this, portfolio diversification is of essence to equity investors, traders and fund managers in order to minimize their risk exposures. Diversification can potentially eliminate risk if returns are uncorrelated (Levy and Sarnat, 1970).

Extreme financial risks are posed by extreme or rare events occurrence at the tails of the returns' distributions. The study of the risk in the tails of equity returns is a focal area in financial markets research. These events may materialize as large positive or negative investment returns, a stock market crash, major defaults, or the collapse of risky asset prices (Uppal, 2013). Extreme events studies include areas such as catastrophic structural failures, extreme heat waves, natural disasters, heavy

rains, storms, currency crises, stock market crashes etc.

This study used both univariate and bivariate structures to model rare events in the financial returns of the four selected African regional markets. The relationship between the markets at the tail region will give a deeper insight of their extremal behaviour, which is termed extremal dependence in the literature. Two or more markets can co-move at the tails of the returns distribution due to contagion effect, economic integration, and stock market characteristics (Pretorius, 2002). Contagion effect is not measurable in itself but rather estimated with the residual from the co-movement that is not explained by fundamentals (Pretorius, 2002). Contagion results from investors rapidly changing their positions, because they think a crisis in one country will eventually affect the others and this may finally result in herd behaviour effects (i.e. panic selling).

Fundamental linkages among stock markets normally emanate from economic linkages that exist between economies. Such economic linkages portray the extent to which countries have close investment, trade and any other forms of economic ties. With strong economic links, bad or good market shock caused by an extremal (rare) event can easily reverberate across regional economies.

Economic integration implies that the more the economies of two countries are integrated, the more interdependent their stock markets will be. Strong multilateral trade relations between two or more nations like trading blocs may augment dependence and bring a higher level of co-movement among their stock markets. Examples of such blocs in Africa include the Common Market for Eastern and Southern Africa (COMESA), Middle East and North Africa (MENA), and the Economic Community of West African States (ECOWAS). COMESA is a free trade area with an operating agreement comprising of nineteen member states that include Egypt and Kenya, and the bloc is one of Africa's economic community trading pillars. The assumption here is that co-movement of equity markets occurs when nations have strong trade links

and have made unanimous economic reforms.

This study examines these four regional African equity markets to determine the degree to which they are integrated within the African continent and ascertain the extremal dependencies among them. Stock markets interdependence and integration through trading blocs indicate improved co-movements of markets which may in turn reduce potential opportunities obtainable from portfolio diversification (Yu and Hassan, 2006). In order to have a successful portfolio diversification, a low or negative co-movement among equity markets is necessary to enable potential poor performance in one side of the stock market to be relatively hedged by good returns on the other side of the paired markets.

Extreme value theory (EVT) has been prescribed by many as a potential tool for tackling tail risk caused by rare events. Thus, EVT models are prominently applied where consideration is given to unusually large or small movements at the tails of the distributions using extreme observations. Volatility is filtered out before the application of EVT models since it is known that the removal of linear dependence and heteroscedasticity from the financial returns makes the EVT modelling more satisfactory (Smith, 2003). McNeil and Frey (2000) also support the use of conditional EVT structure with the combination of GARCH and EVT modelling.

1.2 Research questions

This research should result in responses to the following questions:

1. Which of the markets is most risk-prone?
2. Is asset portfolio diversification feasible among the selected regional markets?
3. What drives the dependence and asymptotic (tail or extremal) dependence of these selected markets?

4. How should investors and all market participants summarize and interpret the dependence structure of the selected markets?

5. Which between the bivariate-threshold-excess model and point process model will be more appropriate to measure the risks and extremal dependence structures of the stated markets?

1.3 Aim and objectives

1.3.1 Aim of the study

The main aim of this study is to model risk and find out if extremal dependence occurs in the pairwise combinations of the selected markets using two-dimensional or bivariate extreme value analysis vis-*à*-vis the bivariate-threshold-excess method of GPD and point process model.

1.3.2 Objectives of the study

The objectives of this study are, to:

1. ascertain the market that displays the greatest volatility among the four. As asserted by Djakovic et al. (2011) and Poon et al. (2004), that performance of the EVT models could be more appropriately studied in market environment with volatility,

2. examine whether tail dependence is more pronounced in equity markets of economies with trading links and also estimate the return levels of rare events in the tails of the distribution in each of the four markets,

3. investigate the type of trend driving the extreme movements of the stated African stock markets,

4. ascertain in hierarchical order the risks and the pairwise combination of markets that are more susceptible to extremal dependence among the four selected ones. The higher (lower) the susceptibility between two markets, the higher (lower) the co-movement and the lower (higher) the opportunity for portfolio diversification,

5. find out the EVT model (between the bivariate-threshold-excess model and the point process model) that will better characterise the risk and tail dependence of the paired markets.

1.4 Significance of the study

The target professionals in need of the findings of this study are investors (both domestic and international) and traders who want to take advantage of the benefits of portfolio diversification if they exist in the selected key African regional markets. In addition, the general markets participants and in particular, policy makers are keenly interested in the outcome of this research for effective policy institution and smooth market control.

The findings will add new facts on risk measurement, management, and extremal dependence observed in the chosen markets to currently available facts in the literature.

1.5 Layout of the project

The rest of the project is structured as follows: Chapter 2 centers on literature review while Chapter 3 presents the methodology that includes data description and the research methods. The data will be analyzed and results interpreted in Chapter 4 using

the proposed models. Conclusions, based on the outcomes, and recommendations will be discussed in Chapter 5.

Chapter 2

Literature Review

2.1 Introduction

In this chapter, various existing articles and monographs in the literature on equity risk management and tail dependence modelling will be scrutinized with up-to-date findings and conclusions of researchers on these.

2.2 Background

The literature is filled with series of studies, information, and narratives on risk measurement and management coupled with extremal behaviours among different equity markets. A range of linear to non-linear methodologies have been applied in such risk analysis and modelling like Value at Risk (VaR), Expected Shortfall (ES), co-integration analysis, Extreme Value Theory (EVT), just to mention a few. Among these methods, the EVT approach is found to outperform the others when compared to modelling financial risk and tail dependence (Dacorogna et al., 1995). Evidence from the literature indicates that most of the studies on risk and extreme tail behaviours have been applied to markets in the developed nations, while markets in emerging economies, especially African equity markets are quite limited and still trail

far behind in the study of extremal dependence. Thus, the dynamic investigation of the risk and tail or extremal dependence characterizing these African markets is of utmost importance to international investors who are considering portfolio diversifications in the continent's markets due to its burgeoning socio-economic features.

A huge number of studies dating back to the 1970s have witnessed low correlation between emerging and mature markets, which in turn create avenues for reasonable portfolio diversification (Ripley, 1973). Recent studies however are beginning to reveal the effect of liberalization and globalization of financial markets, where interdependence and integration of markets are increasing, hence, placing opportunities for portfolio diversifications at a low ebb (Yu and Hassan, 2006). Also research conducted by Haque et al. (2004) concluded that emerging markets like the Middle East and Africa are investment havens for good returns opportunities to investors, considering the trade-off between risk and returns.

Alagidede (2008) used the concept of co-integration analysis with the aim of analyzing the market linkages among South Africa, Nigeria, Egypt, and Kenya. The researcher observed that the markets do not significantly co-move in spite of the subsisting cooperation and economic reforms among them. In line with the markets' indices covered by Alagidede (2008), this study will focus on these four markets with a different approach using a bivariate-threshold-excess model and point process of EVT.

Smith (2003) used extreme value models via the point process approach to model the daily returns for three companies: Pfizer, Citibank and General Electric, from 1982 to 2001. These returns were modelled to analyse the value at risk (VaR) where volatility was filtered using GARCH (1, 1) model. The researcher concluded that it is reasonable to remove ARCH effect and linear dependency to obtain more satisfactory results when modelling using extreme value theory.

The financial returns of S & P 500, Nikkei – 225 and Istanbul Stock Exchange –

100 indices were modelled by Altun and Tatlidil (2015) using GARCH-EVT models. These same returns were also modelled using the GARCH approach jointly with each of the normal, student's t, GED and skew-GED distributions. It was concluded that the GARCH model with each of the distribution performed better than the GARCH-EVT models.

Lipika (2018) evaluated the efficiency of the extreme value theory models with regards to the univariate and multivariate versions. The study was conducted on four Western Cape (South Africa) provinces' climate data from 1965 to 2015 using the trio of Block maxima, point process and threshold excess models. The results indicated that the duo of the point process and threshold excess approaches outperformed the block maxima approach. Furthermore, it was observed that the point process and threshold excess models performed equally well in the univariate analysis, but the former outperformed the latter in the bivariate analysis due to availability of more data.

In other work, two parametric dependence models of point process were compared and the non-parametric approach of polar coordinates also used in a study by Joe et al. (1992). This comparison was done using 504 pairwise combinations of nitrate and sulphate concentrations in the same area in the U.S. The parametric dependence models used are the bilogistic and logistic models. The study concluded that the bilogistic was a better model for modelling the dependence of the pairwise chemical concentrations when compared to the logistic model under both parametric and non-parametric methods.

Bali (2003) investigated extreme movements of the U.S. Treasury yields from mid 1950s through 1998 and recommended EVT based methods for VaR analysis over the conventional normal distribution model based approach. For narratives about analytical proficiency of EVT in risk management, and multivariate approach to extreme value theory, see Longin and Solnik (2001).

2.3 Conclusion

This study uses daily data from the four selected market indices (JSE-ALSI in South Africa, NIGALSH in Nigeria, EGX 30 in Egypt and NSE 20 in Kenya) with emphasis on:

1. knowing the volatility and risk hierarchies of the markets for investors' decision making,
2. modelling the extremal dependence in the pairwise combinations of the markets.

With emphasis focused on these four markets, whose extremal dependencies have been modelled by Alagidede (2008) using co-integration analysis, this study will use different approaches. The two modelling approaches that will be used on the markets are the bivariate-threshold-excess model and point process. To the best of the author's knowledge, these two EVT models have not been used on these four markets before this study.

Chapter 3

Methodology

3.1 Introduction

The various methods used to analyze, measure and manage the risk inherent in the four stated market indices will be examined broadly in this chapter. The risk associated with each of the markets and tail dependence in the paired markets among the selected four indices will be modelled by the bivariate-threshold-excess method and the point process with GPD and the non-homogeneous Poisson process as the limiting distributions respectively. Various R statistical packages including "texmex", "evmix", "ismev", "extRemes", "evd", "eva" and "rugarch" are going to be used for the required modelling. These packages are downloadable via CRAN repository. Also, the maximum likelihood estimation method is used for estimating the parameters, followed by checking for model adequacy and residual analysis.

3.2 Overview of extreme value theory

Following the criticism and challenges of using the VaR (Value at Risk) approach for modelling risk, extreme value theory models have adequately been able to bridge this modelling problem in risk management. This criticism is based on VaR normal

distribution assumption to modelling financial risk, which in turn makes it unsuitable for modelling rare events occurrences at the tails of returns distribution. Actually, financial returns are known to exhibit leptokurtosis or fat tails and the normal distribution is usually not appropriate to estimate such tails behaviour. This deficiency of the normal distribution thus creates a gap bridging opportunity for the application of extreme value theory (EVT) models and that is where the theory becomes very relevant. The EVT based VaR methodology has been developed with comprehensive statistical theory and used for financial data analysis.

The beauty of the use of EVT approach is more pronounced when multiple variables are modelled in what is known as the MEVT (multivariate extreme value theory) method. In this case, two or more variables are used to forecast events movement. This approach is somewhat more robust and it reflects the versatility in the use of the extreme value theory as a model for rare occurrences. In this study, the focus is limited to the bivariate case.

3.3 The history of EVT

The origin of the extreme value limiting distribution as the limits for maxima distribution in samples of independent and identically distributed random variables can be traced to the work of Fisher and Tippett (1928). Classical references on the monographs for fitting the distributions, especially in environmental applications, are Hosking, Wallis and Wood (1985), Prescott and Walden (1980). Modern development in the use of the "peaks-over-threshold" (POT) approach came into limelight with a paper from Todorovic and Zelenhasic (1970). Procedures on the mathematical development of this with regards to extreme order statistics are found in the works of Pickands (1975), Hill (1975), and Weissman (1978). Leadbetter et al. (1983) also contributed to the development of the largest order statistics' limiting distribution.

Smith (1984), and Davidson and Smith (1989) elaborate the GPD as a stable distribution of threshold excesses.

For the joint distribution of largest order statistics, where threshold and the classical approaches are partly combined (see Gomes, 1981 and Tawn, 1988). Sibuya (1960) describes the asymptotic independence and dependence of variables of the multivariate extreme value theory while de Haan (1985), and de Haan and Resnick (1977) avail the multivariate point process. Detailed illustration of these distributions in this study is ensued by considering the following statistical methods of EVT application: the Block Maxima Method (BMM), the bivariate-threshold-excess method and the point process.

3.4 Block maxima method

With block maxima model, extreme losses are divided into identical blocks where maximum loss in each block is the largest observation. This procedure is termed "block maxima" with each local block containing maximum loss. Thus, the extreme losses can be characterized by the local block maxima which are the data to be fitted or modelled using this method.

Theorem 3.4.1 (Fisher-Tippett-Gnedenko)**.** *Assume that* X_1, X_2, ..., X_n *are i.i.d (independent and identically distributed) random variables (losses) with distribution function* F *and let* $M_n = max(X_1, ..., X_n)$*, that is, the maximum of each block of* n *observations. Given* b *blocks with size* m *in each of the block, there will be* $n = b \times m$ *number of observations in total. The maximum of the* n *observations in block* i *can be defined as* $M_{in} = max(X_{i1}, X_{i2}, ..., X_{in}), i = 1, 2, ..., b$*. Now suppose there exist sequence of normalizing constants* $a_n > 0$ *and* $d_n \in \mathbb{R}$ *and some limiting distribution*

function which are non-degenerate such that

$$Pr\left[\frac{M_{in}-d_n}{a_n} \leq x\right] = F^n(a_n x + d_n) \to H(x), \ \ as\ n \to \infty, \tag{3.4.1}$$

where $(M_{in} - d_n)/a_n$ *is a centered and normalized maximum.*

If the distribution of this normalized maximum converges to H as the block size increases, then H is described as belonging to one of these three standard families of extreme value distribution classes

$$\begin{aligned}
\Psi(x) &= e^{-e^{-x}}, \quad x \in \mathbb{R}, && \text{i.e. the Gumbel distribution with } \xi \to 0,\\
\Lambda(x)_\alpha &= \begin{cases} 0, & x \leq 0 \\ e^{-x^{-\alpha}}, & x > 0,\ \alpha > 0 \end{cases} && \text{i.e. the Fréchet distribution with } \xi > 0\ ,\\
\mho(x)_\alpha &= \begin{cases} e^{(-x)^{\alpha}}, & x \leq 0 \\ 1, & x > 0,\ \alpha > 0 \end{cases} && \text{i.e. the Weibull distribution with } \xi < 0.
\end{aligned}$$

When this is the case, F is believed to be in the maximum domain of attraction of H, denoted by $F_X \in D(H)$, and ξ is the shape parameter. These three classes of distributions can be grouped into a single parametric family called the generalized extreme value distributions (GEV), H_ξ (equation 3.4.2).

$$H_\xi(x) = \exp\left[-\left(1 + \xi\frac{x-\mu}{\sigma}\right)_+^{-\frac{1}{\xi}}\right] \tag{3.4.2}$$

such that $\left\{x|1 + \xi\left(\frac{x-\mu}{\sigma}\right) > 0\right\}$

where μ stands for the location parameter that represents the center of the distribution, ξ is the shape parameter or tail index that denotes the tail thickness and it suggests the distribution type, whereas the scale parameter, $\sigma > 0$, gives the size of the deviations about μ. The $+$ operator indicates the positive part of the expression in the parentheses.

This GEV distribution is obtained when the Fréchet distribution is set to $\xi = \alpha^{-1}$ for a positive constant α, the Weibull distribution is $\xi = -\alpha^{-1}$ and the Gumbel is

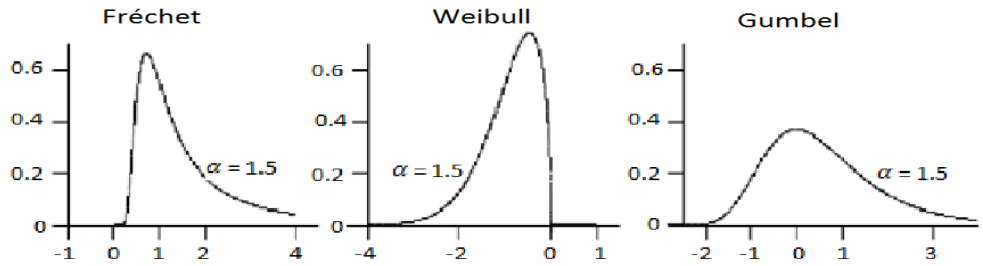

Figure 3.1: The densities for Fréchet, Weibull and Gumbel functions (Gilli and Kellezi, 2006).

interpreted as the limiting distribution at $\xi = 0$ (Gilli and Kellezi, 2006). As asserted by Fisher-Tippett-Gnedenko (1943), for normalized block maxima, the only possible limiting distribution is the generalized extreme value (GEV) distribution and is used in modelling block maxima $M_{n,1}, ..., M_{n,m}$ data.

The shapes of the probability density functions for the standard Fréchet, Weibull and Gumbel distributions are shown in Figure 3.1. It can be seen from the diagrams in Figure 3.1 that the Fréchet distribution will be suitable and adequate for heavy tailed distributions in risk management analysis because of its polynomial decaying (i.e. power-declining) tail. The Weibull distribution is a finite endpoint asymptotic distribution (i.e. distribution with an upper bound), and the Gumbel distribution with its exponentially decaying tails characterizes thin tailed distributions (Gilli and Kellezi, 2006).

As a shortcoming, the BMM can be deficient because it misses some of the needed high observations, and retains some lower ones (Coles, 2001). Based on this, researchers prefer the application of the threshold excess method since it uses all the required high observations, and that makes it more efficient.

3.5 Threshold excess models

The threshold excess model is actualized by selecting a sufficiently high threshold (u). Every point over such a threshold is considered an extreme loss set aside for risk modelling.

3.5.1 The limiting distribution of the threshold excess models: GPD

Balkema and de Haan (1974) and Pickands (1975) independently developed the threshold excess methodology. The theorem for the limiting distribution of the threshold-excess model where $y = x - u$, i.e. the excesses conditioned on $x > u$ is stated as:

Theorem 3.5.1. *If F is a distribution function of return series such that F belongs to the maximum domain of attraction of the generalized extreme value distribution, i.e. $F \in D(H_\xi)$, then as the threshold tends to the right endpoint, $u \to u_{end}$, the generalized Pareto distribution, $G_{\xi,\sigma}(y)$ becomes the only limiting non-degenerate distribution that approximates the distribution of return exceedances $F_u(y)$. That is, as $n \to \infty$*

$$F_u(y) \approx G_{\xi,\sigma}(y)$$

with

$$G_{\xi,\sigma}(y) = \begin{cases} 1 - \left(1 + \frac{\xi}{\sigma} y\right)^{-\frac{1}{\xi}}, & \text{for } \xi \neq 0; \\ 1 - e^{-\frac{y}{\sigma}}, & \text{for } \xi = 0. \end{cases} \tag{3.5.1}$$

where $G_{\xi,\sigma}$ is the GPD (generalized Pareto distribution) for $y \in [0, -\frac{\sigma}{\xi}]$ if $\xi < 0$ and $y \in [0, (x_F - u)]$ if $\xi \geq 0$.

The GPD can also be stated as a function of x as in equation (3.5.2) given that $x = u + y$

$$G_{\xi,\sigma}(x) = 1 - \left[1 + \xi \left(\frac{x - u}{\sigma}\right)\right]^{-\frac{1}{\xi}} \tag{3.5.2}$$

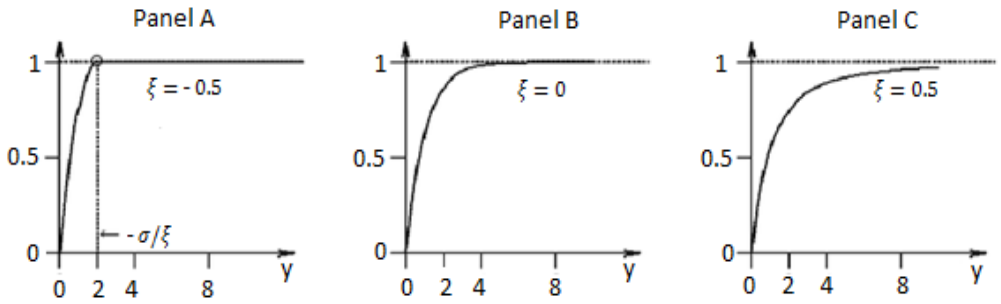

Figure 3.2: Shape of the GPD ($G_{\xi,\sigma}$) with $\sigma = 1$ (Gilli and Kellezi, 2006).

where ξ is the tail index or shape parameter or the extreme value index (EVI), and σ is the dispersion or scale parameter whose value indicates the spread of the tail of the distribution, whether light or heavy tailed. The tail index also characterizes the heaviness of the tail of the returns distribution with fat tailed or power-declining tail distribution in the case of $\xi > 0$, thin-tailed or exponentially declining tail distributions lie in the domain of $\xi = 0$, and finite or no tail distribution when $\xi < 0$ (Longin and Solnik, 2001). As illustrated in Figure 3.2, with scale parameter (σ) fixed at one, the shapes of the GPD as the shape parameter ξ takes negative values ($\xi < 0$), a zero value ($\xi = 0$), and positive values ($\xi > 0$) are displayed.

With positive shape parameter, i.e. when $\xi > 0$, Gilli and Kellezi (2006) report that an upper tail or bound cannot be fixed for financial losses, implying that only distributions with this shape parameter ($\xi > 0$) type as in panel C of Figure 3.2 can be suitable for financial returns modelling.

For a GPD with parameters ξ and σ acting as a suitable model for threshold exceedances u by a variable X, i.e. for $x > u$,

$$\Pr\{X > x | X > u\} = \left[1 + \xi\left(\frac{x-u}{\sigma_u}\right)\right]^{-\frac{1}{\xi}}. \tag{3.5.3}$$

It follows that

$$\Pr\{X > x\} = \eta_u \left[1 + \xi\left(\frac{x-u}{\sigma_u}\right)\right]^{-\frac{1}{\xi}}, \tag{3.5.4}$$

where η_u denotes the probability of an individual observation exceeding the threshold u, i.e. $\eta_u = \Pr\{X > u\}$ (see Coles, 2001).

3.6 Conditional multivariate extreme value modelling

For the multivariate analysis, the Heffernan and Tawn (2004) approach is used for the dependence modelling in this study. This method uses a conditional multivariate approach where the marginal variables are first fitted with the GPD (generalised Pareto distribution) models, after which the dependence structure is estimated. This dependence structure, like the GPD model for exceedance over a threshold, also conditions on a variable exceeding a given threshold. As an illustration on the selected four markets, given the threshold excess of the first market variable, the conditional multivariate approach is able to describe the remaining three markets' conditional distribution, with the use of a regression type model.

The multivariate modelling for this study is limited to the pairwise combination of variables, hence the bivariate-threshold-excess model is used for modelling the selected markets.

3.6.1 Bivariate-threshold-excess model

A class of approximations to the tail of the distribution function F for the threshold u exceedances by a variable X, on the condition that $X > u$ for large enough u, is given by

$$G(x) = 1 - \eta \left[1 + \xi \left(\frac{x-u}{\sigma_u}\right)\right]^{-\frac{1}{\xi}}, \quad x > u \tag{3.6.1}$$

where $\eta = \Pr(X > u)$, $\xi \neq 0$ and $\sigma > 0$ for a family defined on $\{(1+\xi(x-u)/\sigma) >$

0 and $x-u : x-u > 0\}$. Hence $F(x) \approx G(x)$ on $x > u$ for large enough threshold u with parameters η, ξ and σ (Coles, 2001).

Now for the bivariate type, the focus is to obtain a family that approximates a joint distribution $F(x, y)$ on $x > u_x$, $y > u_y$ regions for u_x and u_y that are large enough.

Theorem 3.6.1. *Assume $(x_1, y_1), ..., (x_n, y_n)$ are independent observations of a random vector (X, Y) with variables X, Y and joint distribution function F. For appropriate thresholds u_x and u_y, each margin of F has an approximation of the form in equation* (3.6.1)*, with parameter sets $(\eta_x, \xi_x, \sigma_x)$ and $(\eta_y, \xi_y, \sigma_y)$ respectively ·*

3.6.2 Marginal transformation

Prior to the dependence modelling under the bivariate-threshold-excess model, the regression type dependence model's structure first transforms the margins to the standardised Laplace margins as follow.

Theorem 3.6.2. *Suppose $\boldsymbol{X} = (X_1, ..., X_q)$ for a q-dimensional random variable whose marginal distributions are arbitrary. With $\hat{F}_j$ denoting an estimate of the jth marginal distribution function $(j = 1, ..., q)$, let G indicate the standardised marginal distribution's distribution function to be determined. Transfromation of the original vector variable $\boldsymbol{X}$ is done to obtain $\boldsymbol{Y} = (Y_1, ..., Y_q)$, where $\boldsymbol{Y}$ is a variable which has standardised marginal distributions with the use of the probability integral transform as follows:*

$$Y_j \;=\; (G^{-1}(\hat{F}_j(X_j))),\;\; j = 1, ..., q. \tag{3.6.2}$$

3.6.3 Regression model structure

Theorem 3.6.3. *Let Y_j, $j \in \{1, ..., q\}$ be the conditioning variable. Then $\boldsymbol{Y}_{-j}$ represents the remainder of the vector Y with the exclusion of the jth component. The*

approach used by Heffernan and Tawn conditions on $\boldsymbol{Y}_j$ *being over some high enough threshold* u*. This is used in modelling the dependence of the remaining* $\boldsymbol{Y}_{-j}$ *on the condition that the observed value of* $Y_j > u$*. The regression model's type for the structure of the conditional dependence depends on the exact choice of* G *in equation (3.6.2).*

3.6.4 Laplace margins

The Laplace distribution function is denoted by Laplace margins G, and $\mathbf{Y}$ are the marginally Laplace distributed. On the condition that the $\mathbf{Y}_j$ variable exceeds a high threshold u, the model of Heffernan and Tawn for the remaining variables $\mathbf{Y}_{-j}$ applies the type:

$$\mathbf{Y}_{-j} = \alpha_{|j}\mathbf{Y}_j + (Y_j)^{\beta_{|j}}\mathbf{R}_{|j} \tag{3.6.3}$$

where $\mathbf{R}_{|j}$ is a vector residual and $(q-1)$ dimensional parameter vectors $\alpha_{|j}$ and $\beta_{|j}$ satiate $(\alpha_{|j}, \beta_{|j}) \in [-1,1]^{q-1} \times (-\infty, 1)^{q-1}$. Here, $\alpha_{i|j}$, $\alpha_{|j}$ in association with $\mathbf{Y}_i, (i \in 1, ..., q, i \neq j)$, then $-1 \leq \alpha_{i|j} < 0$ and $0 < \alpha_{i|j} \leq 1$ respectively correspond to negative and positive association between $\mathbf{Y}_i$ and $\mathbf{Y}_j$'s large values (Southworth and Heffernan, 2016).

3.7 Threshold selection

Careful selection of a sufficiently high threshold must be done to avoid having a threshold that is too high, consequently resulting in a high variance estimator. Also it should not be too low to avoid bias where central observations are selected as extreme realizations. Hence, the choice of a reasonable threshold entails a balance of bias and variance. Amongst the various threshold selection methods defined in the

literature, this study will use two threshold selection approaches described in Sections (3.7.1) and (3.7.2) respectively.

3.7.1 Threshold stability plot

As a unique advantage, this threshold selection method provides the conditions for selecting the lowest threshold over which extremal index estimates are approximately constant (Heffernan and Southworth, 2013). Based on this advantage, it will be applied in this study as one of the methods for threshold selection.

The generalized Pareto distribution (GPD) shows threshold stability properties (Hu and Scarrott, 2013), where the shape (ξ) and scale (σ) parameter estimates are plotted against a number of threshold values u, to form a threshold stability plot. That is, the stability plot is aimed at looking for stability of parameter estimates by fitting the GPD at a range of thresholds (Coles, 2001).

The plot is underpinned with the concept that if the excesses (or exceedances) of a high threshold u_o follow a generalized Pareto distribution (GPD) with parameters ξ and σ_{u_o}, then the GPD is valid for exceedances over all thresholds $u > u_o$, with shape and scale parameters ξ_u and σ_u respectively. The two distributions in this situation have identical shape parameters ($\xi_u = \xi$) (Coles, 2001), but the scale parameter is defined as

$$\sigma_u = \sigma_{u_o} + \xi(u - u_o). \tag{3.7.1}$$

The difficulty with the scale parameter σ_u in equation 3.7.1 is that it changes with u except when $\xi = 0$, but this can be alleviated when the generalized Pareto scale parameter is reparametrized (Coles, 2001) as

$$\sigma^* = \sigma_u - \xi u \tag{3.7.2}$$

This new parametrized scale parameter is constant with regards to u by virtue of equation 3.7.1 (Coles, 2001; Bommier, 2014).

The locus of points is used to define the plot as

$$\{(u, \sigma^*);\ u < x_{\max}\} \text{ and } \{(u, \xi_u);\ u < x_{\max}\}, \tag{3.7.3}$$

where the maximum of the observations is denoted by $x_{\max}$ (Bommier, 2014).

Accordingly, estimates of both ξ and σ^* need to be constant above u_o, if u_o is a valid threshold for the GPD asymptotic approximation of the exceedances (Coles, 2001; Bommier, 2014). However, these estimates of the quantities (ξ and σ^*) will not be precisely constant due to sampling variability, but stability should be seen after their sampling errors are allowed (Coles, 2001; Northrop and Coleman, 2014). The argument behind the stability plot therefore recommends that both the ξ and σ^* should be plotted against u, together with each quantity's confidence interval, and the threshold u_o selected as the lowest value of u for which ξ and σ^* remain near-constant (Coles, 2001).

3.7.2 Extreme value mixture models

The mixture model is applied in this study because it is assumed that the traditional fixed threshold approach is subjective and does not account for the uncertainty involved in the choice of threshold. The mixture model approach is built to provide an objective estimate of an appropriate threshold with uncertainty quantification (Scarrott and MacDonald, 2012). The model combines a bulk model under the threshold with GPD above the threshold. The bulk distribution is the high density non-extreme part of the distribution and it contains low information about the tail of the distribution. It is believed that the information in the sample data is spread between this

bulk distribution and the tail observations which are low density with high information (Scarrott and MacDonald, 2012). Some of the distributions used for the bulk model are gamma, Weibull, and normal.

The non-parametric extremal mixture models of MacDonald et al. (2011) will be used to obtain a suitable threshold for this study. It is the Kernel GPD model with GPD for the tail model and the bulk model below the threshold is the standard kernel density. As an advantage over the usual parametric bulk approach, the non-parametric method is more robust to bulk model than the parametric technique (Yang, 2013). The standard Kernel GPD model is presented as

Bulk model based tail fraction approach:

$$F(x|X,\zeta,u,\sigma_u,\xi,\varphi_u) = \begin{cases} H(x|X,\zeta) & x \leq u, \\ (1-\varphi_u) + \varphi_u \times G(x|u,\sigma_u,\xi) & x > u. \end{cases} \tag{3.7.4}$$

where $\varphi_u = 1 - H(u|X,\zeta)$

Parameterised tail fraction approach:

$$F(x|X,\zeta,u,\sigma_u,\xi,\varphi_u) = \begin{cases} (1-\varphi_u)\frac{H(x|X,\zeta)}{H(u|X,\zeta)} & x \leq u, \\ (1-\varphi_u) + \varphi_u \times G(x|u,\sigma_u,\xi) & x > u. \end{cases} \tag{3.7.5}$$

where $H(x|X,\zeta)$ denotes kernel density estimator's distribution function. The GPD's distribution function is $G(x|u,\sigma_u,\xi)$.

To estimate the tail fraction, the bulk model-based tail fraction benefits from borrowing information from the generally ample bulk data. The main challenge with this bulk model however, is that it exposes estimation of the tail to the bulk model's misspecification (Hu and Scarrott, 2018). The parameterized tail fraction approach was introduced by MacDonald et al. (2011) with an extra parameter (φ_u) for the tail fraction, and it can reduce the effect of the misspecification of the bulk model on the tail estimates. Furthermore, the bulk model-based tail fraction is included in

the parameterized tail fraction approach as a special case, where $\varphi_u = 1 - H(u|X, \zeta)$, and it should be clear that either of the two specifications gives a proper density (Hu and Scarrott, 2018).

3.7.3 Estimation of parameters

Parameters of the GPD can be estimated with different statistical methods which include maximum likelihood estimation (MLE), the Bayesian method, and the method of moments. Evidence from the literature shows that MLE is widely used, and adaptable in application to most models and different data types, hence it will be applied in this study for the estimation of parameters. Further evidence indicates that MLE gives good estimates when $\xi > -1/2$, making the method more suitable for estimation of financial return series data with positive tail index of $\xi > 0$ (Bensalah, 2000). However, Smith (1985) stated, as indicated below, the likely difficulty that could be encountered with the use of the likelihood methods. This is connected with the regularity conditions expected for the validity of the asymptotic properties related to the maximum likelihood estimators.

- The maximum likelihood estimators are regular, i.e. they are shown as having the sense of the standard asymptotic properties when $\xi > -1/2$;
- the maximum likelihood estimators can be obtained generally, but they do not have the usual asymptotic properties when $-1 < \xi < -1/2$;
- it is unlikely that the maximum likelihood estimators will be attainable when $\xi < -1$.

When these conditions are violated, the Bayesian estimates are then preferred since they do not depend on these conditions.

For the GPD, let $L(\xi, \sigma|y)$ be the log-likelihood function of the parameter values given the sample $y = y_1, ..., y_n$ which can be expressed as the logarithm of the joint probability density of the n observations.

$$L(\xi, \sigma|y) = \begin{cases} -n\log\sigma - \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^{n} \log\left(1 + \frac{\xi}{\sigma} y_i\right) & \text{if } \xi \neq 0 \\ -n\log\sigma - \frac{1}{\sigma} \sum_{i=1}^{n} y_i & \text{if } \xi = 0. \end{cases} \tag{3.7.6}$$

With $\xi \neq 0$, the log-likelihood is conditional on $(1 + \sigma^{-1}\xi y_i) > 0$ for $i = 1, ..., n$; otherwise $L(\xi, \sigma) = -\infty$. Numerical techniques of the log-likelihood are required because analytical maximization is not possible. Furthermore, care must be taken to avoid numerical instabilities when $\xi \approx 0$ in equation (3.7.6) (where $\xi \neq 0$), and to also ensure that the algorithm does not fail as a result of evaluation outside of the acceptable parameter space (Coles, 2001).

MLE will be used to compute the estimates of the scale and shape parameters ($\hat{\sigma}$ and $\hat{\xi}$) that maximize the log-likelihood function of all the sampled observations in excess of the threshold u. Standard likelihood theory is used to obtain (approximate) standard errors and confidence intervals for the GPD (Coles, 2001). The value-outcome of these estimated parameters will give clues on the heaviness of the tails of the returns distribution.

3.8 Declustering

The conditions for the application of EVT is that the data should be independent and identically distributed (i.i.d). Short-term dependence due to clustering of high level exceedances can be removed from the residual sequence through declustering (Smith, 1989). The data is declustered based on the value of extremal index θ, which indicates the strength (level) of dependence between extremes of sequence that is stationary (Ferro, 2003). The range of extremal index for a process is $0 \leq \theta \leq 1$.

A sequence with an extremal index $\theta = 1$ is an indication of independence in the process. Otherwise, the sequence will need to be declustered at a chosen value of r (i.e. the interval length of clusters), which is the minimum gaps between the clusters (Coles, 2001).

3.9 Return level estimation

The next extreme event of an extreme movement can be measured by return level estimation. This is the level expected of an event to be surpassed on average once in the next "n" number of years (Coles, 2001). This is known as the q-observation return level (R_q) stated as

$$R_q = \begin{cases} \mu + \frac{\sigma}{\xi}\left[(q\eta_u)^{\xi} - 1\right], & \text{for } \xi \neq 0, \\ \mu + \sigma \log(q\eta_u), & \text{for } \xi = 0, \end{cases} \tag{3.9.1}$$

provided q is sufficiently large so that $R_q > u$ (Coles, 2001). If this level is directly given on an annual scale, the return level can be taken as the level expected to be surpassed (i.e. exceeded) once every "n" years. This is the same as q-observation return level if there are n_x observations per year, hence $q = n \times n_x$. The n-year return level is thus,

$$R_n = \begin{cases} \mu + \frac{\sigma}{\xi}\left[(nn_x\eta_u)^{\xi} - 1\right], & \text{for } \xi \neq 0, \\ \mu + \sigma \log(nn_x\eta_u), & \text{for } \xi = 0. \end{cases} \tag{3.9.2}$$

The value of R_q can be obtained by substituting the parameter estimates ($\hat{\sigma}$ and $\hat{\xi}$) from the maximum likelihood estimates into the equation. The sample proportion of observations, η_u, that exceed the threshold u is given as:

$$\hat{\eta}_u = \frac{k}{m} \tag{3.9.3}$$

where k is the number of exceedances, i.e. the number of observations exceeding u and m is the total set of observations.

The confidence interval (via standard error) for R_q can be obtained using profile likelihood intervals or from the delta method (Coles, 2001) given in equation (3.9.4):

$$\text{Var}(\hat{R}_q) \approx \nabla R_q^T V \nabla R_q, \tag{3.9.4}$$

where

$$\begin{aligned}\nabla R_q^T &= \left[\frac{\partial R_q}{\partial \eta_u}, \frac{\partial R_q}{\partial \sigma}, \frac{\partial R_q}{\partial \xi}\right] \\ &= [\sigma q^{\xi}\eta_u^{\xi-1}, \xi^{-1}\{(q\eta_u)^{\xi}-1\}, -\sigma\xi^{-2}\{(q\eta_u)^{\xi}-1\} + \sigma\xi^{-1}(m\eta_u)^{\xi}\log(m\eta_u)],\end{aligned} \tag{3.9.5}$$

estimated at $(\hat{\eta_u}, \hat{\sigma}, \hat{\xi})$. The complete variance-covariance matrix, V, for $(\hat{\eta_u}, \hat{\sigma}, \hat{\xi})$ is approximated as

$$V = \begin{bmatrix} \hat{\eta}_u(1-\hat{\eta}_u)/n & 0 & 0 \\ 0 & v_{1,1} & v_{1,2} \\ 0 & v_{2,1} & v_{2,2} \end{bmatrix} \tag{3.9.6}$$

where $\text{Var}(\hat{\eta}_u) \approx \hat{\eta}_u(1-\hat{\eta}_u)/n$ by the standard properties of binomial distribution, and $v_{i,j}$ represents the (i,j) variance-covariance matrix term of $\hat{\sigma}$ and $\hat{\xi}$ (Coles, 2001).

3.10 Point process

3.10.1 Overview

The risk inherent in the tails of return distribution can also be well characterized by using the notion of point process which provides a structure for extreme value limit results. This modelling approach operates by combining the functionalities of the Block Maxima Method (BMM) and the bivariate-threshold-excess model, where a

threshold will be selected and a region described such that points above the threshold denote the risk or extremal events to be modelled.

The limiting distribution of the point process: Poisson process (Univariate case)

The conventional theory of convergence is required in order to apply point process representation for extreme values modelling.

Theorem 3.10.1. *Let a series of iid (independent and identically distributed) random variables with a common distribution function F be denoted by $Y_1, Y_2, \ldots, Y_n$, where $M_n = max\{Y_1, \ldots, Y_n\}$. The distribution of the normalized maxima, with sequences of constants $\{d_n\}$ and $\{a_n > 0\}$ can be reasonably approximated by a generalized extreme value distribution (GEVD) as*

$$\Pr\{(M_n - b_n)/a_n \leq z\} \to G(z)$$

with

$$G(z) = \exp\left\{-\left[1 + \xi\left(\frac{z-u}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\}, \tag{3.10.1}$$

for z_+ and z_- denoted as the upper and lower endpoints of G respectively, and ξ, μ and $\sigma > 0$ are the parameters to be estimated.

Following this, a sequence of point processes is defined as

$$N_n = \{(i/(n+1), (Y_i - b_n)/a_n)\} \quad \text{for} \quad i = 1, \ldots, n. \tag{3.10.2}$$

The coordinates are chosen such that the first ordinate consistently maps to (0, 1) and the second ordinate is scaled to stabilize the behaviour pattern of extremes as $n \to \infty$ (Coles, 2001).

Now for a sufficiently large value of the threshold u, consider a region of the form $W = [0,1] \times [u, \infty]$ where the point processes converge to a non-homogeneous Poisson process N for any $u > z_-$, i.e. $N_n \xrightarrow{\text{d}} N$ as $n \to \infty$.

This limit exists since random variables Y_i are mutually independent and each of the points in N_n has p probability of falling in the W region, hence $N_n(W)$ follows a binomial distribution.

$$N_n(W) \sim \text{Bin}(n, p)$$

By the standard binomial approximation to a Poisson limit, as $n \to \infty$, the limiting distribution of $N_n(W)$ for any region of the type $W = [s_1, s_2] \times (u, \infty)$, where $[s_1, s_2] \subset [0,1]$ is the Poisson distribution with intensity measure $(\Omega(W))$ i.e. $\text{Poi}(\Omega(W))$ with

$$\Omega(W) = n_x(s_1 - s_2)\left[1 + \xi\left(\frac{u-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}} \tag{3.10.3}$$

where the intensity measure is the average number of events (points) occurrences in any given subset $W \subset \mathbb{W}$ and n_x is the number of years of observation. Hence, the Poisson process is a realistic approximation of the point processes for large but finite sample behaviour on the stated region where the threshold is large enough.

3.10.2 Estimation of parameters

Due to the adaptable application of the maximum likelihood estimation (MLE) to most models and different data types, it will also be applied for the parameter estimation of the point process. In reality, the likelihood of the estimates using the point process approach is synonymous to that of the threshold excess model (Coles, 2001). For both models, the likelihoods are based on all extreme data, i.e. data in excess of the selected threshold. Maximising the likelihood leads to the parameters (μ, σ, ξ)

estimation of the limiting intensity function of the point process, where the likelihood is given as

$$\begin{aligned} L_W(\mu,\sigma,\xi;y_1,...,y_n) &= \exp\{-\Omega(W)\}\prod_{i=1}^{N(W)}\lambda(s_i,y_i) \\ &\propto \exp\left\{-n_x\left[1+\xi\left(\frac{u-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} \\ &\propto \prod_{i=1}^{N(W)}\sigma^{-1}\left[1+\xi\left(\frac{y_i-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}-1} \end{aligned} \tag{3.10.4}$$

This likelihood function will be used to obtain the standard errors and the approximate confidence intervals of the parameters of the model as well as the maximum likelihood estimates. For non-homogeneous Poisson process models, maximization of the likelihood generally requires numerical techniques (Coles 2001).

3.11 Return level estimation

The point process model has both a stationary and a non-stationary version for return level estimation. While the former is easy to calculate, the latter takes the form (Coles, 2001)

$$1-\frac{1}{q}=\Pr\{\max(X_1,...,X_n)\leq R_q\}\approx\prod_{i=1}^{n}p_i \tag{3.11.1}$$

That is, if R_q denotes the q-year return level, and the number of observations in a year taken as n, then R_q satisfies equation (3.11.1), where

$$p_i = \begin{cases} 1-n^{-1}[\aleph]^{-1/\xi_i}, & \text{if } [\aleph]>0, \\ 1, & \text{otherwise,} \end{cases} \tag{3.11.2}$$

for $\aleph = 1 + \xi_i(R_q - \mu_i)/\sigma_i$ and the point process model's parameters for observation i are (μ_i, σ_i, ξ_i). With logarithms taken, this yields

$$\sum_{i=1}^{n} \log p_i = \log(1 - 1/q). \tag{3.11.3}$$

The return level estimate (R_q) is easily obtainable from this, while the standard error estimation is obtained via simulation since neither the profile likelihood nor the delta method can be practically used (Coles, 2001).

3.12 Transformation of the margins

Before modelling extremal dependence using the bivariate point process model, the margins need to be transformed to obtain unit Fréchet margins as

$$\tilde{X} = -\left(\log\left\{1 - \eta_x\left[1 + \frac{\xi_x(X - u_x)}{\sigma_x}\right]^{-\frac{1}{\xi_x}}\right\}\right)^{-1} \tag{3.12.1}$$

and

$$\tilde{Y} = -\left(\log\left\{1 - \eta_y\left[1 + \frac{\xi_y(Y - u_y)}{\sigma_y}\right]^{-\frac{1}{\xi_y}}\right\}\right)^{-1} \tag{3.12.2}$$

This transformation prompts a bivariate variable $(\tilde{X}, \tilde{Y})$ with distribution function $\tilde{F}$ and the margins are approximately standard Fréchet for $X > u_x$ and $Y > u_y$ (Coles, 2001).

With the transformation applied on both margins, the following are obtained

$$\tilde{F}(\tilde{x}, \tilde{y}) = \exp\{-V(\tilde{x}, \tilde{y})\} \quad \text{for } x > u_x \text{ and } y > u_y, \tag{3.12.3}$$

where

$$V(x, y) = 2\int_0^1 \max\left(\frac{\jmath}{x}, \frac{1 - \jmath}{y}\right) dH(\jmath). \tag{3.12.4}$$

H denotes a distribution function on $[0, 1]$ and it satisfies the mean constraint

$$\int_0^1 \jmath \, dH(\jmath) = 1/2. \tag{3.12.5}$$

3.13 Bivariate point process model

Theorem 3.13.1. *Let* $(x_1, y_1), (x_2, y_2)$ ... *be a sequence of observations that are independent bivariate from a distribution with standard Fréchet margins, satisfying the convergence for componentwise maxima* ·

$$\Pr\{M^*_{x,n} \leq x, M^*_{y,n} \leq y\} \xrightarrow{\mathrm{d}} G(x, y) \tag{3.13.1}$$

where G is a non-degenerate distribution function, and it is of the form

$$G(x, y) = \exp\{-V(x, y)\},\ x > 0,\ y > 0 \tag{3.13.2}$$

with $V(x, y)$ as stated in 3.12.4 and 3.12.5 (Coles, 2001).

Let a sequence of point processes denoted by N_n be defined by

$$N_n = \left\{ \left(\frac{x_1}{n}, \frac{y_1}{n} \right), ..., \left(\frac{x_n}{n}, \frac{y_n}{n} \right) \right\} \tag{3.13.3}$$

Then N_n can be reasonably approximated by a non-homogeneous Poisson process (N) on $(0, \infty) \times (0, \infty)$, such that

$$N_n \xrightarrow{\mathrm{d}} N,\ as\ n \to \infty \tag{3.13.4}$$

on region of the form W in equation (3.13.5), bounded from the origin (0, 0).

$$W = \{(0, \infty) \times (0, \infty)\}\ \{(0, x) \times (0, y)\} \tag{3.13.5}$$

The intensity function of the Poisson process N is stated in equation (3.13.6), where H determines the angular spread of points in the limit Poisson process.

$$\urcorner(\ell, \jmath) = 2\frac{dH(\jmath)}{\ell^2} \tag{3.13.6}$$

for

$$\ell = x + y \text{ and } \jmath = \frac{x}{x+y} \tag{3.13.7}$$

The choice of a threshold using the bivariate point process involves a similar consideration as that of the threshold excess model. The values of the selected threshold for the two models intersects the Cartesian x and y axes at the same points and this can be relevant for comparing the two models. However, one salient advantage of using the point process over the threshold excess model is due to an easy transformation to pseudo-polar coordinates from Cartesian, i.e. $(x, y) \to (\ell, \jmath)$. Here, the distance from the origin is measured in ℓ units, while $\jmath$ measures the angle on a scale of [0, 1], where $\jmath = 0$ corresponds to $x = 0$ axis and $\jmath = 1$ corresponds $y = 0$ axis.

Given the region defined in equation (3.13.5), for the Poisson limit,

$$\Pr\{M^*_{x,n} \leq x, M^*_{y,n} \leq y\} \to \Pr\{N(W) = 0\} = \exp\{-\Omega(W)\}, \tag{3.13.8}$$

for

$$\begin{aligned}
\Omega(W) &= \int_W 2\frac{d\ell}{\ell^2} dH(\jmath) \\
&= \int_{\jmath=0}^{1} \int_{\ell=\min\{x/\jmath, y/(1-\jmath)\}}^{\infty} 2\frac{d\ell}{\ell^2} dH(\jmath) \\
&= 2\int_{\jmath=0}^{1} \max\left(\frac{\jmath}{x}, \frac{1-\jmath}{y}\right) dH(\jmath)
\end{aligned} \tag{3.13.9}$$

It is simple for the Poisson limit to reasonably approximate the point process, more especially if a region of the type $W = \{(x, y) : x(n^{-1}) + y(n^{-1}) > \ell_\circ\}$ is chosen, for large enough $\ell_\circ$ (Coles, 2001), by which

$$\Omega(W) = 2\int_W 2\frac{d\ell}{\ell^2} dH(\jmath) = 2\int_{\ell=\ell_\circ}^{\infty} \frac{d\ell}{\ell^2} \int_{\jmath=0}^{1} dH(\jmath) = \frac{2}{\ell_\circ}. \tag{3.13.10}$$

3.13.1 Estimation of parameters

The appropriate parameters estimation using maximum likelihood is stated as

$$\begin{aligned} L(\theta;(x_1,y_1),...,(x_n,y_n)) &= \exp\{-\Omega(W)\}\prod_{i=1}^{N_W}\lambda(x_{(i)}n^{-1},y_{(i)}n^{-1}) \\ &\propto \prod_{i=1}^{N_W} h(\jmath_i) \end{aligned} \tag{3.13.11}$$

where h is the density assumed on H, $\jmath_i = x_{(i)}(x_{(i)}+y_{(i)})^{-1}$, for the N_W i.e. the events or points $(x_{(i)}+y_{(i)})$ falling in region W with the assumption that the data $(x_1,y_1),...,\{(x_n,y_n)\}$ have marginal standard Fréchet distribution (Coles, 2001).

3.14 Diagnostics: Model checking

Quantile plots, probability plots, density plots and return level plots are all suitable for evaluating the quality of a generalized Pareto model fit. These plots are used in this study to check or assess the quality of the various GPD model fits. If the GPD fit is reasonable for modelling exceedances above a threshold u, then both the quantile and probability plots should have points that are approximately linear (Coles, 2010). Furthermore, both the Anderson-Darling and Cramer-von Mises goodness-of-fit tests are used to assess the validity of the GPD parameter estimates.

Anderson-Darling and Cramér-von Mises statistics are in the class of quadratic Empirical Distribution Function (EDF) statistics (Stephens, 1986) that is defined as

$$m\int_{-\infty}^{\infty}(F_m(x)-F(x))^2 v(x)\,dF(x), \tag{3.14.1}$$

where $F(x)$ is the theoretical cdf (cumulative distribution function), with a random sample of size m and $v(x)$ denotes a weighting function.

3.14.1 Cramér-von Mises (CVM) test:

Equation (3.14.1) indicates m times the CVM (Cramér-von Mises Test) statistic for $v(x) = 1$. Computation of the statistic can be done using the sum of squared differences between the EDF and the theoretical CDF defined in equation (3.14.2) (Anderson and Darling, 1954)

$$\text{CVM} = \frac{1}{12m} + \sum_{j=1}^{m} \left(F(x_j, \theta) - \frac{2j-1}{2m} \right)^2 \tag{3.14.2}$$

where $F(x_j, \theta)$ is the theoretical distribution function (Singla et al., 2016).

3.14.2 Anderson-Darling (AD) test:

The Anderson-Darling (AD) test is a modified version of the CVM test, and it gives more weight to the distribution's tails (Singla et al., 2016). When taking $v(x) = [F(x)(1-F(x))]^{-1}$ in equation (3.14.1), the Anderson-Darling test statistic is attained as (Anderson and Darling, 1954)

$$AD = n \int_{-\infty}^{\infty} \frac{(F_m(x) - F(x))^2}{F(x)(1 - F(x))} dF(x) \tag{3.14.3}$$

The formula is alternatively written as

$$AD = -m - 2 \sum_{j=1}^{m} \left\{ \frac{2j-1}{2m} \ln F(x_j, \theta) + \left(1 - \frac{2_j - 1}{2m} \right) \ln(1 - F(x_j, \theta)) \right\} \tag{3.14.4}$$

Here, the null hypothesis is rejected if the AD test statistic value is greater than the p-value. For a completely identified distribution F, the distribution theory of the EDF statistics is well developed and in addition, tables which give the levels of significance are available (Bere, 2016 and Singla et al., 2016). However, for a situation where the location or scale parameter is not specified by the null hypothesis, the exact distributions of EDF statistics are hard to find (Bere, 2016). The Anderson-Darling

statistic belongs to a class of quadratic statistics in which the asymptotic distributions are known.

A modified version of the Anderson-Darling statistic (D'Agostino and Stephens, 1986) where values of the statistic for finite asymptotic significance points' sample size can be compared is given as

$$AD^* = AD\left(1.0 + \frac{0.75}{m} + \frac{2.25}{m^2}\right) \tag{3.14.5}$$

The formulae for approximating the p-values that correspond to this modified version are given in Table 4.9 (page 127) of D'Agostino and Stephens (1986) (Bere, 2016).

3.15 The GARCH model

Before fitting the above-mentioned EVT models to the residuals, volatility will be removed from the returns by fitting GARCH models (Smith, 2003). Engle (1982) proposed the ARCH model but it was later generalized by Bollerslev (1986) as Generalized ARCH (GARCH). These models can be used to ascertain the source and magnitude of volatility. In the GARCH model, the conditional variance is expressed as a linear function of its own lags. The model requires conditional mean and conditional variance equations to be specified. GARCH (1, 1) model is the simplest model specification with the mean equation defined as:

$$\text{Mean equation}: \quad r_t = \mu + \varepsilon_t \tag{3.15.1}$$

where μ is the constant conditional mean, and ε_t denotes the unpredictable part of the time series known as the innovations or residuals. The focus of this study is on this unpredictable part of stock returns in the conditional mean equation, and it will thus be modelled under the assumptions of three distinct error distributions of a normal, student's t and generalized error distribution (GED) and their skewed versions.

The GARCH (P, Q) model's variance equation can be expressed as

$$\varepsilon_t = z_t \sigma_t$$
$$z_t \sim N(0, 1)$$
$$\sigma_t^2 = \omega + \sum_{j=1}^{Q} \alpha_j \varepsilon_{t-j}^2 + \sum_{i=1}^{P} \beta_i \sigma_{t-i}^2 \tag{3.15.2}$$

where z_t is the standardized residual returns and ω is the intercept. The residuals are i.i.d, i.e. independent and identically distributed (McNeil and Frey, 2000) random variables with mean zero and variance one (Smith, 2003).

3.15.1 The standard GARCH model (sGARCH)

The standard GARCH model developed by Bollerslev (1986) may be written as:

$$\sigma_t^2 = \omega + \sum_{k=1}^{M} \zeta_i \upsilon_{it} + \sum_{j=1}^{Q} \alpha_j \varepsilon_{t-j}^2 + \sum_{i=1}^{P} \beta_i \sigma_{t-i}^2 \tag{3.15.3}$$

where ω is the constant term (intercept) with σ_t^2 indicating the conditional variance, ε_t^2 is the squared residuals from the mean (equation) filtration process and υ_{it} denotes the exogenous variables. Also, ζ_i, α_j and β_i are parameters with $k = 1$, ..., M, $j = 1$, ..., Q and $i = 1$, ..., P respectively.

3.16 Test for stationarity

Before choosing time series models in the analysis, tests for unit root in the logarithm prices will be carried out. A unit root existence involves processes which develops through time and can lead to problems in statistical inference that involves models of time series. A process is non-stationary when the linear stochastic process has a unit root i.e. if 1 is the characteristic equation of the process. An integrated order

one non-stationary series can be written as $I(1)$ while a stationary one is $I(0)$. If the data series is not stationary, we difference the logged data to obtain stationarity.

To test for this, the Augmented Dickey Fuller (ADF) test (Dickey and Fuller, 1979) will be performed to know whether the series is stationary or not by testing the null hypothesis that the price series r_t is non-stationary, i.e.

$$H_\circ : I(1), \text{ series is non-stationary}$$
$$H_1 : I(0), \text{ series is stationary}$$

This is executed with the assumption that the data dynamics take an ARMA structure, and the test is centered on estimating the test regression:

$$r_t = \beta' \mathrm{D}_t + \phi r_{t-1} + \sum_{j=1}^{p} \psi_j \Delta r_{t-j} + \varepsilon_t \tag{3.16.1}$$

where the deterministic terms' vector (i.e. constant, trend and so on) is denoted by D_t. The ARMA structure of the errors are approximated by the p lagged difference terms, Δr_{t-j}, with the value of p set to make the error (which assumes homoskedasticity) serially uncorrelated. The ADF t-statistic as presented in equation (3.16.2) is based on the estimates of least squares

$$ADF_t \;=\; t_{\phi=1} = \frac{\hat{\phi} - 1}{SE(\phi)} \tag{3.16.2}$$

where $\hat{\phi}$ and $SE(\phi)$ refer to the least squares and standard error estimates respectively. The null hypothesis is rejected if the test statistic value is less than the critical value, we then conclude that no unit-root is present.

Generally, the asymptotic distributions of the unit root tests are non-normal and non-standard. Also, the critical values are calculated with the use of the methods of simulation and do not have suitable closed form expressions because the distributions are functions of standard Brownian motions (Rossi, 2014).

3.17 Test for serial correlation

After ensuring that the data is stationary, the next step is to fit an ARMA (p, q) model to the stationary data in order to test for serial correlation in the series. This test is aided by plots of the series autocorrelation function (ACF) and partial autocorrelation function (PACF) for visual illustration of the presence of autocorrelation in the residuals.

3.18 Test for ARCH effects

The next step after testing for autocorrelation is to examine the data series for ARCH effect. In other words, the residuals obtained from the conditional mean equation shall be tested for heteroscedasticity before applying the GARCH process on the return series (if ARCH effect is significant). In this research, testing for ARCH effects will be carried out using both the ARCH LM Test and the McLeod-Li test for ARCH effects.

3.18.1 ARCH LM test:

The presence of heteroscedasticity in the residuals of the four markets' return series will be tested for with application of the Lagrange Multiplier (LM) test for ARCH effects introduced by Engle (1982).

To test for ARCH effects in the returns' conditional variance, the test procedure is in two phases: First, we fit an ARMA (p, q) model on the returns series of the indices:

$$r_t = \phi_o + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{j=1}^{q} \varepsilon_{t-j} + \varepsilon_t \tag{3.18.1}$$

The linear regression is run on this model to obtain the residuals $\hat{\varepsilon}_t$. After this, we run a regression of squared ordinary least squares (OLS) residuals ($\hat{\varepsilon}_t^2$) obtained from equation (3.18.1) on k lags of squared residuals to test for ARCH effects of order k as

$$\hat{\varepsilon}_t^2 = \alpha_0 + \left(\sum_{j=1}^{k} \alpha_j \hat{\varepsilon}_{t-j}^2 \right) + \upsilon_t \tag{3.18.2}$$

The Engle's LM test statistic (NR^2) is calculated as the product of the number of observations (N) and the coefficient of correlation (R^2) value from the test regression. In the regression, the number of estimated parameters is represented by k. The LM statistic has an asymptotic χ^2 distribution with k degrees of freedom. The null hypothesis of "no ARCH effect" up to order k in the residual is tested as follows:

$H_o: \ \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$
against the alternative hypothesis
$H_1: \ \alpha_1 \neq 0, \ \alpha_2 \neq 0, \ \cdots, \ \alpha_k \neq 0$
i.e. $\alpha_j \neq 0$ for some j.

The null hypothesis is rejected if p-value < 0.05. The presence of ARCH effect simply means volatility is time varying and hence makes the return series open to GARCH modelling.

3.18.2 McLeod-Li test:

McLeod-Li test can also be used as an alternative to detect the presence of ARCH effects by plotting the p-values of the test using a number of lags. McLeod and Li (1983) proposed a formal test for ARCH effect based on the Ljung-Box test. They considered the autocorrelation functions of the squares of the time series and tested whether the first k squared residual autocorrelations from an ARMA model can be

used to test ARCH effects. The Ljung-Box Q-Statistics of McLeod-Li test is given by

$$Q = S(S+2)\sum_{t=1}^{k}\frac{\hat{\tau}_t^2(\varepsilon^2)}{S-t} \tag{3.18.3}$$

where the sample size is S and $\hat{\tau}_t^2$ is the squared sample autocorrelation of squared residual series at lag t. The test statistics is asymptotically χ^2 distributed with k degrees of freedom, under the null hypothesis of "no ARCH effect in the data".

Chapter 4

Data analysis and discussions

4.1 Introduction

This chapter explains the basic descriptive statistics tests carried out on the return and daily price series of the respective markets with emphasis on measures of locations and dispersions. After this, ARMA-GARCH models are fitted to the returns to filter out linear dependence and heteroscedasticity. Next, the empirical results of the fitted ARMA-GARCH models under three error distributions and their skew-versions are tabulated. The error distribution that best characterize each of the markets is selected using two information criteria: Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).

Modelling of risk in the markets begins with extraction of independent and identically distributed residuals from the markets' volatility-filtered returns. The risk in each market is modelled using the univariate models. Diagnostic checks of applied models are carried out to ascertain model validity under the univariate data analysis. After this, extremal dependence in the markets is modelled using the multivariate version of the underlying models. Attention is restricted to the bivariate, or the two-dimensional case for this study.

4.2 Data description

The raw data used for this project is from Datastream and it includes the daily closing equity indices of the South African, Nigerian, Kenyan and Egyptian stock markets. The periods covered uniformly in the four selected markets is from September 4th, 2000 to August 28th, 2015 with 3910 observations. The markets' indices are JSE-ALSI (Johannesburg Stock Exchange All-Share Index) for South Africa, NIGALSH (Nigeria All-Share Index) for Nigeria, NSE 20 (Nairobi Stock Exchange) for Kenya, and EGX 30 (Egypt Stock Exchange) for Egypt. All indices are in U.S. dollars.

Equity market's price series usually exhibit non-stationarity. Transformation to reasonable stationarity can be achieved by taking the logarithm of the price data as in equation (4.2.1) (Coles, 2001). The return series so generated is then re-scaled by multiplying by 100 for convenience of presentation.

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \times 100 \tag{4.2.1}$$

where P_t denotes the closing stock price index at time t, and P_{t-1} is the previous day's closing market index, t, r_t are the time period in days and the current returns respectively and ln denotes the natural logarithm.

4.3 Exploratory data analysis

As a required first step in the analysis of any typical dataset, exploratory data analysis (EDA) is carried out to get an insight into the dataset and verify whether the data satisfy the underlying distributional assumptions. It may also detect outliers and anomalies in the data, and suggest plausible normalizing transformation. From the visual inspection of the price and return series of the markets in Figure 4.1, it is observed that the plots of the daily equity prices are not stationary while the returns

series show volatility occurrence in bursts. This is volatility clustering with the return series varying around the constant mean while the variance changes with time. The result from the Augmented-Dickey Fuller (ADF) test of unit root in Section (4.4.1) indicates that the returns of the time series for the whole study period is stationary in the four market indices.

The EDA of the residuals as displayed in Figures 4.2, 4.3, 4.4 and 4.5 for the markets indicates that the i.i.d residuals are stationary. The Q-Q plots of the two market indices suggest departure from normality, i.e. the return distributions have tails fatter than that of a normal distribution.

Table 4.1 presents preliminary descriptive statistics for the daily equity prices and returns in each market for the sampled period. Sample means, standard deviations, medians, minimums, maximums, skewness, kurtosis, Jarque-Bera (JB) test and the p-values are obtained for the daily returns series. To start with, the table shows that the Egyptian market index (EGX 30) exhibits the highest mean returns of 0.06, followed by that of South African market (JSE-ALSI) of 0.05, then the Nigerian (NIGALSH) market of 0.03 and lastly the Kenyan (NSE 20) market's average index returns of 0.02. The table also shows that the four markets return series exhibit excess kurtosis (highly leptokurtic since the kurtosis are greater than three) and negative skewness (except for NSE 20 of Kenya). The negative values of the skewness imply that the market indices have long left tails and these further indicate higher probability of large decreases in equity returns for the period sampled in the three markets, except for Kenyan market. On the other hand, daily prices of the JSE-ALSI and NIGALSH indices show positive skewness that indicates long right tail distributions, while those of EGX 30 and NSE 20 reflect negative skewness for long left tail distributions (Table 4.1). Furthermore, the high kurtosis values for the returns indicate that extreme price changes take place more regularly during the sampled period (Sigauke et al., 2010).

Lastly, the Jarque-Bera (JB) test rejects the assumption of normality for prices

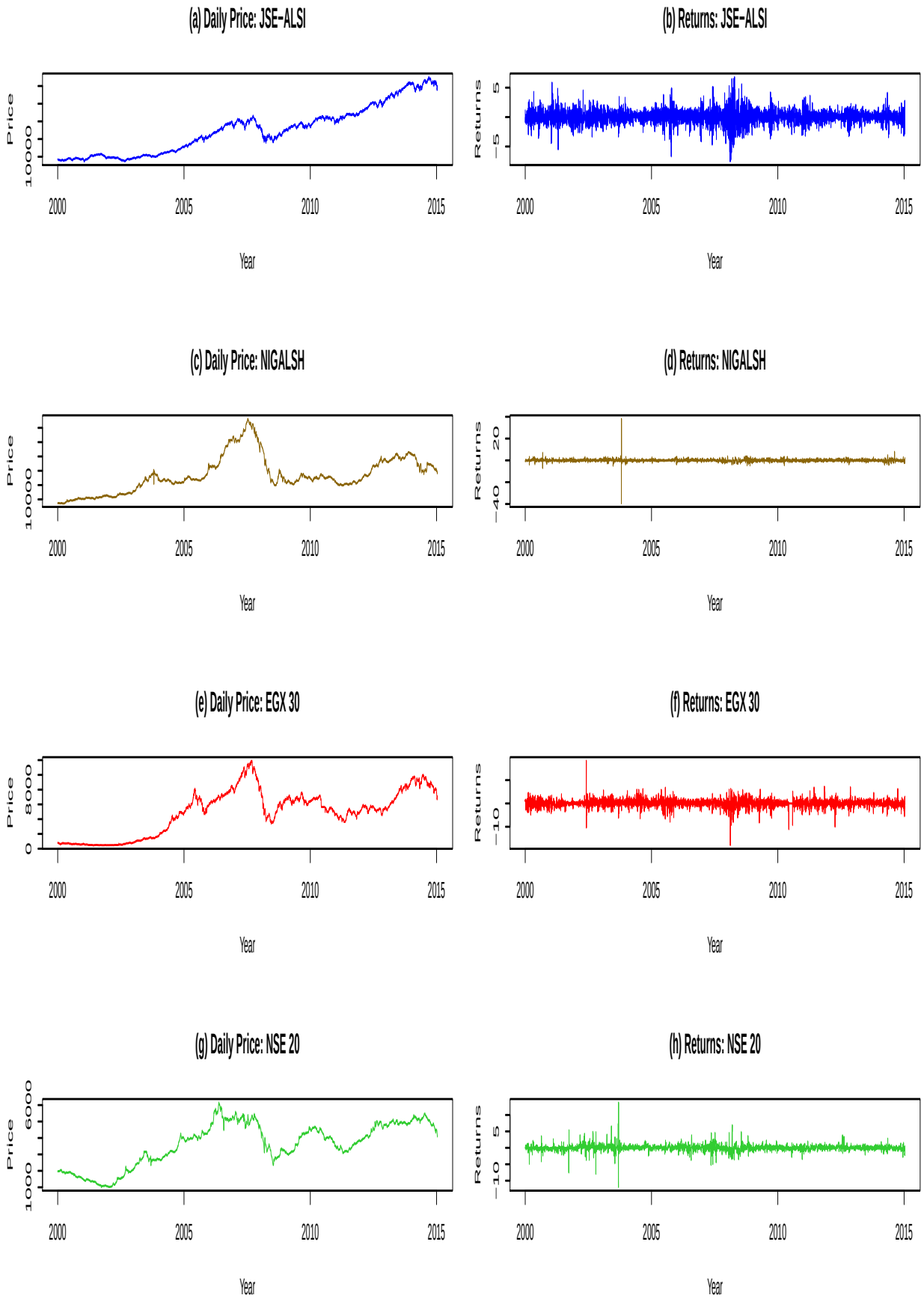

Figure 4.1: Daily price and return for JSE-ALSI, NIGALSH, EGX 30 and NSE 20.

Table 4.1: Descriptive Statistics.

	South Africa (JSE-ALSI)	Nigeria (NIGALSH)	Egypt (EGX 30)	Kenya (KSE 20)
INDEX				
Mean	25118.64	27080.73	4933.66	3611.17
Median	25468.68	24614.97	5367.80	3880.30
Maximum	55188.34	66371.19	11935.67	6161.46
Minimum	7189.99	7093.00	445.53	1001.94
Std. Dev.	13467.44	12614.21	3038.15	1310.94
Skewness	0.43	0.80	-0.06	-0.40
Kurtosis	2.17	3.41	2.03	2.03
Jarque-Bera	233.03	446.99	157.48	258.70
Probability	0.00	0.00	0.00	0.00
Observations	3910	3910	3910	3910
RETURNS				
Mean	0.05	0.03	0.06	0.02
Median	0.02	0.00	0.00	0.00
Maximum	6.83	38.18	18.37	13.81
Minimum	-7.58	-39.57	-17.99	-11.99
Std. Dev.	1.21	1.34	1.69	0.93
Skewness	-0.12	-0.69	-0.43	0.30
Kurtosis	6.65	371.26	13.21	34.66
Jarque-Bera	2181.82	22094541.00	17105.65	163404.00
Probability	0.00	0.00	0.00	0.00
Observations	3910	3910	3910	3910

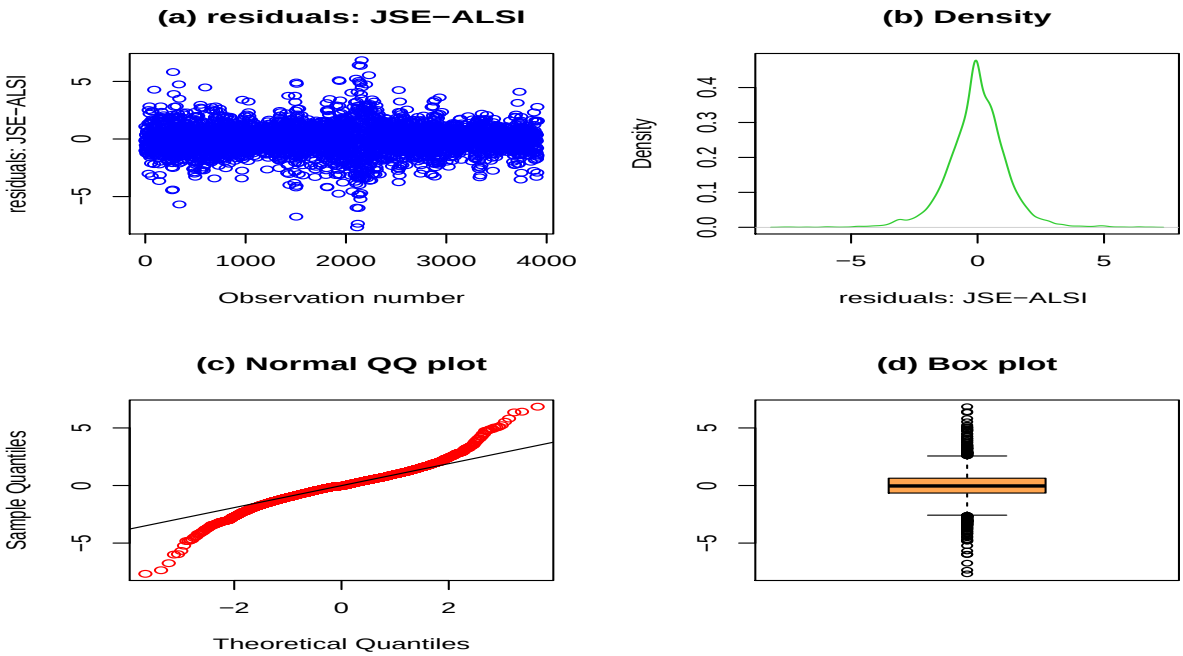

Figure 4.2: JSE-ALSI returns.

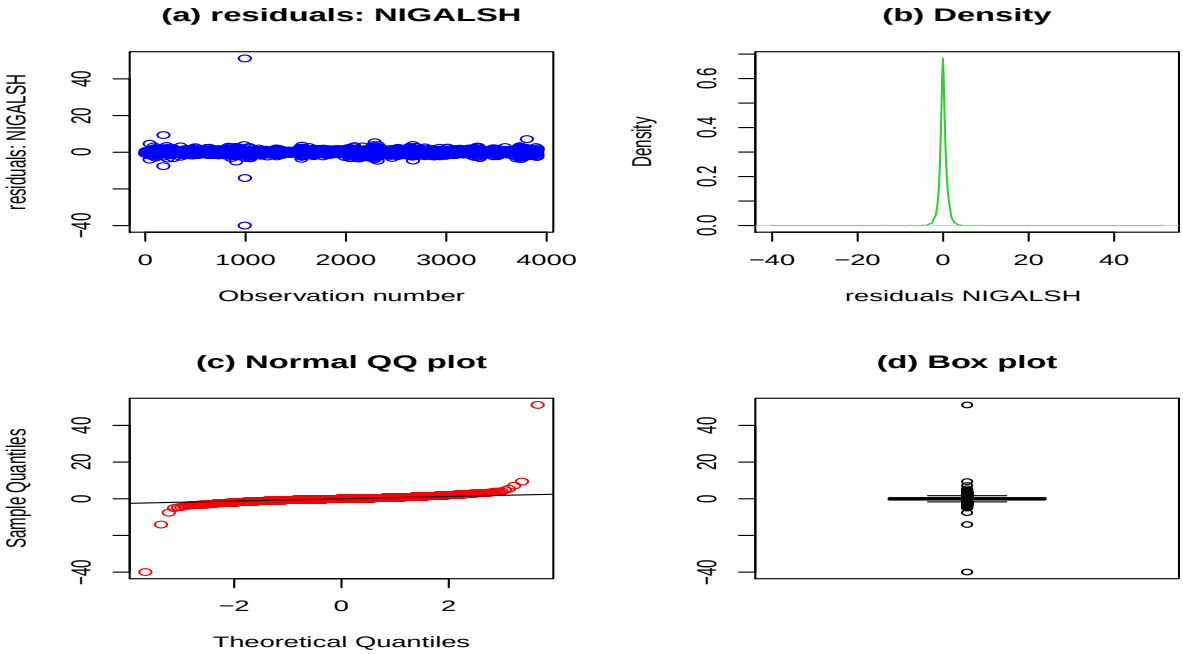

Figure 4.3: NIGALSH returns.

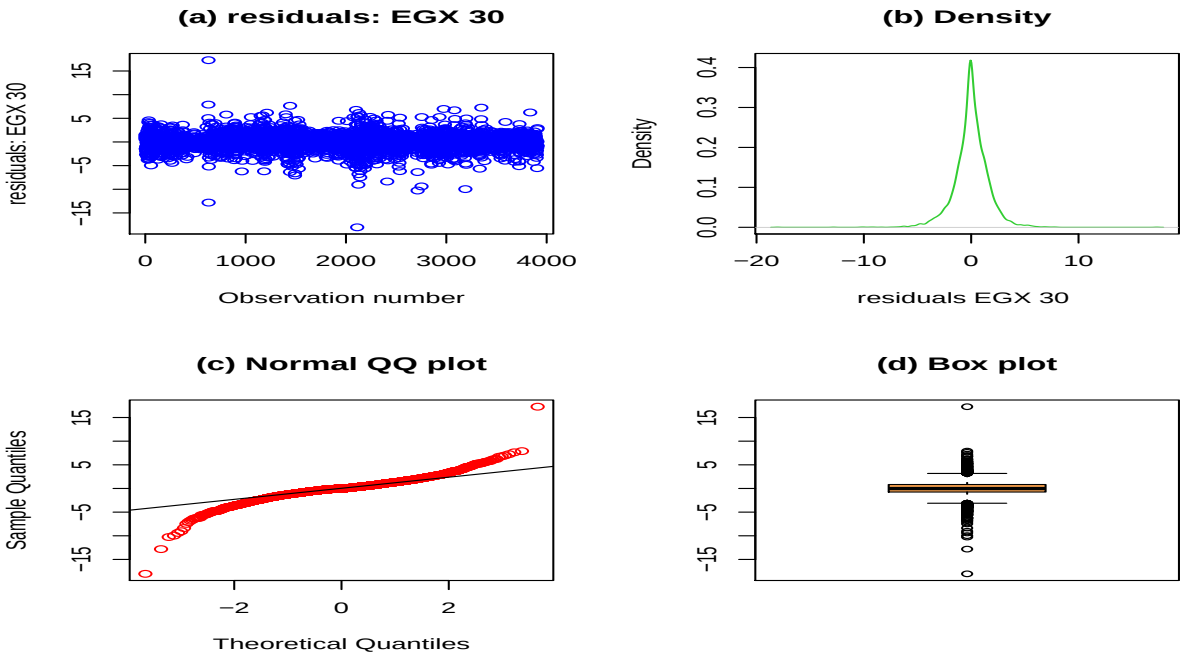

Figure 4.4: EGX 30 returns.

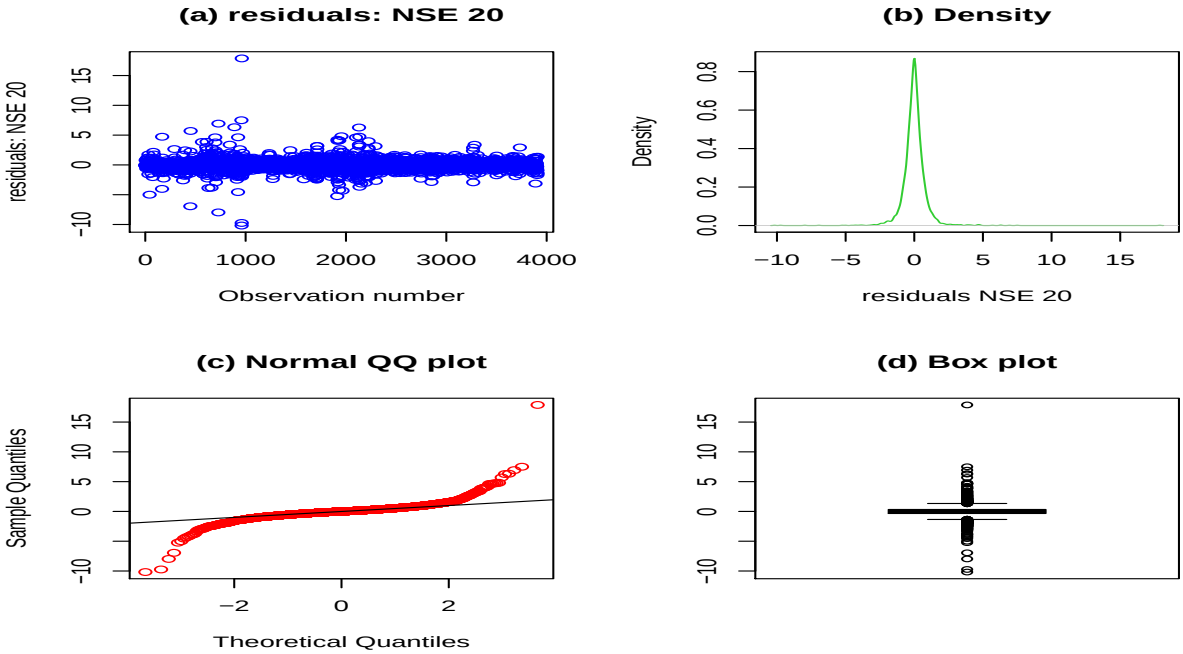

Figure 4.5: NSE 20 returns.

and returns in the four market indices due to the significance of the statistics at 1% level.

4.4 Testing for stationarity, serial correlation and ARCH effects

4.4.1 Output on test for stationarity:

Analysis of the raw price data indicated that the series are non-stationary. The data were then transformed to the return series by taking the difference and the Augmented-Dickey Fuller (ADF) was used to further test for stationarity. The ADF test results in Table 4.2 show that the null hypothesis of the presence of unit root in the return series is rejected since the values of ADF test statistic is less than the critical values at 1%, 5% and 10% significance levels for all the markets' returns. This leads to the conclusion that the returns of the time series for the whole study period is stationary. As a result of this, the time series models can be used to determine the dynamic volatility behaviour of the returns over time.

4.4.2 Output on test for serial correlation:

Serial correlation was observed in the four markets' data and returns series as indicated in the ACF and PACF plots, but adjustment was made to correct it in the four market indices (see Figures 5.1, 5.2, 5.3 and 5.4 in the Appendix). From the Weighted Ljung-Box test output, linear dependency (autocorrelation) was removed after adjustment procedure using ARMA(1, 1) model in equation.

$$r_t = \phi_o + \phi_1 r_{t-1} + \alpha_1 \varepsilon_{t-1} + \varepsilon_t \tag{4.4.1}$$

Since these statistics are no more significant, we conclude that the variance equation is specified correctly.

Table 4.2: ADF unit root test.

	South Africa	
ADF test statistic: -37.7572	1% Critical Value*: -3.4318 5% Critical Value: -2.8621 10% Critical Value: -2.5671	Lag level: 3
	Nigeria	
ADF test statistic: -28.8950	1% Critical Value*: -3.4318 5% Critical Value: -2.8621 10% Critical Value: -2.5671	Lag level: 3
	Egypt	
ADF test statistic: -32.4878	1% Critical Value*: -3.4318 5% Critical Value: -2.8621 10% Critical Value: -2.5671	Lag level: 3
	Kenya	
ADF test statistic: -26.2978	1% Critical Value*: -3.4318 5% Critical Value: -2.8621 10% Critical Value: -2.5671	Lag level: 3

***MacKinnon critical values for rejection of a unit root hypothesis.**

Table 4.3: ARCH-LM Test.

	South Africa	Nigeria	Egypt	Kenya
ARCH-LM test statistic $[NR^2]$	136.1234	979.5028	330.9094	461.2276
Prob. - $\chi^2(1)$	{0.0000}*	{0.0000}*	{0.0000}*	{0.0000}*
ARCH-LM test statistic $[NR^2]$	342.2704	1291.786	343.8845	569.3497
Prob. - $\chi^2(2)$	{0.0000}*	{0.0000}*	{0.0000}*	{0.0000}*
ARCH-LM test statistic $[NR^2]$	456.4641	1445.437	346.3582	581.8817
Prob. - $\chi^2(3)$	{0.0000}*	{0.0000}*	{0.0000}*	{0.0000}*

*** significant at 1% level. Figures in (.) are the lag lengths and in {.} are the p-values. NR^2 defines the LM test statistics where N is the number of observations and R is the sample multiple correlation coefficient.**

4.4.3 Output on ARCH-LM test:

It can be seen from Table 4.3 that the ARCH-LM test statistic is highly significant across-the-board, thus, the test results strongly reject the null hypothesis of "no ARCH effect" in the residuals of the returns series. This result confirms the presence of ARCH effects in the residuals of the four market indices, and it implies that the variances of the return series is time varying with the series exhibiting volatility clustering.

4.4.4 Output on McLeod-Li test:

Another test for ARCH effects was carried out using the McLeod-Li Test. From the plotted p-values, the result shows that the tests are all significant at the 5% level when the number of lags of the squared residuals autocorrelations ranges from 1 to 35 (see Figures 5.12 & 5.13 in the Appendix). This is shown with all the 35 circles enclosed within the broken/dotted lines. The result indicates strong evidence of conditional heteroscedasticity (ARCH effects) displayed in the four markets' return

series. Therefore, GARCH models can be used to capture the heteroscedasticity effects in the series.

Following the outputs of the basic statistics, i.e. since stationarity using ADF test, volatility clustering, and heteroscedasticity (ARCH effects) using ARCH-LM and McLeod-Li tests are confirmed in the return series of the four indices, volatility will need to be filtered (using GARCH model) prior to modelling risk and extremal dependence (Smith, 2003). The volatility of the return series will be modelled using the selected GARCH model under each of the three error distributions and their skewed versions.

4.5 Empirical results of the ARMA-GARCH models

This section describes the volatility dynamics of the four indices to ascertain the market that displays the highest volatility. For the market indices, after running series of candidate models, ARMA(1, 1) model is selected jointly with GARCH(1, 1) model for this study to remove linear dependency and ARCH effects respectively in the returns series. This joint model is run under three distributional assumptions and their skewed versions: the normal, skew-normal, student's t, skew-student's t, Generalized Error Distribution (GED), skew-Generalized Error Distribution (GED).

The parameters of the GARCH models (as indicated in Tables 4.4 and 4.5) are estimated with the use of "rugarch" package developed by Ghalanos, (2015) in R programming language. AIC and BIC respectively denote the Akaike Information Criterion and Bayesian Information Criterion. The results of ARMA(1, 1)-sGARCH(1, 1) under the three error distributions and their skewed versions for the daily series in the four indices indicate that all of the GARCH parameters (ω, α and β coefficients)

Table 4.4: Estimated result of ARMA(1, 1)-sGARCH(1, 1) model.

ARMA(1, 1)-sGARCH(1, 1) under "**normal distribution**" assumption				
Coefficients	South Africa (JSE-ALSI)	Nigeria (NIGALSH)	Egypt (EGX 30)	Kenya (KSE 20)
Mean				
μ (constant)	0.072931*	0.067399*	0.103460*	0.032758**
Variance				
ω	0.018315*	0.001879*	0.112655*	0.016285*
α	0.086618*	0.060561*	0.110707*	0.130562*
β	0.901791*	0.895453*	0.850800*	0.860465*
$\alpha + \beta$	0.920109	0.956014	0.961507	0.991027
AIC	2.9464	2.8403	3.6480	2.1448
BIC	2.9560	2.8499	3.6576	2.1545
ARMA(1, 1)-sGARCH(1, 1) under "**student's t distribution**" assumption				
Coefficients				
Mean				
μ (constant)	0.078241*	0.005332***	0.11094*	0.013122***
Variance				
ω	0.016855*	0.139100*	0.11483*	0.095728*
α	0.084787*	0.420775*	0.17916*	0.302789*
β	0.905037*	0.530336*	0.79821*	0.582234*
$\alpha + \beta$	0.989824	0.951111	0.97737	0.885023
AIC	2.9309	2.3518	3.5281	1.9088
BIC	2.9421	2.3630	3.5393	1.9200
ARMA(1, 1)-sGARCH(1, 1) under "**generalized error distribution" (GED)** assumption				
Coefficients				
Mean				
μ (constant)	0.072171*	0.034649*	0.053762*	0.014412**
Variance				
ω	0.017712*	0.001782*	0.118631*	0.079277*
α	0.085363*	0.050000*	0.157182*	0.258666*
β	0.903465*	0.900000*	0.811050*	0.623865*
$\alpha + \beta$	0.988828	0.950000	0.968232	0.882531
AIC	2.9303	3.3647	3.5268	1.9232
BIC	2.9415	3.3759	3.5380	1.9344

Note: "*", "" and "***" denote that the parameters are significant at 1%, 5% and 10% levels respectively.**

Table 4.5: Estimated result of ARMA(1, 1)-sGARCH(1, 1) model.

ARMA(1, 1)-sGARCH(1, 1) under "**skew – normal distribution**" assumption				
Coefficients	South Africa (JSE-ALSI)	Nigeria (NIGALSH)	Egypt (EGX 30)	Kenya (KSE 20)
Mean				
μ (constant)	0.068455*	0.043267*	0.070841*	0.031544**
Variance				
ω	0.016205*	0.002045*	0.112269*	0.016391*
α	0.083917*	0.088665*	0.116094*	0.130940*
β	0.905718*	0.890871*	0.845806*	0.859912*
$\alpha + \beta$	0.989635	0.979536	0.961900	0.990852
AIC	2.9417	2.6375	3.6401	2.1453
BIC	2.9530	2.6487	3.6513	2.1566
ARMA(1, 1)-sGARCH(1, 1) under "**skew – student's t distribution**" assumption				
Coefficients				
Mean				
μ (constant)	0.066478*	0.042061*	0.099134*	0.018180***
Variance				
ω	0.015466*	0.135854*	0.113232*	0.095364*
α	0.082731*	0.417120*	0.176824*	0.301733*
β	0.907730*	0.536871*	0.800008*	0.583755*
$\alpha + \beta$	0.990461	0.953991	0.976832	0.885488
AIC	2.9278	2.3491	3.5283	1.9092
BIC	2.9406	2.3619	3.5411	1.9221
ARMA(1, 1)-sGARCH(1, 1) under "**skew – GED**" **distribution**" assumption				
Coefficients				
Mean				
μ (constant)	0.062874*	0.034649*	0.06622*	0.018936*
Variance				
ω	0.016328*	0.001782*	0.11984*	0.078309*
α	0.083238*	0.050000*	0.15825*	0.257280*
β	0.906130*	0.900000*	0.81010*	0.627340*
$\alpha + \beta$	0.989368	0.950000	0.96835	0.884620
AIC	2.9285	3.7616	3.5271	1.9235
BIC	2.9413	3.7744	3.5399	1.9364

Note: "*", "" and "***" denote that the parameters are significant at 1%, 5% and 10% levels respectively.**

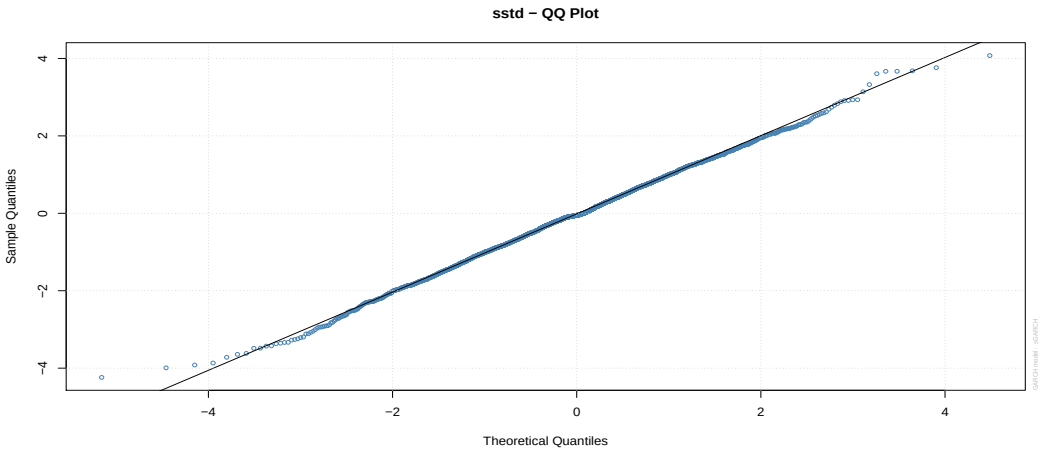

Figure 4.6: Diagnostic: sstd Q-Q plot JSE ALSI (South Africa).

are statistically significant at 1% level.

After fitting these models on each of the four markets, it is concluded that both the South African and Nigerian equity markets can be best described (or characterized) by the skew-student's t distribution (sstd), Egyptian market is best characterized by Generalized error distribution (GED) and the Kenyan market is best described by the student's t distribution. These conclusions are based on the returned values of the information criteria (i.e. the AIC and BIC). The least values of the AIC and BIC are indicative of the best model to describe the markets (see Tables 4.4 and 4.5).

4.5.1 Diagnostic plots of best ARMA-GARCH models

Figure 4.6 above and Figures 5.5, 5.6 and 5.7 in the Appendix display the diagnostic Q-Q plots of the volatility models. The plots show adequate levels of linearity since the volatility points stick very closely to the straight lines, except at the tails, indicating extreme movements. Figures 5.8 to 5.11 in the Appendix show the densities of the residuals under each of the best error distribution.

4.6 The markets' volatility hierarchy

The volatility and persistence of each of the markets is ascertained by non-Gaussian distributions like the student's t and GED. This came to bear since it has been sufficiently shown with time that the normal distribution is not adequate for modelling the fat-tailed (leptokurtic) financial returns (McNeil and Frey, 2000). Hence using student's t and GED assumptions with their skewed versions and the values of volatility persistence $(\alpha+\beta)$, Tables 4.4 and 4.5 show in descending hierarchy that volatility with persistence is highest in the South African market, followed by Egyptian market, then Nigerian market and lastly, the Kenyan equity market. Volatility displays long persistence into the future if the sum $(\alpha+\beta)$ is close to one, hence the closer it is to one, the greater the persistence.

The result is further confirmed by visually inspecting the volatility persistence (i.e. Shock vs Persistence) plots of the markets under student's t and GED assumptions with their skewed versions. Under any of these assumptions, the volatility hierarchy follows the same pattern. However for brevity and space constraint, the volatility persistence plots under skew-student's t assumption in descending hierarchy are attached in Figures 5.14, 5.15, 5.16 and 5.17 in the Appendix. The plots display the volatility persistence of the markets in descending order of magnitudes as follows: South Africa (JSE-ALSI), Egypt (EGX 30), Nigeria (NIGALSH) and Kenya (NSE 20) markets at about 10000, 900, 250 and 240 in magnitudes respectively.

4.7 Univariate analysis

This section focuses on the univariate aspect of the two "risk and extremal dependence" models: the bivariate-threshold-excess model and point process. The risk in each of the four domestic markets is modelled, the results of which will propel the

pairwise modelling of the dependence between markets. That is, the univariate analysis sets a platform for the multivariate analysis, hence it requires special attention. The sequence of procedures involved in the analysis include: threshold selection and diagnostics, sensitivity analysis for a suitable threshold choice, declustering of threshold exceedances, parameters estimation and diagnostics, return levels estimation and diagnostics.

4.7.1 Univariate analysis: The bivariate-threshold-excess model

After extracting the residuals from the returns and obtaining the exploratory data analysis of these, the next thing that comes to bear in this study is the selection of the required thresholds. With four different markets in consideration for this study, the univariate analysis of the residuals of each market under the "bivariate-threshold-excess model" is done one after the other as follows: South Africa, Nigeria, Egypt and Kenya.

4.8 The South African market: JSE-ALSI

The South African market index is called the Johannesburg Stock Exchange All-Share Index, abbreviated - JSE-ALSI, and it contains 3910 observations for the periods covered in this study.

4.8.1 JSE-ALSI: Positive residual observations

It is important to know that the focus of the study is on modelling the risk and extremal dependence of observations in the positive residuals of the entire equity observations. From the original data that contains 3910 number of observations, there are 1890 positive residual observations in this market as shown in Figure 4.7. The

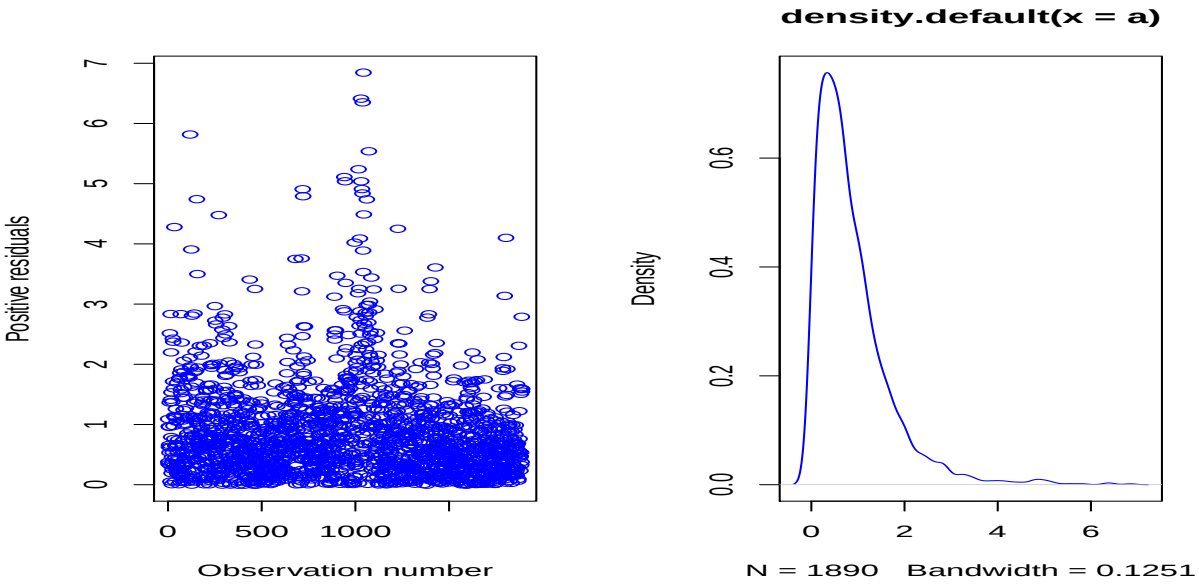

Figure 4.7: JSE-ALSI: Positive residuals.

figure shows the positive residuals represented by the points and their corresponding density.

4.8.2 JSE-ALSI: Threshold selection

For the selection of an appropriate threshold for this market and the rest of the markets in this study, two models are used. The first is the "extreme value mixture models" and the second (used to augment the first) is the "shape threshold stability plot".

The extreme value mixture models used here is the Kernel-GPD model approach. The chosen kernel could be any of Gaussian, gamma, or Weibull distribution as the bulk model below the threshold with GPD above the threshold for the tail modelling. The threshold can be calculated by using either the bulk model based approach or the parameterised approach. For consistency and comparison, this study uses the Gaussian and Weibull kernels with GPD tail model under the parameterised approach

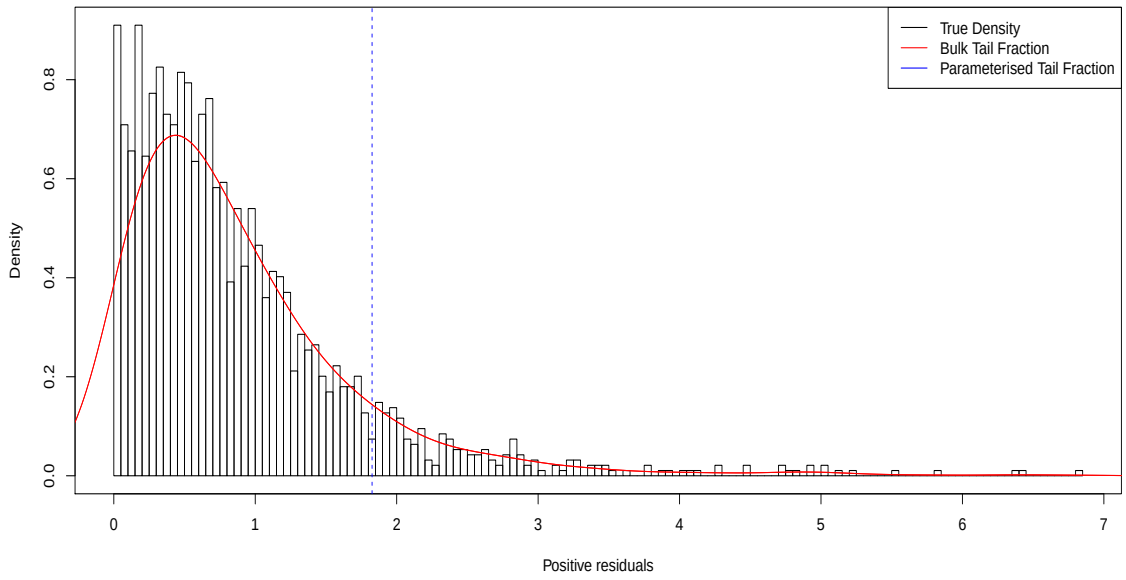

Figure 4.8: JSE-ALSI: Threshold selection with Weibull kernel.

for the four selected markets. For brevity, only the plot for the Weibull kernel with GPD is displayed for all the markets. Figure 4.8 shows the output of the Weibull-GPD fit where the threshold $u = 1.8259$. For the Gaussian kernel with GPD, $u = 1.6034$ is obtained, which is close compared to the former. Both computations were carried out under the parameterised approach.

4.8.3 Diagnostic plots of the Weibull-GPD model

The diagnostic plots of the fitted Weibull-GPD model for the threshold selection as displayed in Figure 4.9 indicate that the fit is adequate. The points are all enclosed within the dashed confidence limits and each set of plotted points is close to linear. Furthermore, the density estimate looks consistent with the histogram.

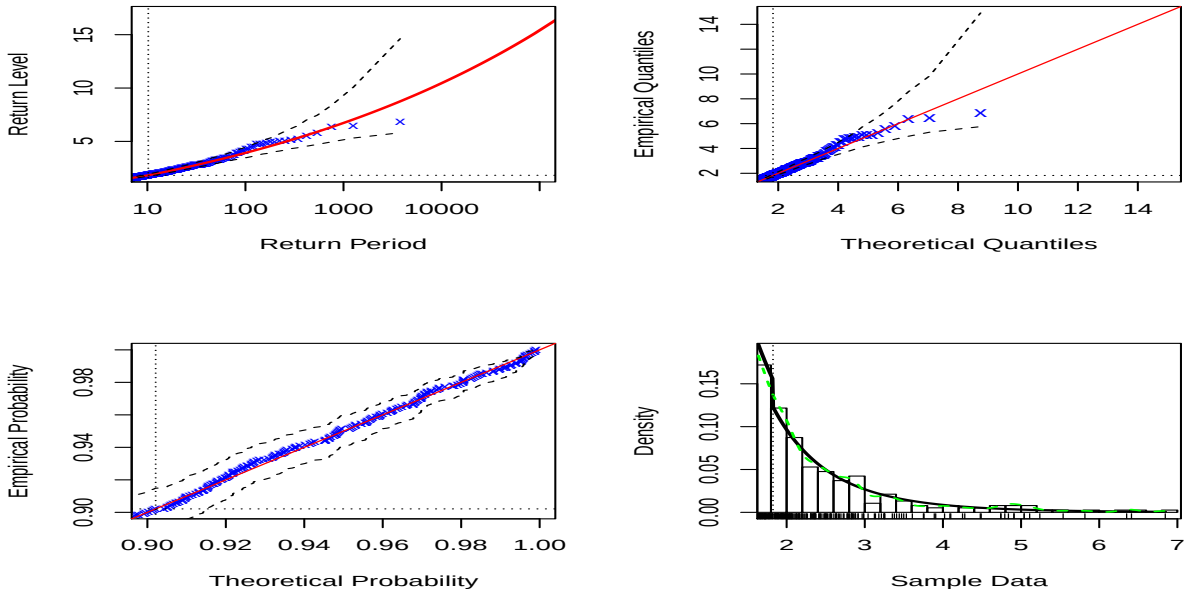

Figure 4.9: JSE-ALSI diagnostic plots: Threshold selection with Weibull-GPD model.

4.8.4 Conditions for final threshold choice

Following Ferro (2003), a sensible choice for a suitable threshold is where the extremal index (θ) is close to 0.5, hence a minimum of $\theta = 0.5$ is used for final threshold selection in this study. Furthermore, normalised "inter-exceedance times" against "standard exponential quantiles" plots as in Figure 4.10 are required to be piecewise-linear with a breakpoint at the $(1-\theta)-$ quantile, $-$log. The vertical line represents the $(1-\theta)-$ quantile (Ferro, 2003).

4.8.5 Sensitivity analysis

It is observed that the extremal index ($\hat{\theta}$) of each of the two obtained threshold values (from the Gaussian and Weibull with the GPD tail model) in Section 4.8.2 is below 0.50. Therefore, to obtain a suitable threshold value with minimum extremal index of 0.50 (Ferro, 2003), a sensitivity analysis was carried out as in Figure 4.10 and

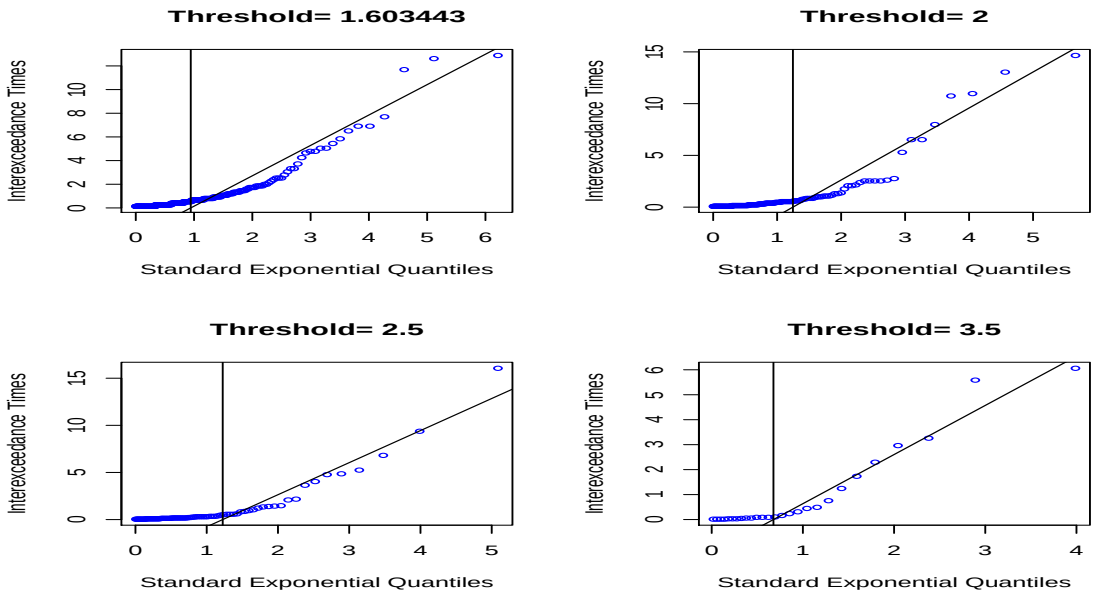

Figure 4.10: JSE-ALSI: Threshold selection sensitivity analysis plots.

threshold $u = 3.50$ with extremal index, $\hat{\theta} = 0.51$ was selected for the JSE-ALSI market.

4.8.6 Shape threshold stability plot

For further support on this threshold choice, the shape threshold stability plot is used. This plot as displayed in Figure 4.11 shows significant departure from linearity beyond $u = 2.7$. This is a reasonable support to the choice of $u = 3.5$ of the Kernel density mixture models.

4.8.7 Declustering

The threshold, $u = 3.5$, generates 28 threshold exceedances. After declustering at this threshold, 15 cluster-maxima are obtained as displayed in Figure 4.12. The cluster-maxima is the maxima of the clusters of exceedances over a high threshold

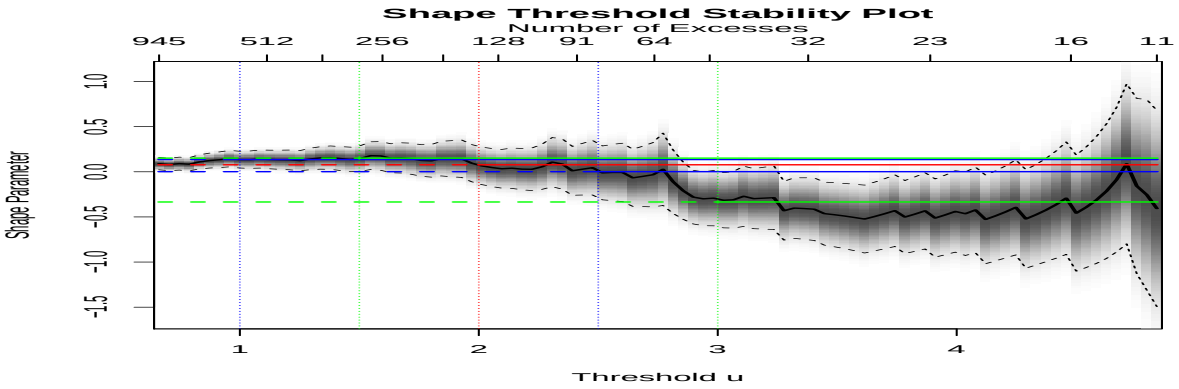

Figure 4.11: JSE-ALSI: Shape threshold stability plot.

(u=3.5 in this case). These are "exceedance residuals" (McNeil and Frey, 2000), i.e. independent extreme tail observations on which the GPD is fit.

4.8.8 Parameters estimation: GPD fit to cluster-maxima

Following the generation of the 15 cluster-maxima, which are collectively classified as the risk to be modelled in the South African JSE-ALSI market at the chosen threshold, the GPD will now be fitted. That is, the risk is modelled by fitting the GPD on the 15 observations and the result in Table 4.6 is deduced. The table also contains the sample proportion, η_u of points (in this case, the cluster-maxima) exceeding the threshold $u = 3.5$. Here, η_u is the number of cluster-maxima k to the total positive residual observations m.

The negative estimate of the shape parameter is a reflection of convexity (Coles, 2001) and it indicates a short tailed distribution. The shape parameter, $\xi < 0$, corresponds to a bounded distribution, hence the risk in this market can be described by the Weibull class of distributions. The reason for this is particularly strong for the 90% and 95% confidence intervals since the ξ is entirely in the negative domain. The strength of evidence from the exceedance data is not too strong on bounded

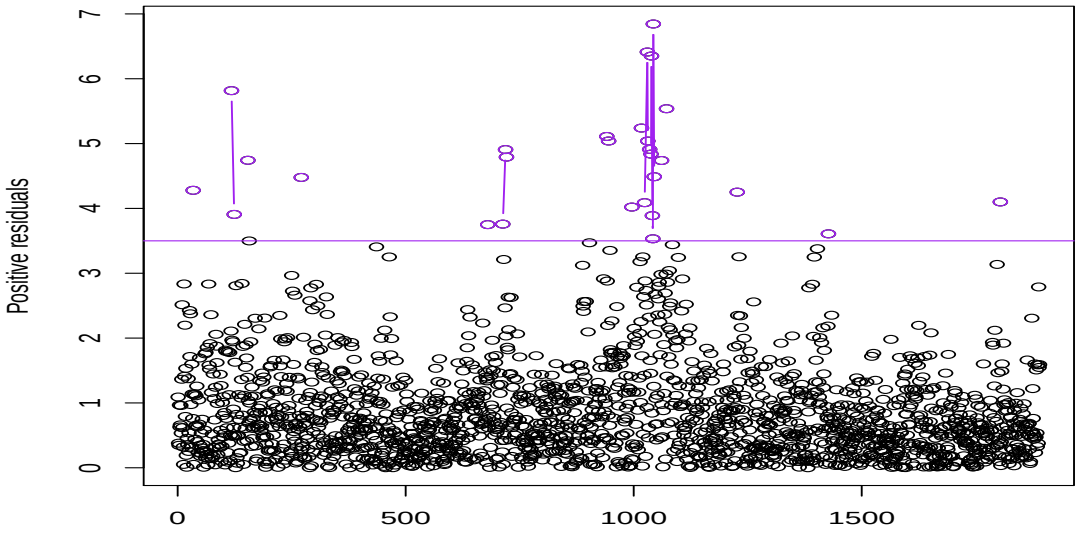

Figure 4.12: JSE-ALSI: Declustered exceedances at $u = 3.5$ (cluster-maxima).

distribution at 99% since the interval extends above zero. An alternative way of obtaining the confidence intervals for the shape parameter (ξ) with better accuracy (Coles 2001) is using the profile likelihood intervals. The profile log-likelihood plots for ξ at the 95% for JSE-ALSI index and the rest of the markets are shown in Figure 5.18 of the Appendix. The plot for the lower limit of the JSE-ALSI index (top left panel) is not fully shown due to the low sample size of the cluster-maxima observations.

Table 4.6: JSE-ALSI: Univariate threshold excess parameter estimates.

	CI	u	m	k	$\eta_u = \frac{k}{m}$	σ	ξ
JSE-ALSI	90%	3.50	1890	15	0.0079	1.89 (0.34)	-0.50 (0.25)
						[1.44; 2.34]	[-0.84; -0.16]
	95%	3.50	1890	15	0.0079	1.89 (0.34)	-0.50 (0.25)
						[1.30; 2.48]	[-0.94; -0.06]
	99%	3.50	1890	15	0.0079	1.89 (0.34)	-0.50 (0.25)
						[1.01; 2.77]	[-1.16; 0.16]

However, the upper limit is approximately 0.12 for the 95% confidence interval.

4.8.9 Diagnostics: Model checking

Model diagnostics are ways of checking or judging the accuracy of a model whether it agrees with the data used in estimating it and to ascertain the validity of the model. Statistical models are fit to data in order to make reasonable conclusions about the population where the data is drawn.

Figure 4.13 shows four diagnostic plots to assess the validity of the GPD fit to the 15 cluster-maxima. Each set of the plotted points of the quantile plot and the probability plot is near-linear; this makes the GPD fit reasonably plausible. The fact is evidently galvanized by the accurateness of the points on the straight line diagonal of the probability plot's tolerance interval. Furthermore, the points in the return level curve also provides an adequate picture of the empirical estimates. In addition, the threshold exceedances' histogram can be compared to the density function of the GPD model's fit.

4.8.10 Goodness of fit of the GPD

This study also uses the goodness of fit tests of Cramér-von Mises and Anderson-Darling to ascertain how well the generalized Pareto distribution fits the cluster-maxima observations in the JSE-ALSI market. The tested hypothesis are:

$$H_\circ : \text{GPD fits the cluster-maxima well}$$
$$H_1 : \text{GPD does not fit it well}$$

Table 4.7 displays the output of the test statistic and critical points (p-values) of the tested hypothesis of the Cramér-von Mises and Anderson-Darling tests. For both

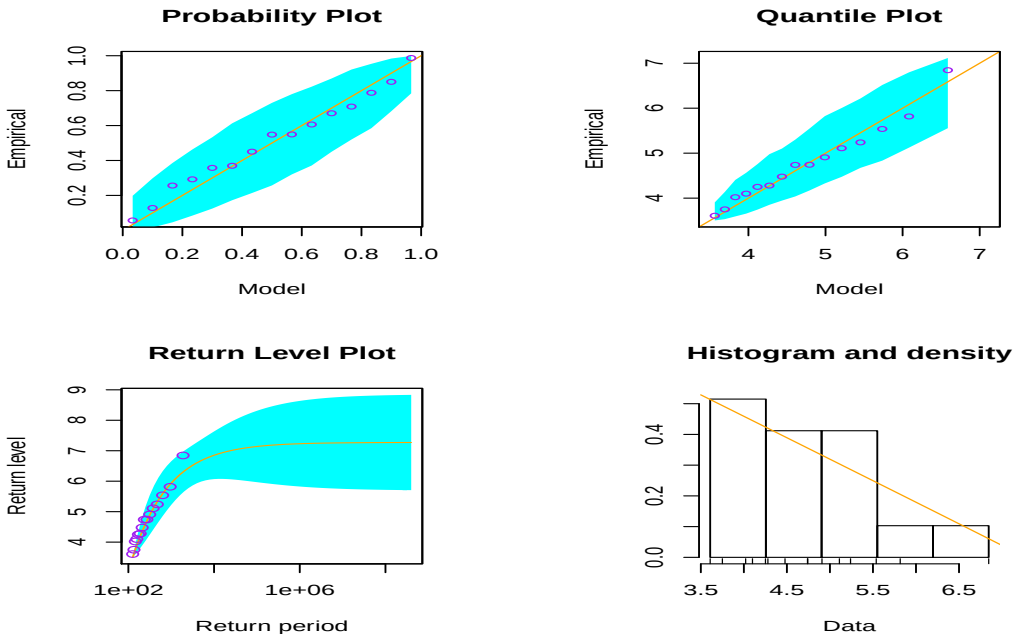

Figure 4.13: JSE-ALSI: GPD diagnostic plots.

Table 4.7: JSE-ALSI: Goodness of fit test.

Test	GPD values	Test values	Statistic values	P-value
Cramér-von Mises	σ: 1.8914 ξ: -0.5014	σ: 1.9244 ξ: -0.5138	0.0679	0.5038
Anderson-Darling	σ: 1.8914 ξ: -0.5014	σ: 1.9244 ξ: -0.5138	0.4633	0.5169

tests, it is observed that the null hypothesis is not rejected since the p-value is large (greater than 0.05). This implies that the GPD fits the generated cluster-maxima observations well.

4.8.11 JSE-ALSI: Return levels

Table 4.8 shows the return levels estimates and their corresponding confidence intervals (in parentheses) for the JSE-ALSI index (South African market). The confidence intervals at 90%, 95%, and 99% are included to increase its informativeness. These

Table 4.8: JSE-ALSI: Threshold excess return level (R_q) estimates.

	$u = 3.50$	2-year	5-year	10-year	20-year
R_q estimates		7.13	7.18	7.21	7.23
90% CI		(6.78; 7.49)	(6.81; 7.56)	(6.67; 7.75)	(6.49; 7.97)
95% CI		(6.67; 7.60)	(6.69; 7.68)	(6.50; 7.92)	(6.26; 8.20)
99% CI		(6.44; 7.83)	(6.45; 7.92)	(6.15; 8.27)	(5.78; 8.68)

confidence intervals are obtained from the delta method's variance formula of equation (3.9.4) to generate the required standard errors (Coles, 2001). For instance, for the 10-year return period, the estimation indicates that on average, a rare event of level (magnitude) 7.21 is expected to be exceeded once over the next 10 years.

4.8.12 Univariate analysis: Point process

The univariate analysis using the point process has the same procedure as that of the bivariate-threshold-excess model analysis. The same threshold is used to enable appropriate comparison between these two models. As a way forward, the already declustered exceedances (i.e. the cluster-maxima) obtained using the univariate analysis of the bivariate-threshold-excess model is applied for the parameters estimation of the point process. The standard errors of the estimates are enclosed in parentheses while the confidence intervals are in the brackets (Table 4.9).

4.8.13 Parameters estimation: Point process fit to cluster-maxima

As it is under the GPD estimation of the parameters, the 15 cluster-maxima are different magnitudes of the risk in the South African equity market at threshold $u = 3.5$. This risk is again modelled by fitting the point process on the 15 cluster-maxima observations and the result in Table 4.9 is obtained. The nature of the result obtained is very similar to that of the GPD estimation with negligible dichotomy.

Table 4.9: JSE-ALSI: Univariate point process parameter estimates.

JSE-ALSI	CI	u	μ	σ	ξ
	90%	3.50	7.08 (0.53) [6.36; 7.79]	0.10 (0.11) [-0.05; 0.25]	-0.50 (0.25) [-0.83; -0.17]
	95%	3.50	7.08 (0.53) [6.14; 8.01]	0.10 (0.11) [-0.10; 0.30]	-0.50(0.25) [-0.94; -0.07]
	99%	3.50	7.08 (0.53) [5.68; 8.47]	0.10 (0.11) [-0.20; 0.40]	-0.50(0.25) [-1.15; 0.15]

The shape parameter estimate is also negative ($\xi < 0$), which reflects convexity (Coles, 2001), indicating a short tailed distribution. This estimate corresponds to a bounded distribution, hence the risk in this market can be described by the Weibull class of distributions. The reason for this is specifically strong for the 90% and 95% confidence intervals since the ξ is exclusively in the negative domain. The strength of evidence from the exceedance data is not too strong on bounded distribution at 99% since the interval extends above zero.

4.8.14 Point process diagnostic plots

The diagnostic probability plot and quantile plot of the point process fit are also satisfactory as the points stick close to the diagonal straight line, except for the outlier at the extreme of the quantile plot (see Figure 4.14).

4.8.15 JSE-ALSI: Return levels

Table 4.10 shows the point process return levels estimates and their corresponding confidence intervals (in parentheses) for the JSE-ALSI index. For instance, for the 10-year return period, the estimation indicates that on average, a rare event of level (magnitude) 7.21 is expected to be exceeded once over the next 10 years.

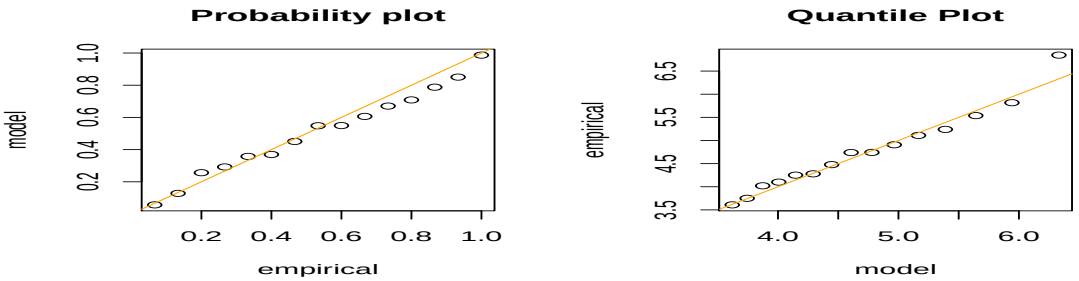

Figure 4.14: JSE-ALSI: Point process diagnostic plots.

Table 4.10: JSE-ALSI: Point process return level (R_q) estimates.

	$u = 3.50$	2-year	5-year	10-year	20-year
R_q estimates		7.11	7.18	7.21	7.23
95% CI		(6.00; 8.21)	(5.92; 8.44)	(5.88; 8.54)	(5.84; 8.62)

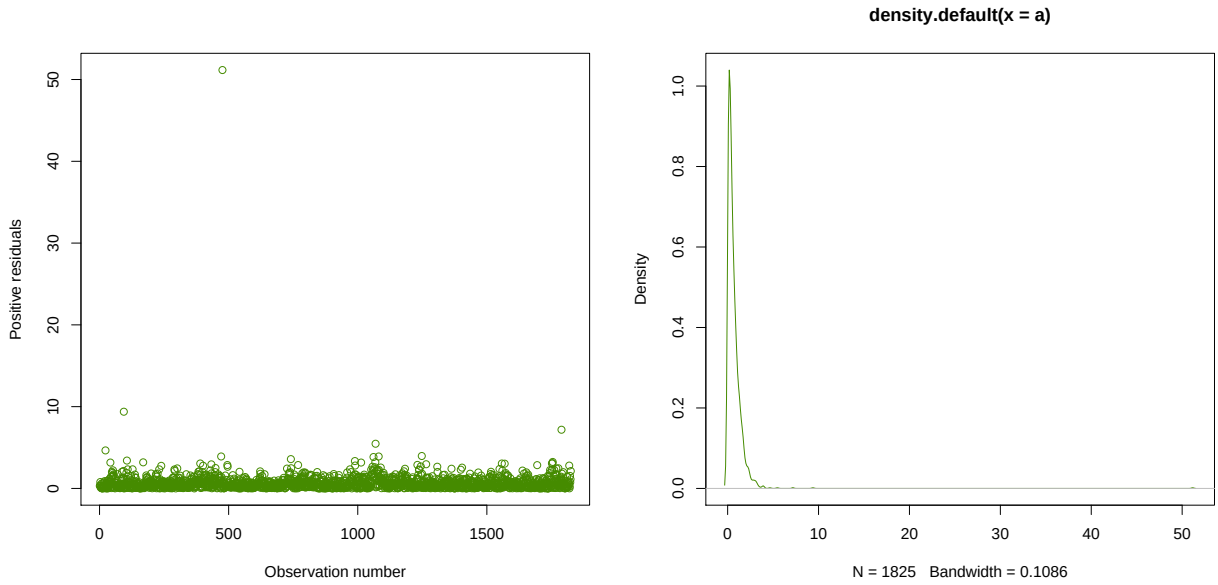

Figure 4.15: NIGALSH: Positive residuals.

4.9 The Nigerian market: NIGALSH

The Nigerian market index is called the Nigeria All-Share Index, abbreviated - NIGALSH, and it contains 3910 observations for the periods covered in this study.

4.9.1 NIGALSH: Positive residual observations

There are 1825 positive residuals observations out of the entire 3910 for the study period. Figure 4.15 displays the positive residuals represented by the points and their corresponding density.

4.9.2 NIGALSH: Threshold selection

The same threshold selection procedure used for the South African JSE-ALSI is applied for the Nigerian NIGALSH, i.e. the Kernel GPD approach. Here, the Gaussian

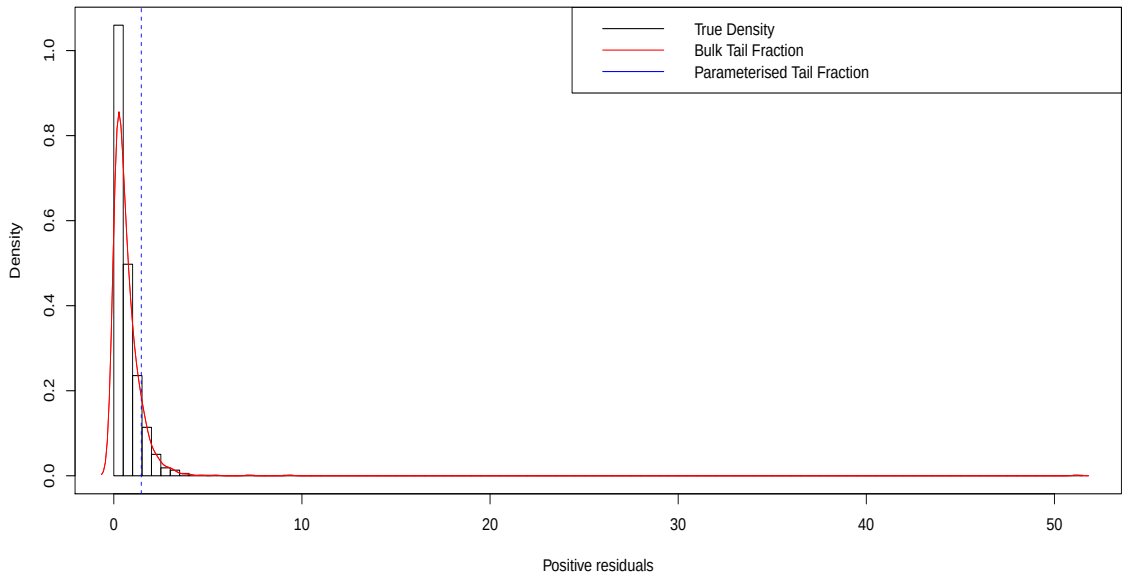

Figure 4.16: NIGALSH: Threshold selection with Weibull kernel.

and Weibull kernels are used one after the other, each with GPD for the tail modelling under the parameterised tail fraction approach. Figure 4.16 shows the output of the Weibull kernel with GPD fit at threshold value $u = 1.4600$. The value of $u = 1.4132$ is obtained for Gaussian kernel with GPD, but the plot is not displayed here for brevity.

4.9.3 Diagnostic plots of the Weibull-GPD model

For the diagnostics, nearly all the points are enclosed within the dashed curves confidence interval and each set of plotted points is close to linear (Figure 4.17). The only exception is in the quantile plot and the return level plot where the largest observation seems to be a clear outlier relative to the rest of the residual data. Furthermore, the density estimate looks consistent with the histogram. All these indicate that the GPD fit is reasonably satisfactory.

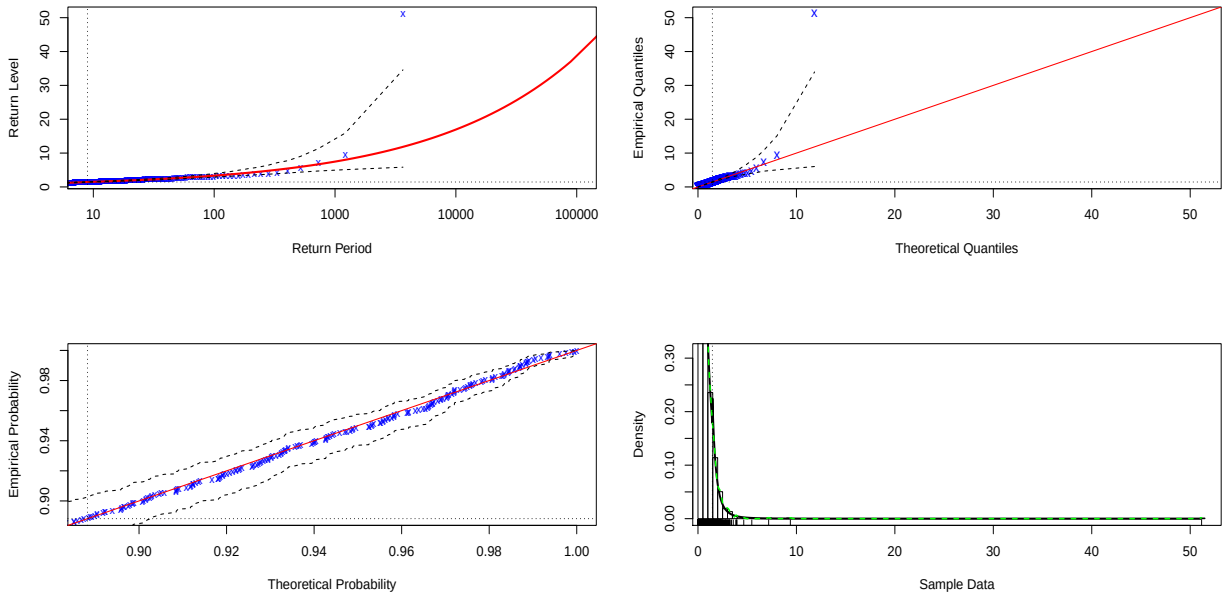

Figure 4.17: NIGALSH diagnostic plots: Threshold selection with Weibull-GPD model.

4.9.4 Conditions for final threshold choice

As a first point of consideration in determining the required threshold, an extremal index θ whose estimate is close to 0.5 (i.e. $\hat{\theta} = 0.5$) is used as a standard (Ferro, 2003). Furthermore, normalised "inter-exceedance times" against "standard exponential quantiles" plots as in Figure 4.18 are required to be piecewise-linear with a break-point at the $(1-\theta)-$quantile, $-$log. The vertical line represents the $(1-\theta)-$quantile (Ferro, 2003).

4.9.5 Sensitivity analysis

The extremal index θ of each of these selected thresholds is greater than 0.5. In particular, it can be observed from the sensitivity analysis plot at threshold $u = 2.0$ that the points stick closest to the straight line with the exception of the final data

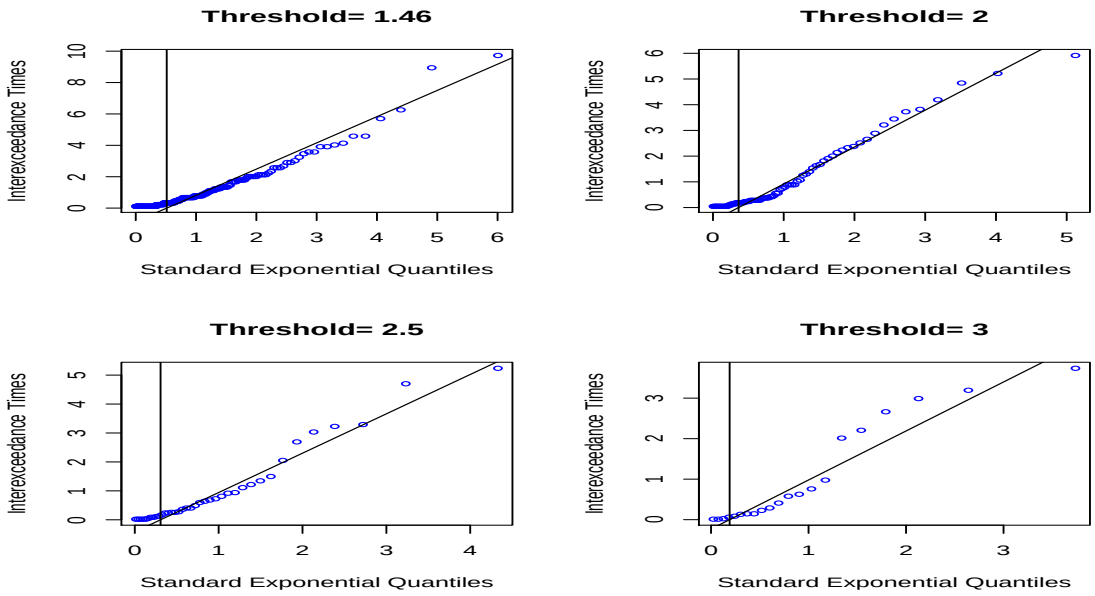

Figure 4.18: NIGALSH: Threshold selection sensitivity analysis plots.

point (see Figure 4.18). The linearity of the points are most observed at this threshold than the others. Hence, it seems intuitively appealing that this is a level ($u = 2.0$) above which the asymptotic assumption of the GPD is most valid for this market, with extremal index $\hat{\theta} = 0.70$.

4.9.6 Shape threshold stability plot

The choice of $u = 2.0$ can further be reinforced using the shape threshold stability plot in Figure 4.19 which shows reasonable linearity up to around $u = 2.0$.

4.9.7 Declustering

The selected threshold $u = 2.0$, generates 85 threshold exceedances. After declustering at this threshold, 55 cluster-maxima are obtained as displayed in the plots in Figure 4.20. The cluster-maxima is the maxima of the clusters of exceedances over a high

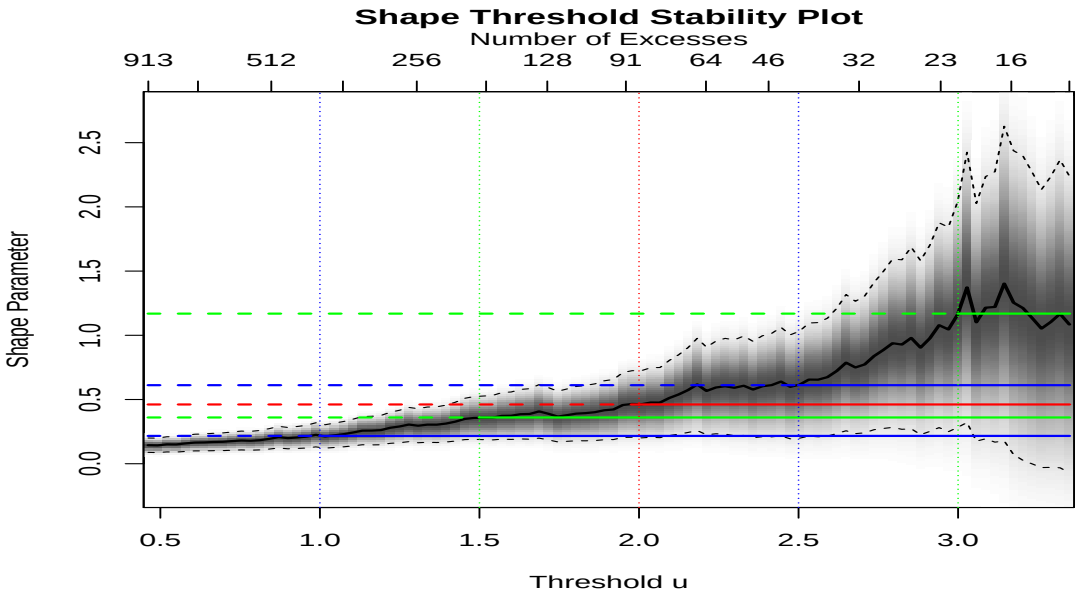

Figure 4.19: NIGALSH: Shape threshold stability plot.

threshold, which in this case is $u = 2.0$. These are "exceedance residuals" (McNeil and Frey, 2000), i.e. independent extreme tail observations on which the GPD and point process are fit respectively.

4.9.8 Parameters estimation: GPD fit to cluster-maxima

The 55 generated cluster-maxima are the extreme observations exceeding the threshold $u = 2.0$ and they are the risk of different magnitudes for modelling in the Nigerian NIGALSH market. The GPD is fit to these cluster-maxima observations and the parameter estimates in Table 4.11 are obtained.

The positive estimate of the shape parameter is a reflection of concavity (Coles, 2001) and it indicates a fat-tailed distribution. This positive shape parameter estimate corresponds to an unbounded distribution, hence the risk in the Nigerian NIGALSH market can be characterized by the Pareto distribution. The strength of

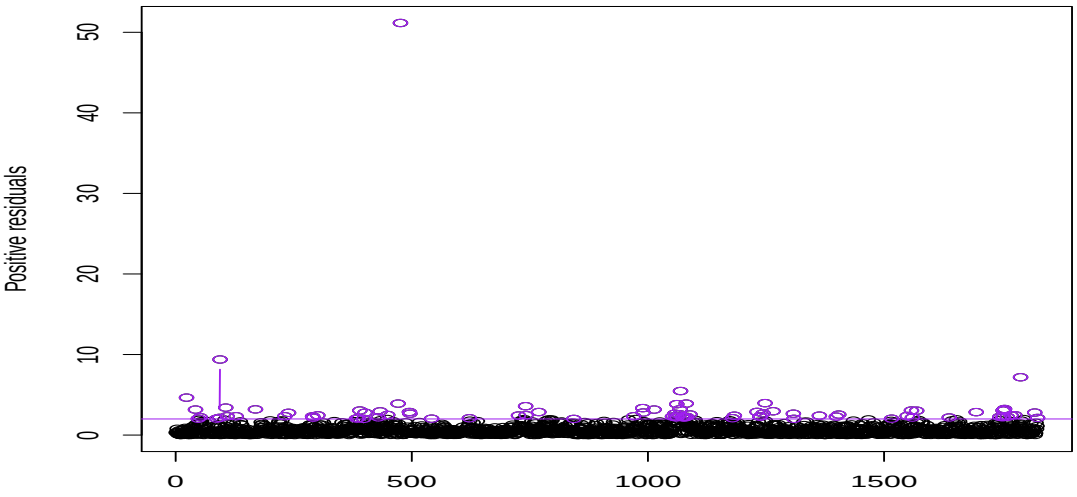

Figure 4.20: NIGALSH: Declustered exceedances at $u = 2.0$ (cluster-maxima).

Table 4.11: NIGALSH: Univariate threshold excess parameter estimates.

	CI	u	m	k	$\eta_u = \frac{k}{m}$	σ	ξ
NIGALSH	90%	2.00	1825	55	0.0301	0.72 (0.20) [0.39; 1.05]	0.49 (0.16) [0.22; 0.75]
	95%	2.00	1825	55	0.0301	0.72 (0.20) [0.33; 1.12]	0.49 (0.16) [0.17; 0.80]
	99%	2.00	1825	55	0.0301	0.72 (0.20) [0.21; 1.24]	0.49 (0.16) [0.07; 0.90]

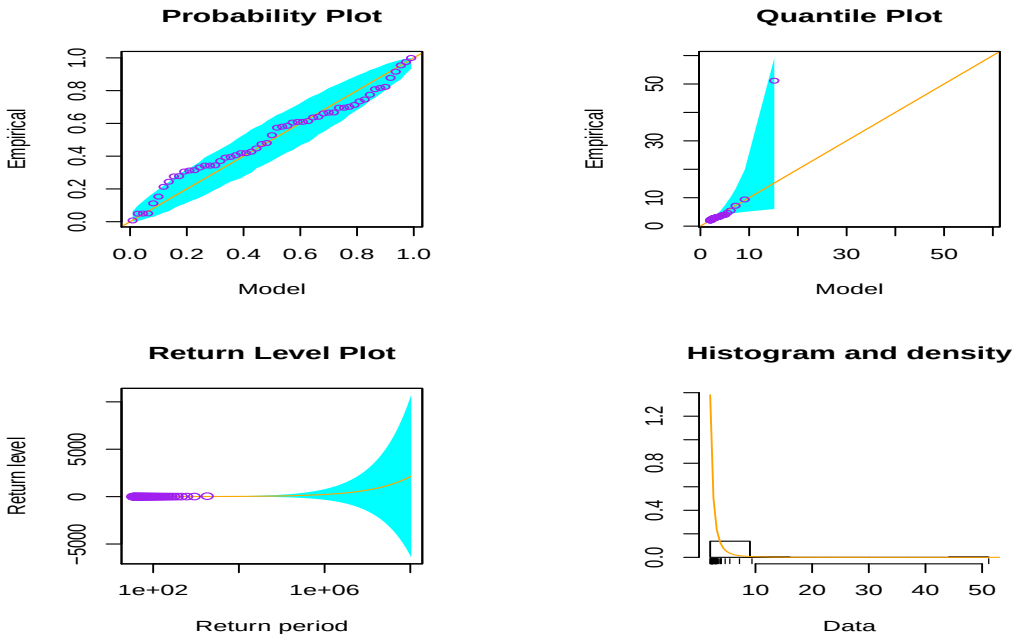

Figure 4.21: NIGALSH: GPD diagnostic plots.

evidence for this is strong at the three confidence intervals of 90%, 95% and 99% where the intervals are exclusively in the positive domain.

The profile log-likelihood plot for ξ for the NIGALSH index is displayed in the "top right panel" of Figure 5.18 in the Appendix. The graph indicates an approximate profile log-likelihood for the 95% confidence interval as [0.20; 0.85].

4.9.9 Diagnostics: Model checking

The accuracy of the fitted GPD model is judged by the diagnostic plots in Figure 4.21. The points on the quantile plot stick closely to the straight line except for the final observation which is a clear outlier. The return level plot displays points stable and well aligned on the straight line, however it shows very large uncertainties that enlarge at extrapolation of the model to higher levels. The density plot is satisfactory and the points on the probability plot are fairly aligned within the tolerance interval

Table 4.12: NIGALSH: Goodness of fit test.

Test	GPD values	Test values	Statistic values	P-value
Cramér-von Mises	σ: 0.7238 ξ: 0.4866	σ: 0.7185 ξ: 0.5096	0.0413	0.6237
Anderson-Darling	σ: 0.7238 ξ: 0.4866	σ: 0.7185 ξ: 0.5096	0.3349	0.5718

curve, hence the fit is reasonable.

4.9.10 Goodness of fit of the GPD

The goodness of fit tests of Cramér-von Mises and Anderson-Darling are further used to determine how well the generalized Pareto distribution fits the cluster-maxima observations in this market. The tested hypothesis are:

$$H_\circ : \text{GPD fits the cluster-maxima well}$$
$$H_1 : \text{GPD does not fit it well}$$

Table 4.12 displays the output of the test statistic and critical points (p-values) of the tested hypothesis of the Cramér-von Mises and Anderson-Darling tests. For both tests, it is observed that the null hypothesis is not rejected since the p-value is large (greater than 0.05). This implies that the GPD fits the generated cluster-maxima observations well.

4.9.11 NIGALSH: Return levels

Table 4.13 shows the return levels estimates and their corresponding confidence intervals (in parenthesis) for the NIGALSH index (Nigerian market).

These confidence intervals are generated using the variance formula of equation (3.9.4) to compute the required standard errors (Coles, 2001). For the 10-year return period, for instance, the estimation indicates that on average, an extreme event of level

Table 4.13: NIGALSH: Threshold excess return level (R_q) estimates.

	$u = 2.00$	2-year	5-year	10-year	20-year
R_q estimates		27.87	43.48	60.97	85.59
90% CI		(25.71; 30.02)	(38.88; 48.08)	(52.93; 69.01)	(71.82; 99.35)
95% CI		(25.30; 30.43)	(37.99; 48.96)	(51.39; 70.55)	(69.19; 101.98)
99% CI		(24.49; 31.24)	(36.27; 50.69)	(48.39; 73.56)	(64.04; 107.13)

60.97 is expected to be exceeded once over the next 10 years. The wide confidence intervals which are obtained for extreme return levels for long return periods are a reflection of the fact that little information is available for future return levels projection with any amount or degree of certainty (Coles, 2001).

4.9.12 Univariate analysis: Point process

The 55 cluster-maxima on which the GPD of the univariate side of the bivariate-threshold-excess model analysis was fitted will also be used for the point process fit since the threshold values are the same. Based on this, the same procedure is followed as that of the bivariate-threshold-excess model analysis.

4.9.13 Parameters estimation: Point process fit to cluster-maxima

The 55 cluster-maxima are the risk of different magnitudes in the Nigerian equity market at threshold $u = 2.0$. Like the GPD univariate estimation, the point process yields the same shape parameter estimate and same positive sign, with averagely the same confidence interval spread as shown in Table 4.14.

The positive estimate of the shape parameter is a reflection of concavity (Coles, 2001) and it indicates a fat-tailed distribution. This positive shape parameter estimate corresponds to an unbounded distribution, hence the risk in the Nigerian NIGALSH market can be characterized by the Fréchet-Pareto class of distributions

Table 4.14: NIGALSH: Univariate point process parameter estimates.

	CI	u	μ	σ	ξ
NIGALSH	90%	2.00	19.92 (7.69)	9.57 (5.84)	0.49 (0.12)
			[7.28; 32.57]	[-0.05; 19.18]	[0.29; 0.70]
	95%	2.00	19.92 (7.69)	9.57 (5.84)	0.49 (0.12)
			[4.86; 34.99]	[-1.89; 21.02]	[0.25; 0.74]
	99%	2.00	19.92 (7.69)	9.57 (5.84)	0.49 (0.12)
			[0.13; 39.71]	[-5.48; 24.61]	[0.17; 0.81]

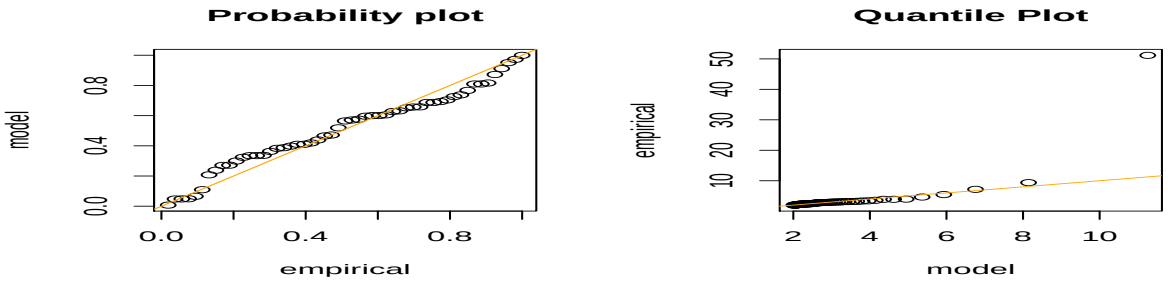

Figure 4.22: NIGALSH: Point process diagnostic plots.

(see Beirlant et. al, 2004). The strength of evidence for this is strong at the three confidence intervals of 90%, 95% and 99% where the intervals are wholly in the positive domain.

4.9.14 Point process diagnostic plots

The probability plots and quantile plots for the model checking indicate that the points are close to the straight lines (Figure 4.22). The only obvious exception is the single outlier of the largest observation displayed in the quantile plot. Hence, the validity of the fit is reasonable.

Table 4.15: NIGALSH: Point process return level (R_q) estimates.

	$u = 2.00$	2-year	5-year	10-year	20-year
R_q estimates		23.86	41.36	59.66	84.89
95% CI		(1.66; 46.05)	(-8.87; 91.58)	(-24.43; 143.76)	(-50.93; 220.70)

4.9.15 NIGALSH: Return levels

Table 4.15 shows the point process return levels estimates and their corresponding confidence intervals (in parentheses) for the NIGALSH index (Nigerian market). For instance, for the 10-year return period, the estimation indicates that on average, a rare event of level (magnitude) 59.66 is expected to be exceeded once over the next 10 years. The wide confidence intervals which are obtained for extreme return levels as displayed in the table and in Figure 5.19 of the Appendix are a reflection of the fact that little information is available for future return levels forecast with any degree of confidence (Coles, 2001) in the NIGALSH index.

4.10 The Egyptian market: EGX 30

The Egyptian market index is called the Egypt Stock Exchange, abbreviated - EGX 30, and it contains 3910 observations for the periods covered in this study.

4.10.1 EGX 30: Positive residual observations

With attention focused on only the positive residuals of the data, for the risk and extremal dependence modelling of this market, there are 1922 positive residuals out of the entire 3910 observations. Figure 4.23 displays the positive residuals represented by the points and their corresponding density.

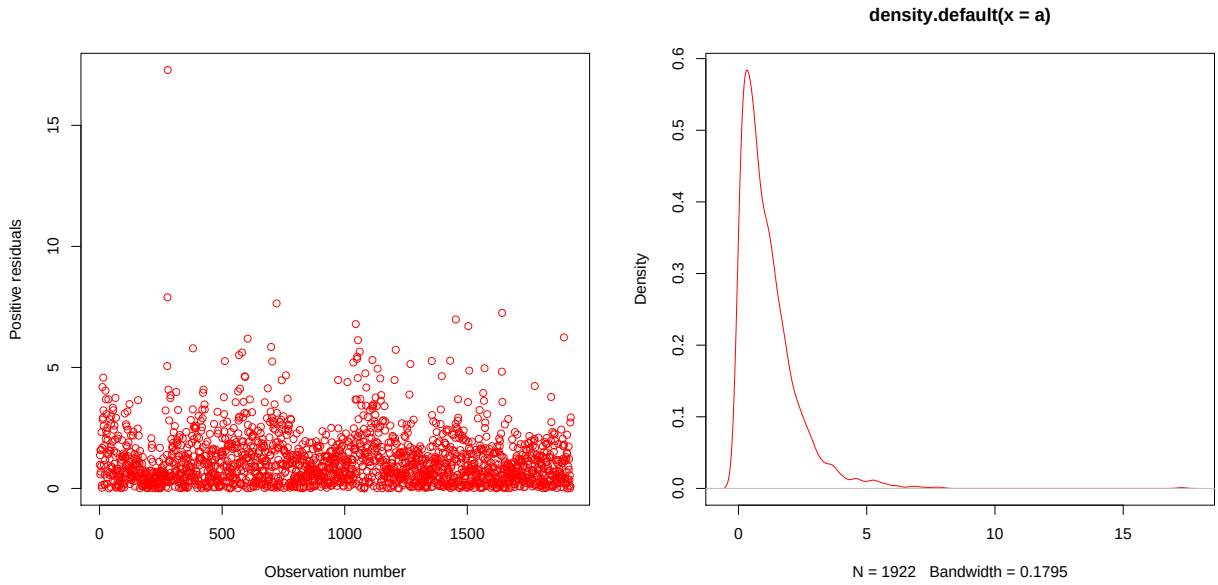

Figure 4.23: EGX 30: Positive residuals.

4.10.2 EGX 30: Threshold selection

As with the first two markets, the Kernel-GPD approach is used for the threshold selection, i.e. Gaussian-kernel with GPD first, followed by Weibull-kernel with GPD tail for comparison purpose. Both methods are applied under the parameterised tail fraction approach with GPD for the tail modelling. Figure 4.24 shows the output of the fit where $u = 2.5087$ is obtained when Weibull kernel is used with GPD. For Gaussian kernel with GPD, threshold value $u = 2.6319$ is obtained but the plot is not displayed for brevity.

4.10.3 Diagnostic plots of the kernel densities

The diagnostic plots of the Weibull Kernel-GPD model in Figure 4.25 indicates that the fit is adequate. The points are all enclosed with the dashed curves confidence interval and each set of plotted points is close to linear. Furthermore, the density

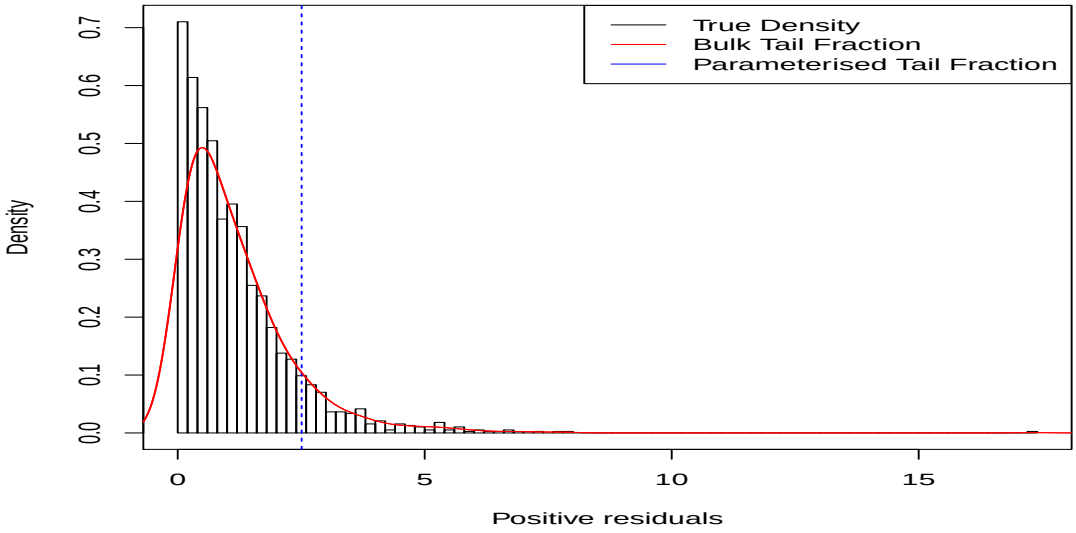

Figure 4.24: EGX 30: Threshold selection with Weibull kernel.

estimate looks consistent with the histogram.

4.10.4 Conditions for final threshold choice

To determine a suitable final threshold, a minimum extremal index, $\theta = 0.5$ is maintained (Ferro, 2003). Furthermore, normalised "inter-exceedance times" against "standard exponential quantiles" plots as in Figure 4.26 are required to be piecewise-linear with a breakpoint at the $(1-\theta)-$quantile, $-\log$. The vertical line represents the $(1-\theta)-$quantile (Ferro, 2003).

4.10.5 Sensitivity analysis

The sensitivity analysis plots in Figure 4.26 indicate that the points stick closest to the straight line for the plot at threshold $u = 3.5$. The linearity of the points is mostly observed at this threshold $(u = 3.5)$ than others, hence this is taken as the suitable

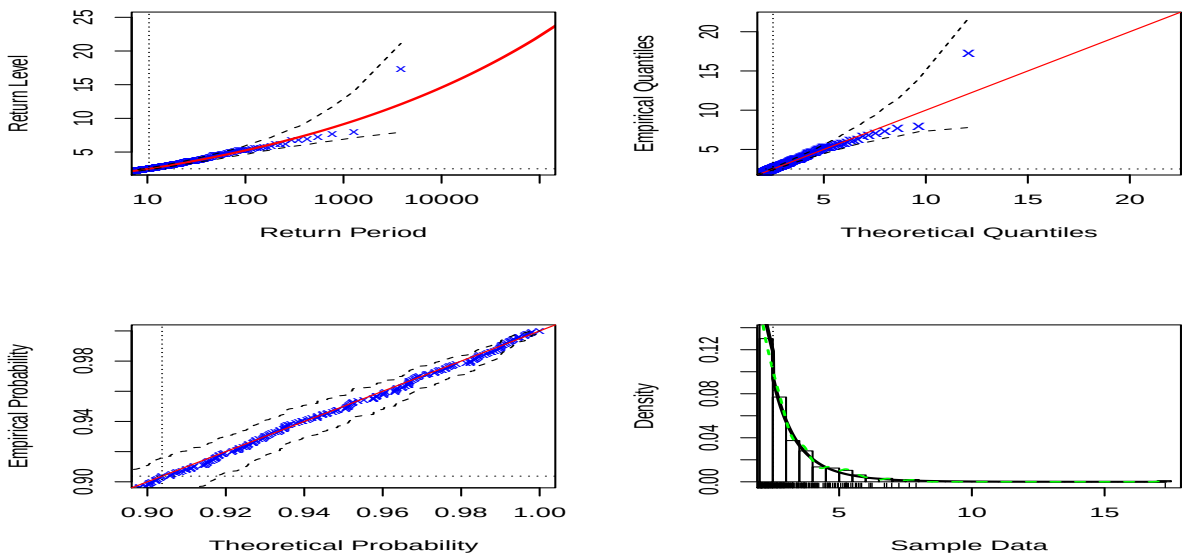

Figure 4.25: EGX 30 diagnostic plots: Threshold selection with Weibull-GPD model.

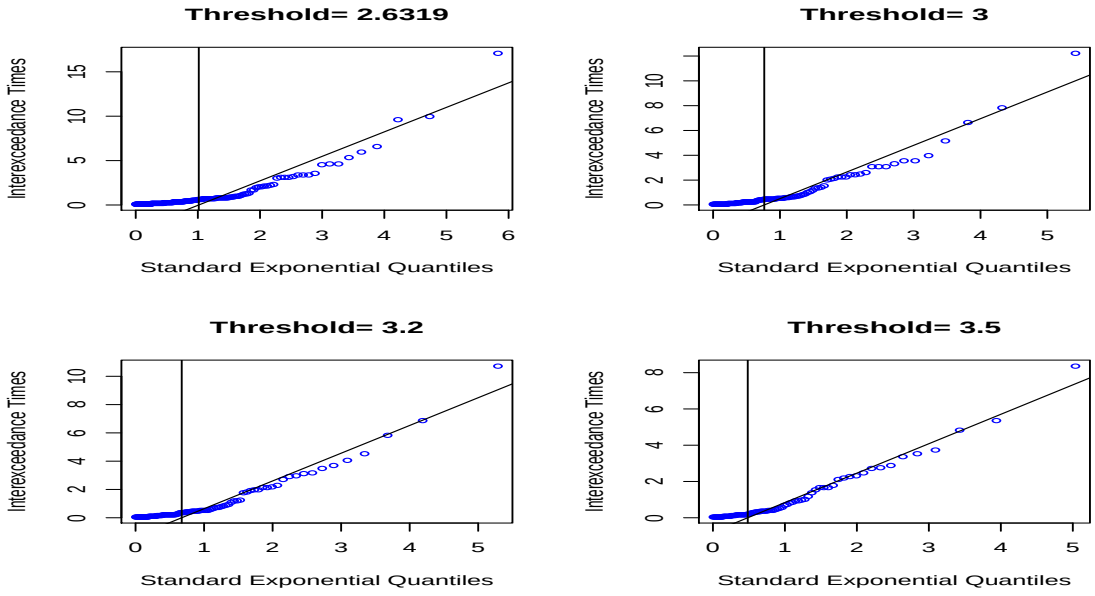

Figure 4.26: EGX 30: Threshold selection sensitivity analysis plots.

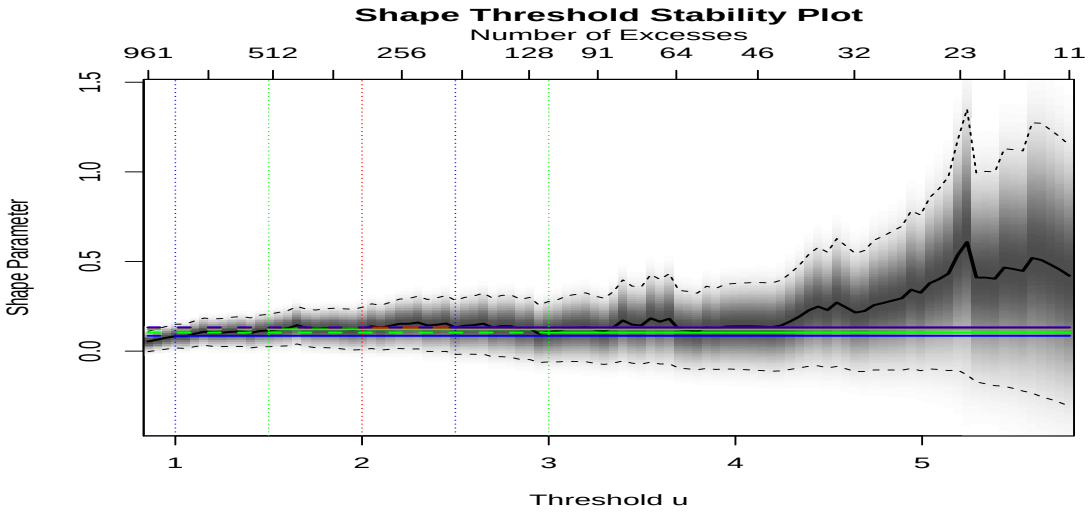

Figure 4.27: EGX 30: Shape threshold stability plot.

threshold for this market, with extremal index $\hat{\theta} = 0.62$.

4.10.6 Shape threshold stability plot

As a support to the adequacy of the chosen threshold after the sensitivity analysis, the shape threshold stability plot is used. Figure 4.27 indicates that the plot is initially linear, but later displays some levels of curvature around the range $3.5 < u < 6$. The change-in pattern around $u = 3.5$ observed in the Weibull-Kernel GPD model in Figure 4.24 is also apparent here. Hence $u = 3.5$ can be a good suggestion for this threshold value, which supports the choice of the value obtained via the sensitivity analysis.

4.10.7 Declustering

The selected threshold $u = 3.5$, generates 78 threshold exceedances. After declustering at this threshold, 49 cluster-maxima are obtained as displayed in the plots in Figure

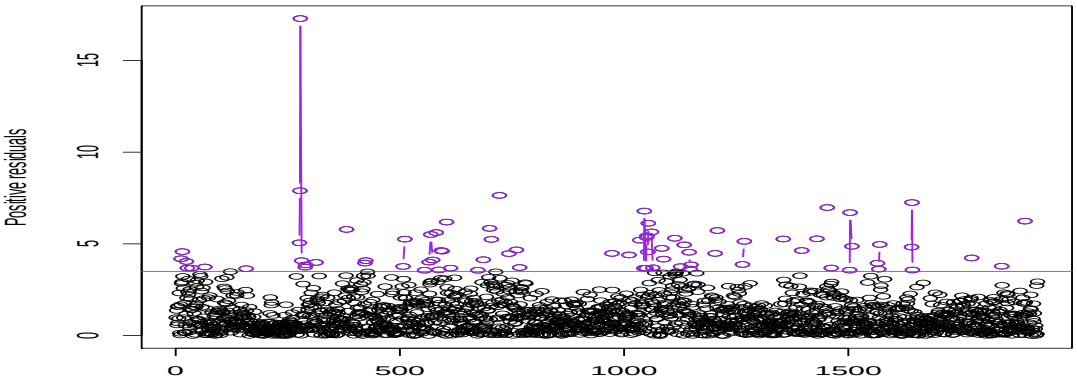

Figure 4.28: EGX 30: Declustered exceedances (cluster-maxima).

4.28. The cluster-maxima is the maxima of the clusters of exceedances over a high threshold ($u = 3.5$ in this case). These are "exceedance residuals" (McNeil and Frey, 2000), i.e. independent extreme tail observations on which the GPD and point process are fit respectively.

4.10.8 Parameters estimation: GPD fit to cluster-maxima

Following the generation of the 49 cluster-maxima (of different magnitudes), which are collectively classified as the risk to be modelled in the Egyptian EGX 30 market at threshold $u = 3.5$. This risk is modelled by fitting the GPD on the 49 observations and the result in Table 4.16 is deduced.

The positive estimate of the shape parameter is a reflection of concavity (Coles, 2001) and it indicates a fat-tailed distribution. That is, the shape parameter (ξ) estimate corresponds to an unbounded distribution, hence the risk in the Egyptian EGX 30 market can be well described by the Pareto distribution. The evidence for this is particularly strong at the 90% and 95% confidence intervals where the intervals are almost entirely in the positive domain. The evidence is not too strong at the 99%

Table 4.16: EGX 30: Univariate threshold excess parameter estimates.

	CI	u	m	k	$\eta_u = \frac{k}{m}$	σ	ξ
EGX 30	90%	3.50	1922	49	0.0255	1.42 (0.19) [1.11; 1.73]	0.12 (0.12) [-0.09; 0.32]
	95%	3.50	1922	49	0.0255	1.42 (0.19) [1.05; 1.79]	0.12 (0.12) [-0.13; 0.36]
	99%	3.50	1922	49	0.0255	1.42 (0.19) [0.93; 1.91]	0.12 (0.12) [-0.20; 0.44]

confidence interval since this extends well below zero.

The profile log-likelihood plot for ξ for the EGX 30 index is displayed in the "bottom left panel" of Figure 5.18 in the Appendix. The graph indicates an approximate profile log-likelihood for the 95% confidence interval as [-0.09; 0.46].

4.10.9 Diagnostics: Model checking

The diagnostic plots in Figure 4.29 do not give any real cause for concern with regards to the quality of the fitted GPD model except for the single outlier that is standing clear from the rest of the data. Also the return level plot shows very large uncertainties that broaden at extrapolation of the model to higher levels. However, the GPD model fit is reasonably adequate; there is no cause to doubt its validity since most of the plotted points are within the curves and close to the straight line. The density plot further gives support to the model's validity.

4.10.10 Goodness of fit of the GPD

As with the two earlier markets, this study uses the goodness of fit tests of Cramér-von Mises and Anderson-Darling to know how well the generalized Pareto distribution fits the cluster-maxima observations in the EGX 30 market. The tested hypothesis are:

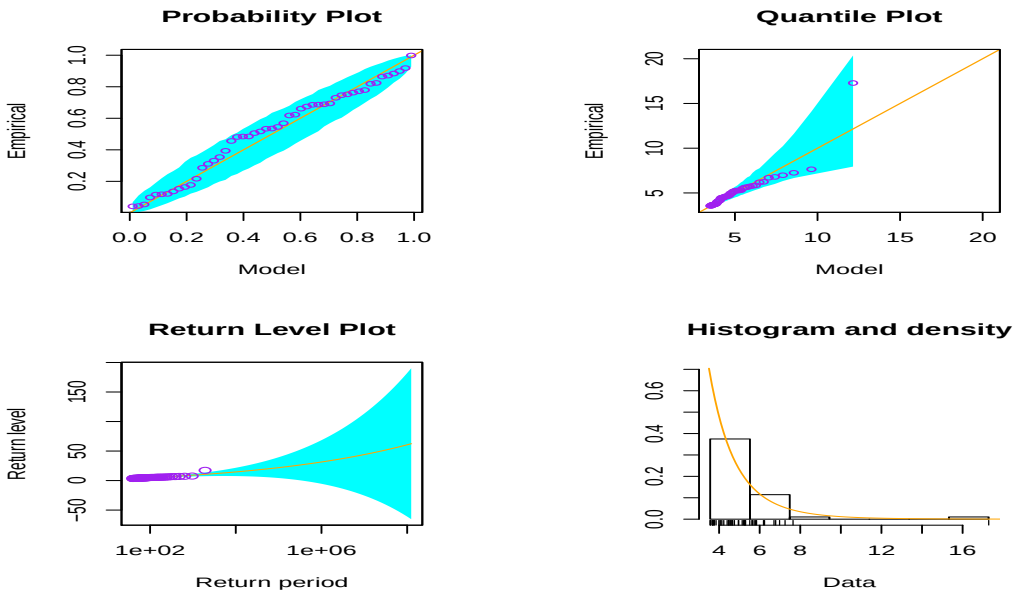

Figure 4.29: EGX 30: GPD diagnostic plots.

Table 4.17: EGX 30: Goodness of fit test.

Test	GPD values	Test values	Statistic values	P-value
Cramér-von Mises	σ: 1.4188 ξ: 0.1191	σ: 1.4191 ξ: 0.1305	0.0423	0.6682
Anderson-Darling	σ: 1.8914 ξ: -0.5014	σ: 1.4191 ξ: 0.1305	0.3400	0.6156

$H_\circ$: GPD fits the cluster-maxima well

H_1 : GPD does not fit it well

Table 4.17 displays the output of the test statistic and critical points (p-values) of the tested hypothesis of the Cramér-von Mises and Anderson-Darling tests. For both tests, it is observed that the null hypothesis is not rejected since the p-value is large (greater than 0.05). This confirms that the GPD model fitted to the generated cluster-maxima is sensible.

Table 4.18: EGX 30: Threshold excess return level (R_q) estimates.

	$u = 3.50$	2-year	5-year	10-year	20-year
R_q estimates		17.71	20.73	23.23	25.96
90% CI		(16.57; 18.85)	(18.67; 22.78)	(20.14; 26.32)	(21.49; 30.42)
95% CI		(16.36; 19.07)	(18.28; 23.18)	(19.55; 26.92)	(20.64; 31.28)
99% CI		(15.93; 19.50)	(17.51; 23.94)	(18.40; 28.07)	(18.97; 32.95)

4.10.11 EGX 30: Return levels

Table 4.18 shows the return levels estimates and their corresponding confidence intervals (in parenthesis) for the EGX 30 index (Egyptian market).

For instance, for the 10-year return period, the estimation indicates that on average, an extreme movement of level (magnitude) 23.23 is expected to be exceeded once over the next 10 years. The wide confidence intervals which are obtained for extreme return levels for long return periods reflect the fact that little information is available for future return levels prediction with any amount of certainty (Coles, 2001).

4.10.12 Univariate analysis: Point process

The univariate analysis using the point process has the same procedure as that of the bivariate-threshold-excess-model analysis. The same threshold is used to enable appropriate comparison between these two models. As a way forward, the already declustered exceedances (i.e. the cluster-maxima) obtained using the univariate analysis of the bivariate-threshold-excess model will be applied for the parameter estimation of the point process (see Table 4.19).

Table 4.19: EGX 30: Univariate point process parameter estimates.

	CI	u	μ	σ	ξ
EGX 30	90%	3.50	15.64 (4.18)	2.86 (1.85)	0.12(0.12)
			[8.76; 22.52]	[-0.18; 5.90]	[-0.08; 0.32]
	95%	3.50	15.64 (4.18)	2.86 (1.85)	0.12(0.12)
			[7.44; 23.83]	[-0.76; 6.49]	[-0.12; 0.36]
	99%	3.50	15.64 (4.18)	2.86 (1.85)	0.12(0.12)
			[4.87; 26.41]	[-1.90; 7.62]	[-0.20; 0.44]

4.10.13 Parameters estimation: Point process fit to cluster-maxima

The generated 49 cluster-maxima are the risk of various magnitudes in the Egyptian equity market at threshold $u = 3.5$. The pattern of the point process estimated output in Table 4.19 is similar to the GPD univariate estimation.

The positive estimate of the shape parameter is a reflection of concavity (Coles, 2001) and it indicates a fat-tailed distribution. That is, the shape parameter (ξ) estimate corresponds to an unbounded distribution, hence the risk in the Egyptian EGX 30 market can be well described by a Fréchet-Pareto class of distributions (Beirlant et. al, 2004). The evidence for this is particularly strong at the 90% and 95% confidence intervals where the intervals are almost entirely in the positive domain. The evidence is not too strong at the 99% confidence interval since this extends well below zero.

4.10.14 Point process diagnostic plots

The probability plots and quantile plots for the point process fit show that the points are close to the diagonal straight lines (Figure 4.30). The only clear exception is the single outlier displayed in the quantile plot. Hence, the fit is reasonably valid.

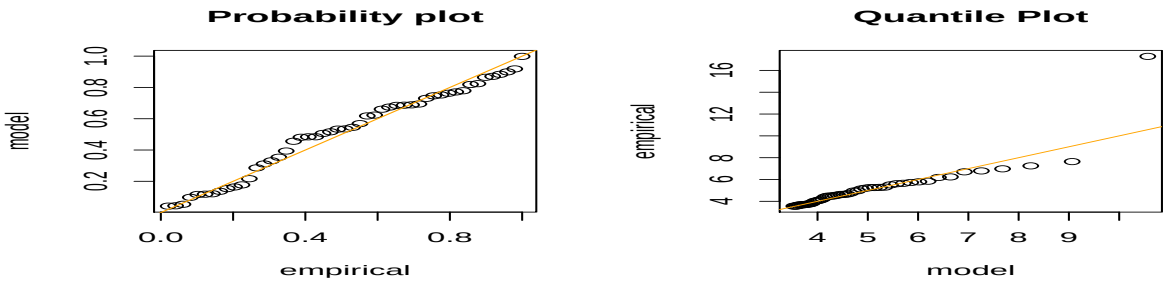

Figure 4.30: EGX 30: Point process diagnostic plots.

Table 4.20: EGX 30: Point process return level (R_q) estimates.

	$u = 3.50$	2-year	5-year	10-year	20-year
R_q estimates		16.72	20.35	23.03	25.85
95% CI		(7.11; 26.32)	(5.41; 35.28)	(3.54; 42.53)	(1.07; 50.62)

4.10.15 EGX 30: Return levels

Table 4.20 shows the point process return levels estimates and their corresponding confidence intervals (in parentheses) for the EGX 30 index (Egyptian market). For instance, for the 10-year return period, the estimation indicates that on average, a rare event of level (magnitude) 23.03 is expected to be exceeded once over the next 10 years. The wide large confidence intervals which are obtained for extreme return levels as displayed in the table and in Figure 5.19 of the Appendix are an indication of a drought of information for future return levels forecast with any degree of certainty (Coles, 2001) in the EGX 30 index.

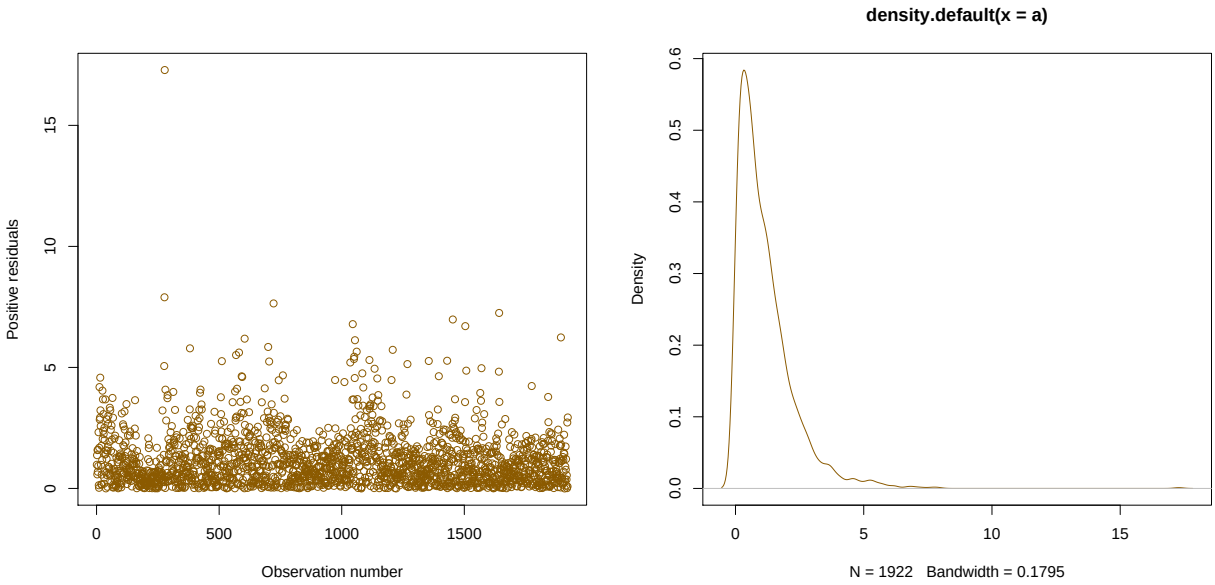

Figure 4.31: NSE 20: Positive residuals.

4.11 The Kenyan market: NSE 20

The Kenyan market index is called the Nairobi Stock Exchange, abbreviated - NSE 20, and it contains 3910 observations for the periods covered in this study.

4.11.1 NSE 20: Positive residual observations

With attention channelled to only the positive residuals of the data, for the risk and extremal dependence modelling of this market, there are 1944 positive residuals out of the entire 3910 observations. Figure 4.31 displays the positive residuals represented by the points and their corresponding density.

4.11.2 NSE 20: Threshold selection

As with the earlier three markets, the Kernel-GPD approach is used for the threshold selection, i.e. Gaussian-kernel with GPD first, followed by Weibull-kernel with GPD

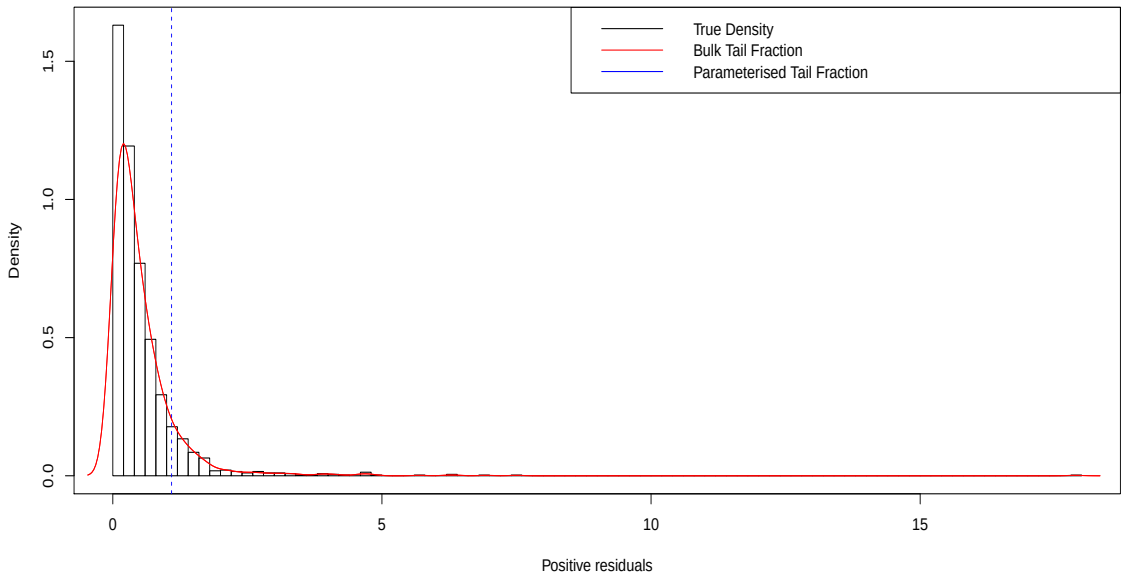

Figure 4.32: NSE 20: Threshold selection with Weibull kernel.

for comparison purpose. Both methods are applied under the parameterised tail fraction approach with GPD for the tail modelling. Figures 4.32 displays the output of the fit where $u = 1.0919$ is obtained when Weibull kernel is used with GPD. For Gaussian kernel with GPD, $u = 0.8121$ is obtained but the plot is not shown for brevity and space constraint.

4.11.3 Diagnostic plots of the kernel densities

The diagnostic plot of the Weibull Kernel-GPD model in Figure 4.33 indicates that the fit is adequate. The points stick close to the straight lines and are all enclosed with the dashed curves confidence interval, hence each set of plotted points is close to linear. Furthermore, the density estimate looks consistent with the histogram.

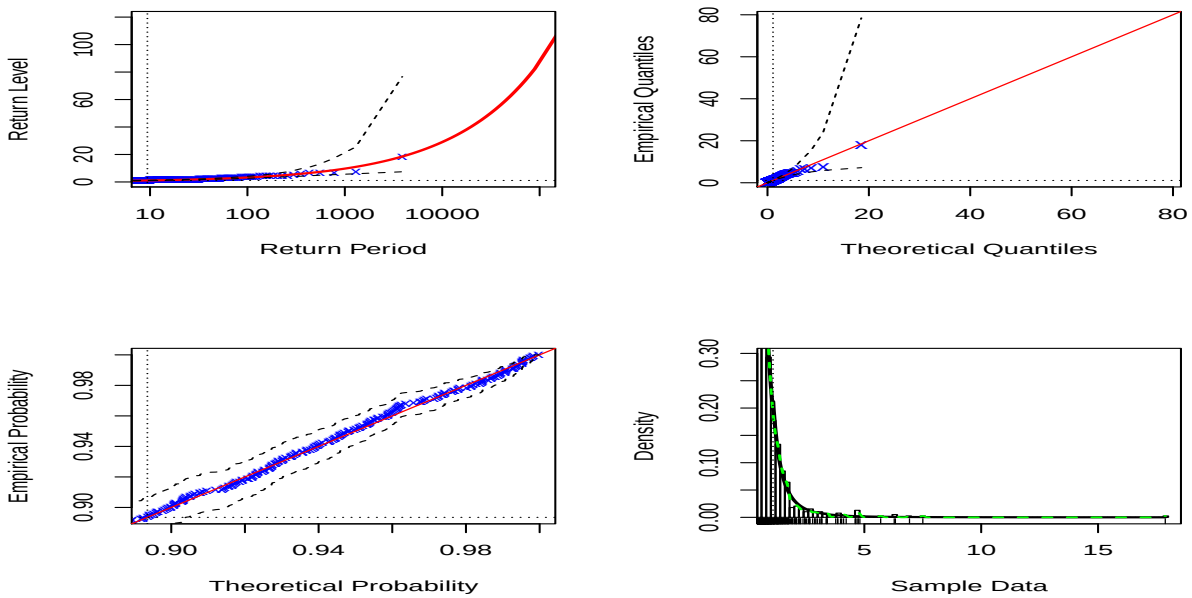

Figure 4.33: NSE 20 diagnostic plots: Threshold selection with Weibull-GPD model.

4.11.4 Conditions for final threshold choice

In line with the other three indices, a minimum extremal index, $\theta = 0.5$ is used to determine the required threshold (Ferro, 2003). Furthermore, normalised "inter-exceedance times" against "standard exponential quantiles" plots as in Figure 4.34 are required to be piecewise-linear with a breakpoint at the $(1-\theta)-$quantile, $-\log$. The vertical line represents the $(1-\theta)-$quantile (Ferro, 2003).

4.11.5 Sensitivity analysis

The sensitivity analysis plots in Figure 4.34 indicate that the points stick closest to the straight line for the plot at threshold $u = 1.09$. The linearity of the points are most observed at this threshold ($u = 1.09$) than the others, hence this is taken as the suitable threshold for this market, with extremal index $\theta = 0.53$.

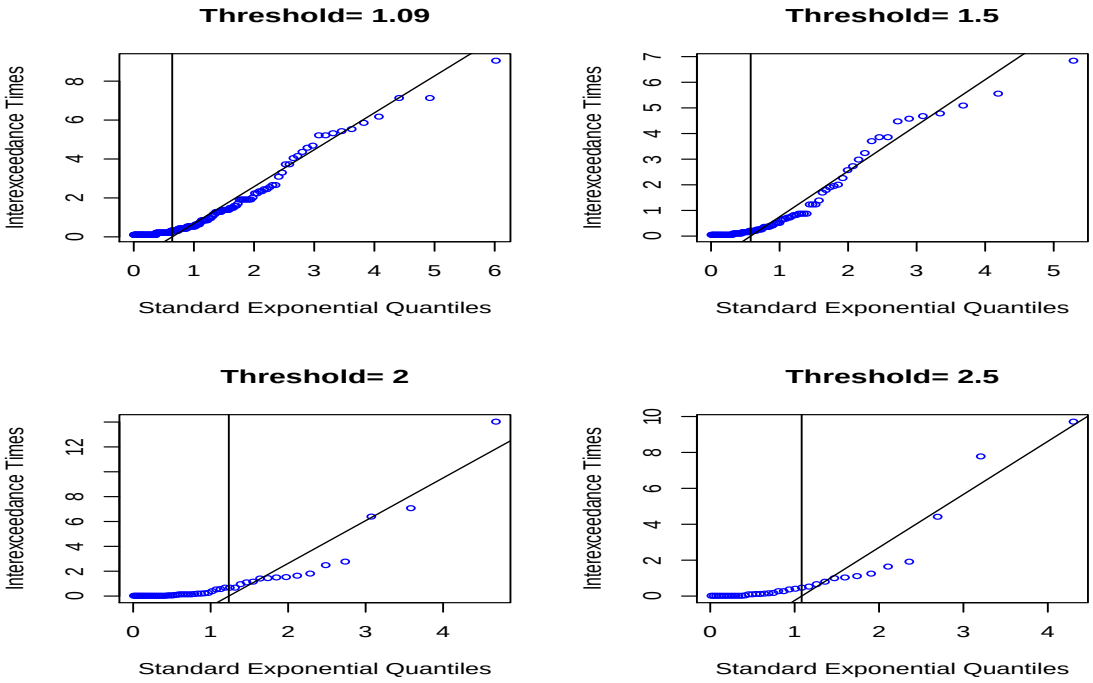

Figure 4.34: NSE 20: Threshold selection sensitivity analysis plots.

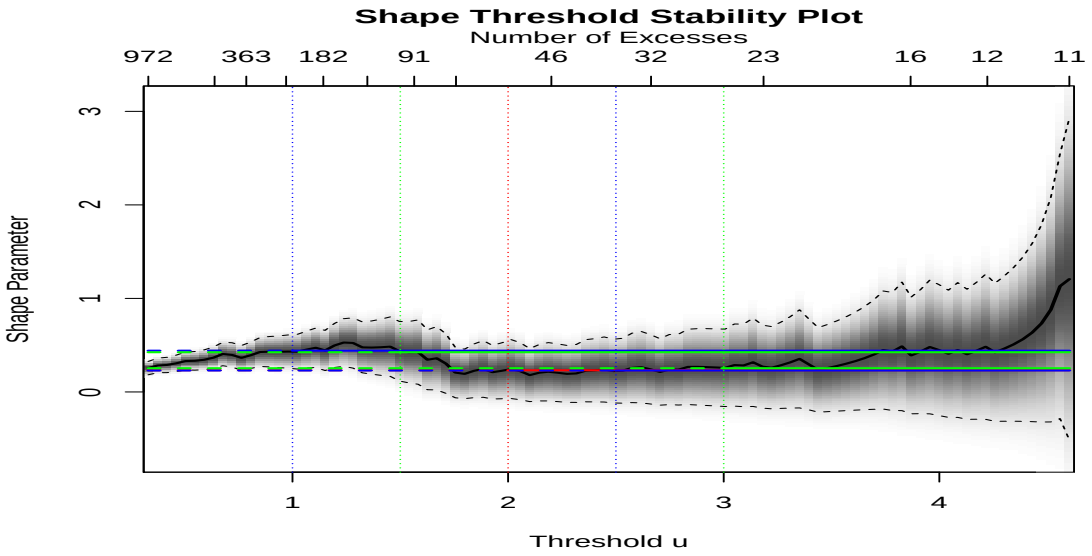

Figure 4.35: NSE 20: Shape threshold stability plot.

4.11.6 Shape threshold stability plot

As an added support to the choice of the Weibull Kernel-GPD threshold, the shape threshold stability plot in Figure 4.35 shows that linearity in the plot is rational to about $u = 1.1$, beyond which departure from linearity is obvious. Hence, the choice of u=1.09 of the mixture models is reasonable.

4.11.7 Declustering

The selected threshold $u = 1.09$, generates 207 threshold exceedances. After declustering at this threshold, 99 cluster-maxima are obtained as displayed in the plots in Figure 4.36. The cluster-maxima is the maxima of the clusters of exceedances over a high threshold (u=1.09 in this case). These are "exceedance residuals" (McNeil and Frey, 2000) for the GPD fit.

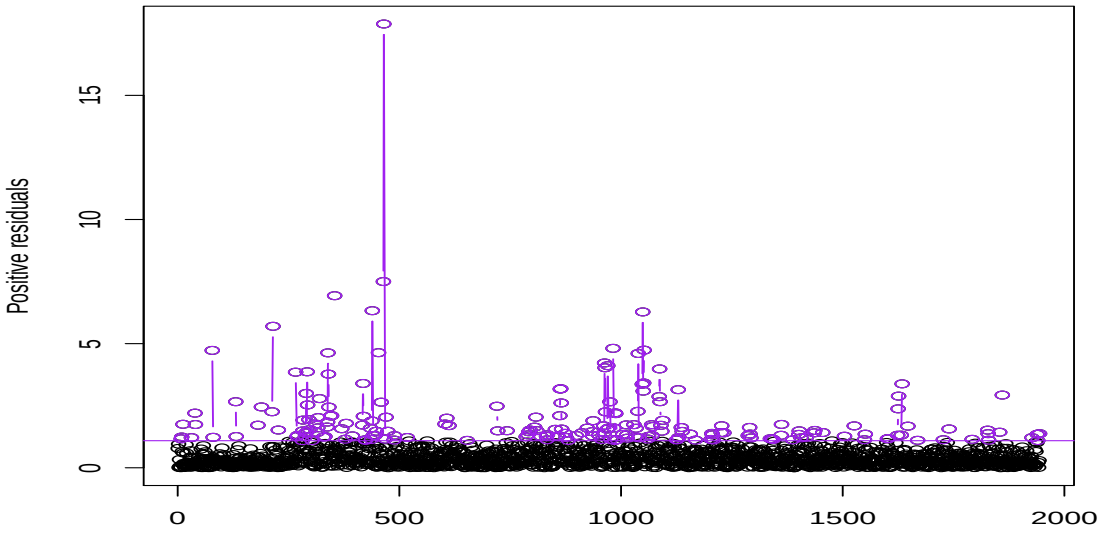

Figure 4.36: NSE 20: Declustered exceedances (cluster-maxima).

4.11.8 Parameters estimation: GPD fit to cluster-maxima

Following the generation of the 99 cluster-maxima (of different magnitudes), which are collectively classified as the risk to be modelled in the NSE 20 market. They are the extreme observations exceeding the threshold $u = 1.09$. This risk is modelled by fitting the GPD on the 99 observations and the parameter estimates in Table 4.21 are obtained.

Table 4.21: NSE 20: Univariate threshold excess parameter estimates.

	CI	u	m	k	$\eta_u = \frac{k}{m}$	σ	ξ
NSE 20	90%	1.09	1944	99	0.0509	0.60 (0.18) [0.30; 0.90]	0.54 (0.16) [0.27; 0.80]
	95%	1.09	1944	99	0.0509	0.60 (0.18) [0.25; 0.96]	0.54 (0.16) [0.22; 0.85]
	99%	1.09	1944	99	0.0509	0.60 (0.18) [0.13; 1.07]	0.54 (0.16) [0.13; 0.95]

The positive estimate of the shape parameter is a reflection of concavity (Coles, 2001) and it indicates a fat-tailed distribution. This positive shape parameter estimate corresponds to an unbounded distribution, hence the risk in the Kenyan NSE 20 market can be well described by the Pareto distribution. The evidence for this is also reasonably strong at the three confidence intervals of 90%, 95% and 99% for ξ where the intervals are all entirely in the positive domain.

The profile log-likelihood plot for ξ for the NSE 20 index is displayed in the "bottom right panel" of Figure 5.18 in the Appendix. The graph indicates an approximate profile log-likelihood for the 95% confidence interval as [0.28; 0.93].

4.11.9 Diagnostics: Model checking

The validity of the fitted GPD model is judged by the diagnostic plots in Figure 4.37. The points on the quantile plot stick closely to the straight line and those of the probability plot are confined within the tolerance interval. The return level plot displays stable points that are well aligned on the straight line. It however shows very large uncertainties that enlarge at extrapolation of the model to higher levels. The density plot is satisfactory and consistent with the histogram, hence the fit is adequate.

4.11.10 Goodness of fit of the GPD

As with the other three markets, this study applies the goodness of fit tests of Cramér-von Mises and Anderson-Darling to determine how well the generalized Pareto distribution fits the cluster-maxima observations in this market. The tested hypothesis are:

$H_\circ$: GPD fits the cluster-maxima well

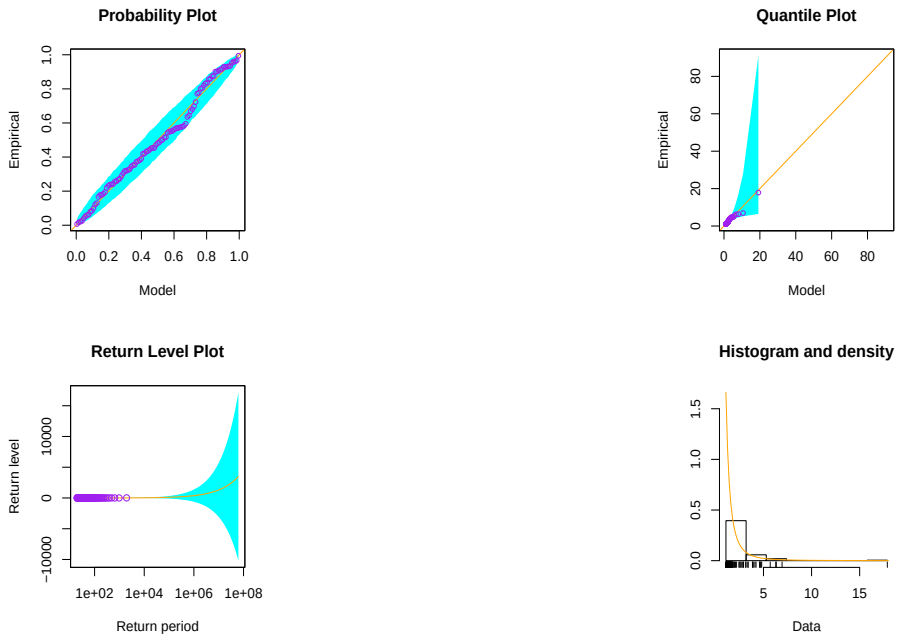

Figure 4.37: NSE 20: GPD diagnostic plots.

Table 4.22: NSE 20: Goodness of fit test.

Test	GPD values	Test values	Statistic values	P-value
Cramér-von Mises	σ: 0.6021 ξ: 0.4866	σ: 0.5973 ξ: 0.5615	0.0414	0.6161
Anderson-Darling	σ: 0.7238 ξ: 0.4866	σ: 0.5973 ξ: 0.5615	0.3356	0.5648

H_1 : GPD does not fit it well

Table 4.22 displays the result of the test statistic and critical points (p-values) of the tested hypothesis of the Cramér-von Mises and Anderson-Darling tests. For both tests, it is observed that the null hypothesis is not rejected since the p-value is large (greater than 0.05). This implies that the GPD fits the generated cluster-maxima observations well.

Table 4.23: NSE 20: Threshold excess return level (R_q) estimates.

	$u = 1.09$	2-year	5-year	10-year	20-year
R_q estimates		38.74	63.41	92.06	133.64
90% CI		(36.11; 41.36)	(57.59; 69.24)	(81.54; 102.58)	(115.04; 152.24)
95% CI		(35.61; 41.87)	(56.47; 70.36)	(79.53; 104.59)	(111.47; 155.80)
99% CI		(34.63; 42.85)	(54.29; 72.54)	(75.59; 108.53)	(104.52; 162.76)

4.11.11 NSE 20: Return levels

Table 4.23 shows the return levels estimates and their corresponding confidence intervals (in parenthesis) for the Kenyan NSE 20 market index. For the 10-year return period for instance, the estimation indicates that on average, an extreme event of level (magnitude) 92.06 is expected to be surpassed once over the next 10 years. The large confidence intervals which are obtained for extreme return levels for long return periods indicate little availability of information for future return levels prediction with any degree of confidence (Coles, 2001).

4.11.12 Univariate analysis: Point process

The univariate analysis using the point process has the same procedure as that of the bivariate-threshold-excess model analysis. The same threshold is used to enable appropriate comparison between these two models. As a way forward, the already declustered exceedances (i.e. the cluster-maxima) obtained using the univariate analysis of the bivariate-threshold-excess model will be applied for the parameter estimation of the point process (see Table 4.24).

Table 4.24: NSE 20: Univariate point process parameter estimates.

	CI	u	μ	σ	ξ
NSE 20	90%	1.09	26.66 (5.75)	14.35 (4.11)	0.54(0.05)
			[17.20; 36.13]	[7.59; 21.10]	[0.46; 0.62]
	95%	1.09	26.66 (5.75)	14.35 (4.11)	0.54(0.05))
			[15.39; 37.94]	[6.30; 22.39]	[0.44; 0.63]
	99%	1.09	26.66 (5.75)	14.35 (4.11)	0.54(0.05))
			[11.86; 41.47]	[3.78; 24.92]	[0.41; 0.66]

4.11.13 Parameters estimation: Point process fit to cluster-maxima

As with the other markets, the generated 99 cluster-maxima of different magnitudes are the risk to be modelled in the NSE 20 market. They are the extreme observations exceeding the threshold $u = 1.09$. This risk is modelled by fitting the point process on the 99 observations and the parameter estimates in Table 4.24 are obtained.

The positive estimate of the shape parameter is a reflection of concavity (Coles, 2001) and it indicates a fat-tailed distribution. This positive shape parameter estimate corresponds to an unbounded distribution, hence the risk in the Kenyan NSE 20 market can be well described by the Fréchet-Pareto class of distributions (Beirlant et. al, 2004). The evidence for this is also reasonably strong at the three confidence intervals of 90%, 95% and 99% for ξ where the intervals are wholly in the positive domain.

The shape parameter estimate of the point process is consistent with that of the GPD with the exception of the wider confidence intervals of the GPD obtained due to higher value of the standard error.

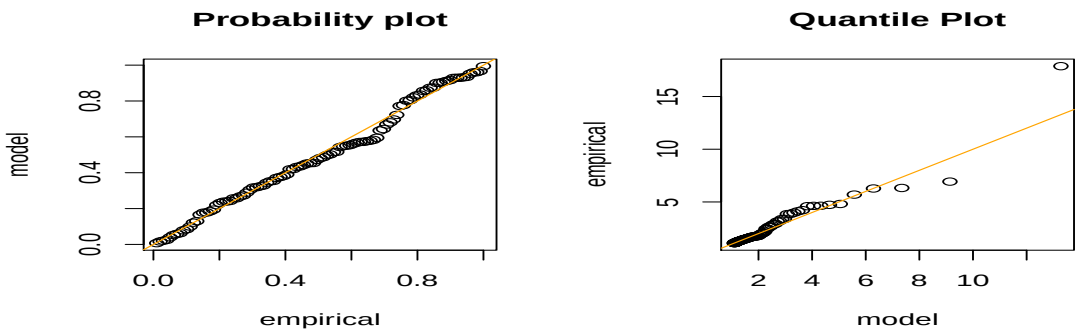

Figure 4.38: NSE 20: Point process diagnostic plots.

4.11.14 Point process diagnostic plots

The probability plots and quantile plots for the point process fit show that the points are close to the diagonal straight lines (Figure 4.38). The main exception are the two outliers displayed at the extreme of the quantile plot. Hence, the validity of the fit is reasonably adequate.

4.11.15 NSE 20: Return levels

Table 4.25 shows the point process return levels estimates and their corresponding confidence intervals (in parentheses) for the NSE 20 index (Kenyan market). The 10-year return period estimate as an example indicates that on average, a rare event of magnitude 87.13 is expected to be exceeded once over the next 10 years. The large confidence intervals which are obtained for extreme return levels as displayed in the table and in Figure 5.19 of the Appendix indicate information deficiency for future

Table 4.25: NSE 20: Point process return level (R_q) estimates.

	$u = 1.09$	2-year	5-year	10-year	20-year
R_q estimates		31.89	58.39	87.13	127.91
95% CI		(13.05; 50.73)	(15.89; 100.89)	(15.48; 158.78)	(10.88; 244.94)

return levels forecast with any degree of certainty (Coles, 2001) in the NSE 20 index.

4.12 GPD bootstrap test

The parametric bootstrap approach can be used to estimate uncertainty and possible bias in the GPD model's parameters. Knowledge of such uncertainty in model parameters is crucial for decision-making in financial data modelling. The bootstrap approach is a great tool especially when only a small data set is used to forecast the behaviour of processes or system. The conventional GPD parameterization are represented by its shape parameter ξ and scale parameter σ. However, it is noticed that when working with the reparameterization of the scale parameter where $\phi = \log \sigma$, the numerical algorithms for log-likelihood optimization tend to converge more consistently (Harry and Heffernan, 2013). Furthermore, the reparameterization also ensures that the scale parameter is always positive (Sigauke and Bere, 2017). Table 4.26 gives a summary of the bootstrap estimation in the markets.

The pattern of the parameter estimates' uncertainty in the markets indicates that the estimated bias and the standard deviations are relatively small in the Nigerian, Egyptian, and Kenyan markets. The uncertainty in the parameter estimate of the South African market is bigger than the rest largely due to the small sample size of 15 cluster-maxima data used for the GPD estimation. These deviations for ξ (denoted by ξ which is "xi") on the right and σ (denoted by ϕ which is "phi") on the left are clearly observed in the bootstrap plots in Figures 4.39, 4.40, 4.41, and 4.42.

Table 4.26: GPD bootstrap tests for the four markets.

	JSE-ALSI		NIGALSH		EGX 30		NSE 20	
	$\hat{\phi}$	$\hat{\xi}$	$\hat{\phi}$	$\hat{\xi}$	$\hat{\phi}$	$\hat{\xi}$	$\hat{\phi}$	$\hat{\xi}$
Original	0.637	-0.501	-0.323	0.487	0.350	0.119	-0.507	0.537
Bootstrap mean	0.851	-0.823	-0.291	0.448	0.373	0.078	-0.499	0.515
Bias	0.213	-0.321	0.032	-0.039	0.023	-0.041	0.008	-0.023
Standard deviation	0.385	0.344	0.235	0.201	0.230	0.181	0.183	0.159
Bootstrap median	0.955	-1.034	-0.282	0.455	0.373	0.072	-0.500	0.514

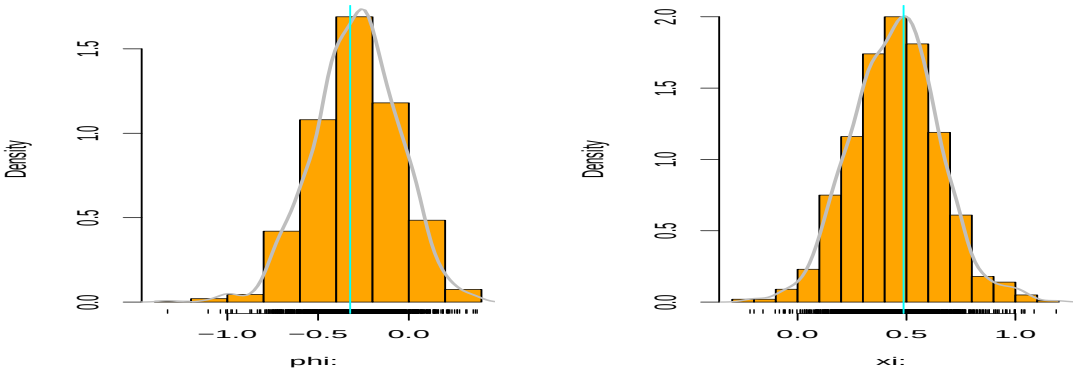

Figure 4.39: NIGALSH: Bootstrap plots.

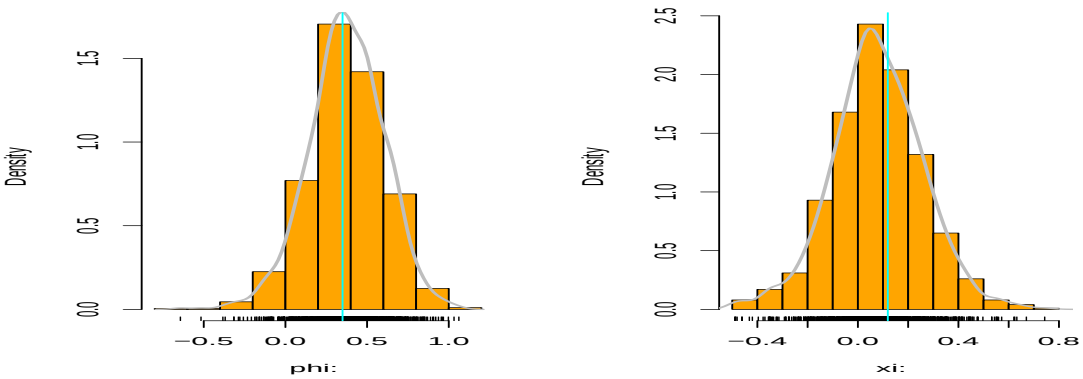

Figure 4.40: EGX 30: Bootstrap plots.

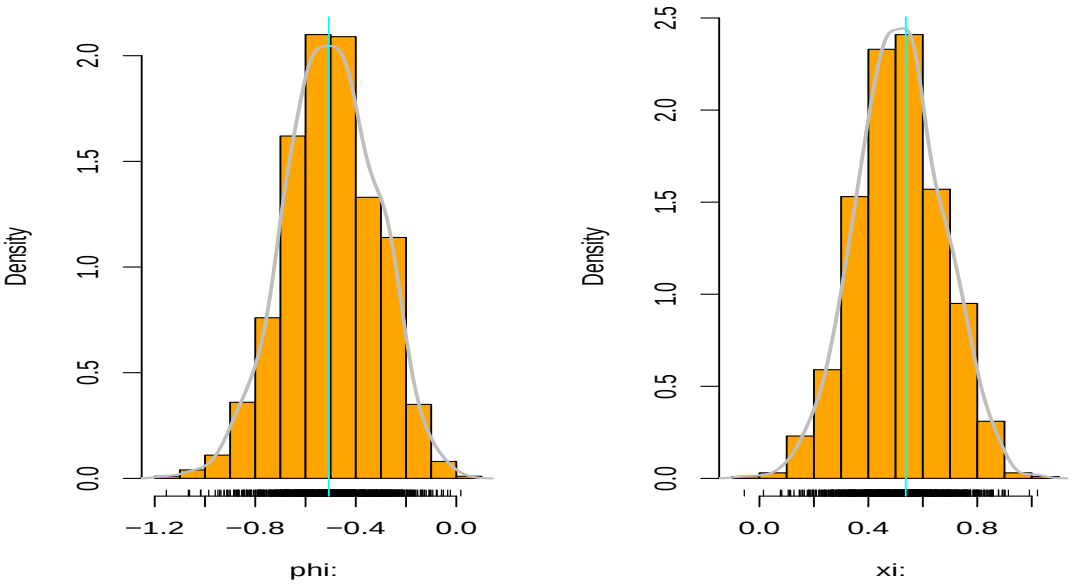

Figure 4.41: NSE 20: Bootstrap plots.

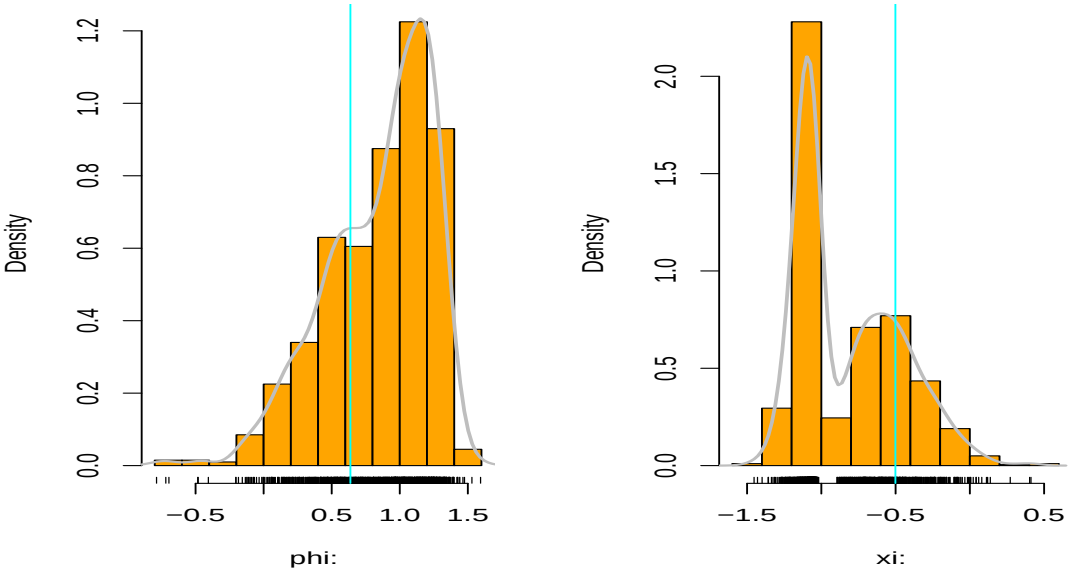

Figure 4.42: JSE-ALSI: Bootstrap plots.

4.13 Risk hierarchy: comparing the markets' risks

The level of risk in the markets can be compared using the boxplot of the risk (the cluster-maxima's) aggregate of each market. The plot of the risk hierarchy is displayed in Figure 4.43, with the associated densities in Figure 4.44. This obviously shows that the Egyptian EGX 30 market is the most risk-prone, followed by the South African JSE-ALSI market, then the Nigerian NIGALSH market and the least risky is the Kenyan NSE 20 market. This conclusion does not correspond with the hierarchy pattern of the volatility of the markets, hence risk is not a brainchild of volatility in these markets. High investment risk may either yield potential high returns as a reward or a huge loss to an investor.

The boxplots further indicate the shapes of the risk (in each of the markets) based on the concentration of the cluster-maxima observations on the scale. From Figure 4.43, the distribution of the risk observations in the JSE-ALSI market is skewed to the left, the NIGALSH is skewed to the right, the EGX 30 is near-symmetric and the NSE 20 is skewed more to the right than the NIGALSH market.

4.14 Multivariate extreme value modelling

Having completed the univariate analysis of the risk and volatility of the markets, attention is now focused on the multivariate dependence of rare movements in the markets. The joint distribution of two or more random variables with dependence can be modelled by the multivariate extreme value theory (MEVT). In reality, multivariate extreme values modelling is more complicated than the univariate aspect (Southworth and Heffernan, 2016). The context of multivariate modelling is relative and could mean modelling multiple random variables at a single location, or a single variable at multiple locations or even multiple variables at various locations. The four

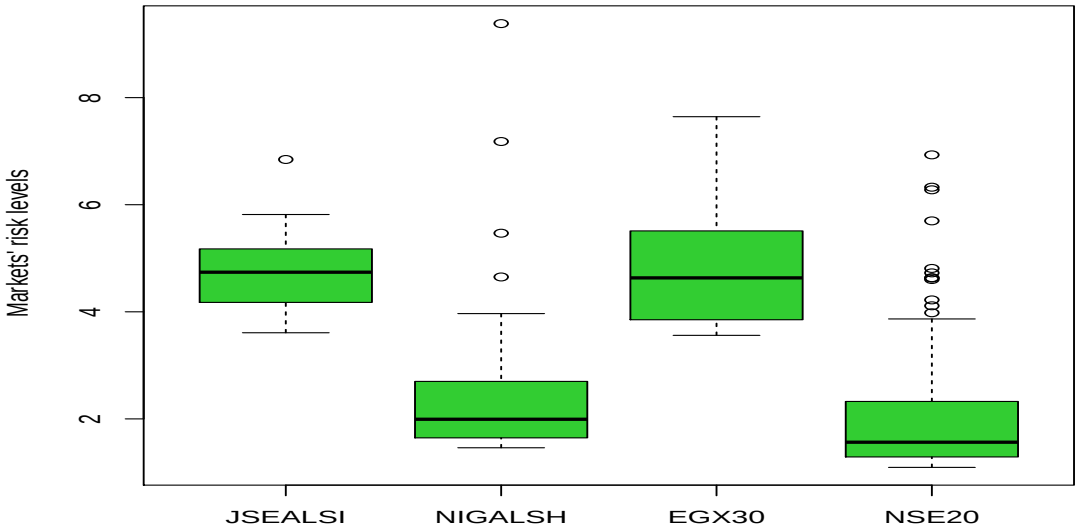

Figure 4.43: Risk hierarchy: comparing the markets' risks.

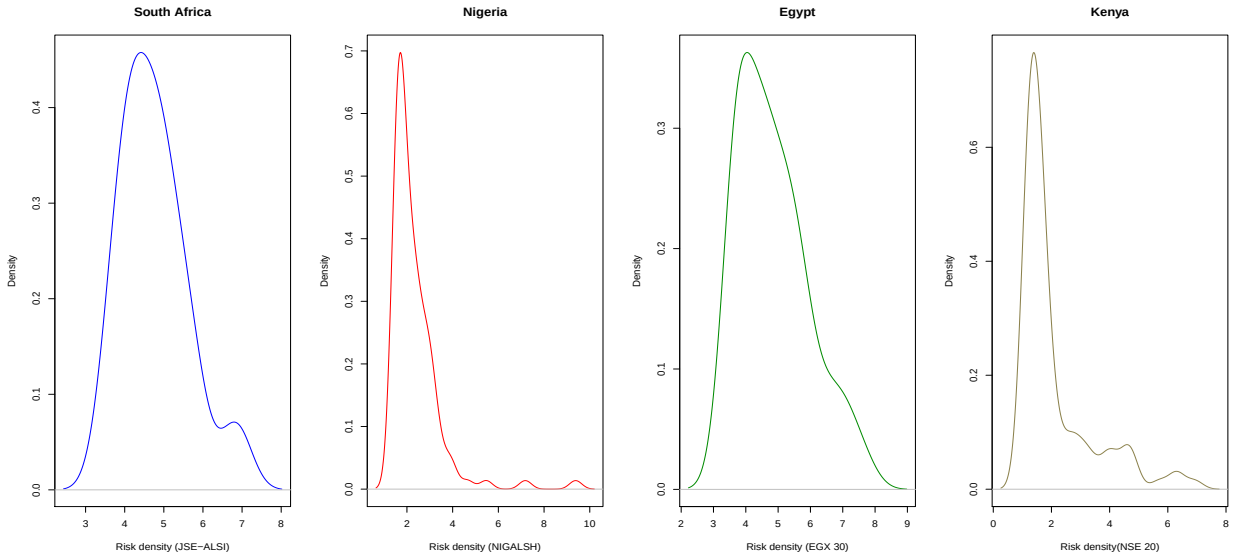

Figure 4.44: Risk densities: comparing the markets' risks.

dimensional financial data in this study will be modelled in a pairwise combination of the markets vis-a-vis the bivariate modelling. The bivariate-threshold-excess model and bivariate point process will be used for the required bivariate modelling.

Like the univariate case, the sequential procedures used for this modelling begins with the exploratory multivariate analysis using exploratory plots, then the pairwise extremal dependence of the variables are examined using χ and $\bar{\chi}$ defined by Coles, Heffernan and Tawn (1999). Following these, the conditional multivariate extreme value modelling is carried out after marginal transformation to standardised margins. Model diagnostics is done to check the model's validity and lastly, the fitted model's parameters are interpreted.

4.15 Multivariate exploratory data analysis

Here, exploratory plots are used to provide an insight into the pairwise dependence of the variables under examination using a pairwise scatterplot. However, it should be known that dependence between paired variables in the data body does not necessarily denote extremal dependence (Southworth and Heffernan, 2016). The exploratory plots of the four markets are first shown at a glance, after which the plots showing each of the market variables against the rest are displayed for better descriptions.

4.15.1 Exploratory plots: The four markets

The scatterplots of the pairwise data displayed in Figure 4.45 reflect lack of extremal dependence between the variables with most of the points clustering roughly around the (0, 0) coordinates. Furthermore, the correlation values are very low (see Figure 4.45) with the highest value of 0.167 shown between Egyptian EGX 30 and South African JSE-ALSI, indicating some levels of extremal dependence.

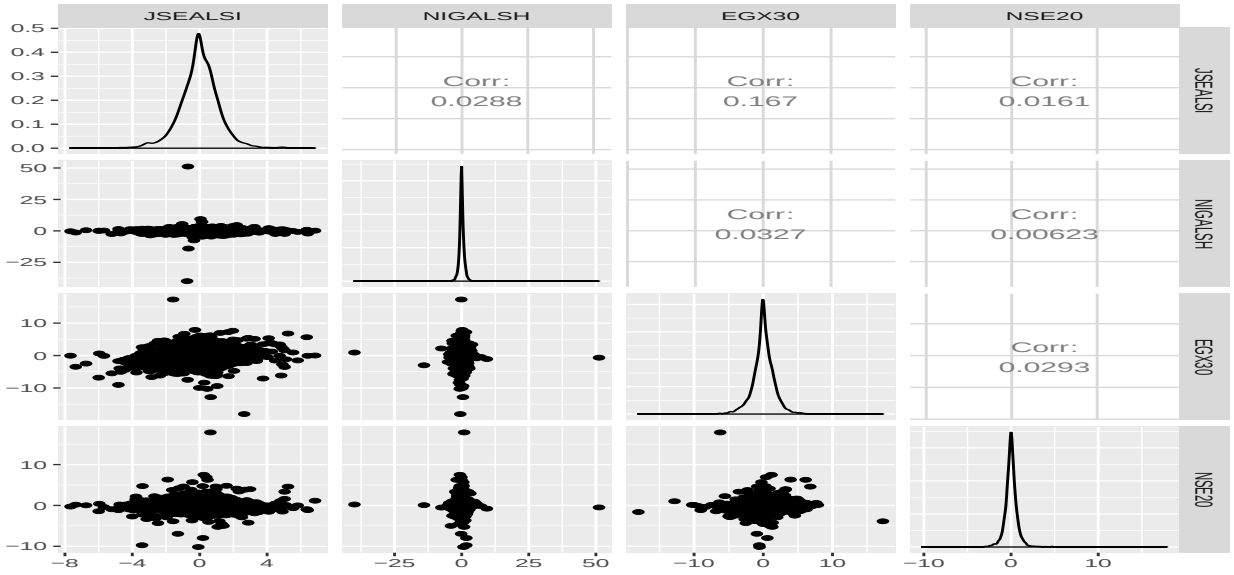

Figure 4.45: Scatterplot of the four indices data.

4.15.2 Exploratory plots: JSE-ALSI against the other markets

A better outlook to the scatterplots is shown when one of the markets is plotted against the others as in Figure 4.46. The two-dimensional historical data of the JSE-ALSI and EGX 30 markets appear to be most elliptically shaped with weak positive correlation compared to the other pairs. The scatterplot of JSE-ALSI and NSE 20 reveals a weaker positive correlation when compared to the JSE-ALSI and EGX 30, while that of JSE-ALSI and NIGALSH suggests near independence.

4.15.3 Exploratory plots: NIGALSH against the other markets

There is a similar pattern disclosed in the shapes of the relationship between the NIGALSH Index and the other three indices (Figure 4.47). The three scatterplots appear to be centered vertically roughly around (0, 0). This further suggests that

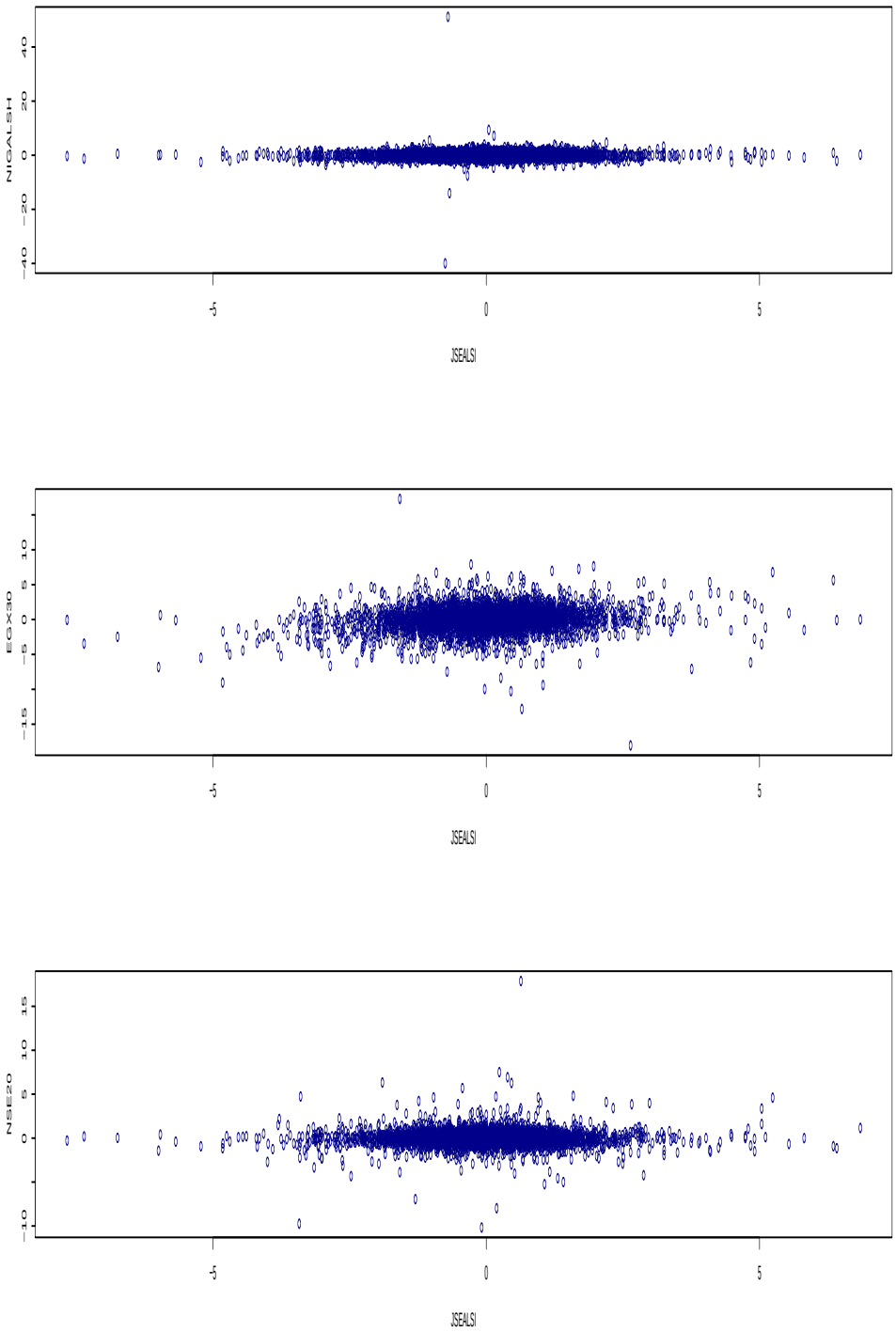

Figure 4.46: Scatterplots of JSE-ALSI against the other markets.

the NIGALSH index may have the same extremal dependence relationship with the remaining three markets.

4.15.4 Exploratory plots: EGX 30 against the other markets

Figure 4.48 shows the elliptical shape of the EGX 30 and JSE-ALSI markets, which reveals some weak positive correlation. The EGX 30 and NSE 20 also disclose a weak positive correlation with most of the points centered roughly around (0, 0), whereas the EGX 30 and NIGALSH markets' shape indicates near independence with points concentrated around (0, 0).

4.15.5 Exploratory plots: NSE 20 against the other markets

Figure 4.49 shows weak positive correlation between the NSE 20 and JSE-ALSI indices and NSE 20 and EGX 30 indices. The points are roughly clustered around (0, 0) in relation to the NSE 20 and NIGALSH indices, indicating near independence.

4.15.6 Exploratory plots of pairwise extremal dependence

To examine extreme association or pairwise extremal dependence between markets, this section is used to plot summary statistics of χ (Chi) and $\bar{\chi}$ (Chi-bar) plots interpreted as follows:

- (a) As the quantile u approaches 1, the limit value of $\bar{\chi}(u)$ shows diagnostics on whether the data displays asymptotic dependence. A $\bar{\chi}(u) = 1$ indicates asymptotic dependence (Southworth and Heffernan, 2016).
- (b) The plot of $\chi(u)$ is used to measure the strength of dependence in the asymptotic dependence area if the limit in (a) is equal to 1. The strength indicates stronger dependence if the value is closer to 1.

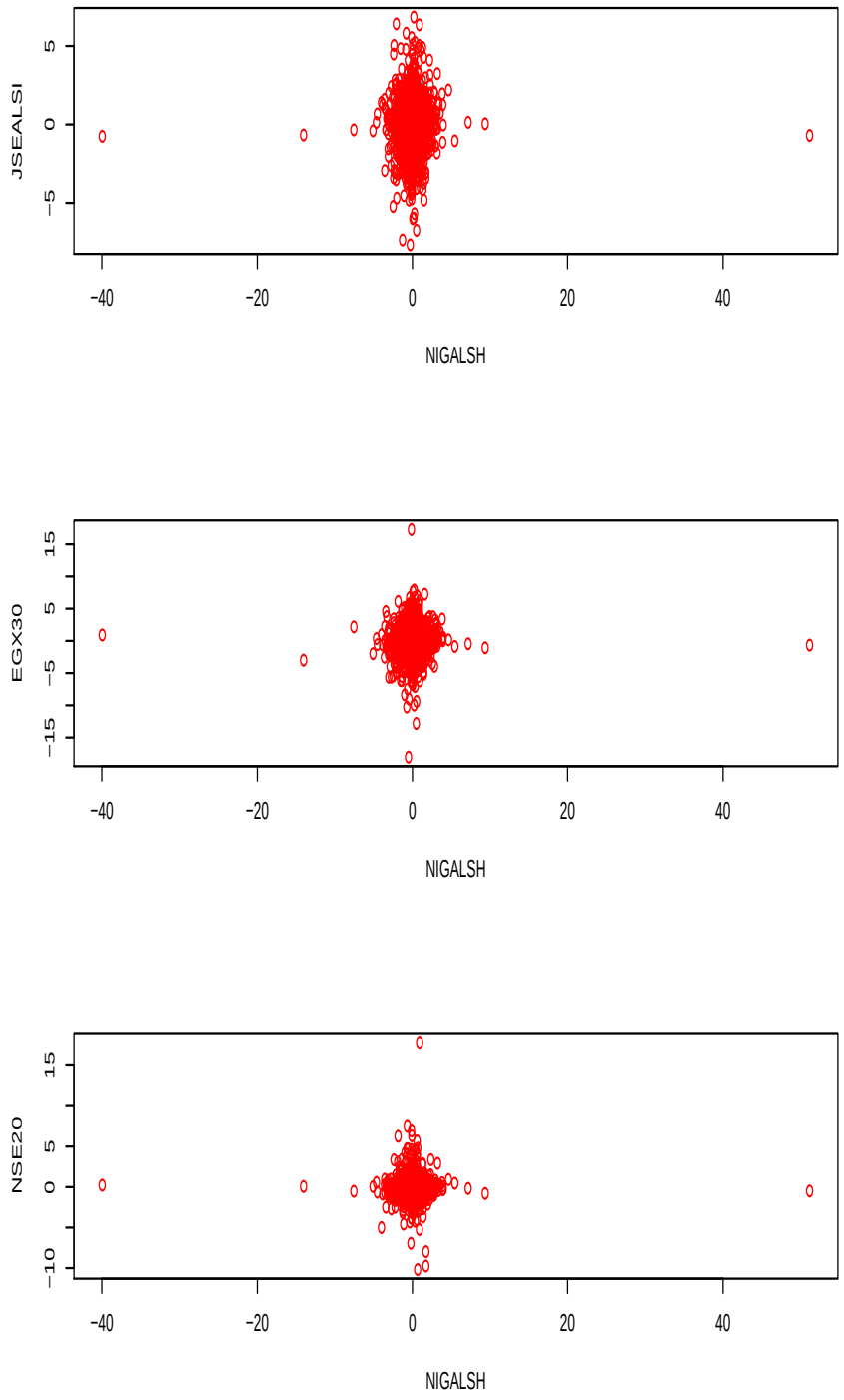

Figure 4.47: Scatterplots of NIGALSH against the other markets.

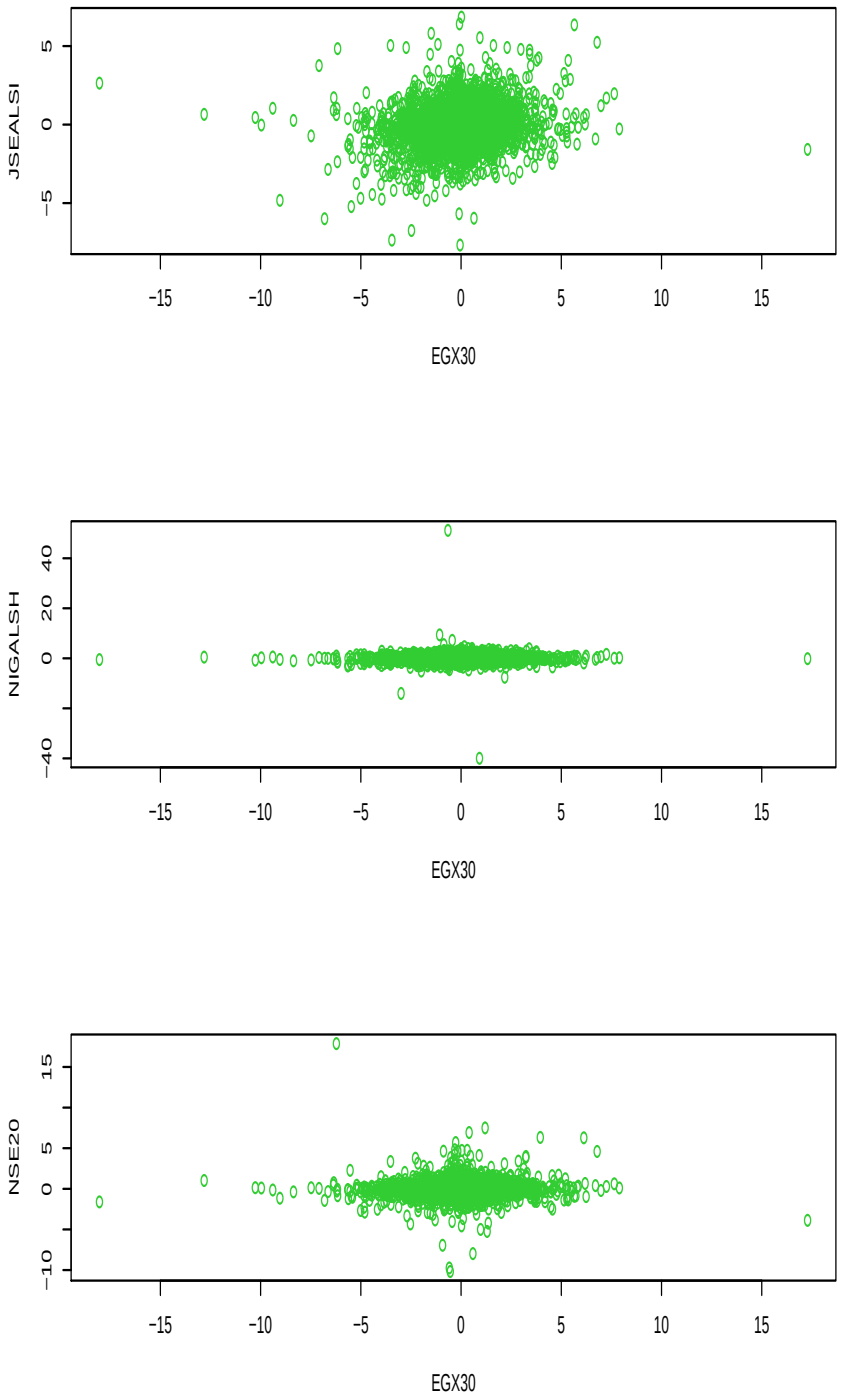

Figure 4.48: Scatterplots of EGX 30 against the other markets.

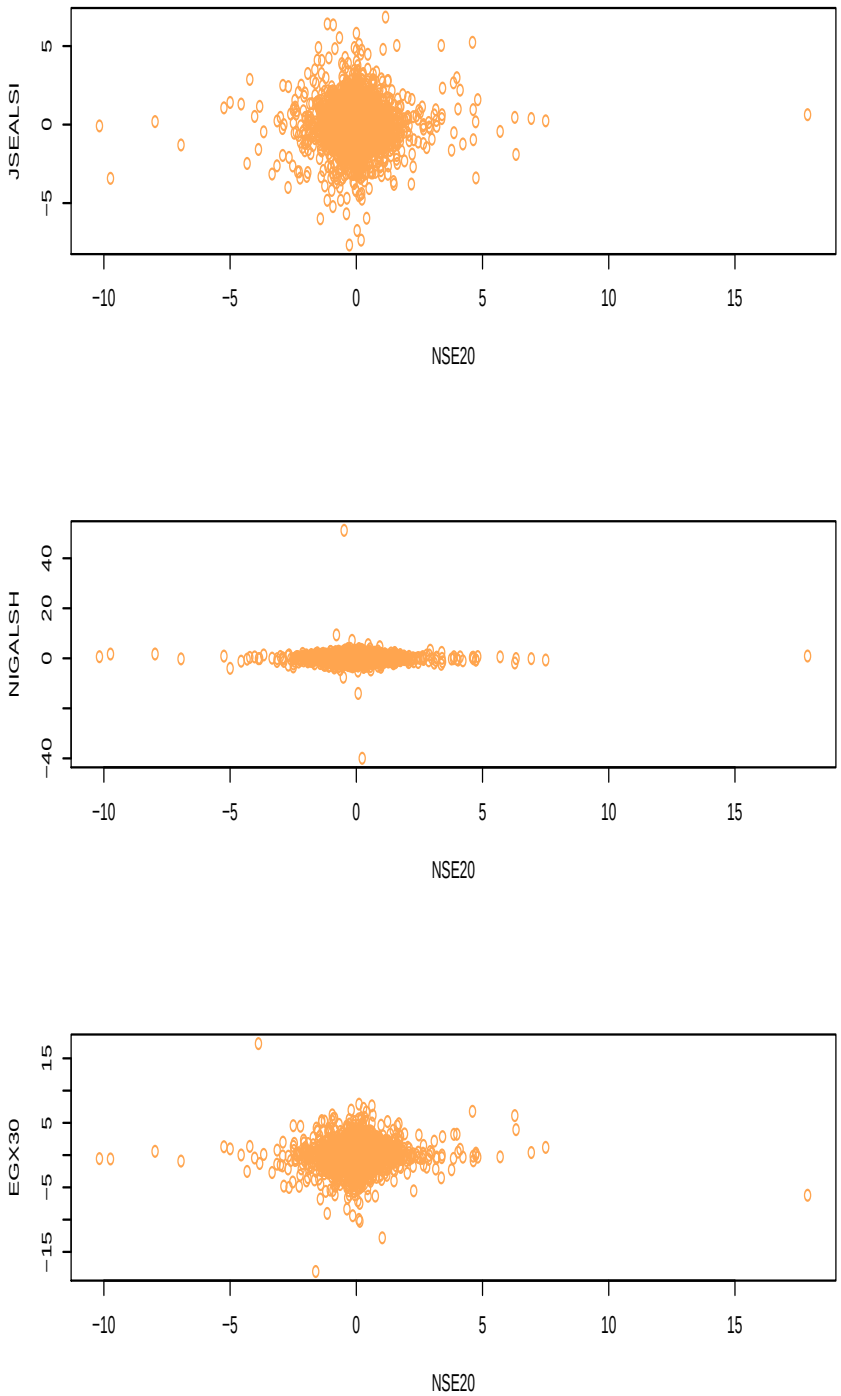

Figure 4.49: Scatterplots of NSE 20 against the other markets.

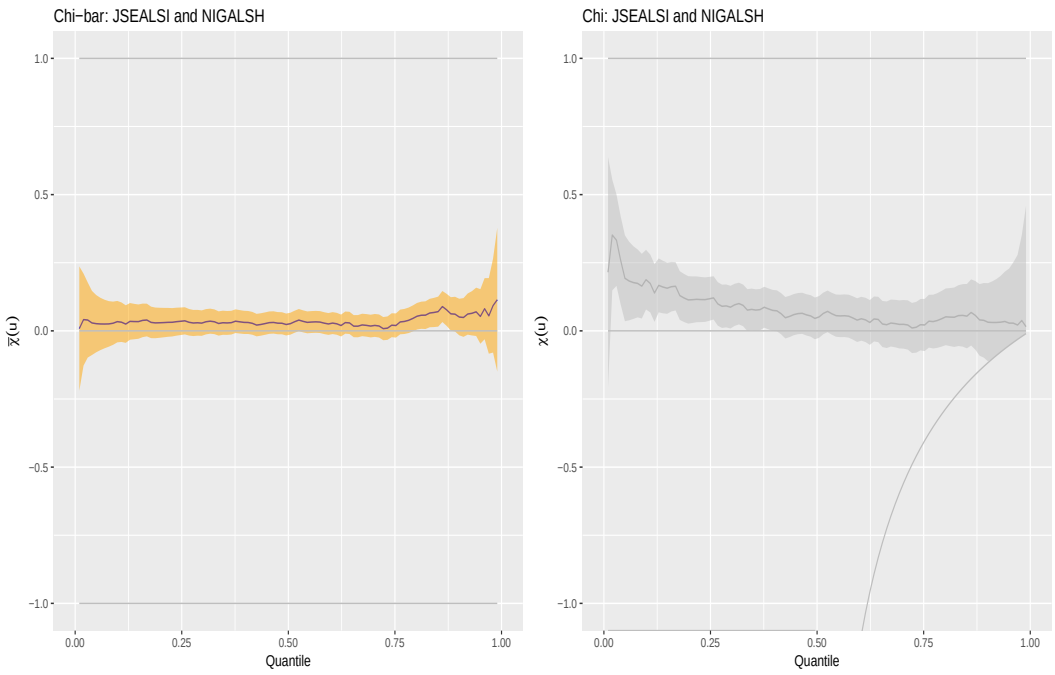

Figure 4.50: Pairwise extremal dependence exploratory plots: JSE-ALSI and NIGALSH indices.

- (c) The plot of $\bar{\chi}(u)$ measures the strength of dependence in the asymptotic dependence area when $\bar{\chi}(u) < 1$. As $u \to 1$, the limiting value indicates the dependence strength, where positive and negative values closer to 1 show stronger positive dependence and stronger negative dependence respectively. Asymptotic near independence is suggested when the values are close to 0 (Southworth and Heffernan, 2016).

Another approach used to check the pairwise extremal dependence is through the multivariate conditional Spearman's (MCS) correlation coefficient of Schmidt and Schmitt, (2007). Both methods shall be used under each market for better clarity.

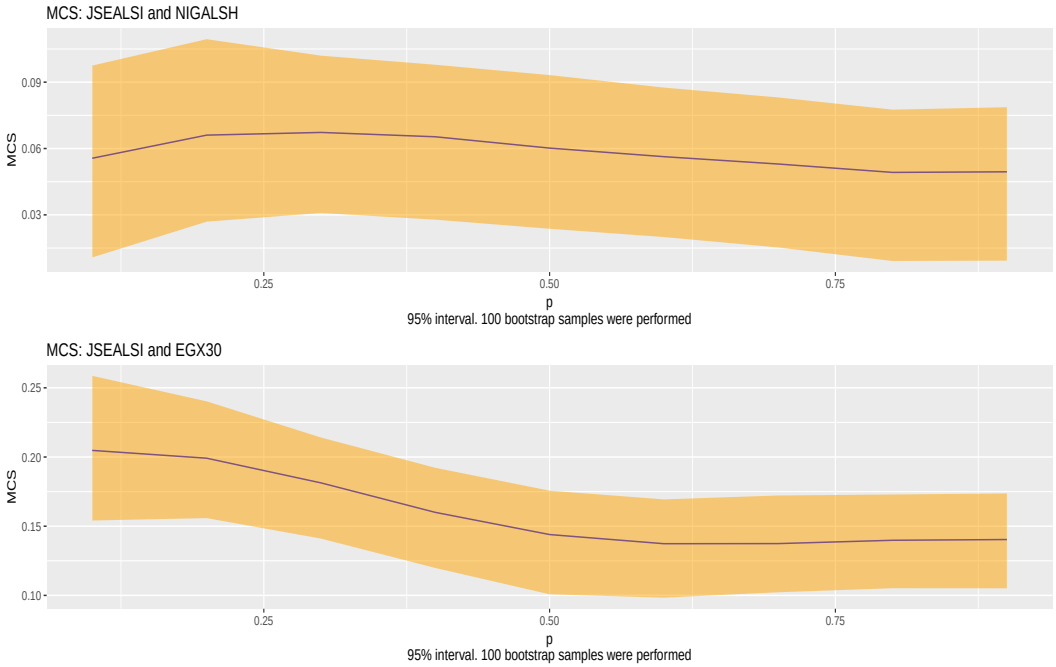

Figure 4.51: MCS plots: JSE-ALSI and NIGALSH indices (top panel) with JSE-ALSI and EGX 30 indices (bottom panel).

4.15.7 Pairwise extremal dependence exploratory plots: JSE-ALSI and NIGALSH indices

The exploratory plot of the $\bar{\chi}$ in Figure 4.50 shows that the South African JSE-ALSI and Nigerian NIGALSH markets are asymptotically dependent with χ indicating weak positive dependence in the class.

The MCS plot with the 95% confidence interval as shown in Figure 4.51 (top panel) indicates weak positive asymptotic dependence.

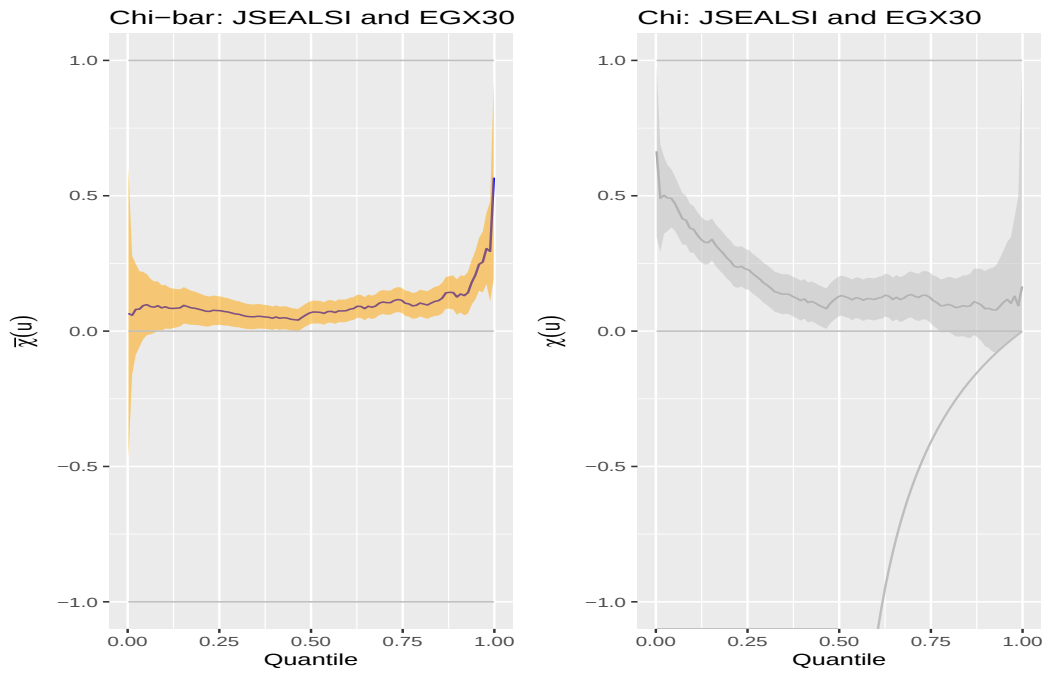

Figure 4.52: Pairwise extremal dependence exploratory plots: JSE-ALSI and EGX 30 indices.

4.15.8 Pairwise extremal dependence exploratory plots: JSE-ALSI and EGX 30 indices

With the value of $\bar{\chi}$ moving towards 1, Figure 4.52 shows that the South African JSE-ALSI and Egyptian EGX 30 markets are more asymptotically dependent than the pair of South African JSE-ALSI and Nigerian NIGALSH markets. However, the χ indicates the dependence as weak positive within the class.

The alternative approach using MCS plot with the 95% confidence interval as shown in Figure 4.51 (bottom panel) also indicates weak positive asymptotic dependence between the markets.

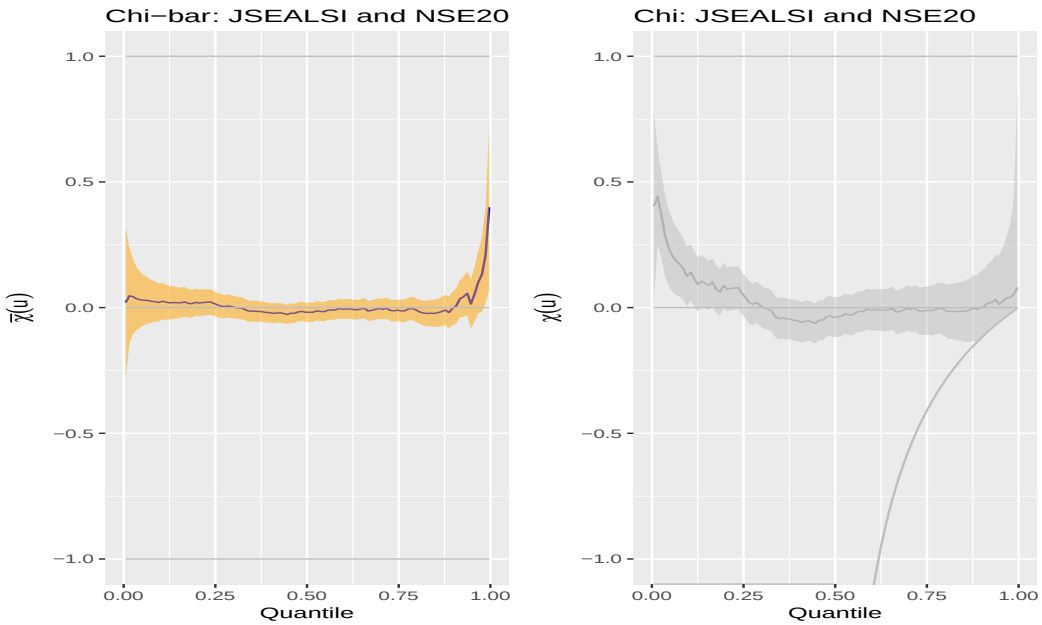

Figure 4.53: Pairwise extremal dependence exploratory plots: JSE-ALSI and NSE 20 indices.

4.15.9 Pairwise extremal dependence exploratory plots: JSE-ALSI and NSE 20 indices

Like the earlier two paired markets, the South African JSE-ALSI and Kenyan NSE 20 display some asymptotic dependence as shown by the $\bar{\chi}$ plot in Figure 4.53 with χ plot indicating weak positive dependence within the class.

Here, unlike the $\bar{\chi}$ and χ plots approach, the alternative MCS plot method with the 95% confidence interval shown in Figure 4.54 (top panel) suggests a slight deviation with weak negative extremal (asymptotic) dependence between the JSE-ALSI and NSE 20 markets for value close to -0.02.

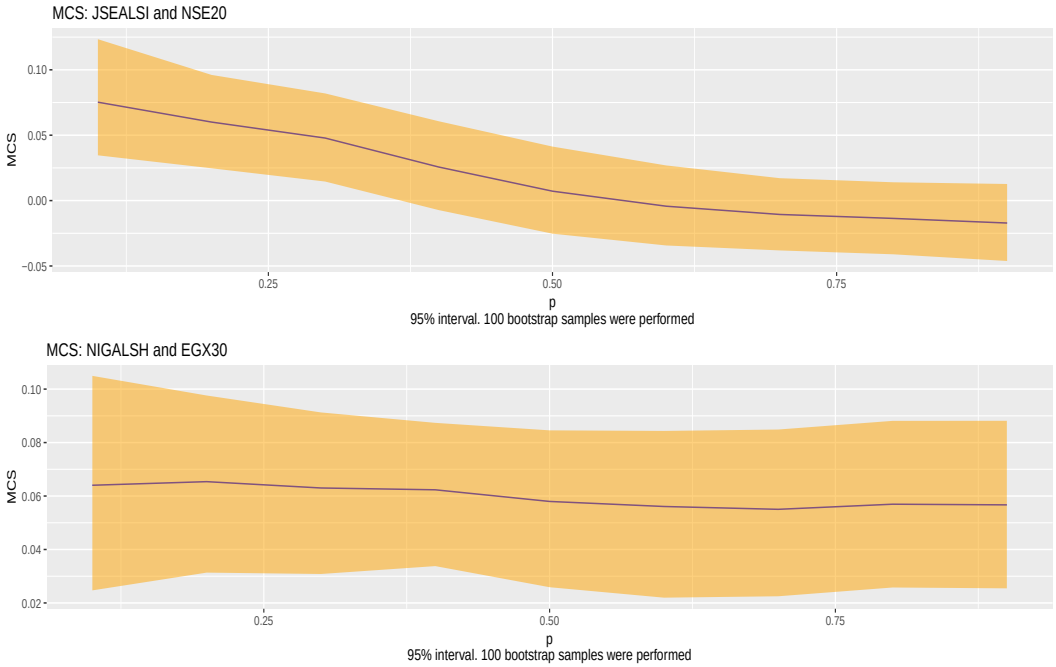

Figure 4.54: MCS plots: JSE-ALSI and NSE 20 indices (top panel) with NIGALSH and EGX 30 indices (bottom panel).

4.15.10 Pairwise extremal dependence exploratory plots: NIGALSH and EGX 30 indices

With the value of $\bar{\chi}$ close to 0 in the plot in Figure 4.55, the Nigerian NIGALSH and Egyptian EGX 30 markets display asymptotic near independence or extremely weak asymptotic dependence. The χ plot indicates very weak positive dependence within the class.

Like the $\bar{\chi}$ and χ plots approach, the MCS plot with the 95% confidence interval also suggests very weak positive asymptotic dependence between the markets (Figure 4.54 (bottom panel)).

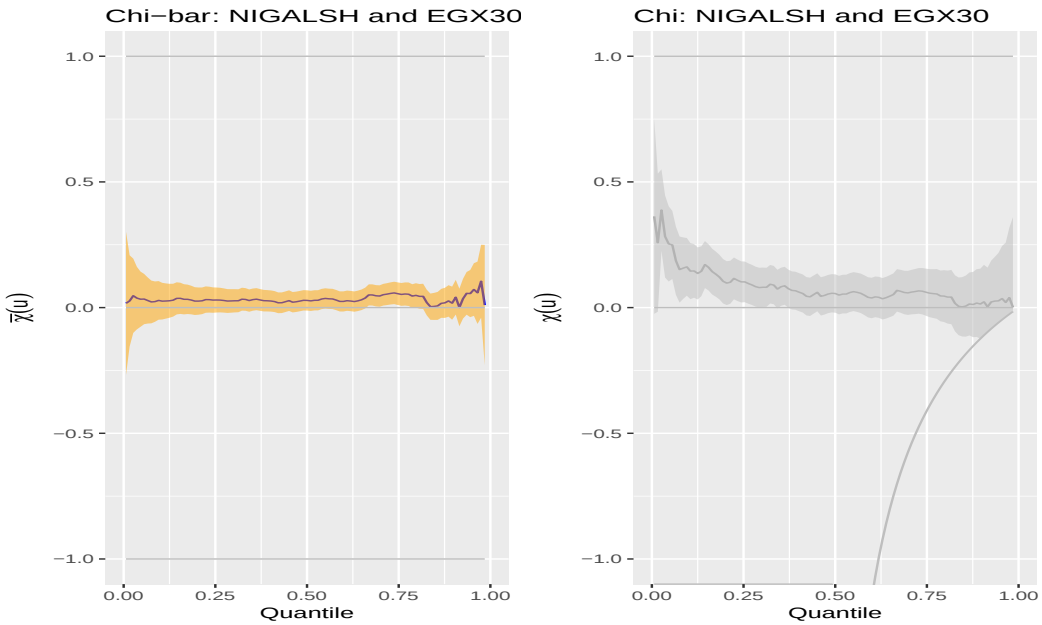

Figure 4.55: Pairwise extremal dependence exploratory plots: NIGALSH and EGX 30 indices.

4.15.11 Pairwise extremal dependence exploratory plots: NIGALSH and NSE 20 indices

These paired markets of Nigerian NIGALSH and Kenyan NSE 20 show some asymptotic dependence as displayed by the $\bar{\chi}$ with a very weak positive dependence within the class as shown by the χ plot (Figure 4.56).

Also, the MCS plot with the 95% confidence interval as displayed in Figure 4.57 (top panel) indicates very weak positive asymptotic dependence between these markets.

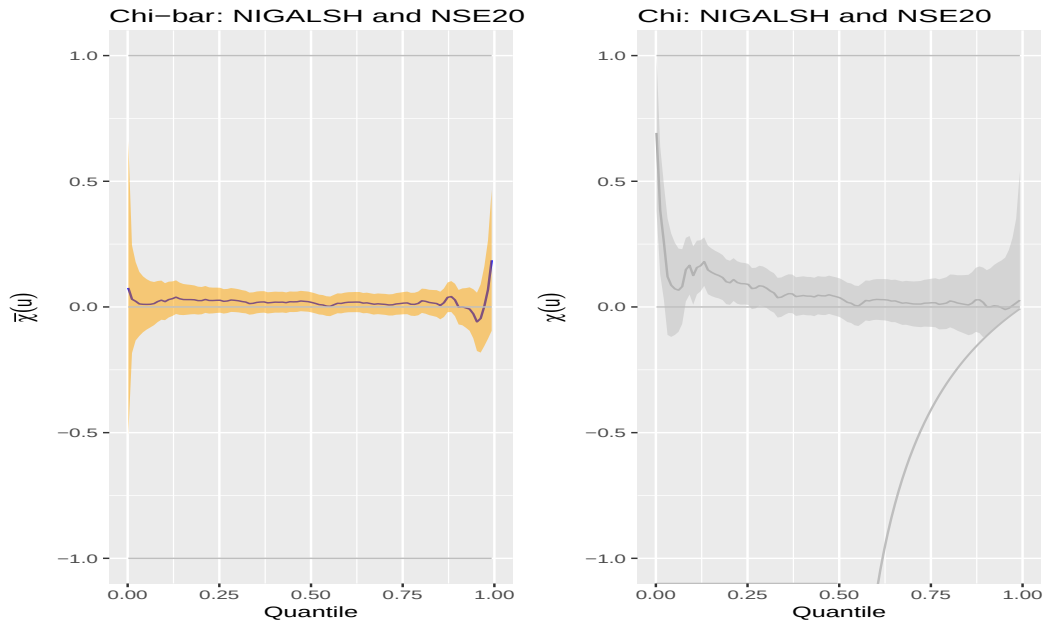

Figure 4.56: Pairwise extremal dependence exploratory plots: NIGALSH and NSE 20 indices.

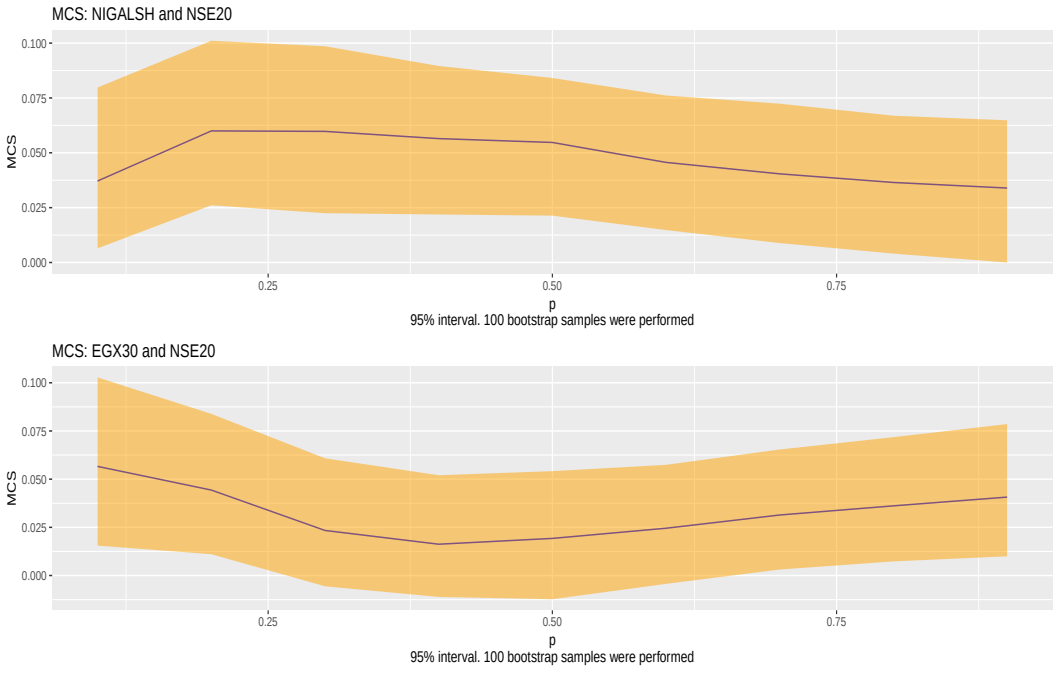

Figure 4.57: MCS plots: NIGALSH and NSE 20 indices (top panel) with EGX 30 and NSE 20 indices (bottom panel).

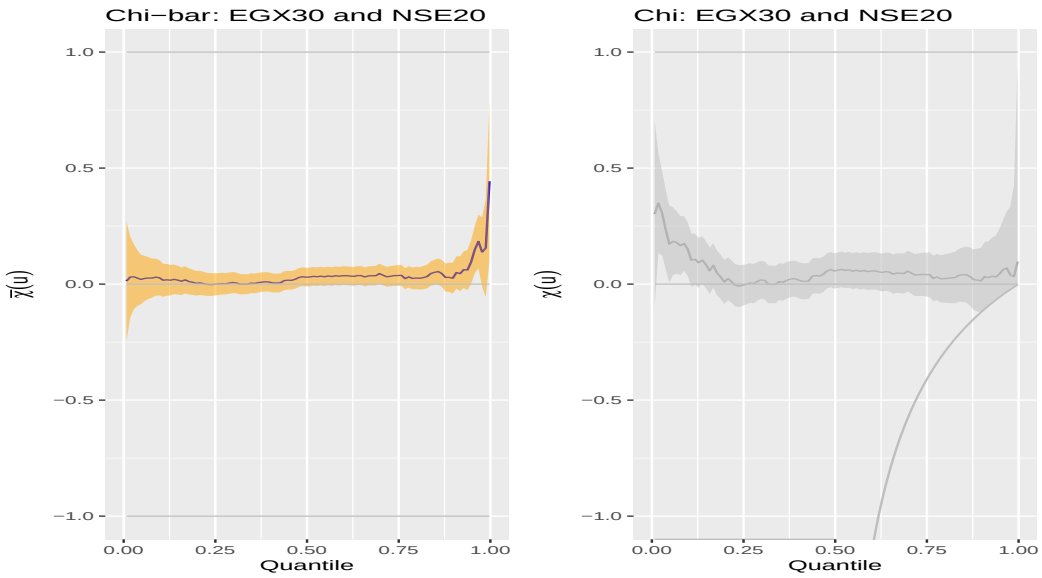

Figure 4.58: Pairwise extremal dependence exploratory plots: EGX 30 and NSE 20 indices.

4.15.12 Pairwise extremal dependence exploratory plots: EGX 30 and NSE 20 indices

As indicated in Figure 4.58, there is an asymptotic dependence between Egyptian EGX 30 and Kenyan NSE 20 markets shown by the $\bar{\chi}$ plot with a weak positive dependence within the class as shown by the χ plot.

The MCS plot with the 95% confidence interval shown in Figure 4.57 (bottom panel) also shows weak positive asymptotic dependence between the markets.

In summary, all the pairwise exploratory plots using the two approaches suggest that modelling of the conditional multivariate extreme value show asymptotic dependencies between the pairwise markets combination with weak or very weak positive dependence and weak negative dependence.

4.16 Model fitting and diagnostics

The bivariate threshold excess model is fitted to the dataset and it is done in turn, conditioning on each of the four marginal variables. To ascertain the required threshold to select for the marginal and dependence modelling, series of candidate marginal and dependence quantiles that define the threshold over which the GPD models will be fitted were run. The tested quantiles were 70th, 80th, 90th and 95th percentiles and the corresponding diagnostic plots to determine the suitable threshold were done.

The condition for this selection is the smoothness of the scatterplot smoothers as displayed in Figure 4.59. If the model is a good fit, the plots' top and centre panels should not show structure with scatterplot smoothers being less or more horizontal (Southworth and Heffernan, 2016). Hence, the more the smoothness of the horizontal lines of the scatterplot smoothers, the better the fit. From the diagnostic plots in Figures 4.59, 4.60, 4.61 and 4.62, it is observed that the parameter estimates are most stable where the horizontal lines are smoothest at the 80th percentile for all the markets.

From top to bottom, the plots show the values (centred and scaled) of the residuals of the dependence model through the extreme conditioning variable's range, then their absolute values, and the original data that was not transformed, where the contours display fitted conditional model's quantiles (Southworth and Heffernan, 2016).

4.16.1 Bivariate-threshold-excess model's parameter estimates

Table 4.27 displays the parameter estimates of the markets' dependence structure conditioning on each of the market variables in turn. That is, given the threshold exceedance by one market variable, the conditional distribution of the three remaining market variables are described. The values of "a" close to -1 or 1 denote strong negative or positive extremal dependence (Southworth and Heffernan, 2016).

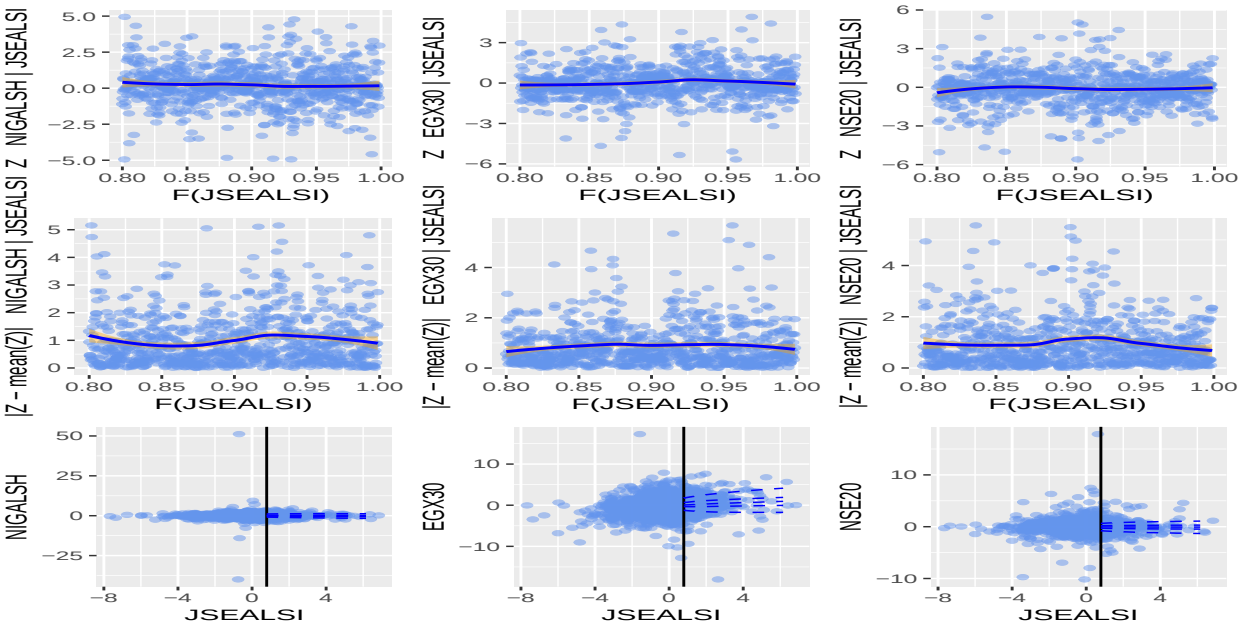

Figure 4.59: Dependence model diagnostics: conditioning on the JSE-ALSI variable.

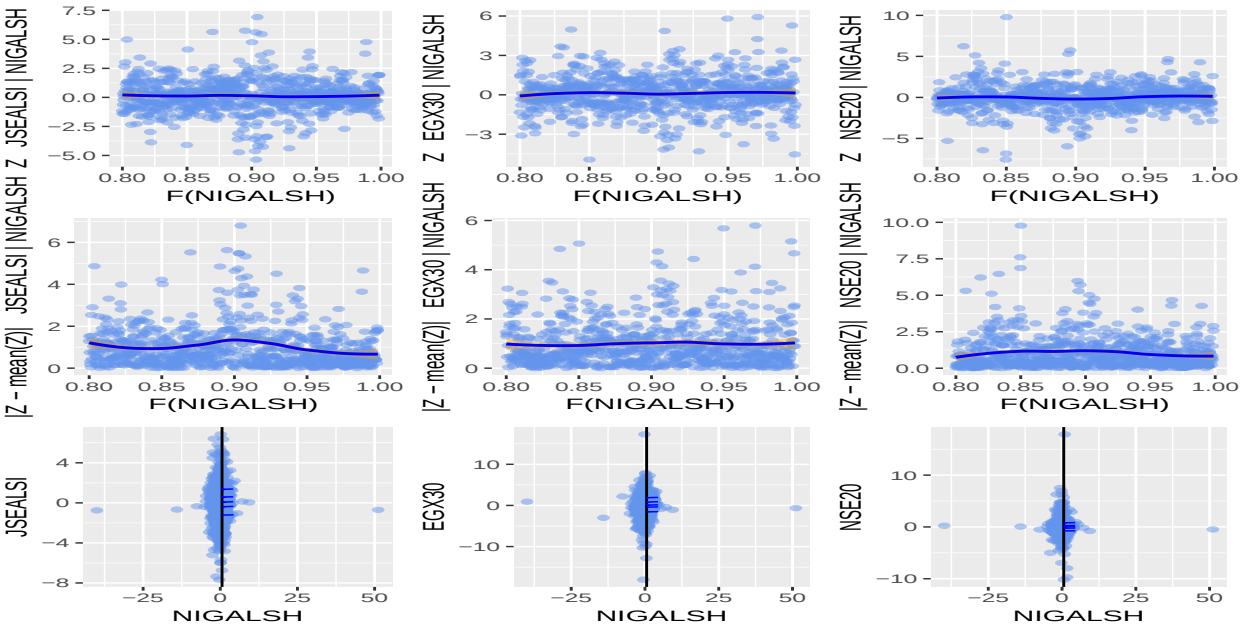

Figure 4.60: Dependence model diagnostics: conditioning on the NIGALSH variable.

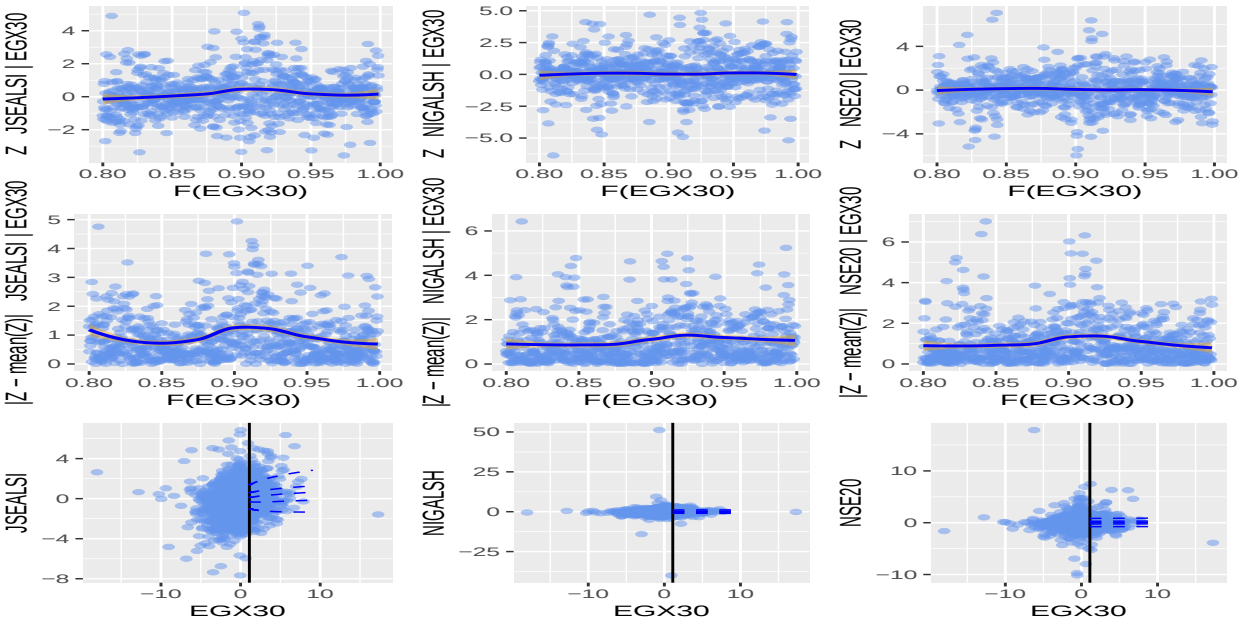

Figure 4.61: Dependence model diagnostics: conditioning on the EGX 30 variable.

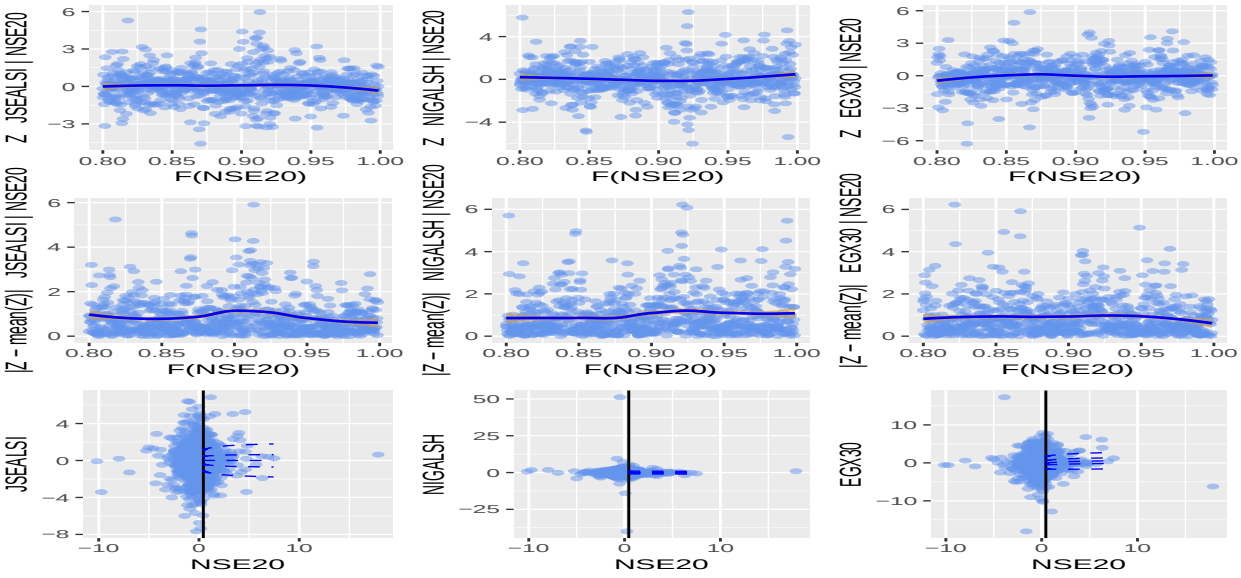

Figure 4.62: Dependence model diagnostics: conditioning on the NSE 20 variable.

Table 4.27: Estimated parameters of the bivariate-threshold-excess model.

	Dependence parameters	NIGALSH	EGX 30	NSE 20
Conditioning on: JSE-ALSI	a	-0.0974	0.1312	-0.0037
	b	0.1762	0.3411	0.2409
	Dependence parameters	JSE-ALSI	EGX 30	NSE 20
Conditioning on: NIGALSH	a	0.0100	0.0100	0.0100
	b	0.0100	0.0100	0.0100
	Dependence parameters	JSE-ALSI	NIGALSH	NSE 20
Conditioning on: EGX 30	a	0.0833	-0.0055	0.0100
	b	0.2914	0.0232	0.0100
	Dependence parameters	JSE-ALSI	NIGALSH	EGX 30
Conditioning on: NSE 20	a	0.2572	-0.2663	-0.0582
	b	0.3668	0.5272	0.7141

4.16.2 Discussion of results: Bivariate-threshold-excess model

Based on the dependence parameter estimates in Table 4.27, the following results are deduced.

1. Conditioning on the JSE-ALSI market: The estimated dependence parameter's values indicate that the Nigerian NIGALSH and Kenyan NSE 20 markets have negative extremal dependence on the South African JSE-ALSI market, the stronger being that of the Nigerian market on the South African market. Egyptian EGX 30 on the other hand, has weak positive dependence on the South African market.

2. Conditioning on the NIGALSH market: The Nigerian NIGALSH market has flat weak positive extremal dependence across the board on the other three markets.

3. Conditioning on the EGX 30 market: Here, it is shown from the values of the dependence parameters that the South African JSE-ALSI and Kenyan NSE 20 markets both have weak positive extremal dependence on large values of Egyptian EGX 30 market, with the stronger of the weak dependence being South African market on the Egyptian market. Nigerian NIGALSH market however has very weak negative dependence on Egyptian market.

4. Conditioning on the NSE 20 market: This dependence parameter's values show that the Nigerian NIGALSH market is negatively dependent on large values of the Kenyan NSE 20 market than does the Egyptian EGX 30 market. However, the South African JSE-ALSI market has a reasonably weak positive level of dependence on the Kenyan market.

4.16.3 Bivariate point process

Now, for comparative purposes, the bivariate point process is also used to model the dependence of the pairwise combinations of the markets. In order to adequately compare the efficiencies of the two models, this dependence approach also applied the marginal thresholds that correspond to the 80% marginal quantiles used for the bivariate-threshold-excess modelling.

With larger bounded coverage region, the point process is able to model more extreme observations than does the threshold excess model since it can use exceedances in either or both of the one margin and in the joint exceedances of both margins unlike the bivariate-threshold-excess model that is limited to the joint exceedances of both margins (Coles, 2001). Using Figure 4.63 as an illustration from Coles (2001), for the bivariate point process modelling, all the observations above the curved region are required, while the bivariate-threshold-excess model is limited to the joint upper quadrant (top right panel), given the same marginal threshold.

The dependence parameters will be estimated using five parametric models and comparison among the models will be made to ascertain the one that gives the best estimation. The models include the logistic, bilogistic, negative logistic, negative bilogistic and the Coles-Tawn dependence models. The Coles-Tawn model is also known as the Dirichlet model.

Tables 4.28 and 4.29 display the dependence parameters α and β with their estimated standard errors in parentheses. The logistic and negative logistic models have lone dependence parameter α, hence the β is shown as "Nil". The Akaike information criterion (AIC) is used to select the model that gives the best fit, where the lowest AIC value indicates best fitting model.

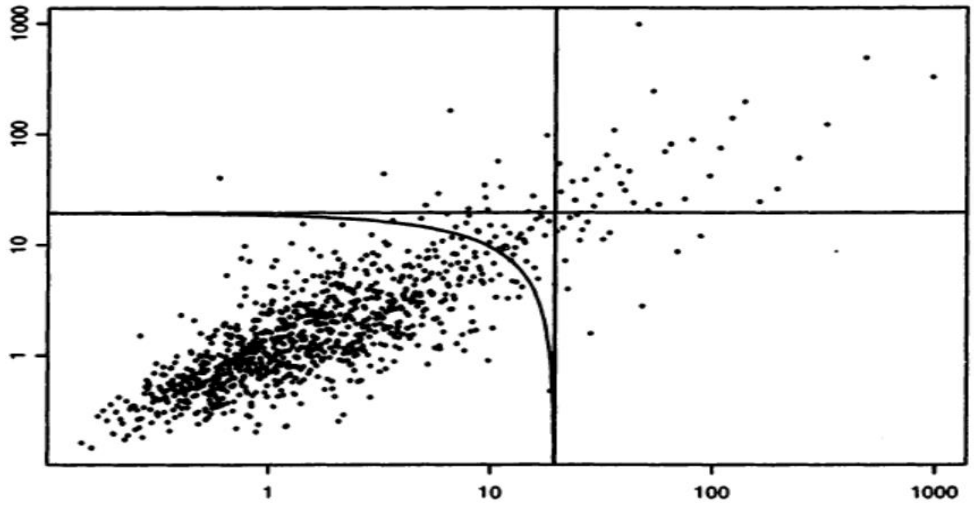

Figure 4.63: Bivariate point process and bivariate-threshold-excess models' thresholds (Coles, 2001).

4.16.4 Discussion of results: Bivariate point process

From Tables 4.28 and 4.29, there are six pairwise combinations of the four underlying African regional markets. Based on the lowest values of the AIC across the paired markets, the result shows that the model that best describes the first paired combination of the South African JSE-ALSI and Nigerian NIGALSH markets is the "negative bilogistic" model. The remaining five markets comprising of the South African JSE-ALSI and Egyptian EGX 30, South African JSE-ALSI and Kenyan NSE 20, Nigerian NIGALSH and Egyptian EGX 30, Nigerian NIGALSH and Kenyan NSE 20, and Egyptian EGX 30 and Kenyan NSE 20 are best fitted by the "negative logistic" model.

For the negative logistic model, complete dependence is obtainable as α tends to ∞, whereas independence is reached when α approaches 0. The likelihood outputs from Tables 4.28 and 4.29 show that α tends closer to 0 than to ∞, indicating independence or very weak dependence in the five affected market pairs.

For the negative bilogistic model on the other hand, complete dependence is attainable as α and β approach 0, whereas independence is attained when the parameters tend to ∞. Table 4.28 indicates that the dependence likelihood parameter estimates for the South African JSE-ALSI and Nigerian NIGALSH indices tend further away from 0 than towards 0. This suggests independence in the extremal dependence modelling of these markets.

Table 4.28: Point process dependence estimates.

JSE-ALSI and NIGALSH			
Parametric model	α	β	AIC
Logistic	0.6563 (0.0065)	Nil	11545.43
Negative logistic	0.7494 (0.0145)	Nil	11339.47
Bilogistic	0.6833 (0.0133)	0.6264 (0.0152)	11542.42
Negative Bilogistic	1.2109 (0.0694)	1.4647 (0.0809)	11338.24
Dirichlet	0.7678 (0.0599)	0.5929 (0.0414)	11409.25
JSE-ALSI and EGX 30			
Parametric model	α	β	AIC
Logistic	0.6428 (0.0067)	Nil	11973.85
Negative logistic	0.7834 (0.0156)	Nil	11801.93
Bilogistic	0.6329 (0.0144)	0.6527 (0.0140)	11975.21
Negative Bilogistic	1.3209 (0.0725)	1.2335 (0.0679)	11803.48
Dirichlet	0.6853 (0.0488)	0.7505 (0.0547)	11854.24
JSE-ALSI and NSE 20			
Parametric model	α	β	AIC
Logistic	0.6637 (0.0063)	Nil	11448.89
Negative logistic	0.7312 (0.0138)	Nil	11218.53
Bilogistic	0.6806 (0.0136)	0.6453 (0.0151)	11449.01
Negative Bilogistic	1.2899 (0.0743)	1.4465 (0.0802)	11219.37
Dirichlet	0.7044 (0.0548)	0.5971 (0.0424)	11305.41
NIGALSH and EGX 30			
Parametric model	α	β	AIC
Logistic	0.6550 (0.0065)	Nil	12071.25
Negative logistic	0.7524 (0.0146)	Nil	11858.67
Bilogistic	0.6418 (0.0150)	0.6674 (0.0139)	12072.27
Negative Bilogistic	1.3697 (0.0778)	1.2883 (0.0751)	11860.34
Dirichlet	0.6479 (0.0478)	0.7097 (0.0553)	11933.33

Table 4.29: Point process dependence estimates.

NIGALSH and NSE 20			
Parametric model	α	β	AIC
Logistic	0.6579 (0.0065)	Nil	11190.27
Negative logistic	0.7460 (0.0143)	Nil	10973.40
Bilogistic	0.6661 (0.0143)	0.6494 (0.0148)	11191.86
Negative Bilogistic	1.3022 (0.0751)	1.3794 (0.0785)	10975.12
Dirichlet	0.6948 (0.0533)	0.6413 (0.0474)	11051.40
EGX 30 and NSE 20			
Parametric model	α	β	AIC
Logistic	0.6538 (0.0065)	Nil	11808.68
Negative logistic	0.7563 (0.0146)	Nil	11600.26
Bilogistic	0.6665 (0.0139)	0.6405 (0.0148)	11809.64
Negative Bilogistic	1.2686 (0.0725)	1.3773 (0.0774)	11601.67
Dirichlet	0.7226 (0.0552)	0.6439 (0.0467)	11671.77

Chapter 5

Discussions and conclusion

5.1 Introduction

Having completed the analysis of the risks and extremal dependencies of the selected markets, this chapter sets the platform for the discussions and conclusion of the outputs of the parameter estimates with regards to the univariate and multivariate modelling.

5.1.1 Discussion: Univariate modelling

For the univariate threshold selection, a more objective approach using the mixture models coupled with sensitivity analysis following Ferro (2003) extremal index requirement was used in place of the subjective traditional fixed threshold approach.

The likelihood estimates of the shape parameters and their corresponding confidence intervals for both threshold-excess model and point process are the same for the four markets. Under both models, the shape parameter estimate for the JSE-ALSI index correspond to a bounded distribution and can be well described by the Weibull class of distributions. For the remaining three indices, i.e. NIGALSH, EGX 30 and NSE 20, the shape parameter estimates correspond to the unbounded distribution and can be characterised by the Pareto distribution for threshold excess model and

Fréchet-Pareto class of distributions (Beirlant et. al, 2004) for the point process.

The model diagnostic checks carried out using probability plots, the quantile plot, return level plots, and histogram and density plots to ascertain the validity of the fitted GPD (generalised Pareto distribution) model were reasonably adequate. For the point process, probability plots and the quantile plot used for the diagnostics indicated that the fits are reasonably satisfactory. Furthermore, goodness of fit test for the GPD fit was carried out using Cramér-von Mises and Anderson-Darling tests, the results showed that the GPD is a good fit in both tests.

The return levels estimates were approximately the same for the threshold excess model and point process under JSE-ALSI, NIGALSH and EGX 30 indices. The NSE 20 market however exhibited a slight difference in that the estimates of the four return periods under threshold excess were a bit more than the estimates obtained using point process (see Tables 4.23 and 4.25).

Under the univariate analysis, the likelihood estimates of both threshold excess and point process models were more or less the same. However, the threshold excess model is easier to estimate since the GPD has got fewer parameters for estimation when compared to the point process approach.

5.1.2 Discussion: Bivariate modelling

For the bivariate modelling, this study believes that the application of Southwouth and Heffernan (2016) approach for the marginal threshold quantile selection using the scatterplot smoothers to obtain diagnostic smooth straight lines in Section 4.16 (for the bivariate-threshold model) facilitates a more objective trade-off between bias and viariance. This argument is based on the fact that the approach does not follow a fixed choice of marginal threshold quantile which can lead to inadequate selection of extreme observations in the joint exceedances and then result to greater estimation uncertainty.

The chosen marginal threshold of 80^{th} percentile was selected for the bivariate-threshold-excess modelling. For comparison, the marginal thresholds that correspond to the 80% quantile were also used for the bivariate point process modelling. The modelling under bivariate point process is usually more informative than that of the bivariate-threshold-excess model because the former uses more extreme observations to model than the latter. Analysis based on the point process model applies greater number of observations due to its curved threshold, which allows wider coverage of exceedances in the individual and joint quadrant, whereas the bivariate-threshold-excess model is limited to the joint quadrant only (Coles, 2001).

For the markets' extremal dependencies, results in Table 4.27 for the bivariate-threshold-excess model and Tables 4.28 and 4.29 for the bivariate point process summarily show how (very) weakly associated or close to independence and negatively associated the African regional markets are to one another. These outcomes jointly describe the fundamental requirements for efficient portfolio diversification between paired markets. A weak or negative co-movement between equity markets is required to enable potential poor performance in one side of the equity market to be relatively hedged by good returns on the other side of the paired markets.

5.1.3 Conclusions and recommendations

The results suggest that the African continent's regional equity markets present a huge investment platform for investors and traders, and offer tremendous opportunity for portfolio diversification and investment synergies between markets. These synergistic opportunities are due to the markets being asymptotic (extremal) independent or (very) weak asymptotic dependent and negatively dependent, which are the basic requirements for business and investment portfolio diversification. This indicates that extreme (rare) movement in one market does not easily spill-over to the other market. Hence, this is a clarion call for investors, especially international investors

and traders with investment interest in African markets to seize on this beneficial opportunity for relevant portfolio diversification between markets.

This study's outcome is consistent with the work of Alagidede (2008) who analysed these same four African regional markets using co-integration analysis. The researcher concluded that the markets do not significantly co-move in spite of the subsisting cooperation and economic reforms among them.

On a broader perspective, this study further concludes by proffering answers to the research questions as follows. In terms of risk hierarchy, the Egyptian EGX 30 market is the most risk-prone, followed by the South African JSE-ALSI market, then the Nigerian NIGALSH market and the least risky is the Kenyan NSE 20 market. However, the study observed that the markets' volatility array does not follow this risk hierarchy pattern. The findings show in descending hierarchy that volatility with persistence is highest in the South African market, followed by Egyptian market, then Nigerian market and lastly, the Kenyan equity market. It is therefore concluded that risk is not a brainchild of volatility in these markets.

In response to the second question, portfolio diversification is highly feasible among the selected regional markets due to the weak extremal dependence and negative dependence observed in the markets.

As an answer to the question on what drives the tail or extremal dependence of the selected market, it is seen from the conclusions that fundamental factors like economic or trade links do not seem to have much effect on extremal dependence of the markets. For instance, Egypt and Kenya are connected by trade through COMESA (Common Market for Eastern and Southern Africa), yet there is a notable weak dependence between their stock markets. In essence, it can be concluded that what may drive extremal dependence between these markets is contagion effect. As stated earlier, contagion results from investors rapidly changing their positions, because they think a crisis in one market will eventually spill-over to other markets, and this may

ultimately result in herd behaviour effects, that is, panic selling.

For the fourth question, the dependence structure of the markets can be summarised by investors and all market participants as weak and it provides a tremendous and flexible opportunity for portfolio diversifications between markets and across the regions.

For response to the last question, both the bivariate-threshold-excess model and point process model are appropriate for modelling the markets' risks. For modelling the extremal dependence however, given the same marginal threshold quantile, the point process has more access to the extreme observations due to its wider sphere of coverage than the bivariate-threshold-excess model.

5.1.4 Future study

In addition to the two extreme value models used in this study, the inclusion of copula approach will be considered for future research. An extension into the use of copula approach may give different perspectives to the research and the results may reflect a broader view of the data analysis.

References

1. Alagidede P., *How integrated are Africa's stock markets with the rest of the world?*, Department of Economics, University of Stirling, (2008).

2. Altun E. and Tatlidil H., *A comparison of extreme value theory with heavy - tailed distributions in modeling daily var*, Journal of Finance and Investment Analysis 4, (2015).

3. Anderson T. W. and Darling D. A., *A Test of Goodness of Fit*, Journal of American Statistical Association, 29, (1954), 765-769.

4. Balkema A. and de Haan L., *Residual life time at a great age.* Annals of Prob., 2, (1974), 792-804.

5. Bali T. G., *An extreme value approach to estimating volatility and value at risk*, Journal of Business, 76, (2003), 83-108.

6. Beirlant J., Goegebeur Y., and Teugels, J, *Statistics of Extremes: Theory and Applications*, Wiley, Chichester, (2004).

7. Bensalah Y., *Steps in Applying Extreme Value Theory to Finance: A Review*, Working Paper (Bank of Canada), (2000), 2000-20.

8. Bere A. *Some non-standard statistical dependence problems*, University of the western cape, (2016).

9. Bollerslev T., *Generalized autoregressive conditional heteroscedasticity*, Journal of Econometrics, 31, (1986), 307-327. Reference

10. Bommier E., *Peaks-Over-Threshold Modelling of Environmental Data*, U.U.D.M. Project Report, Department of Mathematics, Uppsala University, (2014).

11. Coles S., *An Introduction to Statistical Modelling of Extreme Values*, (2001), 81-156.

12. Coles S., Heffernan J., and Tawn J. *Dependence measures for extreme value analyses*, Extremes, 2, (1999), 339-365.

13. Dacorogna M. M., Müller U. A., Pictet O. V., and DeVries C. G., *The Distribution of Extremal Foreign Exchange Rate Returns in Extremely Large Data Sets*, Preprint, O & A Research Group, (1995).

14. D'Agostino R. B. and Stephens M. A., *Goodness-of-fit Techniques*, Marcel Dekker, Inc., New York, (1986).

15. Davidson A. C. and Smith R. L., *Models for exceedances over high thresholds*, J.Roy. Statist. Soc. Ser. B. (1989).

16. Dickey D.A. and Fuller W. A., *Distribution of the estimators for autoregressive time series with a unit root*, Journal of the American Statistical Association, 74, (1979), 427-431.

17. de Haan, L., *Extremes in higher dimensions: The model and some statistics* (stma v27 891), (1985), 1-15.

18. de Haan L. and Resnick S. I., *Limit theory for multivariate sample extremes*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 40, (1977), 317-337.

19. Djakovic V., Andjelic G., and Borocki J. *Performance of Extreme Value Theory in Emerging Markets: An Empirical Treatment*, African Journal of Business Management, 5, (2011), 340-369.

20. Engle R., *Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation*, Econometrica, 50, (1982), 987-1006.

21. Ferro C. A. T., *Statistical Methods for Clusters of Extreme Values*, Lancaster University, (2003).

22. Fisher R. A. and Tippett L. H. C., *Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample*, Mathematical Proceedings of the Cambridge Philosophical Society 24, (1928), 180-190.

23. Gilli M. and Kellezi E., *An Application of extreme value theory for measuring financial risk*, Comput Econ 27, (2006), 1-23.

24. Gnedenko B. V., *Sur la distribution limite du terme maximum d'une série aléatoire*. Ann. Math. 44, (1943), 423-453.

25. Gomes M. I., *An i-dimensional limiting distribution function of largest values and its relevance to the statistical theory of extremes*. In Statistical distributions in scientific work (C. Taillie, G. P. Patil and B. A. Baldessari, eds) 6, (1981), 384-410.

26. Haque M., Hassan M., Maroney N. and Sackley W., *An empirical examination of stability, predictability and volatility of Middle Eastern and African emerging stock markets*, Review of Middle East Economics and Finance, 2, (2004), 19-42.

27. Heffernan J. E. and Tawn J., *A conditional approach for multivariate extreme values*. Journal of the Royal Statistical Society Series B, 56, (2004) 497-546.

28. Hill B. M., *A simple general approach to inference about the tail of a distribution*, Ann Statist. 3, (1975), 1163-1174.

29. Hosking J. R. M., Wallis J. R. and Wood E. F., *Estimation of the generalised extreme-value distribution by the method of probability-weighted moments*, Technometrics 27, (1985), 251-261.

30. Hu Y. and Scarrott C., *Evmix: An r package for extreme value mixture modelling, threshold estimation and boundary corrected kernel density estimation*, (2013).

31. Hu Y. and Scarrott C., *evmix: An R package for Extreme Value Mixture Modeling, Threshold Estimation and Boundary Corrected Kernel Density Estimation*, Journal of Statistical Software, 84, (2018), Issue 5.

32. Joe H., Smith R. L., and Weissman I., *Bivariate threshold methods for extremes*, Journal of the Royal Statistical Society, Series B: Methodological 54, (1992), 171-183.

33. Leadbetter M., Lindgren G. and Rootzen H., *Extremes and related properties of random sequences and processes*, Springer series in statistics, Springer-Verlag, (1983).

34. Levy H. and Sarnat M., *International diversification of investment ponfolios*, American Economic Review, 60, (1970), 668-675.

35. Lipika B., *Multivariate Extreme Value Theory with an Application to Climate Data in the Western Cape Province*, Department of statistical sciences, University of Cape Town, (2018).

36. Longin F. and Solnik B. *Extreme Correlation of International Equity Markets*, Journal of Finance, 56, (2001), 649-677.

37. MacDonald A., Scarrott C. J., Lee D., Darlow B., Reale M., and Russell G., *A*

flexible extreme value mixture model, Comp. Statist. Data Anal., 55, (2011), 2137-2157.

38. McLeod A. I. and Li W. K., *Diagnostic checking ARMA time series models using squared residual autocorrelations*, Journal of Time Series Analysis, 4, (1983), 269-273.

39. McNeil A. J. and Frey R., *Estimation of Tail Related Risk Measure for Heteroscedastic Financial Time Series: An Extreme Value Approach*, J. Empir. Financ., 7, (2000), 271-300.

40. Northrop P. J. and Coleman C. L., *Improved threshold diagnostic plots for extreme value Analyses*, Extremes 17, (2014), 289-303.

41. Pickands J. III, *Statistical inference using extreme order statistics*, Ann. Statist. 3, (1975), 119-131.

42. Poon S., Rockinger M. and Tawn J., *Extreme value dependence in financial markets: diagnostics, models, and financial implications*, Review of Financial Studies 17, (2004), 581-610.

43. Prescott P. and Walden A. T., *Maximum likelihood estimation of the parameters of the generalized extreme-value distribution*, Biometrika 67, (1980), 723-724.

44. Pretorius E., *Economic Determinants of Emerging Market Interdependence*, Emerging Markets Review, 3, (2002), 84-105.

45. Ripley D., *Systematic elements in the linkage of national stock market indices*, Review of Economics and Statistics, 55, (1973), 356-361.

46. Rossi R., *Unit roots Tests*, Fin. Econometrics, 10, (2014), 1-40.

47. Scarrott C. and MacDonald A., *A review of extreme value threshold estimation and uncertainty quantification*, REVSTAT – Statistical Journal, 10, (2012), 33-60.

48. Schmidt F. and Schmitt R., *Multivariate conditional versions of spearman's rho and related measures of tail dependence.* Journal of Multivariate Analysis, 98, (2007), 1123-1140.

49. Sibuya M., *Bivariate extreme statistics, i*, Annals of the Institute of Statistical Mathematics 11, (1960), 195-210.

50. Sigauke C. and Bere A. *Modelling non-stationary time series using a peaks over threshold distribution with time varying covariates and threshold: An application to peak electricity demand*, Energy, 119, (2017), 152-166.

51. Sigauke C., Maposa D., Mudimu E. and Nyamugure, P., *Volatility modeling using ARIMA-GARCH models in a hyper-inflationary economic environment: The Zimbabwean experience*, In Peer-reviewed Proceedings of the Annual Conference of the South African Statistical Association, (2010).

52. Singla N., Jain K. and Sharma S. K., *Goodness of fit tests and power comparisons for weighted gamma distribution*, REVSTAT – Statistical Journal, 14, (2016), 29-48.

53. Smith R. L., *Threshold methods for sample extremes*, In Statistical extremes and applications (J. Tiago de Oliveira, ed.), (1984), 621-638.

54. Smith R. L., *Maximum likelihood estimation in a class of non-regular cases.* Biometrika 72, (1985), 67-90.

55. Smith R. L., *Extreme value analysis of environmental time series: An Application to Trend Detection in Ground-Level Ozone*, Statistical Science 4, (1989),

367-377.

56. Smith R. L., *Statistics of extremes, with applications in environment, insurance, and finance*, Monographs on Statistics and Applied Probability, (2003), 54-60.

57. Stephens M. A., *Tests based on EDF statistics.* In Goodness-of-Fit Techniques (R. B. D'Agostino and M. A. Stephens, Eds.), Marcel Dekker, New York, (1986), 97-193.

58. Southworth H. and Heffernan J. E., texmex: *Statistical modelling of extreme values*, R package version 2.1., (2013).

59. Southworth H. and Heffernan J. E., *Conditional modelling of multivariate extreme value data using R*, (2016).

60. Taleb N. N., *The Black Swan: The Impact of the Highly Improbable.* Random House Trade Paperbacks, 2nd edn. (2010).

61. Tawn J. A., *An extreme value theory model for dependent observations*, J. Hydrology, 101, (1988), 227-250.

62. Todorovic P. and Zelenhasic E., *A stochastic model for flood analysis*, Water resources research 6, (1970), 1641-1648.

63. Uppal J. Y., *Measures of Extreme Loss Risk - An Assessment of Performance During the Global Financial Crisis*, Journal of Accounting and Finance, 13, (2013).

64. Weissman I., *Estimation of parameters and large quantiles based on the k largest observations*, Journal of the American Statistical Association 73, (1978), 812-815.

65. Yang H., *Extreme value mixture modelling with simulation study and applications in finance and insurance*, (2013).

66. Yu J. and Hassan M. K., *Global and regional integration of the Middle East and North African (MENA) stock markets*, The Quarterly Review of Economics and Finance, 48, (2006), 482-504.

Appendix

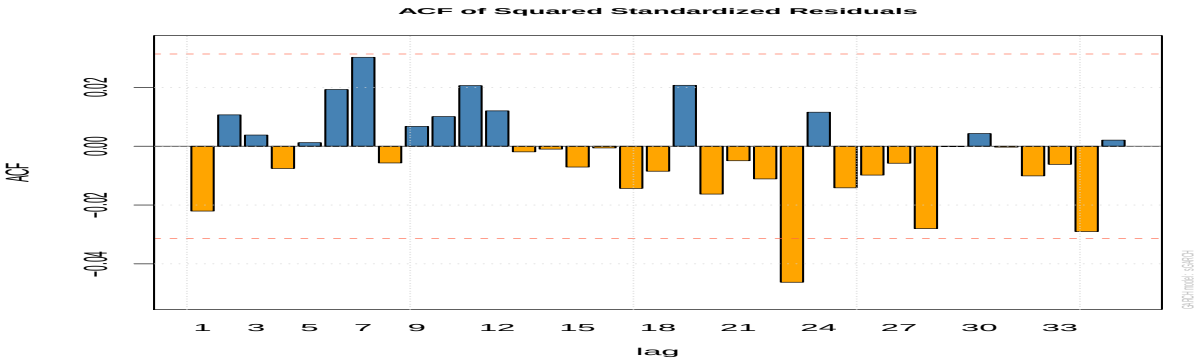

Figure 5.1: ACF of squared standardized residuals for JSE-ALSI.

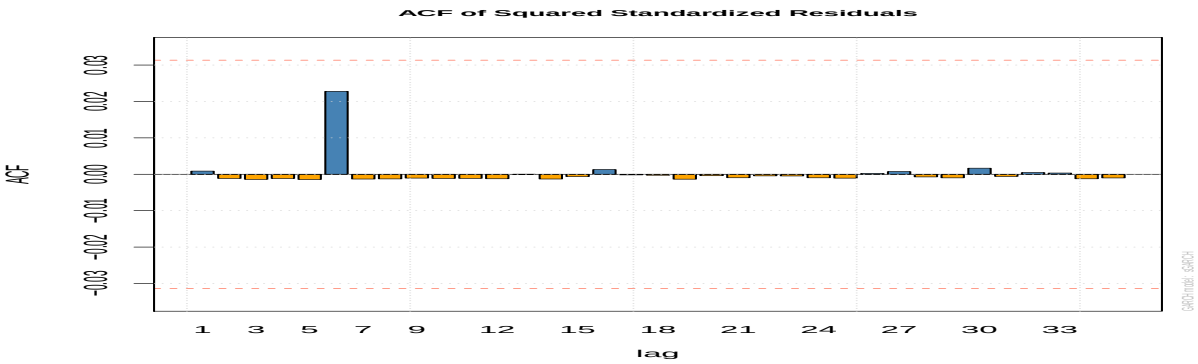

Figure 5.2: ACF of squared standardized residuals for NIGALSH.

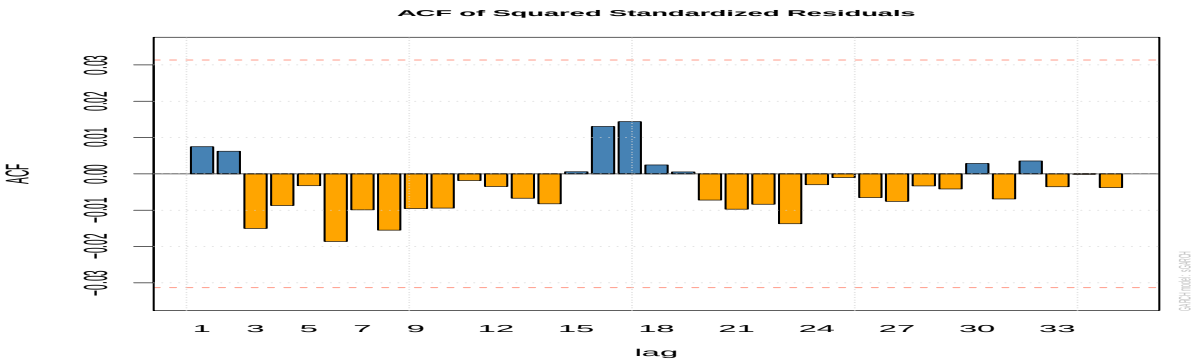

Figure 5.3: ACF of squared standardized residuals for EGX 30.

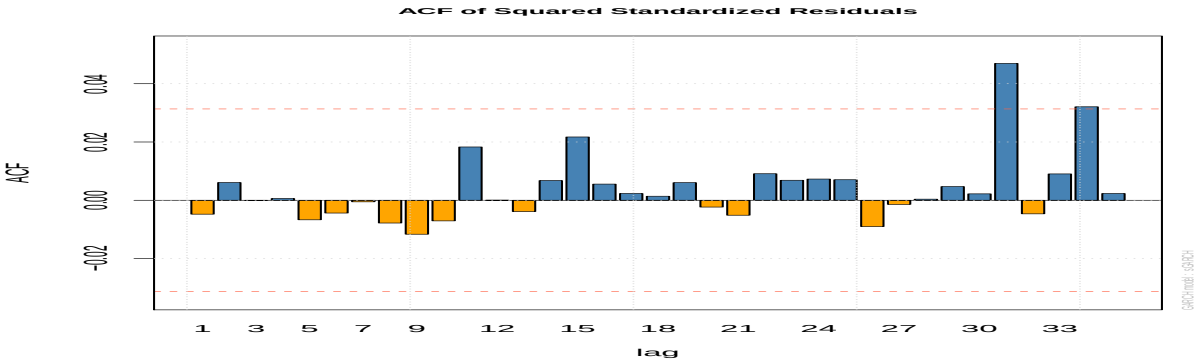

Figure 5.4: ACF of squared standardized residuals for NSE 20.

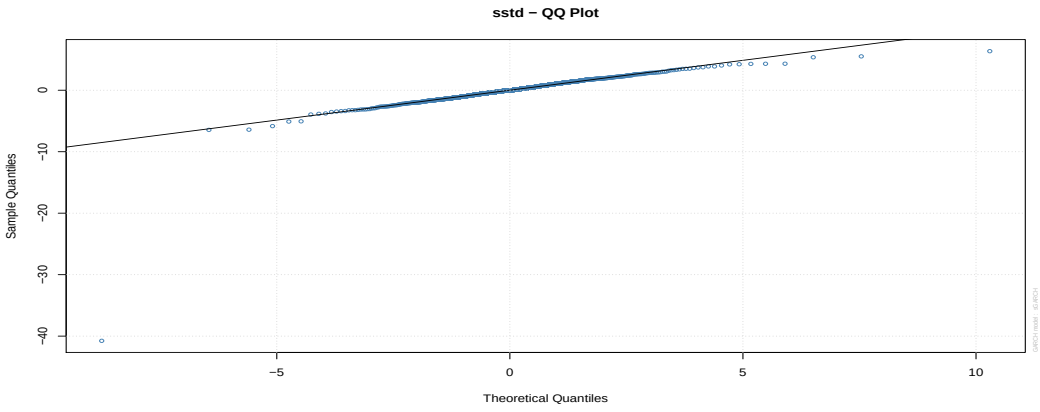

Figure 5.5: Diagnostic: sstd Q-Q plot NIGALSH (Nigeria).

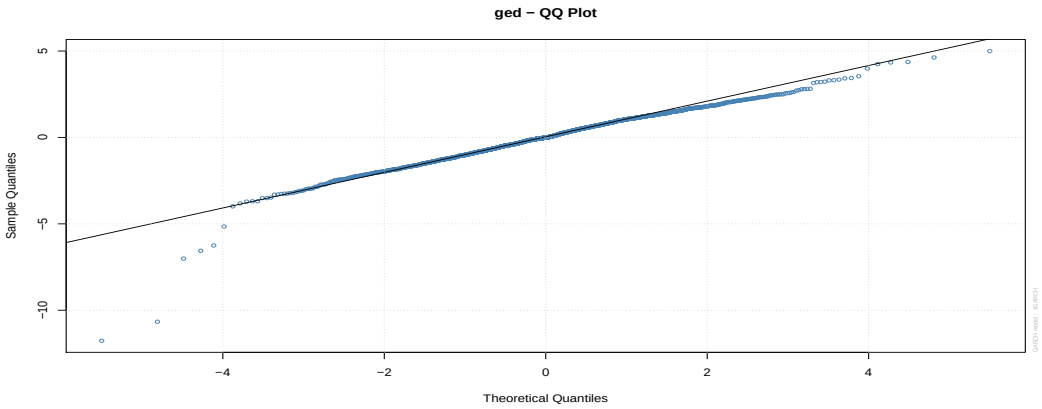

Figure 5.6: Diagnostic: GED Q-Q plot EGX 30 (Egypt).

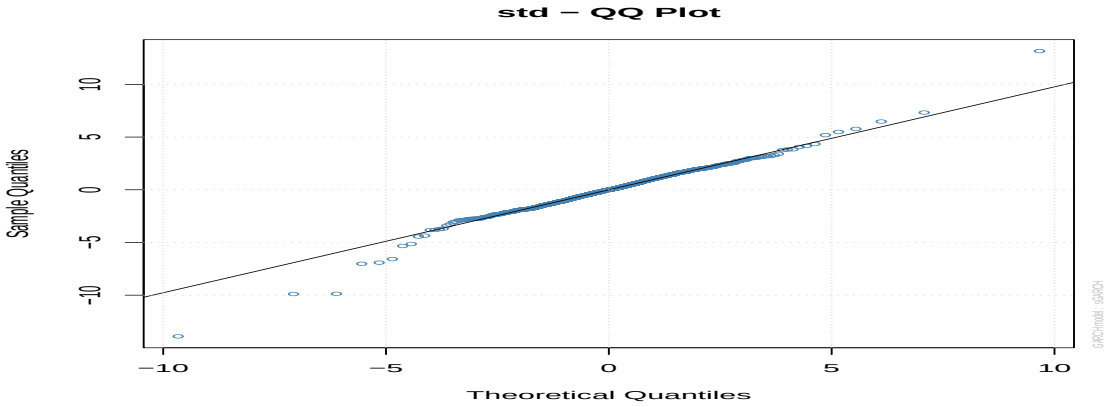

Figure 5.7: Diagnostic: std Q-Q plot NSE 20 (Kenya).

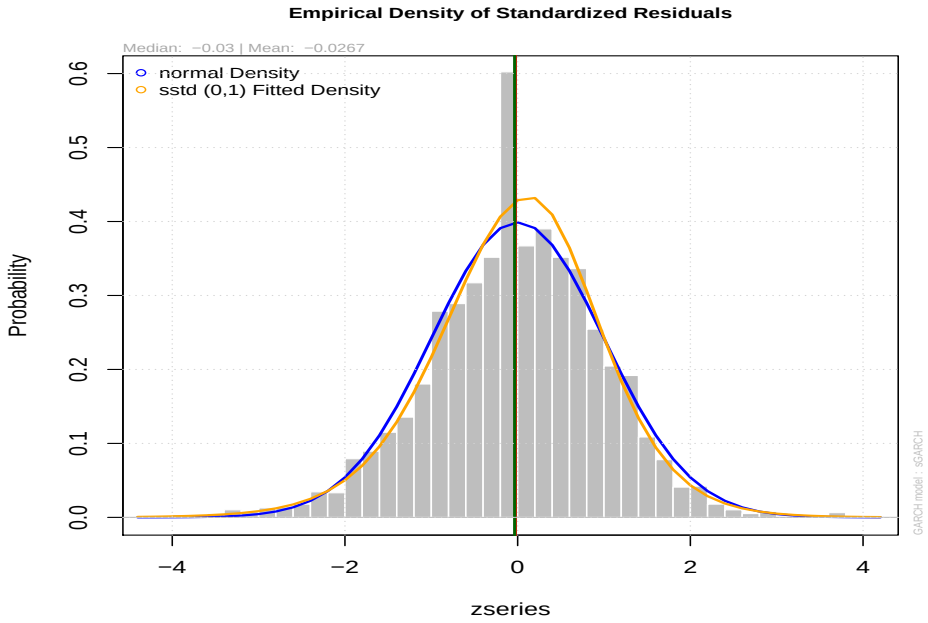

Figure 5.8: JSE-ALSI: Density of standardized residuals.

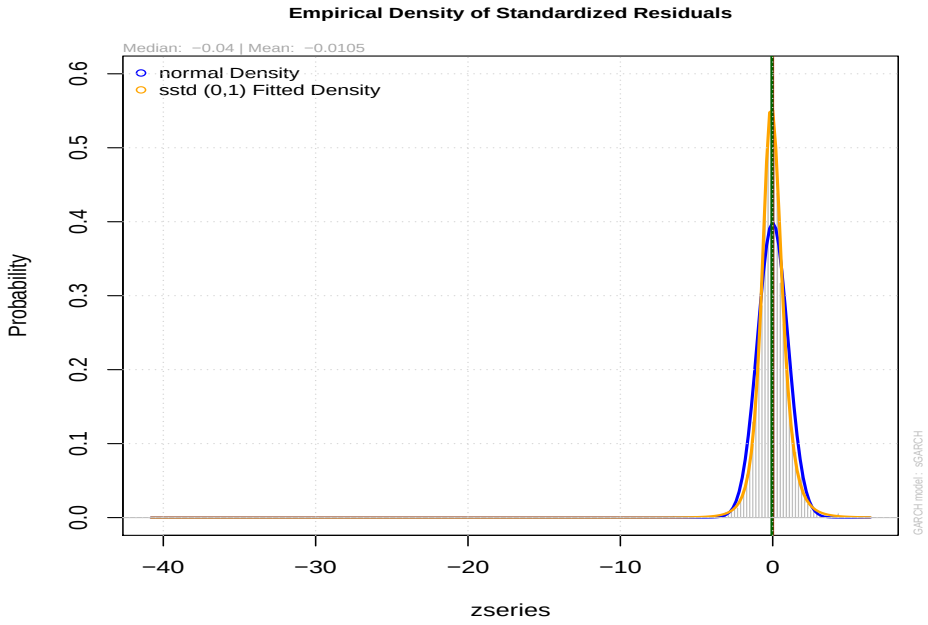

Figure 5.9: NIGALSH: Density of standardized residuals.

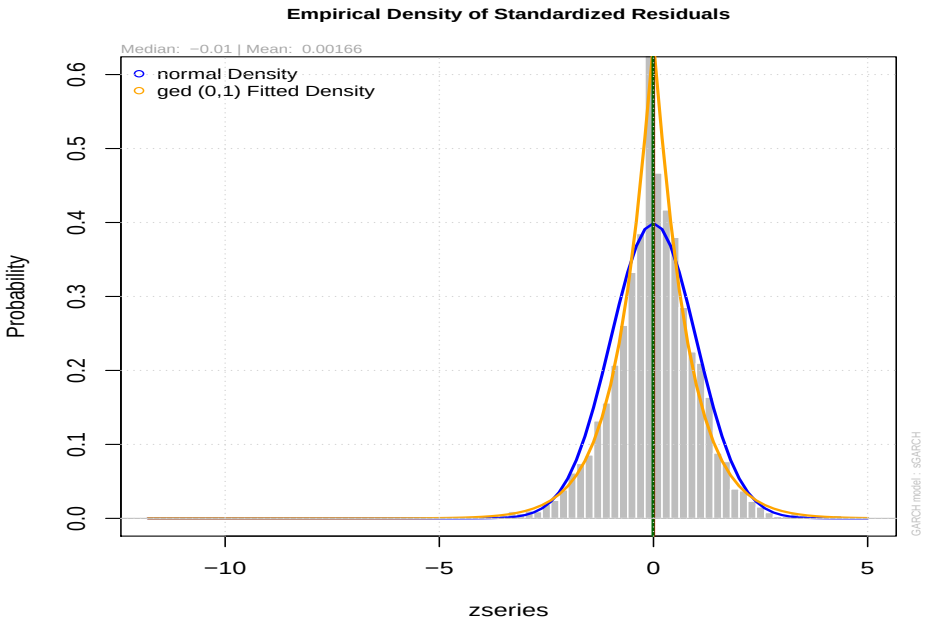

Figure 5.10: EGX 30: Density of standardized residuals.

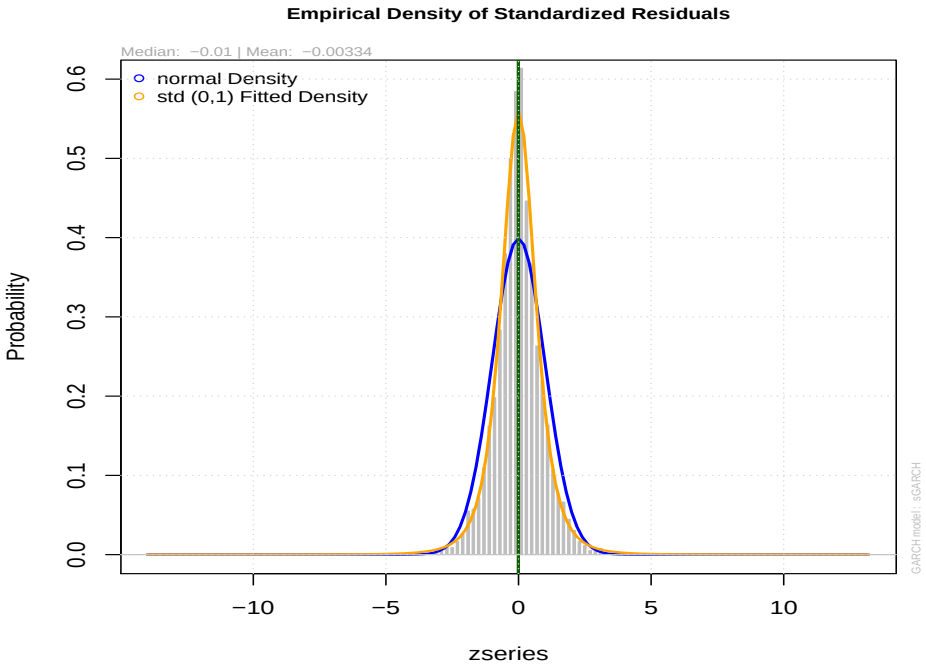

Figure 5.11: NSE 20: Density of standardized residuals.

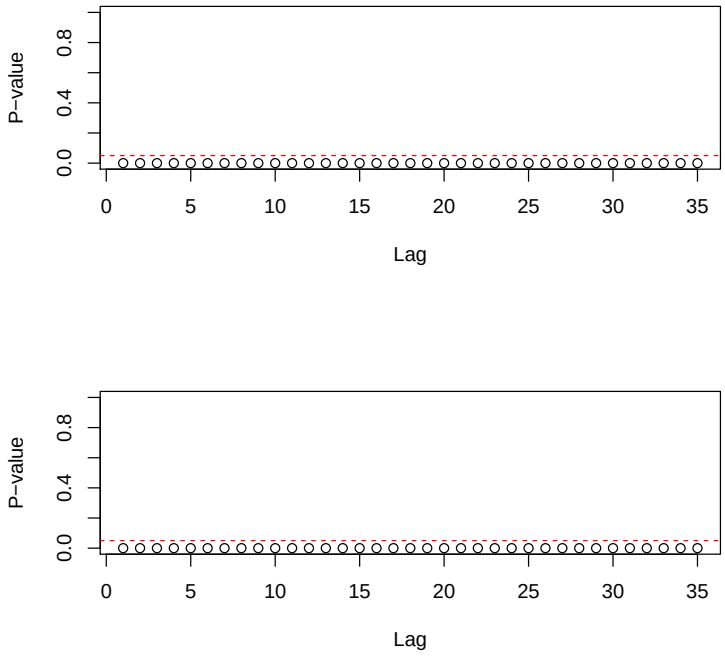

Figure 5.12: McLeod-Li test statistic for daily JSE-ALSI (top panel) and NIGALSH (bottom panel) returns.

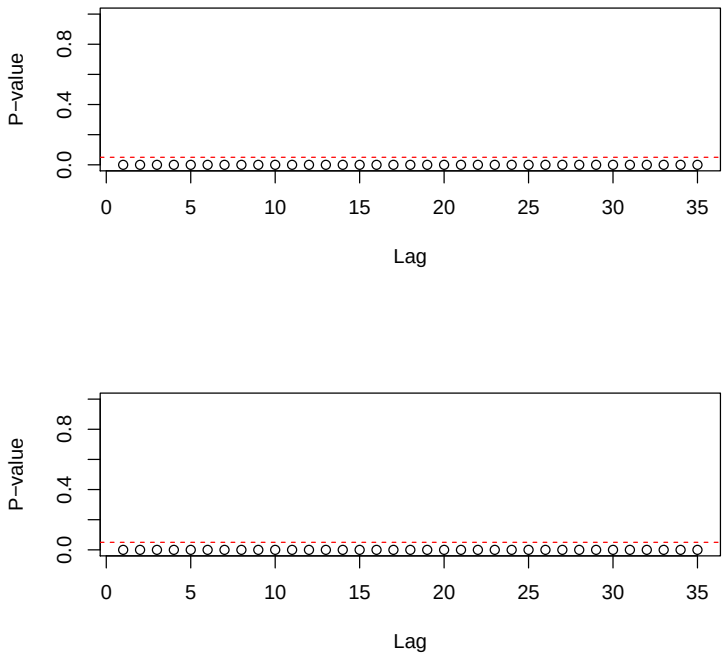

Figure 5.13: McLeod-Li test statistic for daily EGX 30 (top panel) and NSE 20 (bottom panel) returns.

Volatility Persistence Plots

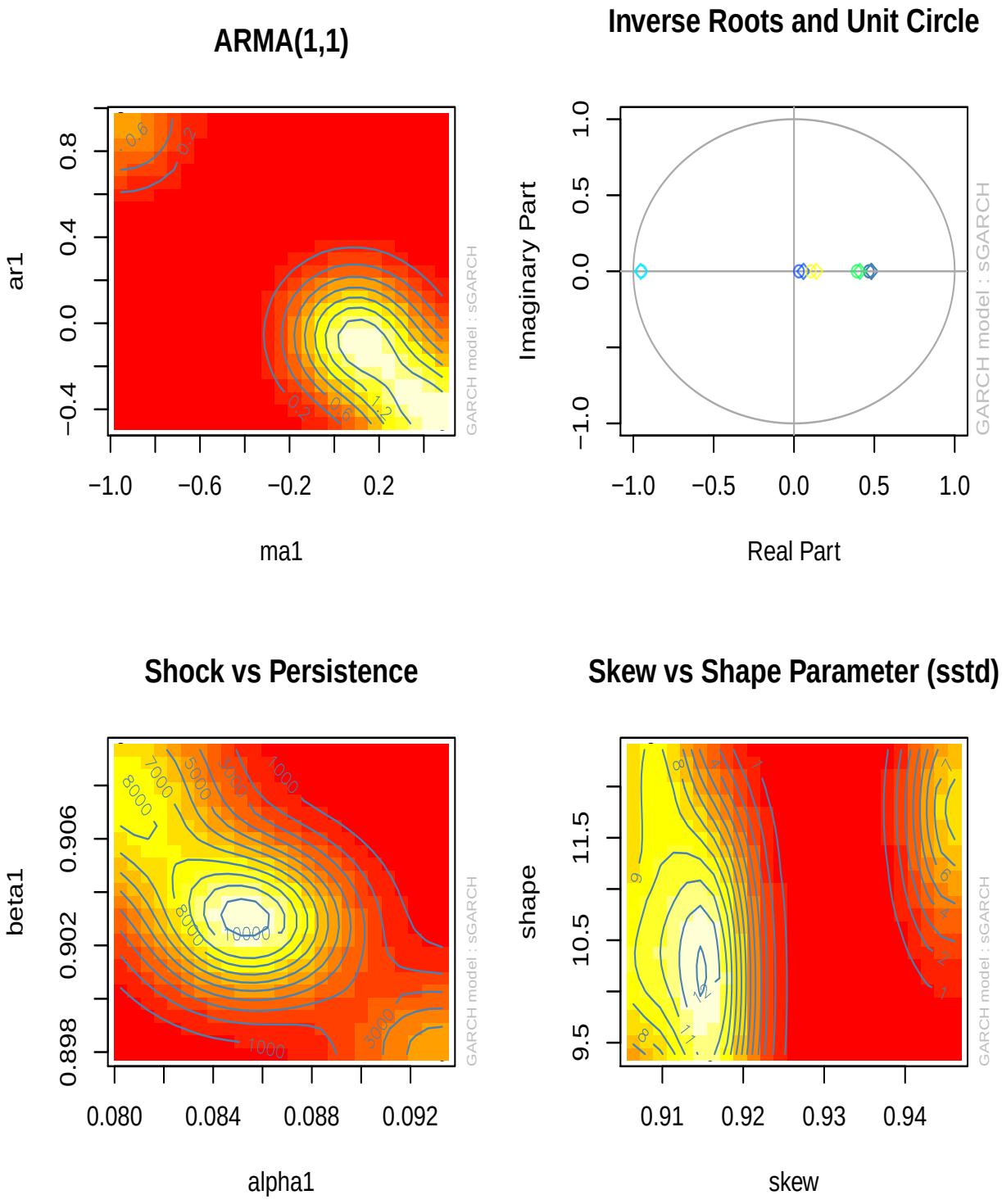

Figure 5.14: Volatility plot: JSE-ALSI.

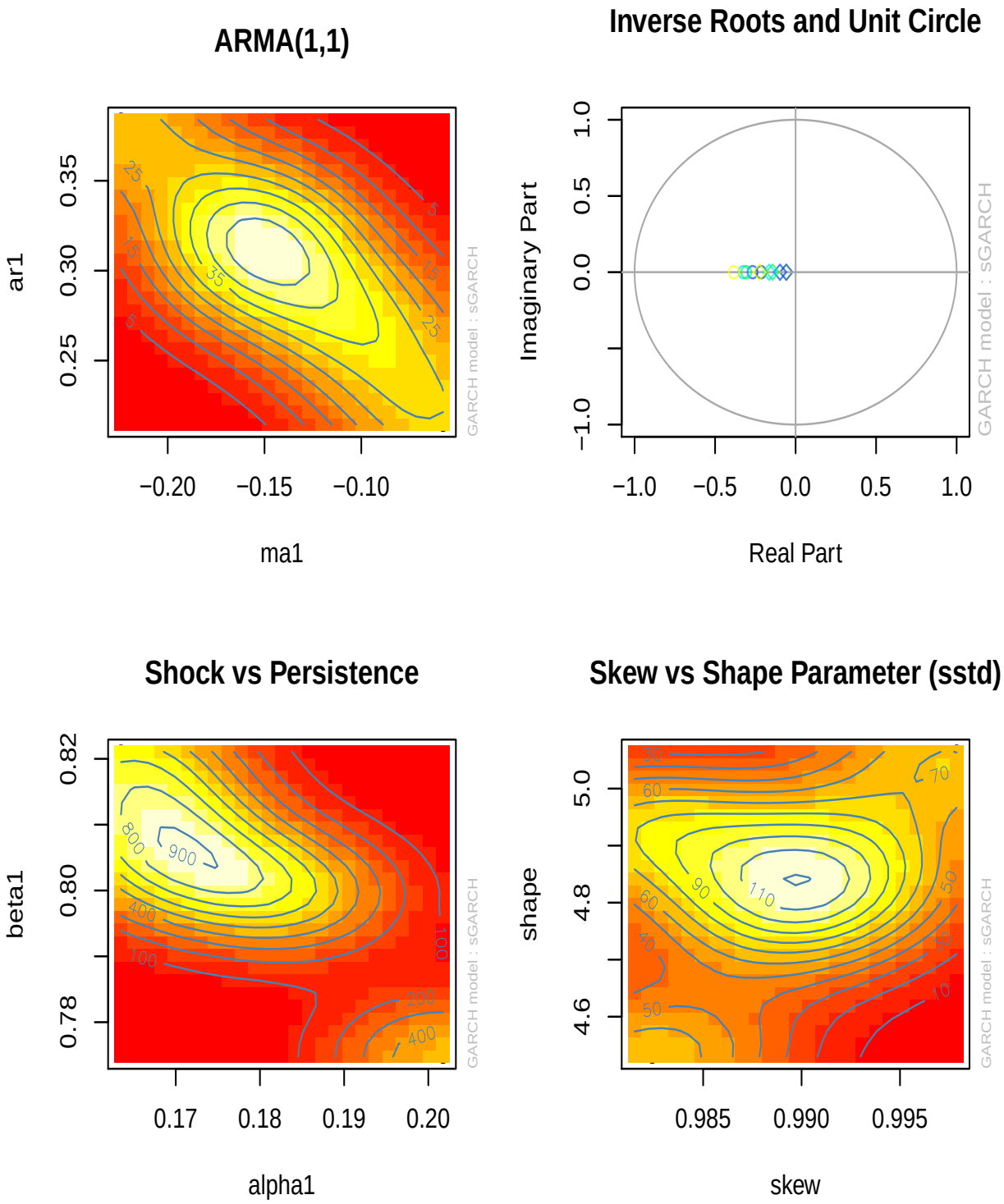

Figure 5.15: Volatility plot: EGX 30.

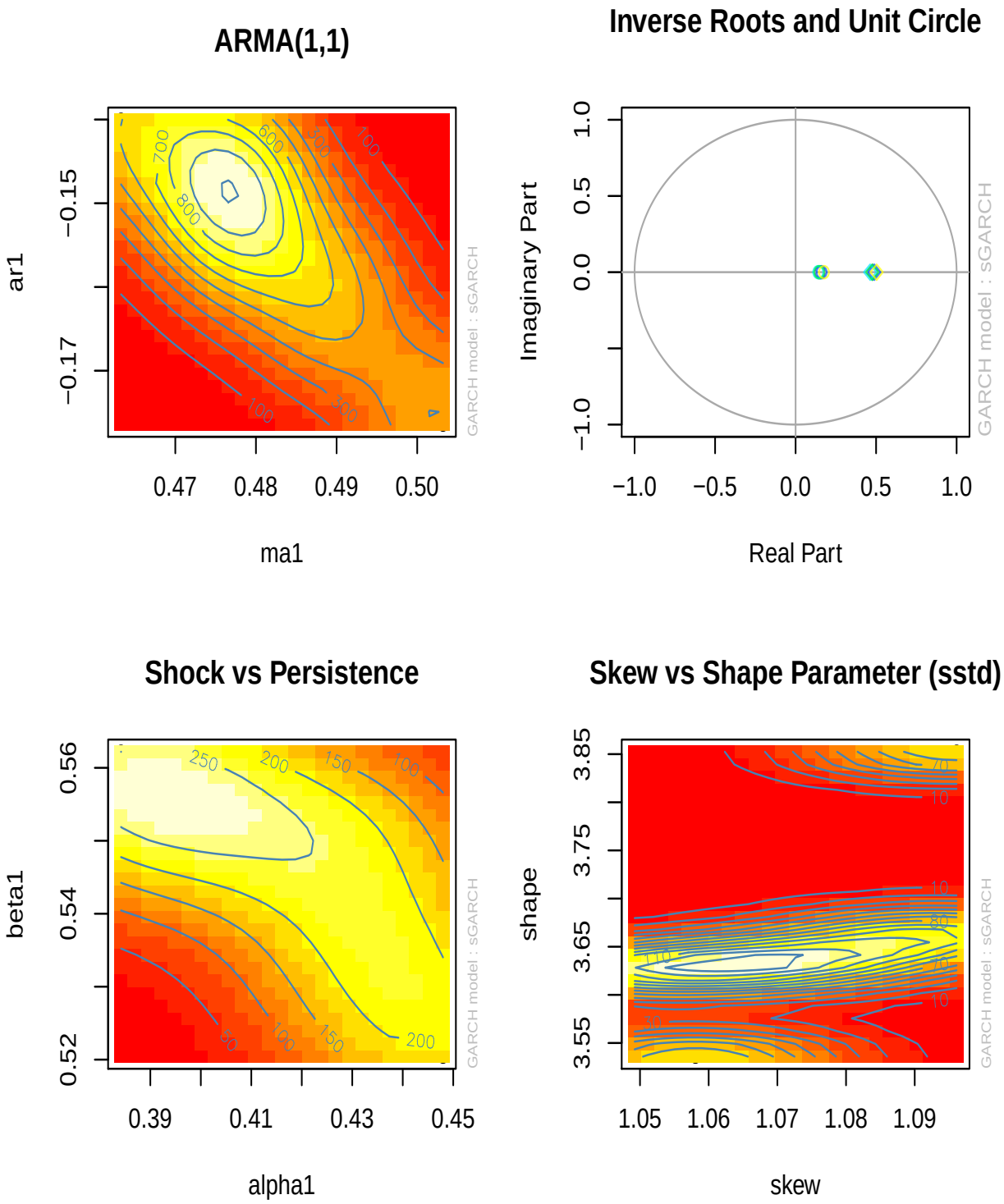

Figure 5.16: Volatility plot: NIGALSH.

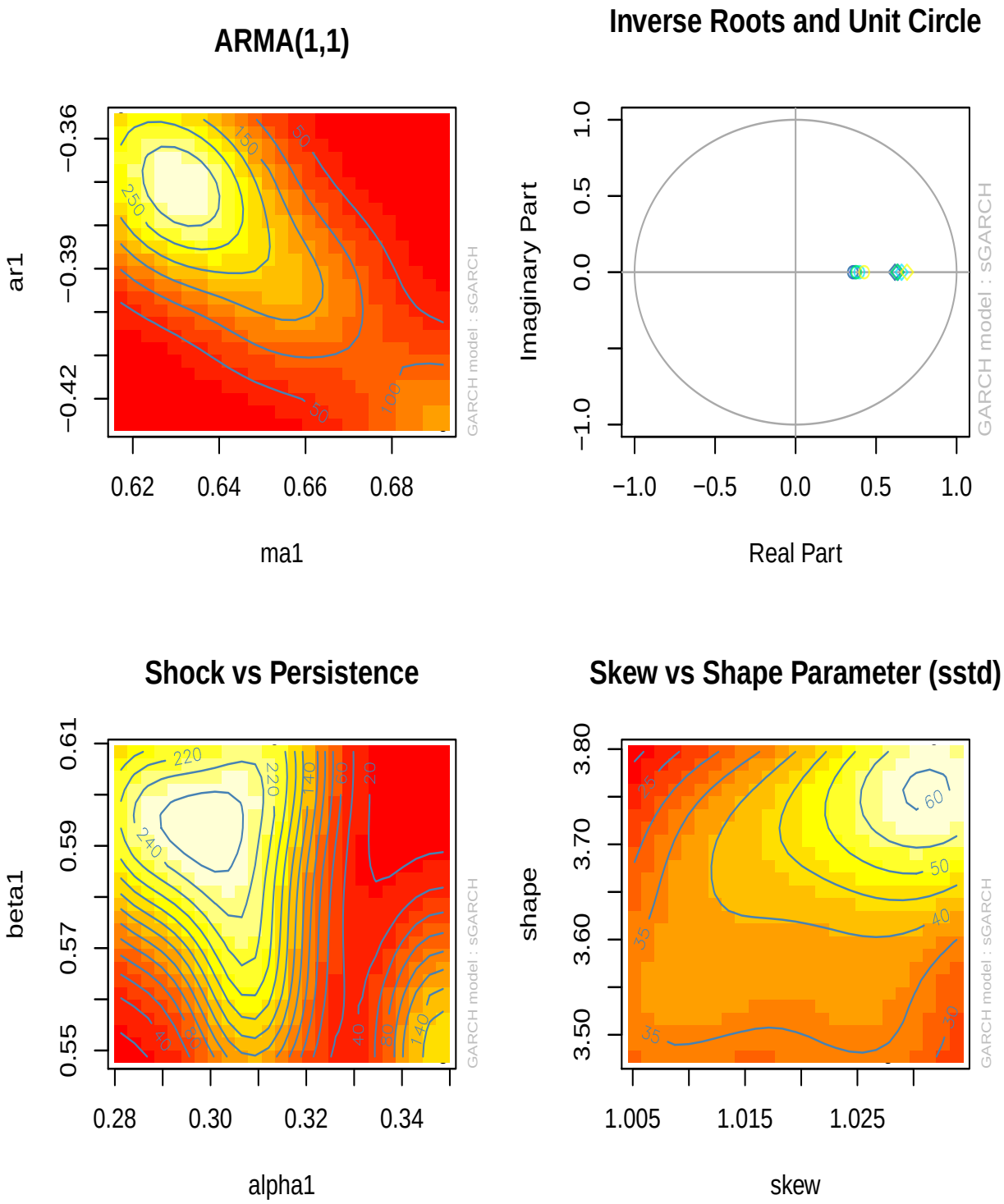

Figure 5.17: Volatility plot: NSE 20.

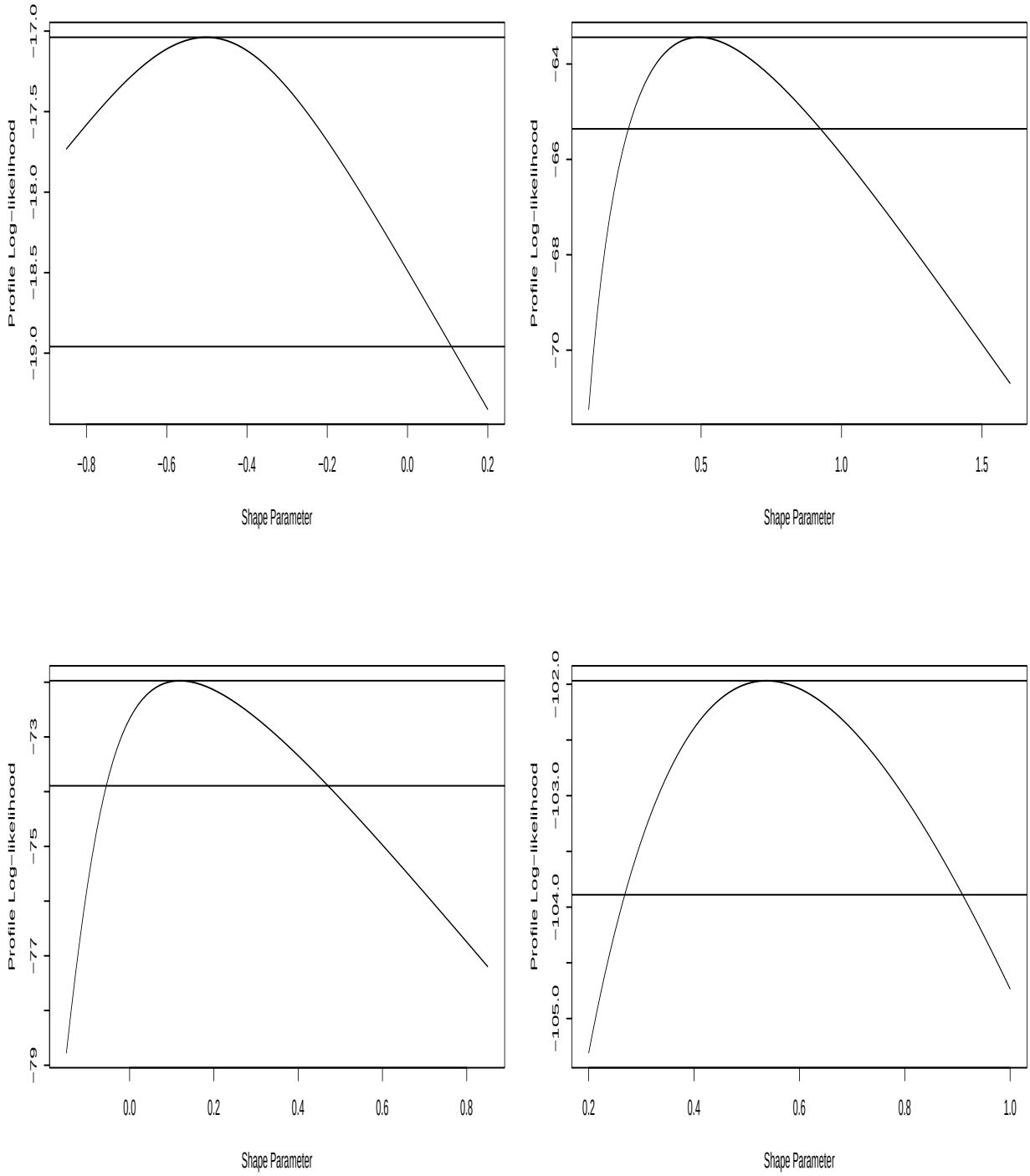

Figure 5.18: Profile log-likelihood intervals of ξ: JSE-ALSI (top left panel), NIGALSH (top right panel), EGX 30 (bottom left panel), NSE 20 (bottom right panel).

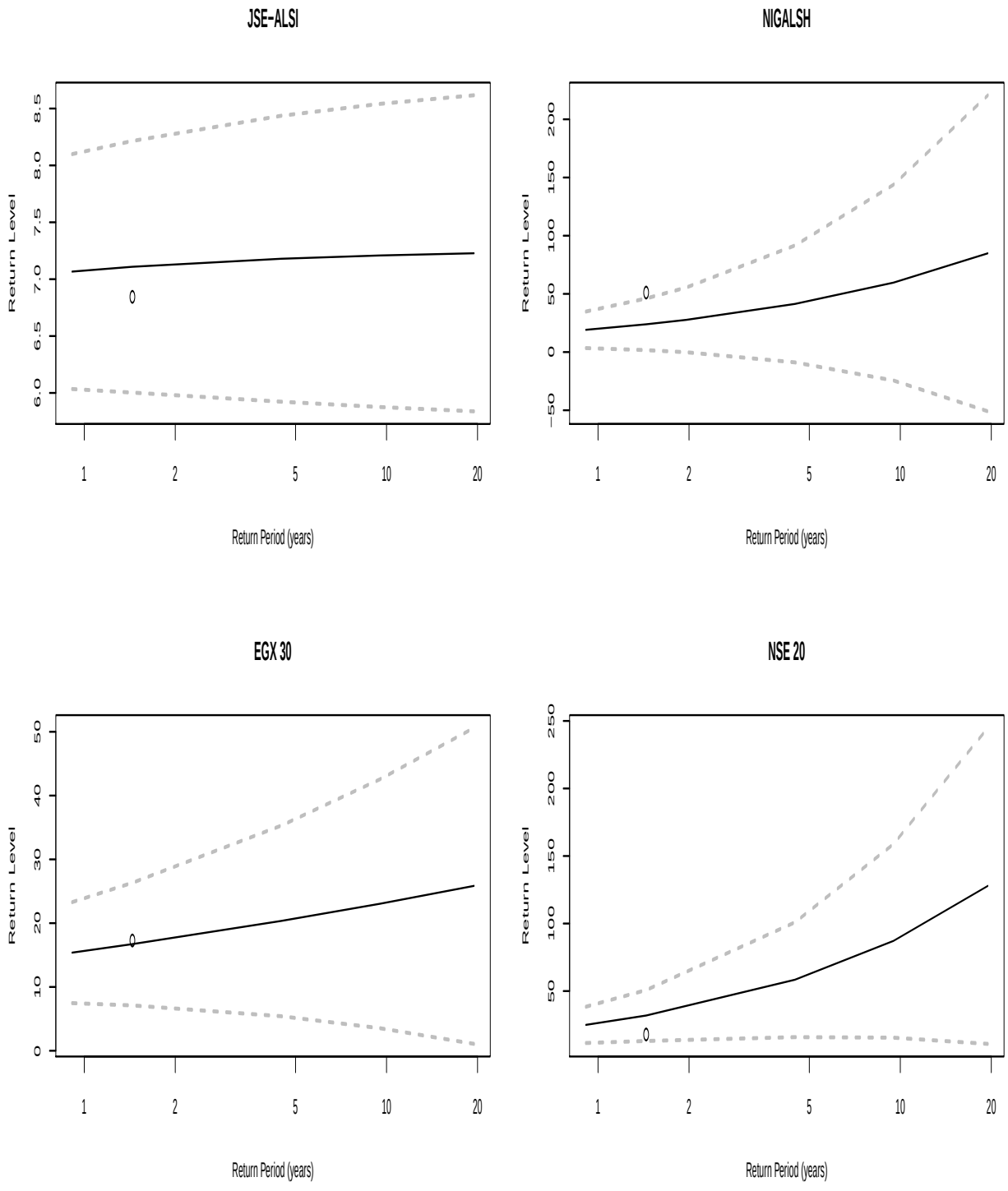

Figure 5.19: The return levels confidence intervals.