UNIVERSITY OF KWAZULU-NATAL
(PMB)

STATISTICS

Financial Modelling of Cryptocurrency: A Case Study of Bitcoin, Ethereum, and Dogecoin in Comparison with JSE Stock Returns

FORBES KASEKE

July, 2022

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

UNIVERSITY OF KWAZULU-NATAL
INYUVESI YAKWAZULU-NATALI

university of kwazulu-natal
School of Mathematics, Statistics and Computer Science
pietermaritzburg campus, south africa

Declaration - Plagiarism

I, Forbes Kaseke, declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowlegded as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then
   (a) their words have been re-written but the general information attributed to them has been referenced, or
   (b) where their exact words have been used, then their writing has been placed in italics and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source being detailed in the thesis and in the reference sections.

[redacted] 25/10/2022

Author: Mr. Forbes Kaseke Date

[redacted] 25/10/22

Supervisor: Prof. S. Ramroop Date

[redacted] 25/10/2022

Co-Supervisor: Prof. H.G. Mwambi Date

Disclaimer

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

ACKNOWLEDGEMENT

Firstly, glory to the almighty God for the grace that carried me through this journey. Secondly, I thank my parents, Mr and Mrs Kaseke, for their continuous support in seeing that I have the best education. I would also like to thank my supervisors, Prof. S. Ramroop and Prof. H. Mwambi, for their continuous support and encouragement in completing this thesis. Special thanks to Rutendo; love, thanks for the support and patience. My siblings, Rease, Rumbi, Fortue, and Ruva, you rock! You guys have been my day ones. To my sister Debra, your financial support was greatly appreciated whenever I was in need. Dr Batidzirai, for the pointers, whenever I would knock on your office door unannounced. I want to thank Dr Chinhamu for helping me overcome my biggest hurdle when I had a brain freeze. Thanks to Dr Mbulisi and Dr Chipunza for reading my first paper and giving me pointers. To Dr Odebiri, the encouragement and support you gave is appreciated. My colleagues from the Department of Statistics, your contributions are not forgotten. My fellow church on campus members, your prayers were heard!

ABSTRACT

The emergency of cryptocurrency has caused a shift in the financial markets. Although it was created as a currency for exchange, cryptocurrency has been shown to be an asset, with investors seeking to profit from it rather than using it as a medium of exchange. Despite being a financial asset, cryptocurrency has distinct, stylised facts like any other asset. Studying these stylised facts allows the creation of better-suited models to assist investors in making better data-driven decisions. The data used in this thesis was of three leading cryptocurrencies: Bitcoin, Ethereum, and Dogecoin and the Johannesburg Stock Exchange (JSE) data as a guide for comparison. The sample period was from 18 September 2017 to 27 May 2021. The goal was to research the stylised facts of cryptocurrencies and then create models that capture these stylised facts. The study developed risk-quantifying models for cryptocurrencies.

The main findings were that cryptocurrency exhibits stylised facts that are well-known in financial data. However, the magnitude and frequency of these stylised facts tend to differ. For example, cryptocurrency is more volatile than stock returns. The volatility also tends to be more persistent than in stocks. The study also finds that cryptocurrency has a reverse leverage effect as opposed to the normal one, where past negative returns increase volatility more than past positive returns. The study also developed a hybrid GARCH model using the extreme value theorem for quantifying cryptocurrency risk. The results showed that the GJR-GARCH with GDP innovations could be used as an alternative model to calculate the VaR. The volatile nature of cryptocurrency was also compared with that of the JSE while accounting for structural breaks and while not accounting for them. The results showed that the cryptocurrencies' volatility patterns are similar but differ from those of the JSE. The cryptocurrency was also found to be an inefficient market. This finding means that some investors can take advantage of this inefficiency. The study also revealed that structural breaks affect volatility persistence. However, this persistence measure differs depending on the model used. Markov switching GARCH models were used to strengthen the structural break findings. The results showed that two-regime models outperform single-regime models. The VAR and DCC-GARCH models were also used to test the spillovers amongst the assets used. The results showed short-run spillovers from Bitcoin to Ethereum and long-run spillovers based on the DCC-GARCH. Lastly, factors affecting cryptocurrency adoption were discussed. The main reasons affecting mass adoption are the complexity that comes with the use of cryptocurrency and its high volatility. This study was critical as it gives investors an understanding of the nature and behaviour of cryptocurrency so that they know when and how to invest. It also helps policymakers and financial institutions decide how to treat or use cryptocurrency within the economy.

KEYWORDS :Cryptocurrency, Bitcoin, GARCH, Extreme value theorem, Value at Risk, Johannesburg Stock Exchange

CONTENTS

LIST OF TABLES

LIST OF FIGURES

RESEARCH OUTPUT

The contents of this thesis are based on the following publications:

1. **Kaseke, F.**, Ramroop, S. and Mhwambi, H. (2021), A Comparison of the Stylised Facts of Bitcoin, Ethereum and the JSE Stock Returns, African Finance Journal 23(2), 50-64.

2. **Kaseke, F.**, Ramroop, S. and Mwambi, H. (2022), 'A comparative analysis of the volatility nature of cryptocurrency and JSE market', Investment Management and Financial Innovations

3. **Kaseke, F.**, Ramroop, S. and Mhwambi, H., Modeling Cryptocurrency Risks Using Hybrid GARCH Models *submitted to African Review of Economics and Finance Journal*

4. **Kaseke, F.**, Ramroop, S. and Mhwambi, H.,Estimating the volatility in Cryptocurrency and the JSE using Markov switching GARCH models. *submitted to Econometrics and Statistics Journal*

5. **Kaseke, F.**, Ramroop, S. and Mhwambi, H.,Investigating the spillover effects between cryptocurrency and Johannesburg Stock Exchange. *submitted to International Journal of Banking, Accounting and Finance*

6. **Kaseke, F.**, Ramroop, S. and Mhwambi, H.,A discussion on factors affecting the growth and adoption of cryptocurrency in Africa and the world. *To be submitted*

The author has also contributed to the following publications during the course of his PhD training:

1. Rusere, W. and **Kaseke, F.** (2021), Modeling South African stock market volatility using univariate symmetric and asymmetric garch models, Indian Journal of Finance and Banking 6(1), 1-16.

1. INTRODUCTION

1.1 Introduction

Investors are always looking for the best investment opportunities to guarantee them profits. The area most invested in is the stock market. Most companies are listed on the stock exchange to raise funds from interested investors. Investors then buy the shares in anticipation of making profits. The need for investment options is so high to the extent that there is a market for derivatives of various financial instruments such as stocks, currencies, commodities, and interest rates. These derivatives also allow investors to reduce their exposure to market movements beyond their control. The key to investment is the ability to forecast future prices and their risk. Making an intelligent investment involves knowing the risky nature of the asset before investing in it. A lack of proper risk analysis may lead to enormous losses for the investors or even liquidation of one's position in the worst-case scenario. For this reason, extensive studies on the stock market are continuously being done. In turn, investors use the knowledge gained to understand the intrinsic underlying risks of their investments.

One of the known early studies of stock market analysis is by French mathematician Jules Regnault, published in Regnault (1863). In the book, he attempts to use mathematics to predict stock markets. His work would then inspire other researchers, such as Alexander (1961) who tested the theory that speculative assets follow the random walk model. The result confirmed the tendency of stock market prices to follow a random walk over time and that once a move occurs, it tends to persist. Fama (1965) tests the same theory and also confirms this finding. The motivation for these studies was the idea that past information on securities contained information that could help predict future behaviour. Such studies are not limited to stocks but also other financially traded assets such as Gold and oil, to name a few.

The last decade has seen a new type of digital currency known as cryptocurrency. It was introduced to challenge the traditional fiat currency. However, unlike traditional currency, which is physical and printed at the government's sole discretion, cryptocurrency is entirely digital with an openly known supply. This new creation has gained various interests from both individual and cooperative investors. Despite its initial purpose as a payment system alternative, cryptocurrency has been labelled as an asset due to its rapid value growth. However, it is not all rosy, as cryptocurrency has been hit by unparalleled volatility, casting doubt on its role as an asset. It is now more of a speculative asset.

The first cryptocurrency was Bitcoin, introduced in 2008 by Nakamoto (2008). Other cryptocurrencies would later be created to rival Bitcoin. At the time of its creation, the purpose of Bitcoin was to be a digital currency meant for transactions. However, it has grown to be more than just a currency. Various studies have been carried out on the nature or class of Bitcoin. There is an ongoing debate about whether Bitcoin meets the conditions of being a currency or if it is instead an asset which is speculative or a store of value like Gold. Such a study is by Glaser et al. (2014). The study focused on the user's intentions with Bitcoin. The results showed that new users of Bitcoin use it as an asset rather than a medium of exchange to purchase and sell goods. Other studies focused on the hedging properties of cryptocurrencies. Shahzad et al. (2019) find that Bitcoin and Gold have weak hedging capabilities in the BRICS (Brazil, Russia, India, China and South Africa) markets. While most studies support hedging with cryptocurrencies, Lahmiri et al. (2018) argues that it cannot be considered suitable for hedging purposes due to its risky nature. Like with all other assets, there is a need to understand the properties of cryptocurrencies. Knowing these properties is essential because financial models are based on them. The ideal model should mimic the statistical properties of data such that if one were to forecast future values, the properties should be similar to those of the original series. The importance of knowing the stylised facts is tackled by Cont (2001). He explains how different stylised facts exist in the same financial markets despite it being all financial data. He further highlights how these stylised facts invalidate our traditional time series models, such as the autoregressive integrated moving average (ARIMA) and other regression models. This invalidation is because these models are built based on certain assumptions, particularly the normal distribution suggested by the central limit theorem and the assumption of constant variance. While these models perform exceptionally well when used with data that adheres to these assumptions, they fail to fully explain the financial data's behaviour. These models fail because the financial data does not adhere to the assumptions. Using such models would lead to wrong forecasts and estimates. Simulations based on these models also show that their data patterns differ from those produced by historical financial data.

Examining the literature reveals extensive studies focusing on stock market data properties. For example, studies by Agbeyegbe (1994), Manahov and Hudson (2013), Delatte and Lopez (2013) amongst others, focus mainly on stock data properties. This revelation is not surprising because stock markets are the most traded and invested. Models have been developed to mimic these properties. For example, the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model was modified from the Autoregressive Conditional Heteroscedasticity (ARCH) model to capture properties such as volatility clustering and heteroscedasticity. However, because some data exhibits leverage effects, models such as the Exponential Generalized Autoregressive Heteroscedasticity model (EGARCH) and the GARCH model of Glosten, Jagannathan, and Runkle which is known as the GJR-GARCH were also developed. With the growth of cryptocurrency and an ever-increasing number of users, it is necessary to identify its properties and develop models that are better suited for it.

In current literature, cryptocurrency is modelled using models used for other financial data. The most used models are the GARCH-type models, as seen in studies by Glaser et al. (2014), Dyhrberg (2016), Bouri, Azzi and Dyhrberg (2017) and Katsiampa (2017). These models often assume the student t and the general error distribution to model the error distribution. Multivariate models such as the Multivariate GARCH (MGARCH) have also been used by Katsiampa et al. (2019) and Cebrián-Hernández and Jiménez-Rodríguez (2021) amongst others. The models have proven adequate for capturing cryptocurrency volatility, but there is still room to improve these models to account for all stylised facts.

Regarding risk modelling, the same models have been combined with the value at risk measure (VaR) in quantifying the risk of cryptocurrencies. These models have been successful in quantifying the risk in stock markets and are, therefore, the go-to models for cryptocurrency.
Other researchers have investigated stylised facts that may cause bias in estimates if ignored. Thies and Molnár (2018) focus on the presence of structural breaks in Bitcoin. The study detected time-series change points and split them into segments. The segments identified several regimes with positive average returns and one with negative average returns. Across regimes, higher volatility is associated with higher average returns, except for the most volatile regime, which is the only regime with negative average returns. Except for the most volatile regime, the other regimes indicate inverse leverage property. The existence of structural breaks suggests the use of regime change models. These models were applied by Ardia, Bluteau and Rüede (2019) who found strong evidence of regime changes in the GARCH process and showed that Markov switching GARCH (MSGARCH) models outperform single-regime specifications when predicting the VaR of Bitcoin.

Because cryptocurrency exists in an integrated market, there may be co-movements between it and other financial assets. Shocks in one market can easily spread to another. This dependency motivates our inclusion of the Johannesburg

Stock Exchange (JSE) returns in this study. Other researchers have studied the spillover effect between cryptocurrency and stock markets. For example, Uzonwanne (2021) uses the multivariate Vector autoregressive moving average GARCH (VARMA-AGARCH) model to model the transmission mechanism of mean return, return spillovers, and volatility spillovers between cryptocurrency and five major stock exchanges. The results showed that there are bi-directional and, in some cases, unidirectional spillover effects of market shocks. Frankovic et al. (2021) using value-weighted cryptocurrency-linked stocks (CLS) return and volatility index studies, Bitcoin, Ethereum, Litecoin, and Ripple and 31 CLS listed on the Australian Securities Exchange. The results reveal significant unidirectional return spillover and weak volatility spillover from the cryptocurrency market to the CLS. A study that used the JSE, Gopane (2021) used the EGARCH model to investigate spillover effects spillovers between the JSE, Bitcoin and the USD/ZAR exchange rate. The results showed independence between Bitcoin and the South African stock market. There was also a bidirectional shock transmission between Bitcoin and USD/ZAR in the mean returns, but not variance. Lastly, there was a bidirectional volatility spillover in the mean and variance between the JSE stock and the United States dollar/ South African rand (USD/ZAR) markets.

Another area of concern is the adoption of cryptocurrency in Africa and the world at large. Due to it being technology-based and virtual, people from developing nations are not better placed to access cryptocurrency. Arias-Oliva et al. (2019) looks into the factors influencing adoption in Spain. The results showed that the performance expectancy for a given cryptocurrency was the most crucial factor for its success, while the risk was not significant. The literature is deprived of in-depth studies that focus on adoption in Africa. The existing ones, such as Walton and Johnston (2018) and Jankeeparsad and Tewari (2022) focus on the attitude and less on the infrastructure and risk evaluation. Foka Nzaha et al. (2022) uses partial Least Square-Structural Equation Modeling (PLS-SEM) to show that compatibility, relative advantage, and attitude towards cryptocurrency are the positive influences for adoption. In another study by Vincent and Evans (2019), the results revealed that cryptocurrency, internet usage, and mobile subscriptions have a significant positive relationship with financial inclusion and financial sector development. This result suggests that countries with higher internet usage and mobile subscriptions are better adapted to cryptocurrency.

1.2 Emergence and advances in cryptocurrency

The first incidence of cryptocurrency occurred in 2008 when Nakamoto (2008) published a white paper introducing the idea of a new digital currency that would be an alternative to the traditional fiat currency. The paper argues for the need for a decentralised currency not controlled by any institution such as governments or central banks but people-led and free from institutional manipulation. The paper emphasizes how conducting electronic transactions requires institutions to serve as the intermediators in the current system. Although it appears efficient, the current system has the flaw of being trust-based, making truly non-reversible transactions impossible. Institutions cannot ignore any disputes as they are transaction intermediators. Such disputes result in additional expenses borne by the parties, on top of the transaction costs already in place. Small transfers are impossible due to these fees because the minimum charges could potentially be greater than the amount transferred. Additionally, more information may be required to use these intermediary organisations than the users are prepared to provide. This is concerning since some of the data may be misused.

In order to overcome this trust-based method, Nakamoto (2008) suggests a trustless mechanism based entirely on cryptography proof. Using this method, the two agreeing parties can make a computationally irreversible transaction without needing a third party. Such mechanisms would protect both the seller and the buyer from fraud. The first Bitcoin was created a year later, in 2009, with Nakamoto receiving the first coin. Other coins started to be created afterwards. As of May 2022, over 18 000 different cryptocurrencies with a combined market capitalisation of more than 2 trillion dollars were listed on CoinMarketCap (n.d.). The number of coins far exceeds those listed on CoinMarketCap (n.d.) as it lists specific projects after vetting them. Bitcoin is the largest coin by market cap, worth around 40% to 45% of the total market cap, followed by Ethereum, which fluctuates from around 12% to 18% of the total market cap. An visual depiction of the market capitalisation in cryptocurrency is shown in Figure 1.1 below.

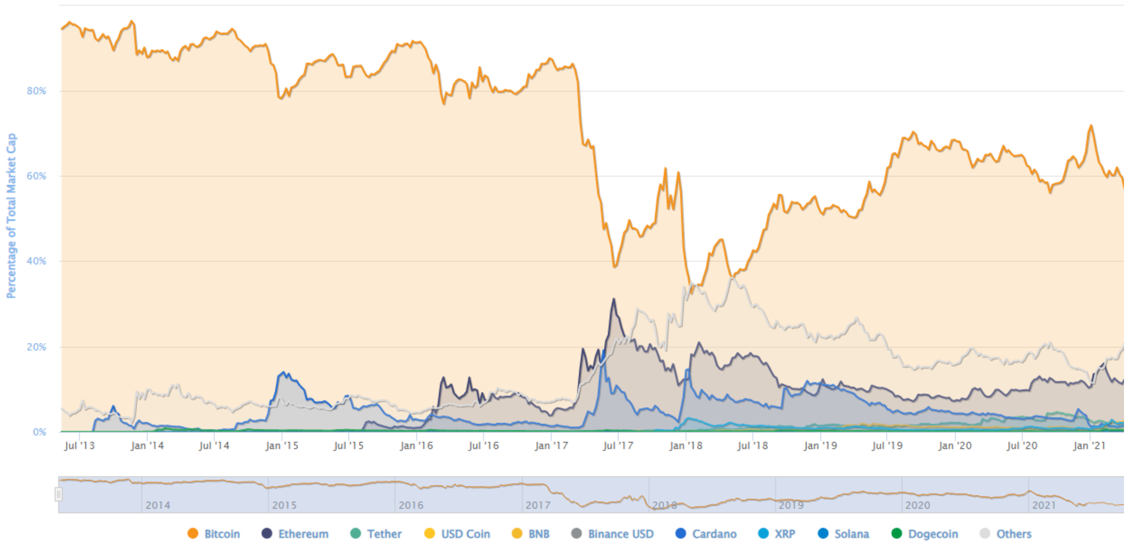

Figure 1.1: Market capitalisation of cryptocurrency as per May 2022: Source coinmaket-cap.com

1.2.1 *How cryptocurrency works: Example using Bitcoin*

In simple, understandable terms, Bitcoin can be described as an electronic cash system. Nakamoto (2008) defines Bitcoin as a chain of digital signatures created on the Bitcoin network. The Bitcoin network is a collection of multiple nodes, which are computers linked to the network, contributing computational power. Each user has two unique keys: a private one known to the user only and a public key derived from and linked to the private key. The public key can be shared with anyone without compromising the owner's assets. To initiate a transaction from user A to user B, user A must possess user B's public key. That is to say, the receiving party shares their public key. User A then notifies the network of the intent to send 1 Bitcoin to user B. He does this by signing a hash function. A harsh function is a mathematical function that converts a numerical input to one unique numerical output. This function will have user A's private key and user B's public key. The network then has to verify this transaction before accepting it as valid. Validation occurs through the mining of blocks. Here, the miner is just a computer linked to the Bitcoin network. The miners compete to solve the mathematical problems that accompany the transaction. The first miner to solve the problems notifies the other network members, and the nodes verify if the solution is correct. If the majority of the network, 51% verifies the proposed solution as correct, the transaction is deemed valid, and the block is closed and added to other completed blocks, creating the chain of blocks. Hence the name "blockchain." After verification, user A no longer owns the Bitcoin but is now owned by user B. The miners are rewarded for providing computing power by giving them bitcoins that are created during the mining process.

One of the stated problems is the prevention of double-spending. For double spending to be prevented, there is a need for a way for the payee to know that the previous owners did not sign any earlier transactions, thereby creating other transactions with already spent Bitcoin. A timestamp saver is used to avoid double-spending. It works by taking a hash of a block of items and publishing them throughout the nodes. The timestamp proves that the data must have existed at the time. Any new timestamp includes the previous timestamps in its hash so that any upcoming timestamp enforces the ones before. In essence, if one tries to double spend, the system ignores the new transaction and only counts the earlier transaction (Nakamoto 2008). In general this is how cryptocurrency works but others may differ depending on the technology used.

1.3 Financial returns

Every investor's goal is to make a profit on their investment. Suppose K is the amount of money invested at time t, then at any time greater than t the value of the investment may change to $K + N$. Here N can be positive, signifying a profit, or it can be negative, signifying a loss on investment. The amount gained or lost is the financial return, or simply the return on investment. It is more common to work with these returns in financial modelling than with actual prices. Reasons being the statistical advantages of returns over prices (Tsay 2014). The main advantages of returns can be summarised as follows:

- Prices are in general not stationary over time which makes them violate most time series data model assumptions
- Asset prices are often quoted in different currency denominations which means comparisons become biased. However, returns are scale free removing the currency difference bias (De Vries and Leuven 1994).
- Campbell et al. (1997) states that there are advantages on the theoretical and empirical side as returns have more attractive statistical properties over prices such as stationarity and others to be discussed in the following section.

Assuming no dividends are paid, and that the price of an asset at time t is P_t. Then the simple net return on investment, R_t is defined as:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1 \tag{1.1}$$

This is known as a single period return or simple period return. In finance when an investment is made, one waits for a certain period before he gains or loses money. That waiting period is known as the holding period. The period varies with the intended purposes but it is usually defined in days, weeks, months or years. Whatever the case might be, the single period return is for the one time period relative to the time scale used. If the investment is held for multiple time periods then we have a multi-period return that is defined as:

$$
\begin{aligned}
1 + R_t[k] &= \frac{P_t}{P_{t-k}} = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times ... \times \frac{P_{t-k+1}}{P_{t-k}} && (1.2)\\
&= (1 + R_t)(1 + R_{t-1})....(1 + R_{t-k+1}) && (1.3)\\
&= \prod_{j=0}^{k-1}(1 + R_{t-j}) && (1.4)
\end{aligned}
$$

From the final representation in equation (1.4), the multi-period return is a just the product of the single period returns. For this reason the multi-period return is also referred to as a compound return. Despite having better properties over the price data, returns still lack some of the desired statistical properties. One of the downside is that the multiperiod returns are a product of single period returns. To overcome this simple returns are transformed to obtain log returns r_t which are defined as:

$$
r_t = \ln(1 + R_t) = \ln \frac{P_t}{P_{t-1}} = \ln p_t - \ln p_{t-1} \qquad (1.5)
$$

These have corresponding multi-period returns that are defined as:

$$
\begin{aligned}
r_t[k] &= \ln(1 + R_t[k])\\
&= \ln(1 + R_t)(1 + R_{t-1})......(1 + R_{t-k+1})\\
&= \ln(1 + R_t) + \ln(1 + R_{t-1}) + ... + \ln(1 + R_{t-k+1})
\end{aligned}
\qquad (1.6)
$$

The biggest gain in using the log returns is that the multi-period return is just a simple summation of the individual simple log returns and not multiples as with multi-period returns. This gives an advantage in the statistical application of the log returns. According to Quigley and Ramsey (2008), the benefit of using log returns is that of standardisation, which allows all variables to be measured in a comparable metric, allowing the core evaluation of two or more variables with different currency bases. Bearing in mind that simple returns are generally assumed to be log-normally distributed. If they are independent and identically distributed (i.i.d.), it follows that the log-returns are i.i.d. normally distributed, allowing great statistical freedom. Having the normality distribution allows the application of most statistical methods, which need the normality or close to normality assumption. Another advantage is that the product of normally distributed variables is not normal; but instead the sum is. With log returns being additive for the multi-period return, the normality distribution assumption remains intact.

1.4 Stylised facts of financial data

For any data, some characteristics are well documented for that particular data. These characteristics are similar across different assets within that asset class. For financial data, the stylised facts are well documented. According to Cont (2001) these stylised facts are:

- The prices of financial data are non-stationary. They tend to fluctuate as per asset supply, inflation and supply.
- Leptokurtosis-financial data tends to have thicker than normal tail distribution
- Price variations tend to have no autocorrelation or have small low order autocorrelations
- Autocorrelations of the squared price returns. Squared returns or absolute returns are generally strongly autocorrelated.
- The return series is not i.i.d but tends to show minor lower order serial correlation.
- Volatility clustering whereby large absolute returns tend to appear in clusters. Likewise small absolute returns occur in clusters.
- Leverage effects- the asymmetry of the impact of past positive and negative values on the current volatility. Given shocks of the same magnitude,negative returns tend to increase volatility by a larger amount than positive returns.

While we know that these stylised facts hold for most financial returns, such as stock returns, we need to investigate whether these facts also hold in cryptocurrency. Understanding stylised facts is the backbone of modelling. This is because models should be designed to capture stylised facts. Capturing the stylised facts results in more accurate models as they can mimic the data in its natural form. In this regard, this thesis will have a chapter that will be dedicated to finding the stylised facts about cryptocurrency.

1.5 Volatility

Various measures are used to measure the riskiness of a financial asset. The most commonly used measure of asset risk is its volatility. Sheppard (2013) defines volatility as simply the standard deviation. While volatility is known, it is, however, not directly observable. Hence, the observable prices are used to estimate volatility. It, therefore, means there is a need for models that can efficiently capture and describe the volatility process based on the price movements. Efficiency is crucial as understanding the volatility enables more accurate forecasts, which will arm market players and investors with knowledge to make informed decisions.
The characteristics of volatility include:

- It is not directly observable
- It tends to show in clusters
- It evolves over time in a continuous manner
- It has a leverage effect, were negative shocks increase it than positive shocks. However for some assets the inverse occurs. In what is termed inverse leverage effect
- It does not diverge to infinity i.e. it is mean reverting.

The study of volatility as a measure of an asset's riskiness has two main advantages. Firstly, it is simply the standard deviation, so it is in the original units, as opposed to variance, which is in units squared. Secondly, it provides a simple way to calculate an asset's risk and can be used in portfolio allocation.
Volatility in cryptocurrency is attributed to the following

- Decentralization-Being decentralised means it means the trading space is full of unknown buyers and sellers.
- Speculation-given the lack of understanding and no real backing, people speculate on the price going up or down.
- Lack of a Tether that is linked with cryptocurrency. Unlike stocks whose value is based on the companies they're attached to, cryptocurrency, in comparison, is not backed by a government or a commodity like gold or silver. Its value is what the trading community decides it is.
- Investor Confidence (or Lack Thereof)
- Security Issues- being an online currency it means the assets can be targeted for theft. This theft can occur via hacking or fake websites that swindle users from their funds.

- Struggle for widespread adoption- Merchants and governments alike are struggling to back cryptocurrency
- Lack of intrinsic value hence they are prone to speculation bubbles.

1.6 Motivation

Cryptocurrency is a relatively new financial asset. Because it is new, more is yet to be known concerning its properties. Likewise, more is still to be known about the models that best model the asset. Literature abounds with research on financial assets and their models. However, the main focus has been on the stock markets, with cryptocurrencies receiving less attention. Of the few studies on modelling cryptocurrency, most focus on applying the traditional models used for other financial data such as stock market data and exchange rates.
Risk measurement is also an essential factor in investing. Investors are interested in extreme case events that may negatively impact their investments. The studies reviewed show varied tail behaviour in cryptocurrency compared to traditional financial data. These findings mean that the VaR obtained with GARCH models can be improved by incorporating tail distributions more appropriate for modelling cryptocurrency. Another issue is the need to study the relationship between cryptocurrency and the African financial markets. Current studies focus on the relationship with developed markets. This leaves developing countries relying on findings derived from developed countries. This generalisation may not be accurate, as we know the states of the developing markets differ from those of developed markets. Against this background, the study seeks to focus our attention on cryptocurrency, but with a developing market perspective.

1.7 Objectives

The objectives of this study were to:

- To perform exploratory data analysis to identify the stylised facts that exist in cryptocurrency
- To make a comparative study on the nature of volatility between cryptocurrency and the JSE market
- To develop a hybrid GARCH model that incorporates Extreme value theorem distributions. This model will be used to quantify the risk of cryptocurrency.
- To investigate the spillover effects between cryptocurrencies and the JSE stock market using the VAR and the DCC-GARCH model.

1.8 Structure of the thesis

This thesis consists of twelve chapters that provide the background of the research, the discussion of models to be used, diagnostic tests, and research articles derived from the study. This thesis contains published and submitted but yet-to-be-published peer-reviewed material in different journals. One important note to add is that from Chapters 6 to 11. Data, introductory paragraphs, and methods will be repeated, despite having been discussed in earlier chapters. This was done to give the reader a continuous flow without having to refer back. Another reason is also due to the thesis being in article format. A brief discussion of the chapters is given below.

1. Chapter 1 introduces the reader to the study. It discusses the coming into being of cryptocurrency and gives the reader an example of how it works.
2. Chapter 2 presents the univariate models used in this thesis. The models include various GARCH-type models and also change point detection methods.
3. Chapter 3 presents the multivariate models used in this thesis. The models include the Vector Autoregressive models and the DCC-GARCH model.
4. Chapter 4 presents the Extreme Value Theorem. It focuses on the Generalised Extreme Value Distribution and the General Pareto Distribution.
5. Chapter 5 discusses the methodology. This methodology does not consider the models as they are discussed in the previous chapters. It, however, focuses on the diagnostics tests to be employed.
6. Chapter 6 is a research article on the stylised facts of cryptocurrency. This chapter has been published in the African Finance Journal
7. Chapter 7 is a research article on development of a Hybrid GARCH model to quantify the Value at Risk for cryptocurrency using the extreme value theorem
8. Chapter 8 is a research article comparing the volatility nature of cryptocurrency and the JSE
9. Chapter 9 is a research article on estimating the volatility of cryptocurrency and the JSE using Markov switching GARCH models
10. Chapter 10 is a research article on the spillover effects between cryptocurrency and the JSE.
11. Chapter 11 is a discussion on factors that hinder cryptocurrency adoption
12. Chapter 12 is the general discussion and conclusion

2. UNIVARIATE MODELS

This chapter discusses the univariate models that were used in this thesis. Firstly The study discusses the univariate volatility models which are the GARCH-type models. Secondly it looks at change point detection methods. These will be crucial in detecting the presence of structural breaks in the data series. Lastly it looks at Markov switching models.

2.1 Univariate volatility models

Traditional time series models such as ARMA are built around specific assumptions. These assumptions included the normality assumption and the assumption that the conditional variance is constant. On the other hand, financial data does not adhere to these assumptions. It tends to be more volatile, and the variance of the error terms is not equal, meaning that it has a non-constant variance. The traditional time series models cannot capture such effects as they are designed to capture constant variance error terms, i.e. (Homoscedasticity). Using these models to capture non-constant volatility data (heteroscedastic data) results in biased standard errors and, consequently, confidence errors. In this section, models that cater for non-constant variance are discussed.

Following the notation of Tsay (2014), the basic volatility model consists of two parts, one that caters for the conditional mean and the second one that caters for the variance of r_t. The conditional equation for the mean is given as

$$\mu_t = E(r_t|F_{t-1}), \tag{2.1}$$

and for the variance by

$$\sigma_t^2 = Var(r_t|F_{t-1}) = E[(r_t - \mu_t)^2|F_{t-1}]. \tag{2.2}$$

Here F_{t-1} is the information set at time $t-1$ and is made up of functions of past returns and a_t is the shock or innovation of an asset return at the time t. In general, r_t follows an ARMA(p,q) model such that $r_t = \mu_t + a_t$. Here μ_t is given as

$$\mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} - \sum_{j=1}^{q} \theta_j r_{t-j}. \tag{2.3}$$

The order of the ARMA model will depend on the frequency of the data series. In the presence of an explanatory variable, then $r_t = \mu_t + a_t$ has u_t as

$$\mu_t = \phi_0 + \sum_{i=1}^{p} \beta_i x_{i,t-i} + \sum_{i=1}^{p} \phi_i y_{t-i} - \sum_{j=1}^{q} \theta_j a_{t-j}, \tag{2.4}$$

where $y_{t-i} = r_{t-i} - \phi_0 + \sum_{i=1}^{p} \beta_i x_{i,t-i}$ is the return series without the effect of the explanatory variables and $x_{i,t-i}$ are explanatory variables available at time $t-j$.

Combining the two models gives the updated equation for the conditional variance so that it becomes

$$\sigma_t^2 = Var(r_t|F_{t-1}) = Var(a_t|F_{t-1}), \tag{2.5}$$

where the square root σ_t is the volatility. It is this σ_t^2 that the volatility models will try to capture. The manner under which σ_t^2 evolves distinguishes one model from another.

2.1.1 The ARCH model

Engle (1982) introduced the model that would revolutionalise the modelling of volatility. He introduced the Autoregressive Conditional Heteroscedasticity model. The idea behind the model was that the current occurrences depended on past information, and the model should, at the current time, incorporate the past shocks. He idealised that the shocks of the returns are serially uncorrelated but dependent, and that this dependence could be modelled by a simple quadratic function of lagged values. The general structure for the ARCH(m) model is as follows.

$$a_t = \sigma_t \epsilon_t, \tag{2.6}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + ... + \alpha_m a_{t-m}^2, \tag{2.7}$$

where ϵ_t is a sequence of iid random variables with mean zero and variance if 1. The conditions that must hold for the model to be valid are that $\alpha_0 > 0$, while $\alpha_i \geq 0 \ for \ i > 0$. The error term ϵ_t can follow any of the distributions, such as the student t, skewed student t and the general error distribution. The student t is given as follows:

The probability density function (PDF) of the standardized Student-t distribution is given by:

$$f(\epsilon, \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{(\nu-2)\pi\Gamma(\frac{\nu}{2})}} \left(1 + \frac{\eta^2}{\nu-2}\right)^{-\frac{\nu+1}{2}} \tag{2.8}$$

where Γ is the Gamma function and ν is the degrees of freedom. To ensure the existence of the second order moment exists we require $\nu > 2$.

For the skewed Student t distribution the pdf is as follows:

$$g(\epsilon|\xi,\nu) = \begin{cases} \frac{2}{\xi+\frac{1}{\xi}} \varrho f[\xi(\varrho\epsilon+\bar{\omega})|\nu] & \text{if } \epsilon < -\bar{\omega}\ /\varrho. \\ \frac{2}{\xi+\frac{1}{\xi}} \varrho f[\xi(\varrho\epsilon+\bar{\omega})/\xi|\nu] & \text{if } \epsilon \geq -\bar{\omega}/\varrho. \end{cases} \tag{2.9}$$

Where f is the pdf of the standardized student t distribution, ν is the degrees of freedom and ξ is the skewness parameter. ϱ and $\bar{\omega}$ are defined as:

$$\varrho^2 = (\xi^2 + \frac{1}{\xi^2} - 1) - \bar{\omega}^2, \quad \bar{\omega} = \frac{\Gamma[(\nu-1)/2]\sqrt{\nu-2}}{\sqrt{\pi}\Gamma(\nu/2)}(\xi - \frac{1}{\xi})$$

The PDF of the standardized generalized error distribution (GED) is given by:

$$f(\epsilon,\nu) = \frac{\nu \exp^{-\frac{1}{2}|\epsilon/\lambda|^{\nu}}}{\lambda 2^{(1+1/\nu)}\Gamma(1/\nu)}, \quad \lambda \equiv \left(\frac{\Gamma(1/\nu)}{4^{1/\nu}\Gamma(3/\nu)}\right)^{\frac{1}{2}} \tag{2.10}$$

The ARCH model introduced two significant advantages over traditional models. First, it removed the unrealistic assumption of constant variance for financial data and allowed error variance to evolve. The second is that it allows the capture of the clustering effect of volatility, as it allows past volatility to depend on the past volatility in its structure, thereby being a more realistic model.

Despite solving the significant problems associated with the traditional models, the ARCH model has its setbacks. The first setback is that it is not a parsimonious model, requiring many parameters to capture the volatility process correctly. Secondly, it fails to capture some of the stylised effects of financial data. These are the asymmetry effects of volatility where it treats both positive and negative shocks equally, yet in reality, the effects are not equal. It is also likely to violate the non-negativity constraints because many parameters increase the likelihood of some of them being negative.

2.1.2 The GARCH model

Seeing the limitations of the ARCH model, Bollerslev (1986) proposed the GARCH model. This model would deal with the major problem of the ARCH model, which requires many parameters to capture the volatility process, leading to over-fitting. As a result, the GARCH model is less likely to violate the nonnegativity constraints. Given that a_t follows the GARCH(m,s) model where $a_t = \sigma_t \epsilon_t$, the conditional variance process for the GARCH model is defined as:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2, \tag{2.11}$$

whereϵ_t is a sequence of Gaussian white noise with a mean of 0 and a variance of 1. The model parameters should have constraints to ensure model stability and non-negative variance. $\alpha_0 > 0$ and $\alpha_i \geq 0, \beta_j \geq 0$ so that we have positive volatility. Another constraint is that $\sum_{i=1}^{max(m,s)}(\alpha_i + \beta_i) < 1$. This implies that the unconditional variance of a_t is finite, i.e. it is covariance stationary, whilst the conditional variance σ_t^2 evolves over time. The parameter α_i for i = 1,2,3..., n measures the reaction to market shocks, whereas the parameter β_i determines the persistence in volatility after a shock. Together with parameter α_0 determine the speed of mean reversion and the long-run GARCH volatility. ϵ_t can be Standard Normal or Standardized Student t or General Error Distribution. The GARCH model reduces to an ARCH(m) model if s=0.

To understand the GARCH model properties, it is represented in the form of an ARMA model of the a_t^2. This is done by letting $\eta_t^2 = a_t^2 - \sigma_t^2$ which can be re-arranged such that $\sigma_t^2 = a_t^2 - \eta_t^2$. Substituting $\sigma_{t-i}^2 = a_{t-i}^2 - \eta_{t-i}^2$ into the conditional variance 2.8 gives;

$$a_t^2 = \alpha_0 + \sum_{i=1}^{max(m,s)} (\alpha_i + \beta_i) a_{t-1}^2 + \eta_t - \sum_{j=1}^{s} \beta_j \eta_{t-j}, \tag{2.12}$$

is martingale difference series with a mean and covariance of zero. Using the conditional mean of the ARMA model Tsay (2005) shows that,

$$E(a_t^2) = \frac{\alpha_0}{1 - \sum_{i=1}^{max(m,s)} (\alpha i + \beta_i)}. \tag{2.13}$$

This is given that the denominator is none zero and positive. By using the simplified GARCH(1,1) Tsay (2005) further shows that the unconditional kurtosis of the GARCH is given by

$$k = \frac{3(1 - \alpha_1 + \beta_1)^2}{1 - 2\alpha_1^2 - (\alpha_1 + \beta_1)^2} > 3, \tag{2.14}$$

given that $1 - 2\alpha_1^2 - (\alpha_1 + \beta_1)^2 > 0$.
As a result, just like in ARCH models, a GARCH(1,1) process has a higher tail distribution than a normal distribution. The model offers a straightforward parametric function that can be used to explain volatility evolution

2.1.3 The exponential GARCH model (EGARCH)

To accommodate the asymmetric effects between positive and negative asset returns, Nelson (1991) introduced the exponential GARCH model. The model can be represented as follows:

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{m} \alpha_i \frac{|a_{t-i}| + \gamma_i a_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^{s} \beta_j \ln(\sigma_t^2), \tag{2.15}$$

where as before $a_t = \sigma_t \epsilon_t$. Using the structure of the model positive and negative a_{t-1} contribute different to the volatility. A negative a_{t-1} will contribute $\alpha_i(1+\gamma_i)|\epsilon_{t-i}|$ whereas a positive a_t will contribute $\alpha_i(1-\gamma_i)|\epsilon_{t-i}|$. This then means the γ_i determines the leverage effect. The normal leverage effect occurs with γ_i being negative while a positive γ_i would an inverse leverage effect. In terms of maintaining positivity of the volatility, no restrictions are required since the EGARCH is based on $\ln(\sigma_t^2)$. This means that even if the parameters are negative, the σ_t^2 will still be positive. However the model requires $|\beta_j| < 1$ to achieve covariance-stationarity.

2.1.4 GJR-GARCH model

Another model that captures the volatility asymmetry is the GJR-GARCH of Glosten et al. (1993). The volatility equation for this model takes the form:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} (\alpha_i + \gamma_i S_{t-i}) a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2, \tag{2.16}$$

$$S_{t-1} = \begin{cases} 1, & \text{if } a_{t-i} \leq 0, \\ 0, & \text{if } a_{t-i} > 0, \end{cases}$$

with $\alpha > 0, \beta > 0, \gamma > 0$ and $\alpha + \beta + 0.5\gamma < 1$. Based on the equation it can be seen that the impact of negative negative shocks is different from that of positive shocks. A positive shock means that S_{t-1} is 0 hence the equation reduces to the standard GARCH(1,1) i.e. it becomes:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} (\alpha_i) a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2. \tag{2.17}$$

While a negative shock increases the volatility by adding the term $\gamma \epsilon_{t-1}^2$. The resulting equation is shown below

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} (\alpha_i + \gamma_i) a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2. \tag{2.18}$$

2.1.5 Maximum likelihood estimation (MLE).

The commonly used method in estimating the parameters is the maximum likelihood method (MLE). In general given $L(\eta|a_1, a_2, ..., a_T)$ as the maximum likelihood function of the data and where η is the set if parameters to be estimated. The data $a_1, a_2, ..., a_t$ are assumed not to be independent such that their conditional density function is given by:

$$f(a_1, a_2, ..., a_t|\eta) = f(a_T|F_{T-1})f(a_{T-1}|F_{T-2})....(f(a_1)), \tag{2.19}$$

where F_t is the information set available at time t and f is the distribution of the error term ϵ_t. Then the likelihood function is

$$L(\eta|F_{t-1}) = f(a_T|F_{T-1})f(a_{T-1}|F_{T-2})....(f(a_1 f(a_1, a_2, ..., a_1|\eta), \tag{2.20}$$

$$= \prod_{t=1}^{T} f(a_t|F_{t-1}). \tag{2.21}$$

Solving the above likelihood results in the maximum likelihood estimates. However, due to the complexity of maximasing the likelihood, the log likelihood is used. The log likelihood is easier to work with and gives the same results as maximizing the likelihood. The log likelihood as as follows:

$$\ell(\eta|F_{t-1} = log[\prod_{t=1}^{T} f(a_t|F_{t-1})]. \tag{2.22}$$

The final estimation will then depend on the assumed error distribution such as the normal and the standardized Student t distributions.

2.2 Change point detection

Studies such as those by Lamoureux and Lastrapes (1990), Ardia, Bluteau and Rüede (2019) have shown that ignoring structural changes in the time series leads to biased estimates, which in turn lead to poor forecasts. Therefore, in the study of time series data, there is a need to be able to identify points of structural change. The structural change splits the series into disjoint segments. Each segment constitutes consecutive observations of similar magnitude or level of events. For example, there are periods of economic boom and crisis in finance, which are characterised by the same level of volatility or returns. The beginning of such periods can then be identified in the data using changepoint detection. Changepoint detection can be defined as a procedure used to identify points of change in the time series structure. The change can be a change in the mean or a change in variance, or both. To say a change has occurred at point B, the statistical properties of the series must have changed at point B. That is, given a series $y_t : t \in 1, ...n$ a changepoint is said to have occurred if the

statistical properties of $y_1, ...y_s$ differ from those of $y_{s+1}, ...y_n$ (Killick and Eckley 2014).
For a change in the mean of the series, the following occurs:

$$\mu_t = \begin{cases} \mu_i & \text{if } t \leq s, \\ \mu_{ii} & \text{if } t > s, \end{cases} \tag{2.23}$$

where $\mu_i \neq \mu_{ii}$. In general, it means the mean of the series from time less than s, should be significantly different from that of period $\geq s$.
For the change in variance we have,

$$\sigma_t^2 = \begin{cases} \sigma_i^2 & \text{if } t \leq s, \\ \sigma_{ii}^2 & \text{if } t > s, \end{cases} \tag{2.24}$$

where $\sigma_i^2 \neq \sigma_{ii}^2$. Like the mean, the variance change occurs if there is a significant change in the variance in the two sub-periods of the time series.

In this thesis, the changepoint detection in the variance is used. Various ways are available to detect the changes in the time series. These can fall into two forms: single-point detection, where only one changepoint is detected, or multiple changepoints, where multiple changepoints are detected depending on the tolerance level. The interest will be in finding structural changes in the time series; hence multiple changepoint methods are considered. The summary of multiple changepoints is given in the next section.

2.2.1 Multiple change points

The procedure of identifying multiple changepoints is not an easy one. As explained by Killick and Eckley (2014), as data sets increase in length, the number of possible solutions to the multiple changepoint problems increases combinatorially. Such increases require more complicated computations, which renders some methods unrealistic. Various algorithms have been presented to solve the multiple changepoint problems. This section briefly discusses some of the search methods used in multiple change point detection. The methods discussed are those used in the R package Changepoint by Killick and Eckley (2014). This are selected as the Changepoint package is to be used for the implementation.

The typical approach to identifying change points is through minimising the function:

$$\sum_{i=1}^{m+1} [C(y_{(\tau_{i-1}+1):ri})] + \beta f(m), \tag{2.25}$$

where C is the cost function for a segment considered and $\beta f(m)$ is the penalty function which helps guard against overfitting i.e. it is the multiple changepoint threshold point.

In the changepoint package, three multiple changepoint algorithms are implemented that minimize the function $\sum_{i=1}^{m+1}[C(y_{(\tau_{i-1}+1):ri})] + \beta f(m))$; binary segmentation (Erdman and Emerson 2008), segment neighborhoods (Auger and Lawrence 1989) and the pruned exact linear time (PELT) (Killick et al. 2012). These methods vary in the accuracy and in computational costs i.e some are exact and some are estimates. The methods are discussed below:

2.2.1.1 *Binary segmentation*

Developed from the work of Edwards and Cavalli-Sforza (1965), Scott and Knott (1974), and Sen and Srivastava (1975), Binary Segmentation is one of the most applied methods. As the name suggests, it works by identifying a change point along the time series, and if found, it splits the time series into two at the location of the change point. It repeats the same procedure iteratively on the split segments. It applies the single change point identification over and over. This iterative process continues until there are no more change points identified. In essence, the method extends the single-changepoint method. A single change point is initially identified across the whole data set. Mathematically we test if a point τ exists that satisfies:

$$C(y_{1:\tau}) + C(y_{(\tau+1):n}) + \beta < C(y_{1:n}).$$

The Binary Segmentation method has the advantage of being computationally efficient, resulting in an $O(nlogn)$ calculation. However, Killick et al. (2012) warns that this comes at a cost as it is not guaranteed to find the global minimum.

2.2.1.2 *Segment neigbourhood*

Introduced by Auger and Lawrence (1989), the method employs the dynamic programming search across the whole series to find m+1 change points. The number of maximum change points is first noted, and the algorithm computes the cost functions for all possible segments. It is an exact method and incorporates an abiotrophy penalty $\beta f(m)$. Given that the maximum number of change points is Q, the algorithm has a downside of high computational; cost, $O(Qn^2)$. As the number of observations increase, the number of change points increases linearly such that Q = O(n) and the method will have a computational cost of $O(n^3)$.

2.2.1.3 Pruned exact linear time (PELT) method

The Pruned Exact Linear Time (PELT) method is an exact method to search for change points. It was developed by Killick et al. (2012) based on the algorithm of Jackson et al. (2005) but introduced a pruning step within the dynamic program which reduces the computational cost of the method without affecting the exactness of the segmentation.

The PELT algorithm identifies the change points using the common approach of minimising costs. It runs through the entire data set, searching for changepoints iterating through the whole data set and down to small and smaller partitions until no change points are identified. The advantage of PELT is that it has a computational cost which at its worst is $O(n^2)$ for linear penalties. The technique is exact and more accurate compared to both other approximate and exact search methods.
The whole algorithm is based on the idea that the method finds a global minimum of the cost function. The pruning part removes these values, which can never be minima from the minimisation performed at each iteration in equation 2.23. The assumption is that adding a change point into a sequence of observations reduces the overall cost, C.

$$C(y_{t+1:T}) \geq C(y_{t+1:s}) + C(y_{s+1:T}). \tag{2.26}$$

This holds for costs based on the negative log-likelihood; and often can be made to hold for costs based on the negative log-marginal-likelihood. Let $0 < t < s < T$ then, if

$$F(t) + C(y_{(t+1:s)}) < F(s), \tag{2.27}$$

then at any future time T $>$ s, t can never be the optimal last change point before T. The condition in the theorem just means that for any T $>$ s the best partition which involves a change point at s will be better than one which has [t,T] as a single segment. Thus t can never be the (optimal) most recent change point before T for all $T > s$.

2.2.2 Choice of penalty function

The choice of penalty option to use rests sorely on the researcher. There exist different options pre-determined and also the manual option. The choices include

- Akaike's Information Criterion (AIC), based on the works of Akaike (1974). The penalty term is constant with regards to the data i.e. the penalty is $\beta = 2p$ where p is the number of additional parameters added in the model. Intuitively the AIC aims to pick the model which is closest to the unknown true model, however, this often leads to it overfitting the data.

- The Bayesian Information Criterion (BIC), also known as the Schwarz Information Criterion (SIC), is from the works of Schwarz et al. (1978). The penalty used for a cost function is $\beta = plogn$ where p is the number of additional parameters introduced by adding a changepoint. The penalty is however susceptible to underfitting the data.

- CROPS (Changepoints for a Range of Penalties). This is used with the PELT method. It gives a range for penalty parameter β such that $\beta \in [Max\beta, Min\beta]$

2.3 Markov switching GARCH models

Given the findings that structural breaks in time series affect the estimates by researchers as Lamoureux and Lastrapes (1990), Ardia, Bluteau and Rüede (2019) amongst others, there became a need to have models that cater for the structural changes. One of the ways to deal with these issues was the use of regime changing/switching models. These regime changing models are models that model different periods with different parameters. Parameters therefore vary across regimes. These regimes are unobservable and are driven by a stochastic process.
One of the simplest ways to deal with the structural changes is the introduction of different regimes governed by a switching parameter. The model allows the process to switch from one of the K fixed regimes whenever necessary. Each regime is characterised by different parameters suitable for modeling the current regime state conditions. The regimes are unobservable and are driven by a stochastic process

Regime switching was introduced by Hamilton (1989) who modelled regime changes of an an autoregressive model. The outcomes were taken as a discrete-state Markov stationary process. Following up Hamilton and Susmel (1994) model the conditional volatility using a regime switching ARCH model. The study showed that they could model the different volatility levels, that is the high and low through the use of the regime switching model. Since then various models have applied the regime switching concept. Our focus on this study is the Markov switching GARCH model (MSGARCH).
Given the the return series r_t and the information set at time t as F_t, then the Markov swithing GARCH model can be expressed as follows:

$$r_t|(s_t = k, F_{t-1}) \sim D(0, h_{k,t}, \xi_k), \tag{2.28}$$

where F_{t-1} is the information set available at time $t-1$. $D(0, h_{k,t}, \xi_k)$ is is continuous with mean zero conditional variance $h_{k,t}$ and xi_k is a vector of shape parameters. s_t defined for $\{1...K\}$ denotes the state of the process at time t. The standardised innovations/shocks are defined as $\epsilon_t = \frac{r_t}{\sqrt{h_{k,t}}} \sim D(0, 1, \xi_k)$.

The state state transition of the model is governed by an unobserved first-order ergodic homogeneous Markov chain with K × K transition probability matrix A:

$$A = \begin{bmatrix} p_{11} & \cdots & p_{1,K} \\ \vdots & \ddots & \\ p_{K1} & & p_{KK} \end{bmatrix}$$

where $P(S_{t=j}|S_{t-1} = i) = p_{ij}$ is the transition probability from state i at time *t-1* to state j at time t. By rules of probability, it suffices that $0 < p_{ij} < 1 \quad \forall i, j \in 1, .., K$ and the sum of all transition probabilities for each regime should sum up to one i.e. $\sum_{j=1}^{k} p_{ij} \quad \forall i \in 1, .., K$.

The conditional variance for the K states follow a GARCH-type model define the GARCH part of the MSGARCH model. Thus for the K states they are K different GARCH models. The models can be the same GARCH models but with different parameters that model the stake k. For the conditional distribution, the model can assume any of the common error types such as the normal, student t, skewed student t and the general error distribution.

2.3.1 Parameter estimation for the MSGARCH model

To estimate the parameters of the MSGARCH model, one can use the maximum likelihood (ML) or by Bayesian Markov chain Monte Carlo (MCMC) techniques. Both methods require the likelihood.
Given that the $\Psi = (\theta, \xi, P)$, is a vector of the model parameters, then the likelihood is given by:

$$\ell(\Psi|F_T) = \prod_{t=1}^{T} f(r_t|\Psi, F_{t-1}). \tag{2.29}$$

where $f(r_t|\Psi, F_{t-1})$ is the density function for r_t given the past observations F_t and the parameter set Ψ. For the MSGARCH model, r_t is given by

$$f(r_t|\Psi, F_{t-1}) = \sum_{i=1}^{K} \sum_{j=1}^{K} p_{i,j} z_{i,t-1} f_D(r_t|s_t = j, \Psi, F_{t-1}). \tag{2.30}$$

where $z_{i,t-1}$ is the filtered state probabilities $P[s_{t-1}|\Psi, F_{t-1}]$ for state i at time t. For the mixture of GARCH models, then the conditionl density function is given as follows:

$$f(r_t|\Psi, F_{t-1}) = \sum_{i=1}^{K} \omega_j f_D(r_t|s_t = j, \Psi, F_{t-1}). \tag{2.31}$$

The maximum likelihood estimates are then obtained by maximising the logarithm with respect to the parameters. This is done numerically.

2.4 Chapter summary

This chapter introduced the univariate models that were employed in the modelling of the data. The models discussed are the GARCH models, markov switching models. The change point detection methods were also introduced.

3. MULTIVARIATE MODELS

In the current world, events rarely occur on their own without being affected by other external factors. This is even more pronounced in finance as the global economy is now integrated, allowing investors to invest in different countries while sitting in the comfort of their homes. Because of this globalisation, events in one country or asset class can easily affect those in another. For example, oil prices in Nigeria and Saudi Arabia considerably affect the fuel prices in South Africa. Hence, knowing the prices in these countries adds value to the modelling of South Africa. With investors also having portfolios with different assets, the need to understand the dependencies of different assets is even more critical. This knowledge lets investors know which assets to use to hedge against risk in their portfolios. According to Caporin and McAleer (2010), the increase in attention to multivariate analysis can be attributed to the ease of obtaining data and the increase in computer power, which allows the solving of complex multivariate algorithms.
The main focus of this chapter is to introduce multivariate models that were used to model the volatility spillover effects amongst the cryptocurrencies and also with the JSE market.

3.1 Vector autoregressive time series

The Vector autoregressive (VAR) is a multivariate model in which each time series is a linear function of its past lags and that of other variables. It is the simplest multivariate time series model and the most studied model. The reasons are that it is relatively easier to estimate. By letting $\boldsymbol{z_t}$ represent the N $\times$ 1 multivariate time series then the VAR(P)

$$z_t = \phi_0 + \sum_{i=1}^{p} \phi_i z_{t-i} + a_t, \tag{3.1}$$

where ϕ_0 is a k dimensional vector of constants, ϕ_1 represents the $k \times k$ matrices for i > 0. a_t is the unpredictable innovation component (the shock) which is sequence of iid random vectors with mean zero and covariance matrix $\boldsymbol{\Sigma_t}$. It follows that at is $VWN(0, \boldsymbol{\Sigma_t})$.

To understand the model the study follows Tsay (2013) were he used a VAR(1) model. The model is as follows;

$$z_t = \phi_0 + \phi_i z_{t-i} + a_t, \tag{3.2}$$

which when represented explicitly is;

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} \phi_{10} \\ \phi_{10} \end{bmatrix} + \begin{bmatrix} \phi_{1,11} & \phi_{1,12} \\ \phi_{1,21} & \phi_{1,22} \end{bmatrix} \times \begin{bmatrix} z_{1,t-1} \\ z_{2,t-1} \end{bmatrix} + \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} \tag{3.3}$$

The linear dependence are given by the (1,2)th elements. where$\phi_{1,2}$ is the linear dependence of z_{1t} on $z_{2,t-1}$ with $z_{1,t-1}$ present. The (2,1) element measures the the linear relationship between z_{2t} and $z_{1,t-1}$ in the presence of $z_{2,t-1}$. Similar interpretations are given to other parameters in the same form.
z_{1t} and z_{2t} are not related if the off diagonal elements $\phi_{1,12}, \phi_{1,21}$ are 0. In such cases the series are said to be uncoupled and can be modeled separately using univariate models.

For cases where one of the off diagonal element is 0 but the other is not zero then only current values of the other depend on past values of the other. For example; If $\phi_{1,12} = 0,\ but\ \phi_{1,21} \neq 0$, then we have

$$z_{1t} = \phi_{10} + \phi_{1,11} z_{1,t-1} + a_{1t}, \tag{3.4}$$
$$z_{2t} = \phi_{20} + \phi_{1,21} z_{1,t-1} + \phi_{1,22} z_{2,t-1} + a_{2t}. \tag{3.5}$$

Here there is a unidirectional relationship where z_{1t} does not depend on the past value of z_{2t}, but z_{2t} depends on the past value of z_{1t}.

3.1.1 Granger causality

Noting the different relationships in the VAR models, Granger (1969) introduced a concept of causality (granger causality), which explains the relations between the time series variables. Two types of causality exist, granger causality and instantaneous granger causality. Given F_t is the informatin available at time t, and $F_{-i,t}$ be F_t with all information concerning the ith component z_{it} removed. Also, given the h step forecast error as $e_t(h)$, then

- Causality: If $\sigma^2[e_{2t}(h)|F_t] < \sigma^2[e_{2t}(h)|F_{-i,t}]$ Then, using the VAR(1) which has $\phi_{1,12} = 0$, but $\phi_{1,21} \neq 0$, we say that z_{1t} causes(granger causes) z_{2t}. This means that we are better at predicting z_{2t} using all the information set F_t available than without information on z_{1t} i.e. $F_{-i,t}$. However, since $\phi_{1,12} = 0$ then $z_{1,t+1}$ does not depend on any past value of z_{2t}.
- Instantaneous causality $\sigma^2[e_{2t}(h)|F_{,t}] < \sigma^2[e_{2t}(h)|F_{,t}]$. This is when $\boldsymbol{\Sigma_t}$ is not a diagonal matrix. We say that then z_{1t} and z_{2t} are instantaneously

correlated. The current value z_{1t} is better predicted if the current value of z_{1t}, is included in the prediction than when it is not.

3.2 Multivariate volatility models

Just like in the univariate case, different models are applied in the presence of heteroscedasticity. The VAR models were applied to a multivariate time series z_t with the innovations a_t which are serially uncorrelated and have a zero mean and a positive-definite covariance matrix $\boldsymbol{\Sigma_t}$ that is time-invariant. However, these assumptions do not hold for financial data in the multivariate case, as the multivariate series consists of heteroscedastic individual time series. The covariance matrix from financial data is time-dependent. It is this time dependence that the volatility models want to capture.
For this section, the general volatility model is defined as

$$z_t = \mu_t + a_t, \tag{3.6}$$

where $\mu_t = E(z_t|F_{t-1})$ represents the conditional expectation of the z_t given the information set F_{t-1}. This μ_t follows the multivariate linear models such as the VAR presented in the last section. a_t is the unpredictable component of z_t. It is unpredictable due to being serially uncorrelated. We define it as:

$$a_t = \boldsymbol{\Sigma_t^{\frac{1}{2}}} \epsilon_t, \tag{3.7}$$

where ϵ_t is a sequence of independent and identically distributed random vectors such that $E(\epsilon_t) = 0$ and $Cov(\epsilon_t) = \mathbf{I}_k$ and $\boldsymbol{\Sigma_t^{\frac{1}{2}}}$ denotes the positive-definite square-root matrix of $\boldsymbol{\Sigma_t}$ (Tsay 2013).
For series with conditional heteroscedasticity, the $\boldsymbol{\Sigma_t} = cov(a_t|F_{t-1})$ which is the conditional covariance matrix, is time dependent. This time dependence is the subject matter of the multivariate volatility models. For a three series case the $\boldsymbol{\Sigma_t}$ is:

$$\boldsymbol{\Sigma_t} = \begin{vmatrix} \sigma_{t,11} & \sigma_{t,12} & \sigma_{t,13} \\ \sigma_{t,12} & \sigma_{t,22} & \sigma_{t,23} \\ \sigma_{t,13} & \sigma_{t,23} & \sigma_{t,33} \end{vmatrix}.$$

The main diagonal contains the variances, while all covariances are the off-diagonal entries. The Cov(x,y) = Cov(y,x) hence, from simple properties of expectations, $\boldsymbol{\Sigma_t}$ is by construction a symmetric matrix. This $\boldsymbol{\Sigma_t}$ is critical in finance; for example, in portfolio creation, the methodology relies on knowledge or estimation of $\boldsymbol{\Sigma_t}$. This is because optimal portfolio shares will also depend on the covariance of asset returns considered in pairs.

Tsay (2013) explains the difficulty of multivariate modelling in two ways. First, he mentions the curse of dimensionality, whereby as the dimension of the multivariate series increases, more variance and covariance elements are required, i.e., the number of elements increases quadratically. The second is the difficulty of maintaining the positive definite assumption on the volatility matrix $\boldsymbol{\Sigma_t}$ as k increases. This assumption is crucial as it ensures that the portfolio variance is always positive regardless of the underlying portfolio. Therefore, there is a need for special attention as k increases.

In short, volatility modelling typically consists of two sets of equations, with the first set governing the time evolution of the conditional mean, μ_t. The second set describes the dynamic dependence of the volatility matrix $\boldsymbol{\Sigma_t}$. These two are referred to as the mean and volatility equations, respectively. In practice, the volatility matrix plays a crucial role in asset allocation and risk management.

Many models can be used to capture the volatility in multivariate time series. These models include the Exponentially Weighted Moving Average (EWMA), Go-GARCH Model, Constant Conditional Correlation (CCC), and the Dynamic Conditional Correlation Models (DCC-GARCH), to name a few. The models differ in their estimation methods and how they capture the conditional variance. For example, the Constant Conditional Correlation (CCC) assumes constant correlation while the Dynamic Conditional Correlation Models (DCC-GARCH) allow the conditional correlation to change over time. The ability to allow correlation to change over time is why this study chose the DCC model to capture the dependencies of our data series for this thesis. The DCC model is described in detail in the next section.

3.2.1 Dynamic conditional correlation models (DCC-GARCH)

The assumption of constant conditional correlations over time is not realistic. Hence, researchers have devised different models to circumvent this shortfall. Engle (2002) proposed a new class of estimator that both preserves the ease of estimation of Bollerslev's constant correlation model yet allows for correlations to change over time. Let $\boldsymbol{\Sigma_t}$ =[$\sigma_{ij,t}$] be the volatility matrix of a_t given $F_{t.1}$, which denotes the information available at time $t-1$. Engle's dynamic conditional correlation structure is defined as follows:

$$\begin{aligned} r_t &= \mu_t + a_t, && (3.8)\\ a_t &= \boldsymbol{\Sigma_t}^{\frac{1}{2}} \epsilon_t, && (3.9)\\ \boldsymbol{\Sigma_t} &= D_t R_t D_t. && (3.10) \end{aligned}$$

where D_t is the diagonal matrix of the volatilities at time t. Each of the volatilities are obtained from the univariate GARCH models.

where :

r_t : n $\times$ 1 vector of log returns of n assets at time t.

a_t : n $\times$ 1 vector of mean-corrected returns of n assets at time t, i.e. $E[a_t]$=0 and $Cov[a_t] = H_t$.

μ_t: n Ą× 1 vector of the expected value of the conditional r_t.

$\boldsymbol{\Sigma_t}$: n $\times$ n matrix of conditional variances of a_t at time t.

D_t: n $\times$ n, diagonal matrix of volatilities at time t. That are obtained from univariate GARCH models with $\sqrt{\sigma_{i,t}}$ on the on the ith diagonal

R_t: n $\times$ n conditional correlation matrix of a_t at time t.

ϵ_t: n $\times$ 1 vector of iid errors such that E[ϵ_t]=0 and E[$\epsilon_t\epsilon_t^T$] = **I**.

For D_t the diagonal elements are the standard deviations as shown below

$$D_t = \begin{vmatrix} \sigma_{1t} & 0 & 0 \\ 0 & \sigma_{2t} & 0 \\ 0 & 0 & \sigma_{3t} \end{vmatrix}.$$

The standard deviations are the volatilities obtained from univariate GARCH models. These GARCH models can be any of the GARCH models. They do not have to be the same GARCH-type and can be of any order. However, the simpler, the better.
Let $\boldsymbol{\gamma_t}$ = $(\gamma_1, \ldots, \gamma_k)'$, be the marginally standardized innovation vector, where $\gamma_{it} = a_{it}\sqrt{\sigma_{ii,t}}$. Then, R_t is the conditional correlation matrix of the standardized disturbances γ_t, i.e: $\boldsymbol{\gamma_t}$= $D_t^{-1}a_t \sim$ N(0,R_t).
Since R_t is a correlation matrix, it is symmetric.

$$R_t = \begin{vmatrix} 1 & \rho_{12,t} & \rho_{13,t} \\ \rho_{21,t} & 1 & \rho_{23,t} \\ \rho_{13,t} & \rho_{23,t} & 1 \end{vmatrix} = \sigma_{it}\mathbf{I}.$$

Hence, elements of $\boldsymbol{\Sigma_t} = D_t R_t D_t$ are as follows : $\boldsymbol{\Sigma_t} = \sqrt{\sigma_{it}\sigma_{jt}\rho_{ij}}$ where ρ_{ii} = 1.

Orskaug (2009) explains that, R_t exists in different forms hence, when specifying a form of R_t two requirements have to be considered:

1. $\boldsymbol{\Sigma_t}$ has to be positive definite because it is a covariance matrix. To ensure $\boldsymbol{\Sigma_t}$ to be positive definite, R_t has to be positive definite (D_t is positive definite since all the diagonal elements are positive).

2. All the elements in the correlation matrix R_t have to be equal to or less than one by definition.

To ensure both of these requirements in the DCC-GARCH model, R_t is decomposed into:

$$\begin{aligned} R_t &= Q_t^* Q_t Q_t^*, && (3.11) \\ Q_t &= (1-\theta_1-\theta_2)\bar{Q} + \theta_1 Q_{t-1} + \theta_2 \gamma_{t-1} - \gamma_{t-1}^T, && (3.12) \end{aligned}$$

where $\bar{Q}$ = Cov$[\gamma_t, \gamma_{t-1}^T]$ = E$[\gamma_{t-1}, \gamma_{t-1}^T]$ is the unconditional covariance matrix of the standardized errors γ_{t-1}. $\bar{Q}$ can be estimated as :

$$\bar{Q} = \frac{1}{T}\sum_{t=1}^{T} \gamma_{t-1}\gamma_{t-1}^T. \tag{3.13}$$

The parameters θ_1 and θ_2 are scalars, and Q_t^* is a diagonal matrix with the square root of the diagonal elements of Q_t at the diagonal. Q_t^* rescales the elements in Q_t to ensure the second requirement $|\rho_{ij}| = |\frac{q_{ijt}}{\sqrt{q_{iit}q_{jjt}}}| \leq 1$ In addition to the conditions for the univariate GARCH model to ensure positive unconditional variances, given earlier, the scalars θ_1 and θ_2 must satisfy: $\theta_1 \geq 0$, $\theta_2 \geq 0$ and $\theta_1 + \theta_2 < 1$.

Tse and Tsui (2002) proposed a second type of DCC models which are given as:

$$R_t = (1-\theta_1-\theta_2)\bar{R}_t + \theta_1 R_{t-1} + \theta_2 \psi_{t-1} \tag{3.14}$$

where $\bar{R}_t$ is the unconditional correlation matrix of γ_t. θ_i are non-negative real numbers satisfying additional constraint $0 < \theta_1 + \theta_2 < 1$, and ψ_{t-1} is a local matrix that depends on $\{\gamma_{t-1}, ..., \gamma_{t-m}$ for some positive integer m.

Tsay (2013) explains that both models start with the unconditional covariance matrix of $\boldsymbol{\gamma_t}$. However, they differ in how local information at time *t-1* is used. The DCC model of Engle (2002) uses $\boldsymbol{\gamma_t}$ only so that Q_t must be re-normalized at each time index t. On the other hand, the DCC model of Tse and Tsui (2002) uses local correlations to update the conditional correlation matrices. Tsay (2013) adds that DCC models are extremely parsimonious as they only use two parameters θ_1 and θ_2 to govern the time evolution of all conditional correlations regardless of the number of assets k. This simplicity is both an advantage and a weakness of the DCC models. It is an advantage because the resulting models are relatively easy to estimate. It is a weakness of the model because it is hard to justify that all correlations evolve in the same manner regardless of the assets involved. Tsay (2013) goes on to say "experience albeit limited, indicates that a fitted DCC model is often rejected by diagnostic checking".

3.2.1.1 Estimation of DCC-GARCH

Orskaug (2009) gives the estimation procedure for DCC models. When the standardized errors, $\boldsymbol{z_t}$, are multivariate Gaussian distributed, the joint distribution of $z_1, ..., z_T$ is

$$f(\boldsymbol{z_t}) = \prod_{T=1}^{T} \frac{1}{(2\Pi)^{\frac{1}{2}}} exp\{-\frac{1}{2} z_t z_t^T\}. \tag{3.15}$$

Here t = 1, ..., T is the time period used to estimate the model.

Using the rule for linear transformation of variables , the likelihood function for $a_t = \boldsymbol{\Sigma_t}^{\frac{1}{2}} z_t$ is

$$L(\theta) = \prod_{T=1}^{T} \frac{1}{(2\Pi)^{\frac{1}{2}}} |\boldsymbol{\Sigma_t}^{\frac{1}{2}}| exp\{-\frac{1}{2} a_t^T \boldsymbol{\Sigma_t}^{-1} a_t\}, \tag{3.16}$$

where θ denotes the parameters of the model. Let the parameters, θ, be divided in two groups; (ϕ, ψ) = $(\phi_1, ..., \phi_n, \psi)$, where ϕ_i= $(\alpha_{0i}, \alpha_{1i}, ..., \alpha_{qi}, \beta_{1i}, ..., \beta_{pi})$ are the parameters of the univariate GARCH model for the ith asset series, i = 1, ..., n. $\psi = (\theta_1, \theta_2)$ are the parameters of the correlation structure in 3.12.

By taking the logarithm of 3.16 and substituting $\boldsymbol{\Sigma_t} = D_t R_t D_t$ we get the log-likelihood:

$$\begin{aligned} ln(L(\theta)) &= -\frac{1}{2} \sum_{t=1}^{T} (nln(2\Pi) + ln(|\boldsymbol{\Sigma_t}| + a_t^T \boldsymbol{\Sigma_t}^{-1} a_t)), \\ &= -\frac{1}{2} \sum_{t=1}^{T} (nln(2\Pi) + ln(|D_t R_t D_t|) + a_t^T D_t^{-1} R_t^{-1} D_t^{-1} a_t, \\ &= -\frac{1}{2} \sum_{t=1}^{T} (nln(2\Pi) + 2ln(|D_t|) + ln(R_t) + a_t^T D_t^{-1} R_t^{-1} D_t^{-1} a_t. \end{aligned} \tag{3.17}$$

Due to complex calculations required for the exact likelihood estimation, DCC allows a two-step estimation. In the first step, the univariate GARCH models are estimated, and the parameters ϕ are saved. The likelihood used in the first stage results in replacing R_t with the identity matrix I_n. In the second stage, the parameter ψ is estimated using the correctly specified log-likelihood, given the parameter ϕ. Detailed steps for estimating the two stages are given by Orskaug (2009) for the multivariate Student's t-distributed errors and multivariate skew Student's t-distributed errors.

3.3 Chapter summary

This chapter introduced the multivariate models. The fist model discussed is the VAR model. This model allows for the testing of the Granger Causality which was used to test the causality between the different assets. The chapter then discussed the multivariate volatility models. Here focus was on the DCC model which allows the conditional correlation to change over time.

4. EXTREME VALUE THEOREM

4.1 Introduction

This chapter reviews the theoretical foundation of the Extreme Value Theorem (EVT). The extreme value theorem focuses on the extreme tail of the distribution. This translates to the extreme events. In essence it is a study of extreme events occurrence. The chapter also introduces the Value at risk (VaR) which is used to quantify risk.

4.2 Extreme Value

In the real world, data is rarely uniformly distributed. Instead, it can take different distributions. One common occurrence is having most values centred around a typical central value and fewer values near the maximum and the minimum. The normal distribution usually captures this kind of distribution. However, in some cases, the values at the extreme are less or more than that of the normal distribution. In finance, it is the latter that is common. The extreme values tend to be more than those expected under the normal distribution. Such data distribution often negatively impacts results from models that assume the normal distribution due to the tail distribution distortion. These extreme values are rare but are very important. For example, geologists may be interested in modelling extreme weather patterns, and public health officers may want to know the maximum number of expected hospital visits during a pandemic to prepare enough resources. In finance, investors and business owners want to know what their positions will be like in extreme cases of financial stress. On a small scale, they lead to business closure, and on a larger scale, they may lead to a financial crisis such as the Stock market crash of 1987 and the Global financial crisis of 2008.

One of the most common ways to deal with these extreme values is the exclusion of these values from analysis. This approach seems natural and takes care of the problem without complications. However, researchers over time have realised that extreme values are crucial in understanding the market behaviour in normal conditions and under abnormal turbulent conditions. This realisation has increased interest in these extreme values in the last few decades, creating a new area of study known as the Extreme Value Theory. To describe this theory in mathematical terms, assume we have the losses on investment on a fixed period, such as a day given by x_t. x_t can be our daily log returns which can be negative or positive. If we have returns for n days then we have a sequence of $x_t = x_1, x_2, ..., x_n$ returns. From the sequence of x_t we have the maximum loss/return, which is defined as M_n. The extreme value theory is the theory that studies the properties of M_n as n increases. It can be viewed as the central limit theorem for extreme values.

The theorem distribution of M_n can be derived as follows:

$$\begin{aligned} P\{M_n \leq x &= PX_1 \leq x, ..., x_n \leq x\}, \\ &= P\{X_1 \leq x \times ... \times PX_n \leq x\}, \\ &= \prod_{i=1}^{n} P(X_1 \leq x), \\ &= [P(X_1 \leq x)]^n, \\ &= [F(x)]^n. \end{aligned} \tag{4.1}$$

where $F(x)$ is the cumulative distribution of x_t. In practice this is not directly helpful because the distribution function of F is unknown which in turn means that M_n is unknown. As n increases, $[F(x)]^n$ approaches zero if x < μ and approaches 1 if x $\geq \mu$. This problem of a problematic degenerated M_n can be avoided through linear renormalization of M_n as follows:

$$M_n^* = \frac{M_n - b_n}{a_n}, \tag{4.2}$$

where $a_n > 0, b_n$ are constants

4.2.1 Extremal types theorem

Theorem: if there exist sequences of normalisation constants $a_n > 0$ and b_n such that

$$\lim_{u \to \infty} P(M_n^* \leq x) \to F_*(x). \tag{4.3}$$

where $F_*(x)$ is a non-degenerate distribution which belongs to either of the families of Jenkinson (1955) which are as follows

- The Gumbel family known as Type 1 where $\xi = 0$, hence a cumulative distribution function (CDF) becomes

$$F(x) = \exp[-\exp(-x)], for \infty < x < \infty. \tag{4.4}$$

- The Fr'echet family known as Type 2 where $\xi > 0$, hence the CDF is

$$F(x) = \begin{cases} exp[-(1+\xi x)]^{-\frac{1}{\xi}} & \text{if } x > -\frac{1}{\xi}. \\ 0 & \text{if } otherwise. \end{cases} \tag{4.5}$$

- The Weibull family known as Type 3 where $\xi < 0$, hence CDF is

$$F(x) = \begin{cases} exp[-(1+\xi x)]^{-\frac{1}{\xi}} & \text{if } x < -\frac{1}{\xi}. \\ 1 & \text{if } otherwise. \end{cases} \tag{4.6}$$

The implication of the theorem is that the normalised maxima will always have a limiting distribution which can be either of the three extreme value types regardless of the population distribution. This feature hence gives the equivalence of the Central limit theorem of extreme values (Coles et al. 2001).

4.3 The generalized extreme value distribution

The three distributions Gumbel, Frechet and Weibull are made distinct by their tail distribution. The Weibull has a finite tail, the Frechet and Gumbel have a tail that goes to infinity. The distinction between Gumbel and Frechet is that density of decays exponentially for the former and polynomially for the later. Therefore the density decay is slower on the Frechet. The three distributions can be combined into one representation for better analysis as follows;

$$F_*(x) = \begin{cases} exp[-(1+\xi x)]^{-\frac{1}{\xi}} & \text{if } \xi \neq 0, \\ \exp[-exp(-x)] & \text{if } \xi = 0. \end{cases} \tag{4.7}$$

Here $x < -1/\xi$ if $\xi < 0$ and $x > -1/\xi$ if $\xi > 0$. The parameter ξ is the shape parameter responsible for the tail behaviour of the limiting distribution while $1/\xi = \alpha$ is the tail parameter responsible with the tail behaviour of the distribution. This limiting distribution $F_*(x)$ is what is known as the generalized extreme value distribution (GEVD) of Jenkinson (1955) for the maximum. The name arises as it generalises all the extreme value limiting distributions into one equation.

4.3.1 The block maximum approach

Because parameters of the distribution are unknown, there is a need to estimate the parameters from the data. The general extreme value theorem only has three parameters that need to be estimated. These are the shape parameter, the location parameter and the scale parameter. The block maxima method involves blocking the data into equally sized blocks and fitting the GEVD to each block.

Given observations $X_1, X_2, ...X_N$, and assume that the data is split into a sequence of n non-overlapping blocks of size s. Then EVT is applied to each block to obtain the maximum for each block. In general, the data distribution in each sub-period is difficult to know. However, for n large enough and assuming that the maximum for each block i is $M_{n,i}$, then the distribution of the normalised maximas from each block $y_{n,i} = (M_{n,i} - \mu)/\sigma$ converges asymptotically to a generalised extreme value distribution. The collection of the normalised maximas $y_{n,i}$ from the n subperiods are taken as the new data set. From this new data set, the unknown parameters, the shape parameter ξ, the location parameter μ, and the scale parameter σ are estimated. The choice of the block size is of paramount importance as it can be the difference between a reliable model and a biased model. Coles et al. (2001) states that choice amounts to a trade-off between bias and variance. If blocks are too small, then the approximation of the limiting distribution will likely be poor, leading to biased estimates. On the other hand, larger blocks give fewer maxima, leading to larger variance estimates. In a real-world application, the choice of block size s depends on the problem at hand. Often, the number of blocks n considers the natural periods. For example, for monthly data, one may consider 21, which corresponds with the number of trading days, and if the asset is traded daily, then n can be 30.

4.3.2 Maximum likelihood estimation of the GEVD

Given that $x_1, x_2, ...x_m$ (notation changed from M_n for maxima for easy reading), are the obtained block maximas that are independent and follow the GEVD, then to obtain the parameter estimates the following log-likelihood is used:
For $\xi \neq 0$,

$$\ell(\mu,\sigma,\xi) = -m\log\sigma - (1+1/\xi)\sum_{i=1}^{m}\log[1+\xi(\frac{x_i-\mu}{\sigma})] - \sum_{i=1}^{m}[1+\xi(\frac{x_i-\mu}{\sigma})]^{-\frac{1}{\xi}}, \tag{4.8}$$

provided that

$$1+\xi(\frac{x_i-\mu}{\sigma}) > 0 \; for \; i = 1,...,m.$$

If a combination of the parameters violets the provision above the likelihood will be zero and hence the log-likelihood will go to infinity. For $\xi = 0$,

$$\ell(\mu,\sigma) = -m\log\sigma - \sum_{i=1}^{m}(\frac{x_i-\mu}{\sigma}) - \sum_{i=1}^{m}\exp\{-(\frac{x_i-\mu}{\sigma})\}. \tag{4.9}$$

The estimates (μ, σ) are obtained by maximising the likelihoods. For μ we differentiate with respect to μ

$$\frac{d\ell}{d\mu} = -(1+\frac{1}{\xi})\sum_{i=1}^{m}\frac{1}{1+\xi(\frac{x_i-\mu}{\sigma})}*(-\frac{\xi}{\sigma}) - \left[1+\xi(\frac{x_i-\mu}{\sigma})\right]^{-\frac{1}{\xi}}\left[-\frac{1}{\xi}\frac{1}{1+\xi(\frac{x_i-\mu}{\sigma})}*(-\frac{\xi}{\sigma})\right],$$

$$\frac{d\ell}{d\mu} = (1+\frac{1}{\xi})\sum_{i=1}^{m}\frac{\xi}{\sigma+\xi(x_i-\mu)} - \left[1+\xi(\frac{x_i-\mu}{\sigma})\right]^{-\frac{1}{\xi}}\left[\frac{1}{\sigma+\xi(x_i-\mu)}\right], \tag{4.10}$$

For σ we differentiate with respect to σ

$$\begin{aligned}\frac{d\ell}{d\sigma} &= -\frac{m}{\sigma}(1+\frac{1}{\xi})\sum_{i=1}^{m}\frac{1}{1+\xi(\frac{x_i-\mu}{\sigma})}.(-\xi)(\frac{x_i-\mu}{\sigma^2}) \\ &= \sum_{i=1}^{m}\left[1+\xi(\frac{x_i-\mu}{\sigma})\right]^{-\frac{1}{\xi}}\left[-\frac{1}{\xi}\sum_{i=1}^{m}\frac{1}{1+\xi(\frac{x_i-\mu}{\sigma})}.(-\xi)(\frac{x_i-\mu}{\sigma^2})\right], \qquad (4.11) \\ \frac{d\ell}{d\sigma} &= -\frac{m}{\sigma}\sum_{i=1}^{m}\frac{\xi(x_i-\mu)}{\sigma[\sigma+\xi(x_i-\mu)]} - \sum_{i=1}^{m}\sum_{i=1}^{m}\left[1+\xi(\frac{x_i-\mu}{\sigma})\right]^{-\frac{1}{\xi}}\frac{\xi(x_i-\mu)}{\sigma[\sigma+\xi(x_i-\mu)]}.\end{aligned}$$

For σ we differentiate with respect to σ. The differential is done implicitly to obtain

$$\begin{aligned}\frac{d\ell}{d\xi} &= \frac{1}{\xi^2}\sum_{i=1}^{m}\ln\left[1+\xi(\frac{x_i-\mu}{\sigma})\right] - (1+\frac{1}{\xi})\sum_{i=1}^{m}\frac{(x_i-\mu)}{\sigma+\xi(x_i-\mu)}, \\ &= -\frac{1}{\xi^2}\sum_{i=1}^{m}\sum_{i=1}^{m}\left[1+\xi(\frac{x_i-\mu}{\sigma})\right]^{-\frac{1}{\xi}}\ln\left[1+\xi(\frac{x_i-\mu}{\sigma})\right], \qquad (4.12) \\ &= +\frac{m}{\sigma}\sum_{i=1}^{m}\sum_{i=1}^{m}\left[1+\xi(\frac{x_i-\mu}{\sigma})\right]^{-\frac{1}{\xi}}\frac{\xi(x_i-\mu)}{\sigma[\sigma+\xi(x_i-\mu)]}.\end{aligned}$$

There equations 4.12,4.13 and 4.14 have no analytical solution, but can be solved numerically for each given data. The optimisation algorithms must be set to avoid violating the parameter assumptions.

4.4 *Generalized pareto distribution (GPD)*

While the distribution of the block maxima asymptotes to the GEVD, the distribution of the exceedances asymptotes to the General Pareto Distribution (GPD). The GPD is derived from the approximation attained in the Peaks over threshold (POT) method in section 4.4.1. The probability distribution is as follows:

$$G_{\xi,\sigma(\eta)}(y) = \begin{cases} 1 - [1 + \frac{\xi y}{\sigma(\eta)}]^{-1/\xi} & \text{if } \xi \neq 0, \\ 1 - exp(-y/\sigma(\eta)) & \text{if } \xi = 0, \end{cases} \tag{4.13}$$

where $\sigma(\eta) > 0$, y $\geq$0 when $\xi \geq 0$ and $0 \leq y \leq -\sigma(\eta)/\xi$ when $\xi < 0$. This conditional distribution is the POT approximation result of equation 4.10 with parameters $\sigma(\eta = \sigma + \xi(\eta - \mu)$
An important property for the GPD is that if the excess distribution of threshold η_0 is a GPD with shape parameter ξ and scale parameter $\psi(\eta_0)$. Then for any $\eta > \eta_0$ the excess distribution over η also follows a GPD with shape parameter ξ and scale parameter $\sigma(\eta) = \sigma(\eta_0) + \xi(\eta - \eta_0)$.

4.4.1 Peaks over threshold (POT) approach

The Peaks over threshold method models the extremes using exceedances rather than the block maximas. The block method tends to be wasteful and inefficient with data as some blocks may have more extremes that would be considered extremes in other blocks but fail to be used due to not being maximas in their block. The obtained maximas also depend on the subperiods used, and the results are affected if a wrong choice is made. The POT considers all the data points and only focuses on those exceeding a certain threshold, thereby not wasting any extreme data points.
Assume our series is X_t and the threshold value is denoted by the η. Then the exceedance is defined by y= $X_{ti} - \eta$. It is the exceedance over the threshold. If the exceedance occurs at time t_i then the data point is (t_i, y). The frequency of occurrence of the extreme events is given by the times at which exceedance occurs $\{t_i\}$, while the exceedance itself gives the exceeding quantity of interest. The value of the threshold must be known before application. Each threshold leads to different parameter estimates. Tsay (2014) states that the choice of the threshold is a subject of debate, but researchers agree that it can not only be a statistically estimated value but should also make financial sense. Therefore the ideal value should cater for both. For risk analysis, the value for the threshold η will depend on the risk the investor wants. The common value for daily log returns is around 2.5%, but it can go as high as 10% for a volatile time series.
For the POT approach, the exceedance is determined by the probability law. It is found as the conditional distribution of the exceedance given that the return X_t is greater than the threshold. The conditional probability is as follows:

$$P(x \leq y + \eta \mid x > \eta) = \frac{P(\eta \leq x \leq y + \eta)}{P(x > \eta)} = \frac{P(x \leq y + \eta) - P(x \leq \eta)}{1 - P(x \leq \eta)}. \tag{4.14}$$

By using the CDF from equation (4.7), the equation is manipulated to give

$$\begin{aligned} P(x \leq y+\eta \mid x > \eta) &= \frac{F_*(y+\eta) - F_*(\eta)}{1 - F_*(\eta)} \\ &= \frac{\exp\left[-(1+\frac{\xi(y+\eta-\mu)}{\sigma})^{-\frac{1}{\xi}}\right] - \exp\left[-(1+\frac{\xi(\eta-\mu)}{\sigma})^{-\frac{1}{\xi}}\right]}{1 - \exp\left[-(1+\frac{\xi(\eta-\mu)}{\sigma})^{-\frac{1}{\xi}}\right]} \\ &\approx 1 - (1 + \frac{\xi(\eta-\mu)}{\sigma + \xi(\eta-\mu)})^{-\frac{1}{\xi}} \end{aligned} \tag{4.15}$$

where $y > 0$ and $1 + \xi(\eta - \mu)/\sigma > 0$. This approximation connects the POT method to the tradition EVT (Tsay 2014).

4.4.2 Threshold selection

Before applying the GPD, there is a need to determine the threshold value. For each threshold, a different estimate of the shape parameter ξ and the tail index $1/\xi$ will be obtained (Tsay 2014). Such different parameters for different thresholds means the threshold must be decided with care. The threshold value should consider bias and variance resulting from it. The value must not be too low or too high. A threshold value that is too low is more likely to violate the asymptotic nature of the model, which could result in biased estimates.
On the other hand, very high leads to few values exceeding the threshold. This means that less data is available for calculating the model parameters, resulting in a large variance. A standard procedure is to obtain a low threshold determined whilst considering the limiting model. Coles et al. (2001) gives two ways to achieve this: making use of exploratory techniques before the model estimation and assessing the stability of the parameter estimates by fitting the model using different threshold values. The techniques are discussed in the following sections.

4.4.3 Mean residual life plot

Given the threshold η_0 and that the excess $y - \eta_0$ is a valid distribution GPD with parameters ξ and σ with $0 < \xi < 1$. The mean excess for the excess is given by:

$$E(x - \eta_0 \mid x > \eta_0) = \frac{\sigma_\eta}{1 - \xi}, \tag{4.16}$$

where $\xi < 1$.
If the GPD holds for η then it suffices to say that for any $\eta > \eta_0$, subject to the

new relevant scale parameter σ_η the GDP should also hold such that;

$$E(x-\eta \mid x>\eta) = \frac{\sigma_\eta}{1-\xi}, \quad (4.17)$$

$$= \frac{\sigma_\eta + \xi\eta}{1-\xi}. \quad (4.18)$$

This shows that for an $\eta > \eta_0$, the mean excess function for a fixed ξ is a linear function of η. This gives an empirical estimation of the mean of the threshold excess which given as

$$e_T(\eta) = \frac{1}{N_\eta}\sum_{i=1}^{N_\eta}(x_{ti}-\eta). \quad (4.19)$$

Here x_{ti} are the values of returns that exceed the threshold η, and N_η are the number of returns that exceed η. A graphical plot termed the mean residual plot or the mean excess plot can then be derived by plotting $e_T(\eta)$ against η. This plot should be linear for $\eta > \eta_0$ if the GDP is the distribution followed. Confidence intervals can also be added to the plot.

Interpreting the mean excess plot is not as straightforward in real life. To interpret, one has to consider the graph, which tends to have a curvature or nonstable section, but once it gets to a section where it gets linear, then that is usually taken as the threshold. In cases where this is a higher point, leaving very few values above that threshold for meaningful inference, one can look for another change to linearity before that point, even if it is followed by drops or curvatures later.

4.4.4 Parameter stability method

Given the complexities of threshold selection using the mean residual plot, the parameter stability method can be an alternative. This method requires fitting the GPD to different threshold values and observing the stability of the parameters.
The previous section showed that if the GPD is the distribution function for the threshold, then any excess above the threshold also follows the GPD. The shape parameters for these are also identical. Recall that the scale parameter is given by $\sigma(\eta) = \sigma(\eta_0) + \xi(\eta - \eta_0)$. Here the scale parameter changes as the threshold changes unless the shape parameter $\xi = 0$. By reparameterizing the GPD scale parameter, we can avoid this change. We reparemetarize the scale parameter to:

$$\sigma^* = \sigma(\eta) - \xi\eta. \quad (4.20)$$

This will be a constant with respect to η and also estimates of σ^* and ξ should be constant above threshold η_0 given η_0 follows the GPD. Because of sampling variability, the estimates will not be exactly constant but may differ slightly but stable enough after allowing for sampling error. A plot of $\hat{\sigma}^*$ and $\hat{\xi}$ against η is thus expected to be near-constant. The threshold is selected at the stable point.

4.4.5 Maximum likelihood estimation of the GPD

Given the threshold μ has been determined the next step is to estimate the scale and tail parameters. Given the the exceedances as $\{y_i\}$ the for a sub-sample of $y_1 - \mu, ..., y_m - \mu$ follows the GPD.
We derive the log of the exceedances y_i as

$$\ln_{G_\mu}(y_i - \mu) = \begin{cases} -\ln(\sigma) - \frac{1+\xi}{\xi}\ln[1 + \xi(\frac{y_i-\mu}{\sigma})] & \text{if } \xi \neq 0, \\ -\ln(\sigma) - \frac{1}{\sigma}(y_i - \mu) & \text{if } \xi = 0. \end{cases} \tag{4.21}$$

Hence the conditional loglikelihood $\ell(\xi, \sigma)$

$$\ell(\xi, \sigma|(y_i - \mu) = \begin{cases} -m\ln(\sigma) - \frac{1+\xi}{\xi}\sum_{i=1}^{m}\ln[1 + \xi(\frac{y_i-\mu}{\sigma})] & \text{if } \xi \neq 0, \\ -\ln(\sigma) - \frac{1}{\sigma}\sum_{i=1}^{m}(y_i - \mu) & \text{if } \xi = 0. \end{cases} \tag{4.22}$$

To obtained the parameter estimates equation is differentiated with respect to the parameters. For σ we get,

$$\frac{d\ell}{d\sigma} = \begin{cases} -\frac{m}{\sigma} + \frac{1+\xi}{\sigma\xi}\sum_{i=1}^{m}\frac{\xi(y_i-\mu)}{\sigma+\xi(y_i-\mu)})] & \text{if } \xi \neq 0, \\ -\frac{m}{\sigma} + \frac{1}{\sigma^2}sum_{i=1}^{m}(y_i - \mu) & \text{if } \xi = 0. \end{cases} \tag{4.23}$$

For ξ we get

$$\frac{d\ell}{d\xi} = \begin{cases} \frac{1}{\xi^2}\sum_{i=1}^{m}\ln[1 + \xi\frac{(y_i-\mu)}{\sigma}] - \frac{1+\xi}{\sigma\xi}\sum_{i=1}^{m}\frac{y_i-\mu}{1+\xi(\frac{y_i-\mu}{\sigma})} & \text{if } \xi \neq 0, \\ 0, & \text{if } \xi = 0. \end{cases} \tag{4.24}$$

Just like in the GEV, the equations do not have analytical solutions and therefore require numerical methods to be solved.

4.5 Value at risk (VaR)

One of the most important aspects of investing is risk analysis. Investors need to know the riskiness associated with their investments and decide if it is worth the risk. Three broad categories of risk exist market risk, credit risk, and operational risk. Market risk is defined as the risk associated with market price movements such as stock prices, interest rates, commodity prices, and volatility risk. Due to readily available associated data, market risk is thus the most researched area. Credit risk is a risk associated with default on payments by creditors. The data for credit risk is usually collected by credit companies and do not make public. This results in the area being less studied. Operational risk results from failed internal systems and other external occurrences. Examples are legal issues that the companies may face and also the political atmosphere. In this thesis, focus is on the market risk through the use of what is termed the Value at Risk (VaR).

This VAR measure was introduced by JP Morgan in their publication RiskMetrics with an academic paper being published in JP Morgan (1995). Choudhry (2013) defines VaR as the maximum loss which can occur with given percentage confidence over a holding period of t days. In other terms, it is a point estimate of the possible amount the investment in the risk category can lose due to market movements. Investors use this value to understand how to set up their financials so that they remain in business in the event of extreme/catastrophic events.

Given that investors usually hold two different positions in the market, the long position and the short position, then the VaR depends on the position held. For the long position, the investor holds stocks hoping for an increase in prices to make profits. Whilst for the short position, the investor, borrows stocks and sells them immediately, betting on a drop in prices so that he can buy back the stocks and return them while making a profit. Therefore for the long position, the VaR is associated with the lower tail distribution while the short position is associated with the upper tail of the distribution. Suppose the risk associated with the financial position for the next ℓ periods must be estimated. For ease of notation, we define a general loss variable at time t as $L_t(\ell)$ and the value of the investment at time t as V_t . The value of $L_t(\ell)$ will be a positive or negative value of $V_{t+\ell} - V_t$ depending on the position as being long or short. Then the VaR for the financial position is over ℓ periods is given by;

$$VaR_{1-p} = \inf\{x|F_\ell(x) \geq 1-p\}, \tag{4.25}$$

where p is the loss probability which is always a small value as it calculates an extreme loss that is supposed to be rare. $F_\ell(x)$ is the cumulative distribution of the loss variable $L_t(\ell)$. Inf represents the smallest real number x that satisfies the equation.
VaR is concerned with the upper tail probability of the loss distribution *(1-p)=q*. Hence VaR is simplified to be

$$VaR_p = F^{-1}(1-q). \tag{4.26}$$

where F^{-1} is the inverse of the quantile function of the loss distribution. To note is that VaR is simply a quantile with the upper tail probability p. It does not describe the actual tail behaviour of the loss random variable. Hence, one can have the same VaR for different variables for the same probability p, but the tail behaviour may differ.

Tsay (2014) gives the VaR for the GEVD and GPD as follows;

For GEVD

$$VaR_q = \begin{cases} \mu_n - \frac{\sigma_n}{\xi_n}\{1 - [-n\ln(1-p)]^{-\xi_n}\} & \text{if } \xi \neq 0, \\ \mu_n - \sigma_n \ln[-n\ln(1-p)] & \text{if } \xi = 0. \end{cases} \tag{4.27}$$

where n is the subperiod length.

$$F_{\xi,\beta}(x) = \begin{cases} 1 - [1 + \frac{\xi x}{\beta}]^{-1/\xi} & \text{if } \xi \neq 0, \\ 1 - \exp[-x/\beta] & \text{if } \xi = 0. \end{cases} \tag{4.28}$$

For GPD

$$VaR_q = \eta - \frac{\sigma(\eta)}{\xi}\{1 - [\frac{T}{N_\eta(1-q)}]\}, \tag{4.29}$$

where, T is the sample size, N_η is the number of exceedances, and the rest are parameters as defined under the GPD distribution.

4.6 Chapter summary

This chapter introduced the extreme value theorem. It discussed different extreme value distributions and their structure. The chapter also discussed the Value at risk (VaR).

5. REVIEW OF TOOLS FOR DESCRIPTIVE ANALYSIS AND DIAGNOSTIC TESTS

This chapter discusses the methods that were used in the pre modelling checks and diagnostic checks of the data. The first section will be the methods used in data checks before modelling. Here various checks that evaluate the suitability of the data to be modelled with the models that are to be used are discussed. These checks also help us decide the suitable models. In some cases the checks also reveal the need for data transformations before applying the models. The second part will be the model selection techniques and the model diagnostics. Lastly the chapter discusses the diagnostic techniques for the Value at Risk (VaR).

5.1 Pre modeling data checking test

5.1.1 Stationarity tests

For most time series models, the key property that allows for reliable modelling is the stationarity property, particularly weak stationarity. The advantage of a stationary time series is that the statistical properties do not change over time. This does not mean the values will be the same; rather, the overall behaviour is constant. Such a property allows predictions to be made with less error. A non-stationary series leads to biased estimates. Mathematically, given a random series of $x_t = x_1, x_2, ..x_N$ then the series x_t is said to be weakly stationary if: $\mathbb{Z}$

$$\begin{aligned} E(x_t) &= \mu \; \forall t \in \mathbb{Z}, \\ E(x_t - \mu)^2 &= \nu_0 \; \forall t \in \mathbb{Z}, \\ Cov(x_t, x_s) &= Cov(x_{t+h}, x_{s+h}) \; \forall s, t, h \in \mathbb{Z}. \end{aligned}$$

Here, the mean and variance are time-invariant. The weak stationarity is essential because it provides the basic framework for prediction. This study will use the following stationarity tests: Augmented Dickey-Fuller (ADF) test (Dickey and Fuller 1981), Phillips-Perron (PP) test (Phillips and Perron 1988), Kwiatkowski, Phillips, Schmidt and Shin (KPSS) test (Kwiatkowski et al. 1992). These tests are discussed below.

5.1.1.1 Augmented Dickey-Fuller (ADF) test

The ADF test is also known as the Unit root test. Given X_t as the log return series then to test for unit root we employ the models

$$X_t = \phi_1 X_{t-1} + \epsilon_t, \quad (5.1)$$
$$X_t = \phi_0 + \phi_1 X_{t-1} + \epsilon_t. \quad (5.2)$$

where ϵ_t is the error term. Here the hypothesis is such that $H_0 : \phi_1 = 1$ vs $H_1 : \phi_1 < 1$, i.e. the null hypothesis that the time series has a unit root vs the alternative that it is stationary. The test statistic is given by the t ratio as follows:

$$t = \frac{\hat{\phi}_1 - 1}{SE(\phi_1)}. \quad (5.3)$$

The ADF test can be conducted allowing for an intercept, or an intercept and deterministic trend, or neither, in the test regression.

5.1.1.2 Phillips-Perron (PP) test

Phillips and Perron (1988) developed a unit root test that takes into account the serial correlations and heteroscedasticity Herranz (2017). The null hypothesis is that there exist a unit root in the series, and the alternative is that the series is stationary. The test is carried out on an time series of the form:

$$x_t = \phi_0 + \phi_1 x_{t-1} + \epsilon_t. \quad (5.4)$$

5.1.1.3 Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test

The KPSS test was developed to complement tests for non-stationarity as the ADF and PP tests have low power for near non-stationary and long-run trend processes. KPSS has stationary as the null hypothesis. Assuming there is no trend, then it follows that:

$$x_t = \beta' \boldsymbol{D}_t + \mu_t + \mu_t, \quad (5.5)$$

$$\mu_t = \mu_{t-1} + \epsilon_t \textit{ and } \epsilon_t \sim N(0, \sigma_\epsilon^2).$$

where $\boldsymbol{D}_t$ has the deterministic components which can be a constant or a constant and time trend. The null hypothesis that x_t is $\boldsymbol{I}(0)$ is stated as $H_0 : \sigma_\epsilon^2 = 0$ indirectly implying that μ_t is constant. The alternative is that $H_1 : \sigma_\epsilon^2 > 0$. The test statistic is given by

$$KPSS = (T^{-2}\sum_{t=1}^{T}\hat{S}_t^2)/\hat{\lambda}^2. \tag{5.6}$$

where $\hat{S}_t = \sum_{j=1}^{t}\hat{\mu}_j$, μ_t is the residual of a regression of x_t on $\boldsymbol{D}_t$ and $\hat{\lambda}^2$ is a consistent estimate of the long-run variance of μ_t using $\hat{\mu}_t$

5.2 Tests for ARCH effects

One of the stylised features of financial data is that it has a non-constant variance. This is termed conditional heteroscedasticity, which is also known as the ARCH effect. Due to the non-constant variance, the traditional time series models, such as the autoregressive models and the moving average models, become less inadequate as they assume constant variance. The study will use two methods to test the presence of the ARCH effects: the Ljung-Box test and the Lagrange multiplier test of Engle (1982). Given the residuals of the mean as $a_t = r_t - \mu_t$, then the residuals a_t^2 will be used to check for the presence ARCH effects.

5.2.1 Ljung-Box test

Box and Pierce (1970) developed the Portmanteau test also known as the Box-Pierce test that tests for the significance of lag autocorrelations. The test statistic is given by:

$$Q*(m) = T\sum_{\ell}^{m}\hat{\rho_\ell^2}. \tag{5.7}$$

This test tests if the joint hypothesis that all m of the ρ_1 correlation coefficients are simultaneously equal to zero i.e. the null hypothesis $H_0 : \rho_1 = ... = \rho_n = 0$ against the alternative hypothesis $H_1 : \rho_i = 0$ for some i $\in \{1, ..., m\}$. This means that one autocorrelation coefficient has to be significant for the null hypothesis rejection

However, the Box-Pierce test lacks the power to deal with small samples efficiently, often leading to wrong conclusions. Ljung and Box (1978) modified the Box-Pierce statistic, increasing its power in small samples. The new modified test known as the Ljung-Box statistic Q(m) is given by

$$Q(m) = T(T+2)\sum_{\ell=1}^{m}\frac{\hat{\rho_\ell^2}}{T-\ell}. \tag{5.8}$$

The test statistic is chi-square distributed hence the decision rule is to reject H_0 if Q(m) χ_α^2. From the equation of Q(m), it can be seen that as the sample size increases to infinity, the (T + 2) and (T - ℓ) terms will cancel out, leaving the Ljung-Box test to be equivalent to the Box-Pierce test.

5.2.2 Portmanteau test

The portmanteau test also exists for the multivariate time series case. Given the innovation of the multivariate series z_t is $a_t = \sum_t^{\frac{1}{2}} \epsilon_t$. If a_t has no conditional heteroscedasticity, then its conditional covariance matrix $\sum_t$ is time-invariant. This implies that $\sum_t$, and therefore also, a_t^2, does not depend on the a_{t-i}^2 for i>0. Hence, we test the hypothesis

$$H_0 : \rho_1 = ... = \rho_m = 0,$$

against the alternative hypothesis

$$H_a : \rho_i \neq 0,$$

for some i $(1 \leq i \leq m)$, where ρ_i is the lag-i cross-correlation matrix of a_t^2. The test statistic is a Ljung box statistic

$$Q*(m) = T^2 \sum_{i=1}^{m} \frac{1}{T-i} \mathrm{b}_i'(\hat{\rho_0^{-1}} \bigotimes \hat{\rho_0^{-1}})\mathrm{b}_i, \tag{5.9}$$

where T denotes the sample size, k is the dimension of a_t, and b_i = vec($\hat{\rho}_i$), $\bigotimes$ is the kronecker product. $Q*k(m)$ is asymptotically distributed as $\chi^2_{k^2m}$.

5.2.3 Rank based test

Due to the heavy-tailed nature of financial returns, the portmanteau test results can be biased. To deal with the problem, Dufour and Roy (1985) developed a rank-based test that considers the rank series of the standardized series ϵ_t. Given R_t as the rank of ϵ_t. Then the lag $_\ell$ rank autocorrelation of ϵ_t is defined as follows:

$$\tilde{\rho_\ell} = \frac{\sum_{t=\ell+1}^{T}(R_t - \bar{R})(R_{t-\ell} - \bar{R})}{\sum_{t=1}^{T}(R_t - \bar{R})^2}, \ \ell = 1,2,....,\tag{5.10}$$

where $\bar{R_t} = \sum_{t=1}^{T} R_t/T = (T+1)/2$ and $\sum_{t=1}^{T}(R_t - \bar{R})^2 = T(T^2-1)/12$
The test static being

$$Q_R(m) = \sum_{i=1}^{m} \frac{[\tilde{\rho_i} - E(\tilde{\rho_i})]^2}{Var(\tilde{\rho_i})}. \tag{5.11}$$

This is distributed χ^2_m asymptotically if ϵ_t has no serially dependance.

5.2.4 ARCH-LM test/ Lagrange multiplier test

The ARCH-LM test, also known as the Lagrange Multiplier test, is used to test for the presence of autocorrelation in the squared series. The presence of autocorrelation shows that there are conditional heteroscedasticity/ ARCH effects. The test was developed by Engle (1982). It tests the null hypothesis H_0: $\alpha_0 = ... = \alpha_m = 0$, against the alternative that at least one α is not zero in the linear equation

$$a_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + ... + \alpha_m a_{t-m}^2 + \epsilon_t, t = m+1, ..., T,$$

where ϵ_t is the error term, T is the sample size and m is some positive integer. The test statistic is as follows:

$$F = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(T-2m-1)}, \tag{5.12}$$

where $SSR_0 = \sum_{t=m+1}^{T}(a_t^2 - \bar{\omega})$, where $\bar{\omega}$ is the sample mean of a_t^2 and $SSR_1 = \sum_{t=m+1}^{T}\hat{\epsilon_t^2}$ where $\hat{\epsilon_t^2}$ is the least squares residual of the prior linear regression. The test statistic is asymptotically distributed as a χ^2 distribution with m degrees of freedom. The null hypothesis is rejected for the alternative if the test statistic is larger than the critical value.

5.3 Diagnostic checks of volatility models

Diagnostics are done on the residuals $\hat{a}_t = z_t - \hat{\mu}_t$ so as to check the adequacy of the model. Where $\hat{\mu}_t$ is the fitted conditional mean of z_t.

5.3.1 Ling and Li statistics

Ling and Li (1997) developed a method of diagnosis of multivariate time series, which unlike other approaches such as (Box and Pierce) and (Mcleod and Li), did not use residuals autocorrelations. However, it used the sum of the squared residual autocorrelations to develop several new portmanteau statistics.
Assume that the innovation ϵ_t also satisfies
(a) $E(\epsilon_{it}^3) = 0$ and $E(\epsilon_{it}^4) = c1 < \infty$ for i=1, . . . , k, and
(b) ϵ_{it} and ϵ_{jt} are mutually uncorrelated up to the fourth order for $i = j$. Ling and Li (1997) employed $\hat{a}_t$ to propose a model checking statistic for volatility models. Let

$$\hat{\epsilon}_t = \hat{a_t'}\hat{\boldsymbol{\Sigma}_t}^{-1}\hat{a}_t, \tag{5.13}$$

be a transformed quadratic residual series. If the fitted model is correctly specified, then, by the ergodic theorem,

$$\frac{1}{T}\sum_{t=1}^{T}\hat{\epsilon}_t = \frac{1}{T}\sum_{t=1}^{T}\hat{a}_t'\hat{\boldsymbol{\Sigma}}_t^{-1}\hat{a}_t \rightarrow a.s \;\; E(a_t'\boldsymbol{\Sigma}_t^{-1}a_t = E(\epsilon_t'\epsilon_t) = k, \tag{5.14}$$

where $\rightarrow a.s$ denotes almost sure convergence or convergence with probability 1.
The lag-l sample autocorrelation of ϵ_t, therefore, can be defined as

$$\hat{\rho_\ell} = \frac{\sum_{t=\ell+1}^{T}(\hat{\epsilon}_t - k)(\hat{\epsilon_{t-1}} - k)}{\sum_{t=1}^{T}(\hat{\epsilon}_t - k)^2}. \tag{5.15}$$

If the model is correctly specified

$\frac{1}{T}\sum_{t=1}^{T}(\hat{\epsilon}_t - k)^2 \rightarrow$ a.s $E(a_t'\boldsymbol{\Sigma}_t^{-1}a_t - k)^2$ as n $\rightarrow \infty$,

and

$$E(a_t'\boldsymbol{\Sigma}_t^{-1}a_t - k)^2 = E(\epsilon_t'\epsilon_t)^2 - k^2 = E(\epsilon_{it}^4 - 1)k = ck,$$

where $c = E(\epsilon_{it}^4) - 1$ Since the denominator of $\hat{\rho_\ell}$ in Equation 5.15 converges to a constant, it suffices to consider the numerator in studying the limiting properties of $\hat{\rho_\ell}$. By letting

$$\hat{C_\ell} = \frac{1}{T}\sum_{t=\ell+1}^{T}(\hat{\epsilon}_t - k)(\hat{\epsilon_{t-1}} - k) \tag{5.16}$$

be the lag ℓ sample autocovariance of the transformed residual $\hat{\epsilon}_t$ and C_ℓ be its theoretical counterpart with $\hat{\epsilon}_t$ replaced by $\epsilon_t = a_t'\boldsymbol{\Sigma}_t^{-1}a_t$. To investigate the properties of $\hat{\rho_\ell}$ as a function of the estimate $\hat{\theta}$, Taylor series of expansion is used on $\hat{C_\ell} \approx C_\ell + \frac{\partial C_\ell}{\hat{\theta}}$ as seen in Ling and Li (1997) and Tsay (2013).

Finally, the test static is given as

$$Q_{ll}(m) = T\rho_m'\hat{\Omega}^{-1}\hat{\rho_m} \tag{5.17}$$

where Ω is as defined and derived in Tsay (2013).

5.4 Model selection

Model selection is a critical aspect of modelling. This is because different models tend to give different results/predictions. It then follows that one model is better than the other. Another issue is that other models may outdo other models but use too many model parameters. These models are said to be not parsimonious (having few parameters). The best model should make a compromise on accuracy and parsimony. Various model selection criteria exist, and we will discuss some of them here.

5.4.1 The Akaike Information Criterion (AIC)

Perhaps the most used information criterion in academic papers. Developed by Akaike (1974), the AIC is defined as:

$$AIC = \frac{-2}{T} ln(likelihood) + \frac{2}{T}(p), \quad (5.18)$$

$$AIC = \ln(\tilde{\sigma}^2) + \frac{2p}{T}. \quad (5.19)$$

where T is the sample size and p is the number of parameters in the model. The likelihood function is evaluated at the maximum likelihood estimates. The model with the least AIC value is the better model. The criterion seeks the model with minimum $\tilde{\sigma}^2$. However, this value will always decrease as the number of parameters in the model increase, i.e. variance decreases with sample size increase. To cater for that, the model adds a penalty function that is dependent on the number of parameters p. Ideally, the best model will have the lowest AIC and the lower number of parameters.

5.4.2 The Bayesian information criterion (BIC)

The Bayesian information criterion (BIC) also known as the Schwarz Bayesian information criterion (SBIC) was developed by Schwarz et al. (1978).

$$BIC = -2\ln(likelihood) + p \ln T, \quad (5.20)$$

$$BIC(\ell) = \ln(\tilde{\sigma}^2) + \frac{p\ln(T)}{T}. \quad (5.21)$$

The BIC is similar to the AIC but differs in the penalty function. It has a larger penalty function and penalises models with more parameters more.

5.4.3 Hannan-Quinn information criteria (HQIC)

The Hannan-Quinn criterion was introduced by Hannan and Quinn (1979). It is given as follows;

$$HQIC = 2\ln[\ln(T)(P - 2\ln(likelihood), \quad (5.22)$$

$$HQIC = \ln(\tilde{\sigma}^2) + \frac{2\ell\ln(T)}{T}. \quad (5.23)$$

HQIC is similar to the AIC and BIC but also differs in the penalty function. The penalty function is less than that of the BIC but more than that of AIC.

In summary, BIC has a stricter penalty term than AIC, while HQIC has a larger penalty than the AIC but less than the BIC. BIC is also more consistent but less efficient, while the AIC is consistent it tends to select models which are not parsimonious as it does not penalise for parameter increase.

5.5 Model diagnostics

Models are created with assumptions required to be met by the data to give reliable results. These assumptions have negative consequences when violated; hence there is a need to check that the data meets these assumptions. Sometimes, when the data violates the assumptions, it can be transformed to meet the assumptions. One example is a non-stationary series, which can be made stationary by taking logs or by differencing the data. Ideally, a model works by capturing the stylised facts of the data in question. Hence, when developing models, one must be wary of the stylised facts of the data that will be modelled. This section will discuss the diagnostic tests used in this thesis.

5.5.1 Normality test

The majority of model assumptions assume that the data follows a normal distribution. One of the goodness of fit tests to be utilised is the test for normality in the data and in the standardised residuals. The tests to be considered are the Jarque-Bera (JB), Shapiro-Wilk, and Anderson-Darling (AD), which are numerical statistical tests. The study will also consider visual tests using the probability plot and quantile-quantile plot.
First the skewness which is the normalized third moment and also Kurtosis which is the normalised fourth moment of X are defined.

$$S(x) = E\left[\frac{(X-\mu_x)^3}{\sigma_x^3}\right], \qquad K(x) = E\left[\frac{(X-\mu_x)^4}{\sigma_x^4}\right]. \tag{5.24}$$

where S(x) is the skewness and K(x) is kurtosis. Skewness describes the asymmetry of the distribution; for the normal distribution, this value is zero, implying perfect symmetry. A negative and positive value for skewness means that the distribution is negatively or positively skewed, respectively. On the other hand, kurtosis describes the peakedness of the distribution. It essentially measures the extent to which observed data falls near the centre of the distribution or the tails thereof. It is a measure of peakedness. A distribution that is more peaked than the normal distribution will have more kurtosis (Karlsson 2002). The normal distribution has K(x) = 3 hence a new measure quantity K(x) - 3, known as the excess kurtosis, is defined as a way of ease of comparison to the normal distribution excess kurtosis. The excess kurtosis for the normal distribution is thus 0. If the distribution has positive excess kurtosis, it has more data points on the tails compared to the normal distribution. The opposite is also true.

5.5.2 Jarque-Berea (JB) test

Introduced by Jarque and Bera (1987), the JB test checks for normality by comparing the skewness and kurtosis of the data to that of the standard normal distribution. The JB variable is defined as:

$$JB = T[\frac{S(\hat{x})^2}{6} + \frac{(\hat{K}(x) - 3)^2}{24}], \tag{5.25}$$

where the $S(\hat{x})^2$ and $\hat{K}(x) - 3)$ are the skewness and the excess kurtosis respectively. The null hypothesis is that the data conforms to the normality assumption, while the alternative rejects normality. The JB variable is asymptotically distributed as a chi-squared random variable with 2 degrees of freedom. Here we reject H_0 if JB $> \chi^2_{\alpha,2}$ or if the p-value is less than α.

5.5.3 Shapiro-Wilk test

The Shapiro-Wilk test was proposed in 1965 by Shapiro and Wilk (1965). It tests if a random sample X conforms to the normal distribution. The test calculates a W test statistic which is defined as

$$W = \frac{(\sum_1^n a_i x_i)^2}{\sum_1^n (x_i - \bar{x})^2}. \tag{5.26}$$

where the x_i are the ordered sample values starting with the smallest as x_1 and the a_i are mean generated constants, variances and covariances of the order statistics from a normal distribution. The null hypothesis of normality is rejected if the calculated p-value is less than α, the significance level.

5.5.4 Anderson-Darling test (AD test)

$$An = -T - \frac{1}{T}\sum_{i=1}^{T}(2i-1)\{\ln F^*(x_i) + \ln(1 - F^*(x_{T+1-i}))\}, \tag{5.27}$$

where $F^*(x_i)$ is the cumulative function of the distribution and the $x_i's$ are the ordered data.

The test statistic, An is defined under the null hypothesis of standardized residuals following the specified distribution. It is a one-sided test, and the null hypothesis is rejected, if An is greater than the critical value (given by the table of AD test) or the p-value is less than α significant level

5.5.5 Quantile-Quantile (Q-Q) plot

This is a graphical diagnosis method. The Q-Q plot is used to check if data comes from the assumed distribution, for example, the normal distribution. For the normal distribution, the Q-Q plot is known as the normal Q-Q plot. It is a scatter plot of the quantiles of the first data set and the second data, with the second being a pre-assumed distribution.

The graph also has a 45-degree straight line from the y and x-axis intersection, which acts as a reference point. If both data sets are from the same distribution, then the points should roughly lie on that straight line. If both sets of quantiles came from the same distribution, we should see the points lying close to the 45-degree straight line. An example of a normal Q-Q plot with almost perfect data is shown in Figure 5.1.

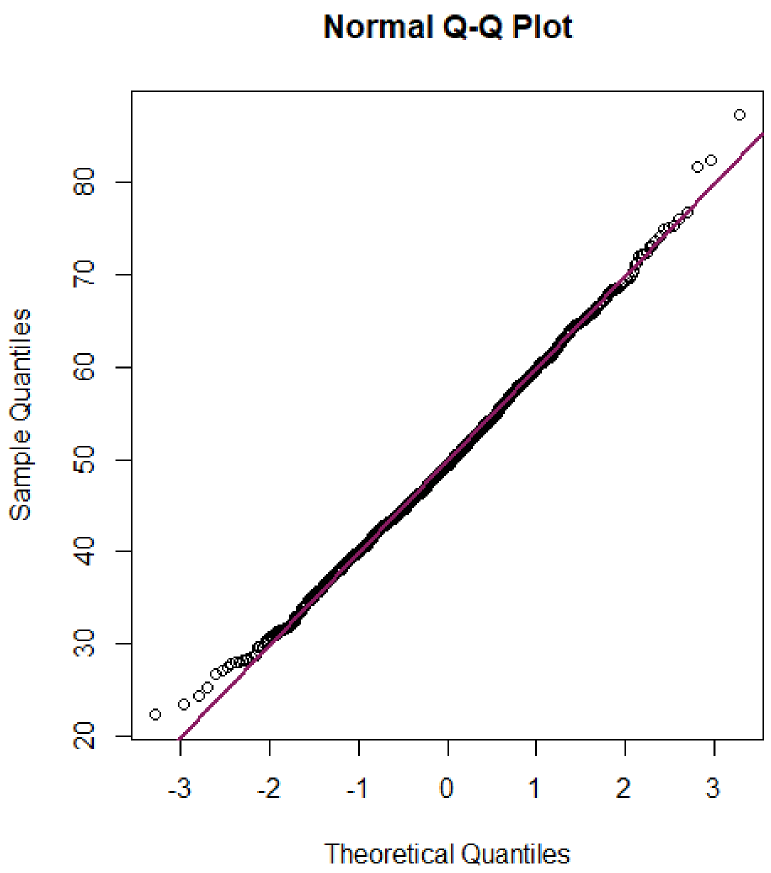

Figure 5.1: Example of Normal QQ plot

5.5.6 Probability plot

Chambers et al. (2018) introduced the probability plot. It works similar to the Q-Q plot. It assesses if a data set follows a given theoretical distribution. A scatter plot is created by plotting the data against the theoretical distribution of choice. Like in the Q-Q plot, if the data fit the distribution, then the points should lie roughly on the straight line, with deviations from the straight line showing a deviation from the assumed distribution.

5.6 Diagnostic procedure for VaR

After fitting the models and making VaR estimates, there is a need to evaluate the model's predictive capabilities. One of the most commonly used methods is the mean square of forecast errors (MSFE) in out-of-sample data. It is also affectionately known as "backtesting" in the financial space. This procedure is discussed below.

5.6.1 Backtesting

Backtesting is the process of comparing the predicted losses by the VaR model and the actual losses experienced in the testing period. It helps the risk managers and investors understand the model's capabilities. The data is divided into two groups, i.e., the estimation sub-sample and the forecasting sub-sample. The choice of the division point is up to the modeller, but it should allow enough data to create a reliable/accurate model. The model estimation is done using the first sub-sample, and the forecast is done using the remaining sub-sample. The backtesting ensures that actual losses do not exceed expected losses at a given level of confidence. The number of observations/exceedances that are acceptable to be above the expected levels should not exceed one minus the confidence level. For example, for VaR with 99% confidence, then less than 1% of the observations are allowed as exceptions.
To carry out the backtesting, the Kupiec likelihood ratio unconditional coverage test of Kupiec et al. (1995) is used.

5.6.2 Kupiec likelihood ratio test

Kupiec et al. (1995) test is the earliest proposed VaR backtesting technique. It examines the observed frequency of losses along the tail and compares them with the frequency of tail losses predicted by the model. Under the null hypothesis of good estimation, the number of tail losses x follows a binomial distribution (Dowd 2007)

$$Pr(x|n,p) = \binom{n}{x} p^i (1-p)^{n-x}, \tag{5.28}$$

where n is the number of observations and p, the predicted frequency of tail losses, is equal to 1 minus the VaR confidence level.
The test statistic of the the Kupiec statistic, given as follows:

$$LR = 2\ln((\frac{x^{\alpha}}{n}^{x^{\alpha}}.(1-\frac{x^{\alpha}}{n})^{n-x^{\alpha}} - 2\ln(\alpha^{x^{\alpha}}(1-\alpha)^{n-x^{\alpha}}), \tag{5.29}$$

which follows a $\chi^2(1)$.

5.7 Chapter summary

The chapter introduced various methods used to diagnose the data. This was followed by model selection techniques which allow us to select the best model for the data amongst a group of different models. The study also discussed the model checks, which evaluate the goodness of fit of the models. Lastly it discussed the techniques to evaluate VaR. In the next chapter the modelling of the data will commence.

6. A COMPARISON OF THE STYLISED FACTS OF BITCOIN, ETHEREUM AND THE JSE STOCK RETURNS

This chapter is a data exploratory chapter. The chapter focuses on understanding the characteristics of cryptocurrency while including the JSE stock return data for comparison. Knowing that cryptocurrency is still relatively new compared to stock market data, it is necessary to constantly evaluate the characteristics as they keep changing as mare information gets available.

6.1 Introduction

Financial data is different and exists in different markets. Decades of research have shown that some assets have very similar statistical properties, although varying to some degree. The homogenous properties are mostly seen on assets within the same class or markets. These types of properties are known as "stylised facts." While stylised facts are not the same, they help develop a general understanding of assets, which in turn aids in model development and decision-making. An adequate model depends on the unobservable data generation process for any data. While the process is unobservable, we can observe the patterns within the data. It is this pattern that the stylised facts reveal.

Various scholars first reported the unique characteristics of financial data early in the 20th century. Mitchell (1938) was the first to report the existence of fat tails in financial data; other scholars, such as Mitchell (1938) and Mills (1927) also made such reports. Kendall and Hill (1953) and Houthakker (1961) were among the first to give empirical evidence that prices were non-stationary. This finding would lead to most models requiring transformed data prices, such as the returns, which are stationary. Cont (2001) discuss the various stylised facts that exist in financial markets and further explain how these invalidate the use of traditional time series models. The perfect model would require taking into account these various stylised facts. Thompson (2011) explains that if a model reasonably approximates the actual data generating process for a stock price, then the fitted model should exhibit the stylised facts. In essence, the stylised facts further act as a model check.

In 2008, Nakamoto (2008) proposed a new form of a financial asset known as cryptocurrency. This new currency was paperless and an online virtual currency based on computer blockchain technology. The paper argued that this peer-to-peer transaction was purely computer-based, requiring proof of work to prevent double-spending, which had been a problem with other online payments. In 2009, the creation of the first cryptocurrency, Bitcoin, occurred. Since then, many other cryptocurrencies have been created. Among the new cryptocurrencies is Ethereum, which is currently the second-biggest cryptocurrency by market value as per CoinMarketCap (n.d.) data.

Literature is abounding with studies of the stylised facts of stock returns. This is not surprising, as stocks are the most traded asset class and have been in existence for decades. The emergence of cryptocurrency meant a new form of financial data that also had to be understood. The author notes that other studies have investigated the stylised facts in cryptocurrency, and some of the existing literature is considered here. Zhang et al. (2018) studied the stylised facts of 8 cryptocurrencies, namely Bitcoin, Ethereum, Litecoin, NEM, Stellar, Monero and Ripple. Using various data exploratory methods such as the Hurst exponent and models such as the GARCH and GJR-GARCH, their main findings were that heavy tails, volatility clustering, leverage effects, and long-range dependence exist across the coins, as evidenced by the Hurst exponent. The presence of fat tails is also reported by Takaishi (2018) using the 1-min return of Bitcoin. The study also showed that the kurtosis differs from that of the normal distribution, but for large time scales, it converges slowly to that of the normal distribution. Bouri, Molnár, Azzi, Roubaud and Hagfors (2017) using daily and weekly data, they found that Bitcoin is negatively skewed and was associated with higher volatility compared to the stock return data from the DAX 30 Index, SP500, and FTSE 100. Dyhrberg (2016) demonstrated that positive and negative shocks have no asymmetric effect on Bitcoin returns and gold, allowing Bitcoin to be used as a hedge. Shanaev and Ghimire (2021) noted that Bitcoin and Ethereum relatively stabilise with time and can be modelled using asymmetric power-law distributions. Drożdż et al. (2018) showed that the Hurst exponent for Bitcoin using 1-minute data was slightly less than 0.5, meaning that Bitcoin did not exhibit persistence for earlier years. However, after 2018, the Hurst exponent hovered around .5 and above, showing persistence. Zhang et al. (2020) using Bitcoin, Ethereum, and Litecoin reject the presence of autocorrelation and conclude that the hourly data of cryptocurrency is in line with an inefficient market. Costa et al. (2019) using daily data, found that Bitcoin and Ripple tend to act as efficient financial assets, with Ethereum and Litecoin showing persistence.

As noted from past studies, there is no consensus on all the findings. The properties of cryptocurrency need to be further studied as more and more data becomes available. This additional data would help in more accurate modelling and forecasting. Our study seeks to expand the current knowledge base on the properties of cryptocurrencies by using Bitcoin and Ethereum whilst also comparing with Gold and the South African stock exchange data represented by the FTSE/JSE Top 40 index. The reason for using Gold is that Bitcoin has been named by most as the new digital Gold, while JSE is chosen to make a case for local investors to be able to compare and contrast this new asset to the well-known FTSE/JSE Top 40. This may be handy in portfolio allocation and diversification. Since cryptocurrency is a relatively new form of currency, it would not be surprising to discover new or different properties from those found in past studies using older data sets. The study follows the methodology used in the seminal work of De Vries and Leuven (1994) where he examined the stylised facts of exchange rates. It, however, make provision for bull and bear markets to further investigate the effect of the bullish and bearish markets on the results. The study uses the daily return data of these assets to investigate the tail behaviour, the volatility clustering, autocorrelation nature, the long-run dependency, and leverage effects of the assets.

6.2 Methodology

The study will follow the methodology used by De Vries and Leuven (1994). However, it will only focus on examining the following: fat tail phenomenon, skewness, autocorrelation, long-range dependence, volatility clustering and the leverage effect.

6.2.1 Skewness, kurtosis

One of the most common ways to check for the normality assumption is through the third and fourth-moment measures, i.e. skewness and kurtosis. The skewness is concerned with the symmetry of the distribution, whilst the kurtosis is for the tail distribution. The skewness and kurtosis are defined by:

$$S(x) = E\left[\frac{(X-\mu_x)^3}{\sigma_x^3}\right] \qquad K(x) = E\left[\frac{(X-\mu_x)^4}{\sigma_x^4}\right], \tag{6.1}$$

where S(x) is the skewness and K(x) is the kurtosis. For the Gaussian distribution, the kurtosis is known to be 3. Hence, a new measure called excess kurtosis is defined as K(x)-3. With the Gaussian distribution having an excess kurtosis of zero. Therefore, any distribution with negative excess kurtosis has thinner tails than the normal distribution (platykurtic distribution), while those with positive excess kurtosis have thicker tails (leptokurtic distribution).

6.2.2 Hurst exponent

Hurst (1951) introduced a measure to determine the long-range dependence in hydrology. This measure was later applied to other fields and generalised by researchers such as Mandelbrot and Van Ness (1968) and then popularised in financial time series by Peters (1996), Peters (1994) where it became a standard a measure for long-term memory of a time series. The measure ranges between 0 and 1, with values close to 0 having the least long-range dependence and those closer to 1 having the most long-range dependence. The Hurst exponent is linked to the Efficient Market Hypothesis (EMH), where market efficiency is achieved when prices of the different assets that make up the market incorporate all relevant information entirely and instantaneously. That is to say, the current price incorporates all the information available. This means that the return movements are independent over time and that there is no memory in the financial market with the prices behaving as a random walk.
Mitra (2012) gives this interpretation of the Hurst Exponent,

- When H = 0.5 is for an uncorrelated series, hence suggesting an independent process. It is associated with a series with with low autocorrelations that in small time delays may initiate a completely uncorrelated series, but where the absolute values of autocorrelations decrease exponentially and rapidly to zero

- When $0.5 < H < 1$ it indicates long-term autocorrelation, hence implying a persistent time series i.e. they show long-term memory.

- When $H < 0.5$ it indicates anti-persistence time series. Where high values can be followed by a low value with not much consecutive values of the same magnitude.

6.2.3 GARCH(1,1) model

The GARCH model was developed by Bollerslev (1986) as an improvement to the ARCH model of Engle (1982). The model estimates current conditional variance as the sum of past conditional variances and the past squared shocks. Using the notation of Tsay (2014), let r_t be the log return at time t. Then r_t follows the GARCH(1,1) model if

$$r_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2, \tag{6.2}$$

where ϵ_t is a sequence of Gaussian white noise with mean 0 and variance 1. $\omega > 0, \alpha \geq 0, \beta \geq 0$, and $\alpha + \beta < 1$. Of interest will be the significance of the α and β parameters. β determines the persistence of volatility after a shock, and together with the parameter α they determine the speed of mean reversion.

GJR-GARCH model

Glosten et al. (1993) adopted the GARCH model to be able to capture the asymmetric volatility. The GJR-GARCH(1,1) is then as follows:

$$r_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma S_{t-1} \epsilon_{t-1}^2, \tag{6.3}$$

where,

$$S_{t-1} = \begin{cases} 1, & \text{if } \epsilon_t \leq 0, \\ 0, & \text{if } \epsilon_t > 0. \end{cases}$$

with $\alpha > 0, \beta > 0, \gamma > 0$ and $\alpha + \beta + 0.5\gamma < 1$. Based on the equation it can be seen that the impact of negative returns is different from that of positive returns. A positive return means that S_{t-1} is 0 hence the equation reduces to the standard GARCH(1,1) i.e. it becomes:

$$\sigma_{t+j}^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2. \tag{6.4}$$

While a negative return increases the volatility by adding the $\gamma \epsilon_{t-1}^2$ term. The resulting equation is shown below.

$$\sigma_{t+j}^2 = \omega + (\alpha + \gamma) \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2. \tag{6.5}$$

It is important to note that the increase in volatility after a negative return will occur if the gamma is positive. This is the standard leverage effect, but sometimes a negative gamma is witnessed, giving rise to the reverse leverage effect.

6.3 Data

The data used for this analysis are the daily prices for Bitcoin, Ethereum, Gold and the FTSE/JSE 40. The FTSE/JSE 40 index represents the top 40 companies by market capitalisation on the Johannesburg Stock Exchange. The data is from 18 September 2017 to 27 May 2021 and was retrieved from the website *Financial News and Stock quotes* (n.d.). The data period was selected based on data availability and also when cryptocurrency started gaining momentum. Because price data has a trend and is not stationary, daily prices are converted to daily log returns. By letting the price be P_t where t = 1, 2, . . . , N, P_t is converted to log-returns using the formula $r_t = \ln(P_t / P_{t-1})$.

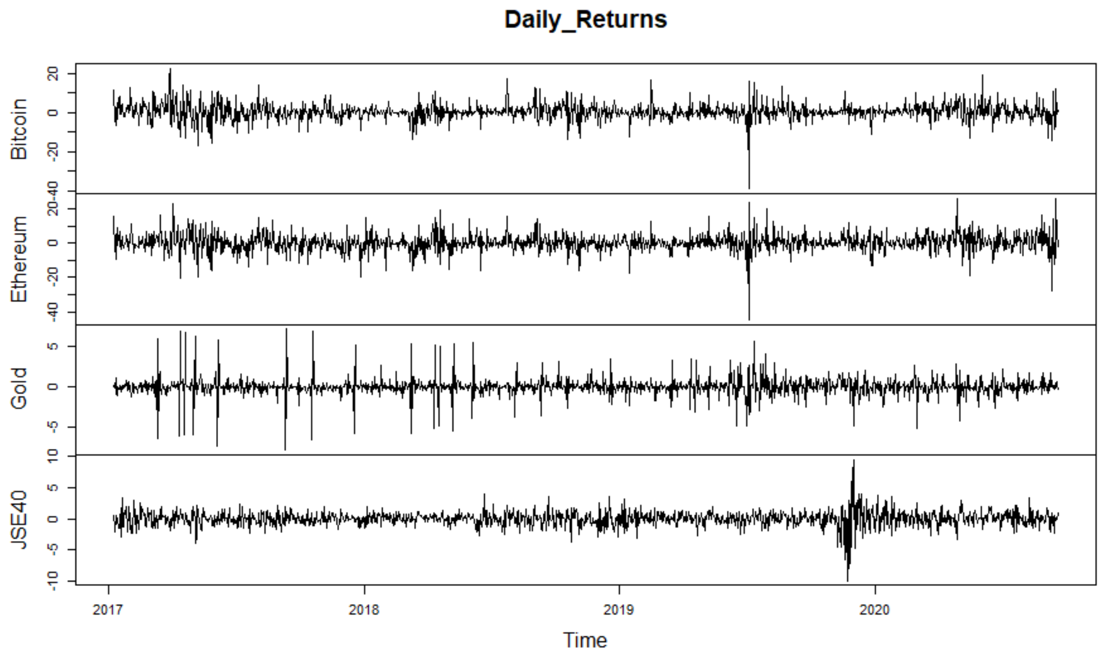

Figure 6.1: Log return series plots from 17 September 2017 to 27 May 2021

Figure 6.1 above shows the log return plots. It can be seen that the Gold and JSE are mainly bound within -5 and 5, whilst Bitcoin and Ethereum returns are spikier and are mainly bound within -10 and 10. This observation shows that Gold and JSE are relatively stable compared to the two cryptocurrencies.

6.4 Data Exploratory Analysis

Table 6.1 below shows the descriptive statistics for the daily log returns for the whole period. The maximum daily log return is highest for the cryptocurrencies, with Ethereum having 25.96% and Bitcoin with 22.55%. The JSE had a maximum return of 9.48% while Gold was the lowest with 7.14%. The corresponding minimum returns were lowest for the cryptocurrencies, with Ethereum having -44.55% and Bitcoin having -39.18%. On the other hand, the JSE index had a -9.92% return and Gold -7.8% return. It is interesting to note that for Gold and the JSE index, the maximum and negative magnitude are almost identical, i.e. -7.8% to 7.14% for Gold and -9.92% to 9.48% for the JSE. For the average daily log returns, Ethereum has the highest average with 0.32% and also the highest standard deviation of 5.3%. Bitcoin has the second-highest return of 0.26% and standard deviation of 4.2%, while the Gold and JSE index had average daily returns of 0.02%, 0.03%, and standard deviation of 1.21%, 1.23% respectively. The skewness value was similar across the four assets. All assets had a negative skewness of -0.277,-0.288, -0.106 and -0.302 for Bitcoin, Ethereum, Gold and JSE, respectively. The excess kurtosis for Bitcoin, Ethereum, Gold and FTSE/JSE40 is 7.8, 6.18, 12.008 and 10.13 respectively, all greater than 0 indicating that the returns have thicker tails than the normal distribution.

Table 6.1: Descriptive summary for the full period

Asset	Minimum	Maximum	Mean	Median	Stdev	Skewness	Kurtosis
Full period							
Bitcoin	-39.1816	22.5512	0.2674	0.1622	4.2422	-0.2777	7.8431
Ethereum	-44.5472	25.9572	0.3254	0.1400	5.3682	-0.2886	6.1876
Gold	-7.8056	7.1446	0.0244	0.0266	1.2155	-0.3023	12.0756
JSE	-9.9229	9.4798	0.0306	0.0739	1.2337	-0.1061	10.1323

Table 6.2 shows the descriptive summary of the data series separated into the bullish periods of Cryptocurrency. The mean return is greater in cryptocurrencies, where the average mean return is around 0.5%. This is almost double the mean return of the entire period data. The daily return for Gold and JSE is more than 15 times less than that of the cryptocurrencies. It also did not change much from the mean return observed throughout the period.

Table 6.2: Descriptive summary for Bullish Periods

Asset	Minimum	Maximum	Mean	Median	Stdev	Skewness	Kurtosis
Bull Period Mar 2016-Dec 2017							
Bitcoin	-31.3422	25.8599	0.6580	0.0000	6.5683	0.0242	3.1154
Ethereum	-18.6939	22.7602	0.4857	0.3554	4.1178	-0.0815	4.9994
Gold	-6.9991	7.8220	0.0111	0.0218	1.6283	0.2759	8.8454
JSE	-4.0493	2.2837	0.0318	0.0718	0.8645	-0.3162	1.0927
Bull Period Jan 2019-Jun 2019							
Bitcoin	-17.4669	14.4330	0.4395	0.1760	4.4931	-0.1512	2.8372
Ethereum	-14.6123	15.8967	0.5914	0.3362	3.5605	0.2548	4.6498
Gold	-5.0236	5.4414	0.0573	0.0773	1.3734	0.2685	6.7278
JSE	-3.0783	2.1976	0.0908	0.1034	0.8623	-0.5674	1.2970
Bull Period Mar 2020-Apr 2021							
Bitcoin	-58.9639	23.0189	0.5888	0.5972	5.7545	-2.6116	28.1984
Ethereum	-49.7278	17.7424	0.4872	0.4447	4.4541	-3.2230	39.1615
Gold	-5.1213	5.6266	0.0338	0.0363	1.3107	-0.1412	3.0078
JSE	-10.4504	9.0570	0.1013	0.1672	1.8354	-0.6392	8.5006

Table 6.3 shows the descriptive summary during the cryptocurrency bear period. The mean returns for the cryptocurrencies are negative, indicating the negative market movements during the bearish period. Interestingly, Gold and the JSE are also slightly negative in the first bearish period. In the second bearish period, Gold has a positive mean return. The bear and bull periods considered were for cryptocurrency, but we note that the negative return also occurred in

the other two assets. This revelation suggests that there may be a spillover effect in the markets.

Table 6.3: Descriptive summary for Bullish Periods

Asset	Minimum	Maximum	Mean	Median	Stdev	Skewness	Kurtosis
Bear Period Jan 2018-Dec 2018							
Bitcoin	-22.0368	17.6031	-0.5871	-0.3130	5.5666	-0.2753	1.2042
Ethereum	-16.9752	12.9743	-0.3173	-0.0412	4.1538	-0.3367	1.5507
Gold	-7.7700	7.1446	-0.0124	-0.0257	1.6580	-0.2002	10.1875
JSE	-3.8292	3.8977	-0.0469	0.0220	1.1806	0.0864	0.7664
Bear Period Jul 2019-Feb 2020							
Bitcoin	-19.0467	14.5378	-0.1210	-0.1395	4.0506	-0.6071	3.6696
Ethereum	-14.0443	15.3976	-0.0968	-0.1921	3.4114	0.0778	3.3924
Gold	-4.8772	3.4664	0.0530	0.0755	1.1159	-0.2788	3.5214
JSE	-4.6471	1.9630	-0.0767	-0.0198	1.0338	-1.2233	3.3420

6.5 Stationarity

An essential check in any time series data is assessing the stationarity of the data. If data is non-stationary, it is unpredictable and cannot be forecasted accurately. Instead, the results will be non-consistent and unreliable. The price series of the assets is non-stationary, hence the reason for using the log returns. The study use the Argumented Dicky Fuller test to test the null hypothesis that the log return series is non-stationary. The alternative hypothesis is that the log return series is stationary. The test gave an Augmented-Dickey Fuller statistic of -13.612 with a p-value of $0.01<0.05$, thus rejecting the null hypothesis at 5% significance level, meaning that the log-returns are stationary. The results shown in Table 6.4 below show that all the time series are stationary at all three levels of significance.

Table 6.4: Stationarity tests for the returns

	Test statistic	1%	5%	10%	Conclusion
Panel A daily returns					
Bitcoin	-24.8279	-2.58	-1.95	-1.62	Stationary
Ethereum	-24.5003	-2.58	-1.95	-1.62	Stationary
Gold	-32.6818	-2.58	-1.95	-1.62	Stationary
JSE	-25.1035	-2.58	-1.95	-1.62	Stationary

6.6 Non-Gaussian nature and the heavy tails

In finance, the tail distribution of asset returns is of paramount importance. One cannot just discard the extreme positive and negative values as outliers, as

they have a role in explaining market risk. Against this background, measures such as the Value at Risk (VaR) are established. Early studies have shown that the normal distribution is insufficient for financial data, but more heavy-tailed distributions are suitable. Cont (2001) study revealed that for high-frequency data, i.e., 1-min returns, the distribution was thick-tailed and deviated significantly from the normal distribution. The study revealed a slow convergence to the normal distribution for larger sampling periods.
From Table 6.1 above, there is positive excess kurtosis for all assets under both time scales. The JSE has a more positive excess kurtosis than the two cryptocurrencies. These results suggest that cryptocurrencies have thicker than normal tails but are less thick than stock return data.
The density plots in Figure 6.2 below give a visual comparison between the expected normal distribution, which has the mean and standard deviation of the data, and the corresponding density plot of the actual data. All the data series have thicker tails and are more peaked than the normal distribution. This finding further affirms our findings of deviation from the normal distribution. The same observations are made on the bull and bear period data.

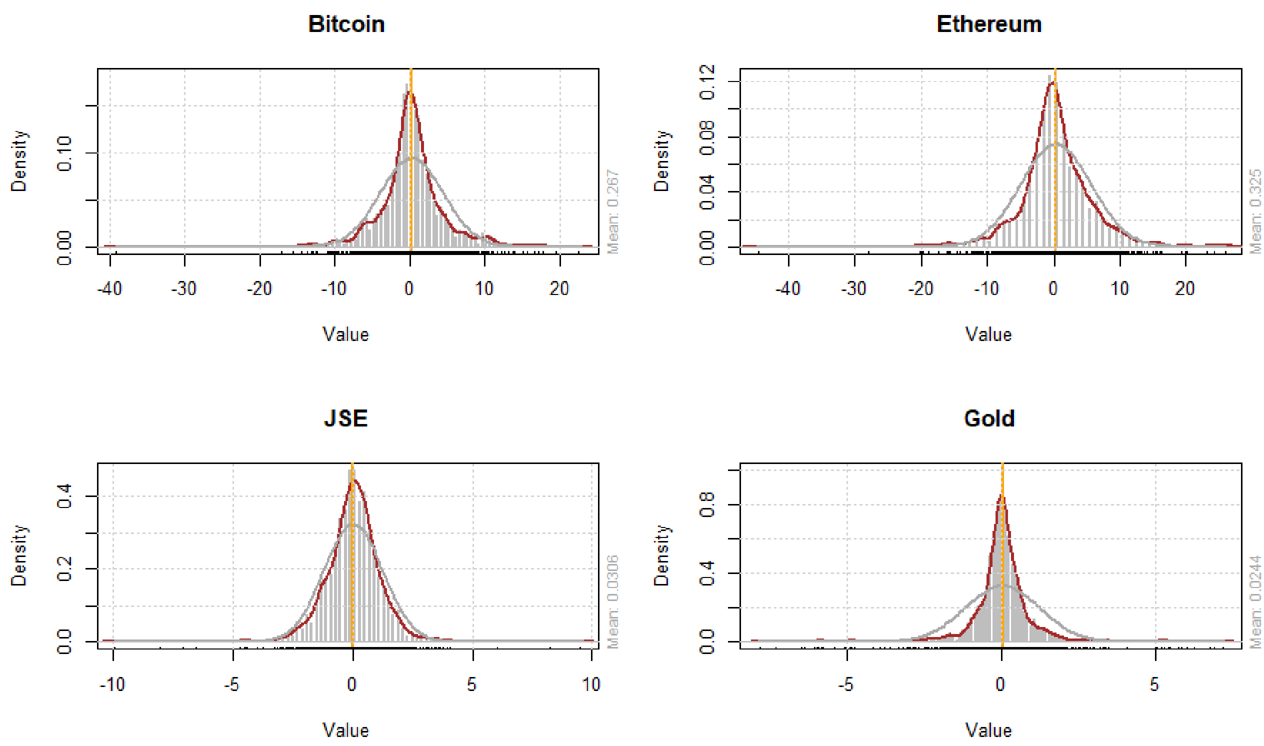

Figure 6.2: Density plots for the daily log returns

6.7 Long range dependency

Another common stylised fact of data is the long-range dependence, whereby some time series exhibit more persistent behaviour than others. We calculated various forms of the Hurst exponent for the asset returns. The results are presented in Table 6.5 below. All measures point to persistent behaviour in the cryptocurrencies, with an average Hurst exponent greater than 0.6 for Bitcoin and Ethereum. However, for Gold and the JSE, the Hurst exponent range was between 0.47 and 0.54, which is very close to the value of 0.50, suggesting that these form an efficient market with their movements resembling a random walk.

Table 6.5: Hurst Exponent measures for long range dependence

Hurst measure	Bitcoin	Ethereum	Gold	JSE
Panel A daily returns				
Simple R/S Hurst estimation	0.5789	0.5931	0.4811	0.5400
Corrected R over S Hurst exponent	0.6579	0.6552	0.5065	0.5564
Theoretical Hurst exponent	0.5367	0.5367	0.5367	0.5367

A known fact is that markets tend to depart from their usual overall behaviour, with efficient markets departing from efficiency now and then. The typical way to check that is to use sub-divisions of the whole series. This study considers the long-range dependence of the Bull and Bear periods. Table 6.6 below shows the Hurst measures for the Bullish periods. The first bull period has the cryptocurrencies exhibit long-range persistence, as the Hurst exponent above 0.54. Gold and JSE still exhibit an efficient market.

Table 6.6: Hurst Exponent for bullish periods

Hurst measure	Bitcoin	Ethereum	Gold	JSE
Bull Period Mar 2016-Dec 2017				
Simple R/S Hurst estimation	0.5809	0.5681	0.4426	0.4975
Corrected R over S Hurst exponent	0.6403	0.5951	0.4317	0.4989
Theoretical Hurst exponent	0.5422	0.5422	0.5424	0.5480
Bull Period Jan 2019-Jun 2019				
Simple R/S Hurst estimation	0.4875	0.5485	0.4824	0.5594
Corrected R over S Hurst exponent	0.5259	0.6135	0.5370	0.6258
Theoretical Hurst exponent	0.5576	0.5576	0.5256	0.5249
Bull Period Mar 2020-Apr 2021				
Simple R/S Hurst estimation	0.4780	0.5226	0.5511	0.5446
Corrected R over S Hurst exponent	0.4923	0.5574	0.6045	0.6230
Theoretical Hurst exponent	0.5620	0.5494	0.5535	0.5484

Table 6.7 shows the Hurst exponent measures during the bear period. The results did not reveal significant changes in the Hurst exponent compared to the bullish period. The Hurst exponent hovered around the same value for all the assets.

Table 6.7: Hurst Exponent for bearish periods

Hurst measure	Bitcoin	Ethereum	Gold	JSE
Bear Period Jan 2018-Dec 2018				
Simple R/S Hurst estimation	0.4890	0.5091	0.4268	0.4852
Corrected R over S Hurst exponent	0.5308	0.5479	0.4102	0.4755
Theoretical Hurst exponent	0.5613	0.5613	0.5717	0.5646
Bear Period Jul 2019-Feb 2020				
Simple R/S Hurst estimation	0.5666	0.4900	0.4654	0.5639
Corrected R over S Hurst exponent	0.6478	0.5053	0.4915	0.5987
Theoretical Hurst exponent	0.5723	0.5723	0.5779	0.5802

Overall, the results show that the Hurst exponent for the whole period considered is around 0.5-0.6 for all the assets. However, fluctuations in the Hurst exponent values obtained for the sub-periods are observed. This tallies with the results of Mitra (2012) where he considered 12 stock series and found the Hurst exponent value of the entire series to be around 0.50, pointing towards market efficiency. Nevertheless, the Hurst exponent value varied wildly once the series was split into 60 trading days.

6.8 Autocorrelation

Just like the Hurst exponent from the previous section, the autocorrelation considers the long-range dependence of the data. The absence of significant autocorrelation or minor low order correlations is one of the known documented stylised facts for stock data. Autocorrelation is defined as the correlation between the return series and its past values. The lag ℓ autocorrelation is defined as

$$\rho_\ell = \frac{Cov(r_t, r_{t-\ell})}{\sqrt{Var(r_t)Var(r_{t-\ell})}} \tag{6.6}$$

where the property $Var(r_t) = Var(r_{t-\ell})$ for weakly stationary series is used. The absence of autocorrelation occurs when $\rho_\ell = 0$ for all $\ell > 0$.

It is in this background that theories such as the efficient market hypothesis are born. This is also used in the Capital asset pricing model, which assumes that there are no autocorrelations in the return series, hence the series being unpredictable.

The Autocorrelation Function (ACF) plots are shown in Figure 6.3. The plots indicate that the return sample ACF are within the 5% bounds with a few just crossing, suggesting that there may be some negligible weak serial correlations. The JSE has slightly more lags that cross the 5% bound. Gold has a significant first lag. Using the Ljung box test, the p-value obtained for all the four assets' daily returns was less than 0.05, suggesting the presence of autocorrelations. These observations are consistent with most studies where the daily and short-period stock return data are autocorrelated. Kaseke (2015) showed that the monthly stock returns for JSE-based companies were not autocorrelated, but the daily stock returns had lower-order serial autocorrelations.

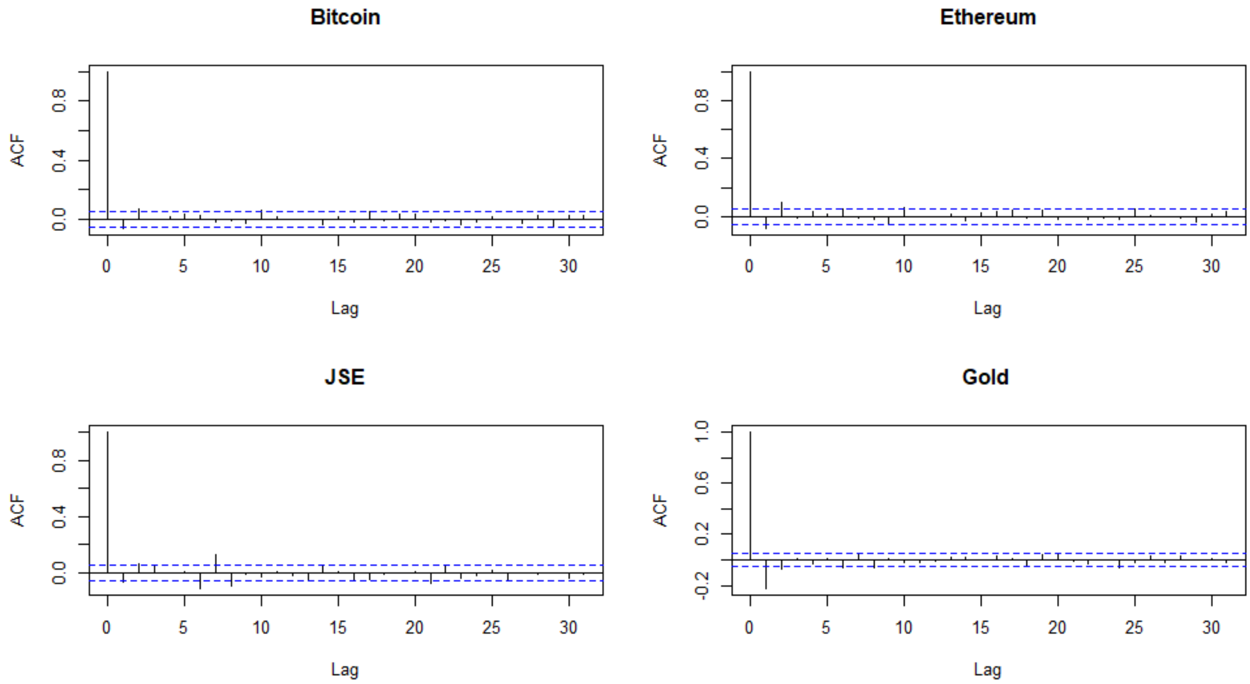

Figure 6.3: ACF plots for period from 17 September 2017 to 27 May 2021

The ACF plots for the Bull periods are shown in Figure 6.4. The results show that there are either no serial autocorrelations or low-order ones for cryptocurrency. This contrasts with the JSE and Gold, where the low-order correlations are slightly more significant.

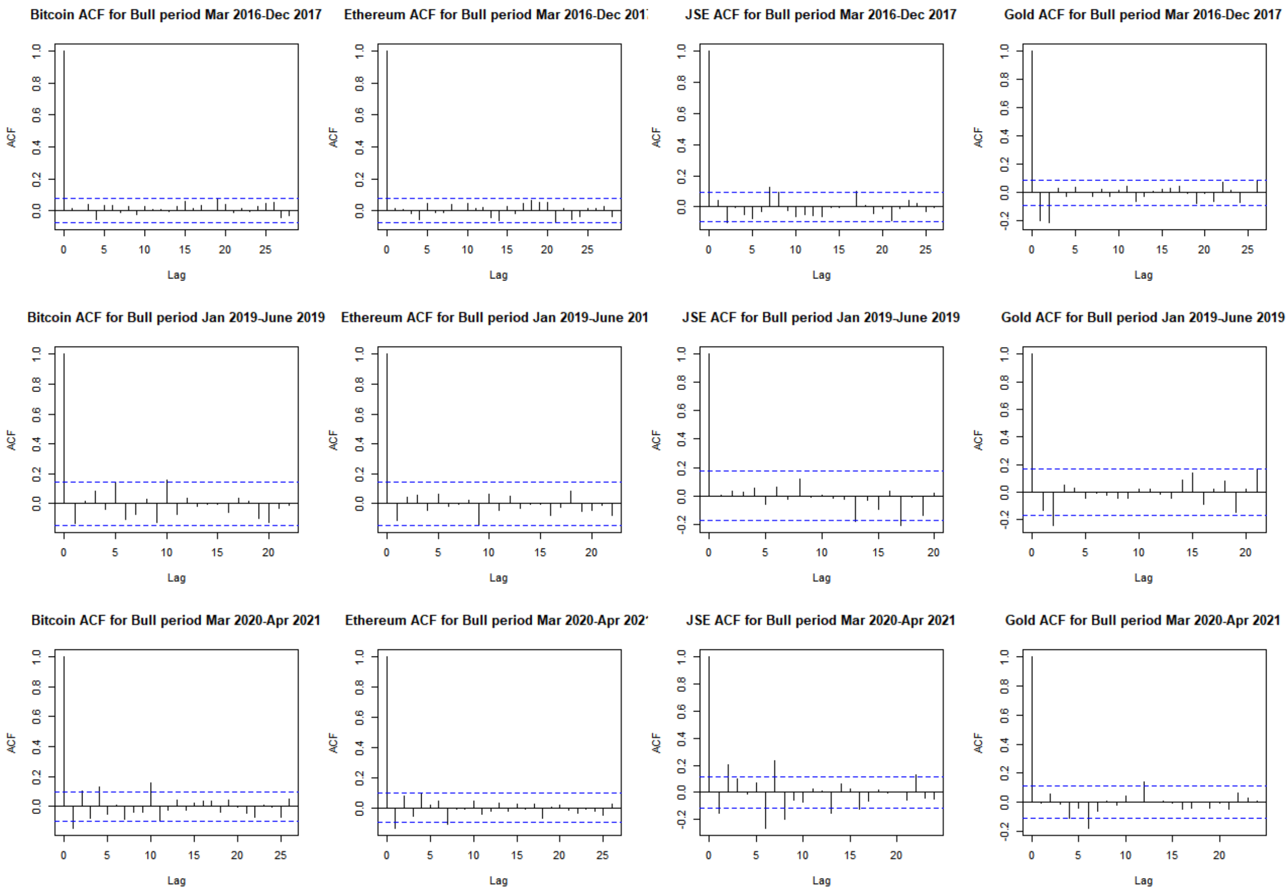

Figure 6.4: ACF plots for the Bull period

Figure 6.5 shows the ACF plots for the Bear periods. The results show no significant lags for the cryptocurrency, while there are minor lower-order significant lags in the JSE and Gold. This suggests that the series for cryptocurrencies are not dependent on cryptocurrency.

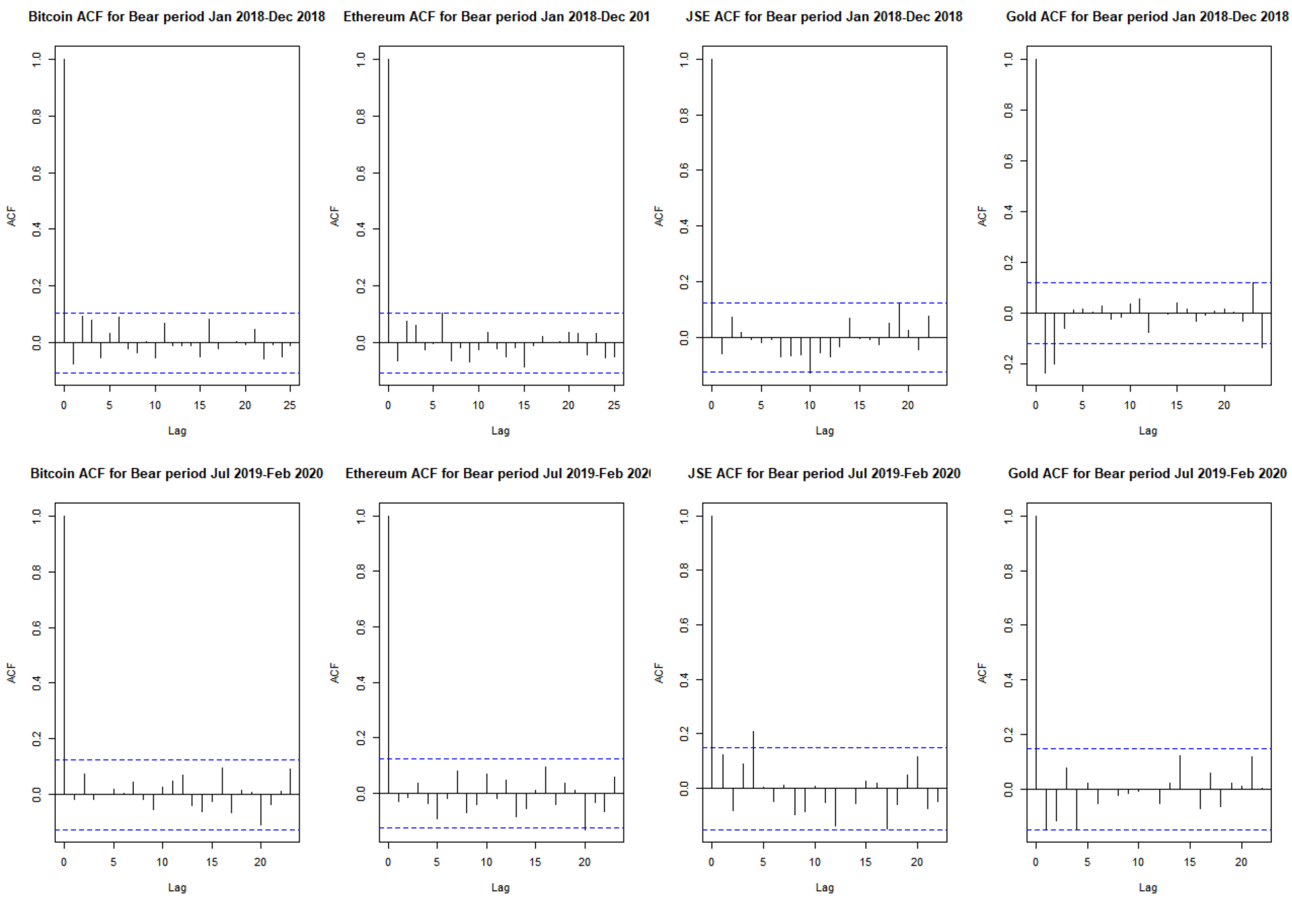

Figure 6.5: ACF plots for the Bear period

Overall, the study showed that the cryptocurrency has no autocorrelations or it has low order serial correlations. This result was further confirmed using numerical test using the Ljung box autocorrelation test. The significant lags are less than those witnessed in the JSE and Gold data. The explanation may be that the bear and bull periods under consideration are for cryptocurrency. The JSE and Gold may have been in a different circle period altogether.

6.9 Volatility clustering

Having no autocorrelations does not mean the presence of independent increments. Other nonlinear functions of the series without autocorrelations may exhibit autocorrelations. This is one of the observations by Mandelbrot (1997). While tests on the return series serial dependence did not detect any dependency, there was a dependency on the squared returns. This dependence is known as volatility clustering, a phenomenon where periods of high return fluctuations are followed by periods of high return fluctuations. This observation also holds for periods of low return fluctuations. This phenomenon is visible in the return plots in Figure 1. Periods of high return seem to occur in clusters; likewise, the periods of low return are also in clusters. The GARCH model of Engle and Bollerslev (1986) as well as other variants of the conditional heteroscedastic models make use of this volatility clustering (also known as ARCH effects) in modelling the dependence in the data. They do this by letting the conditional variance depend on the past squared innovations, directly capturing the volatility clusterings. A way to test for volatility clustering is by using the autocorrelations of the squared returns or the absolute returns as explained by Tsay (2005). He states that one can use the squared series a^2 (where a is the log-returns less the mean of the series, i.e. $a = r_t - \mu_t$) to check for the ARCH effects by applying the Ljung-Box statistics test to the series. Table 6.8 below shows that ARCH effects are present for all assets under daily data.

Table 6.8: Test for ARCH effects

	χ^2 value	df	P-value	Conclusion
Panel A daily returns				
Bitcoin	65.589	12	$<$0.001	ARCH effects present
Ethereum	85.455	12	$<$0.001	ARCH effects present
Gold	140.99	12	$<$0.001	ARCH effects present
JSE	1499.9	12	$<$0.001	ARCH effects present

Zhang et al. (2018) employs the GARCH model of Bollerslev (1986) to test for the volatility clustering. This study uses the same method to test for volatility clustering. A significant α and β parameters indicate that past volatility influences current volatility. If $\alpha + \beta$ is close to 1, it means that the volatility is persistent. The results are presented in Table 6.9 below. For daily data, all the cryptocurrencies have significant α and β parameters. The same is observed for the JSE but not for Gold which has an insignificant alpha parameter. All assets are highly persistent, with $\alpha + \beta$ close to 1 i.e. all above 0.96. The results for Bitcoin and Ethereum are consistent with the findings of Zhang et al. (2018).

Table 6.9: Results of estimation using GARCH(1,1) model

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha} + \hat{\beta}$
Panel A daily returns			
Bitcoin	0.0720***	0.9270***	0.9999
Ethereum	0.1097***	0.8739***	0.9836
Gold	0.0000	0.9990	0.9999
JSE	0.0890***	0.8788***	0.9678

Note: *,**,*** is statistical significance at the 1%,5% and 10% critical level.

The study then considered the Bullish and bearish periods separately. The Bullish periods are shown in Table 6.10 below. All assets had a significant β

parameter. There was no consistent behaviour on the α parameter. In some cases, the α parameter was estimated to be zero for five decimal places.

Table 6.10: Results for the bullish periods

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}+\hat{\beta}$
Bull Period Mar 2016-Dec 2017			
Bitcoin	0.2771***	0.7219***	0.9999
Ethereum	0.2025***	0.7965***	0.9999
Gold	0.0000	0.9990***	0.9999
JSE	0.1069*	0.8084***	0.9153
Bull Period Jan 2019-Jun 2019			
Bitcoin	0.0000	0.9990***	0.9999
Ethereum	0.0854***	0.9136***	0.9999
Gold	0.0000	0.8535***	0.8535
JSE	0.0000	0.9990***	0.999
Bull Period Mar 2020-Apr 2021			
Bitcoin	0.1019**	0.7997***	0.90156
Ethereum	0.0926***	0.9064***	0.9999
Gold	0.1675	0.6616***	0.8291
JSE	0.1941***	0.7757***	0.9697

Note: *,**,*** is statistical significance at the 1%,5% and 10% critical level.

The bear periods did not seem to give any distinguishable estimates from the bull periods. The β parameters were significant for all the assets, while the α parameter was sometimes estimated to be zero, as noted in the bullish periods. Overall persistence was high, above 0.99, except in the second bearish period where Ethereum and JSE had persistence of 0.7638 and 0.6996, respectively.

Table 6.11: Results for the bearish periods

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}+\hat{\beta}$
Bear Period Jan 2018-Dec 2018			
Bitcoin	0.1613**	0.8086***	0.9699
Ethereum	0.0713***	0.9278***	0.9999
Gold	0.0000	0.9990***	0.9999
JSE	0.0000	0.9989***	0.9999
Bear Period Jul 2019-Feb 2020			
Bitcoin	0.0000	0.9990***	0.9999
Ethereum	0.1966	0.5670***	0.7638
Gold	0.0000	0.9990***	0.9999
JSE	0.0408	0.6589***	0.6996

Overall, the β, which is the GARCH parameter, was significantly different from zero in all cases, supporting the presence of volatility clusters. There is no distinct difference in the volatility clustering of the cryptocurrency and the JSE and Gold.

6.10 Leverage effect

Black (1976) first noted that for most financial data, the returns are negatively correlated with the volatility measures of the returns. Since then, many studies using stock return data have confirmed the same results. Research showed that negative shocks at time *t-1* have a larger effect on the variance at time *t* than positive shocks. This phenomenon is what is termed the leverage effect. An explanation for this is that when volatility is high, most investors avoid investing in that stock due to higher risk and increased leverage.

Using the squared returns as a measure of volatility, The study considers the correlations between the returns of our assets and their squared returns. The results are presented in Table 6.12 below. The daily returns of all assets were negatively correlated with their squared returns. This is in agreement to Black (1976) and also Cont (2001). This indicates the existence of leverage effects.

Table 6.12: Correlations between returns and squared returns

Daily	Bitcoin	Ethereum	Gold	JSE
Bitcoin	-0.0485			
Ethereum		-0.0587		
Gold			-0.0176	
JSE				-0.0727

Note: *,**,*** is statistical significance at the 1%,5% and 10% critical level.

The study further explores the leverage effect using a more sensitive GJR-GARCH model. The focus here is the leverage parameter γ. Table 6.13 below shows the obtained results for the assets considered. JSE and Gold had a positive gamma, with JSE being statistically significant at 1% level. The positive gamma indicates the leverage effect on the returns. It means that negative returns result in greater volatility in the next period. For Bitcoin and Gold the γ is negative. This negative has been termed the reverse leverage effect. Stavroyiannis (2018) explains using Gold that, in terms of a crisis, investors are likely to move funds to other assets that are negatively correlated to the market or weakly related to the market, hence creating the negative gamma. These assets act as a safe haven during a crisis. Indeed, it can be seen in Table

6.14 that the two cryptocurrencies are negatively correlated to the JSE returns. However, these results are different compared to previous studies like Zhang et al. (2018) where Bitcoin and seven other cryptocurrencies had a negative γ while Ethereum was positive. Tempering with different GJR-GARCH models shows that the gamma of the Gold changes to negative in some instances. This behaviour is also noted by Stavroyiannis (2018) and where Gold had negative gamma. Stavroyiannis (2018) also discusses the effect of software on the value of gamma obtained. Some software has put a constraint on the gamma value such that it can only take a positive value. However, the condition $\alpha + \beta + 0.5\gamma < 1$ should hold in the case of a negative gamma just as it must hold for a positive gamma. For all assets, we have $\alpha + \beta + 0.5\gamma$ positive and less than 1, meaning that our models are stable with the volatility mean-reverting and fluctuating around the standard deviation.

Table 6.13: Estimates from the GJR-GARCH.

	$\hat{\alpha}$	$\hat{\beta}$	γ	$\hat{\alpha} + \hat{\beta} + 0.5\gamma$
Panel A daily returns				
Bitcoin	0.0773***	0.9326***	-0.0218	0.9999
Ethereum	0.1128***	0.8806***	-0.0163	0.9853
Gold	0.5820***	0.1650***	0.5038*	0.9990
JSE	0.0000	0.8988***	0.1440***	0.9708

Note: *,**,*** is statistical significance at the 1%,5% and 10% critical level.

Table 6.14: Pearson Correlations

Panel A: Pearson correlation

	Bitcoin	Ethereum	Gold	JSE
Bitcoin	1			
Ethereum	0.7476*	1		
Gold	0.0536*	0.0363	1	
JSE	-0.0145	-0.0364	-0.0246	1

Note: * is statistical significance at 5% critical level.

We also considered the GJR-GARCH for the Bullish and Bearish periods separately as shown in Table 6.15 and Table 6.16. For the Bullish periods, we observe that all assets has a negative gamma signifying the presence of the normal leverage effect. There is one incident of Bitcoin in the Mar 2016-Dec 2017 where it exhibited inverse leverage effect. For the Bearish periods we observe inverse leverage effect for the cryptocurrency as compare to the JSE with the normal leverage effect.

Table 6.15: GJR-GARCH for the Bullish periods

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\alpha}+\hat{\beta}+0.5\hat{\gamma}$
Bull Period Mar 2016-Dec 2017				
Bitcoin	0.2735***	0.7218***	0.0074	0.999
Ethereum	0.2524***	0.8056***	-0.1183**	0.999
Gold	0.1496***	1.0000***	-0.3041***	0.998
JSE	0.1799**	0.8854***	-0.1568**	0.987
Bull Period Jan 2019-Jun 2019				
Bitcoin	0.0212	1.0000***	-0.0603***	0.991
Ethereum	0.0764***	1.0000***	-0.1620***	0.995
Gold	0.4005***	0.9421***	-0.9999***	0.843
JSE	0.04304	1.0000	-0.0881	0.999
Bull Period Mar 2020-Apr 2021				
Bitcoin	0.1250*	0.8288***	-0.0581	0.925
Ethereum	0.1174***	0.9319***	-0.1005***	0.999
Gold	0.1005	0.9594***	-0.1218	0.999
JSE	0.2356***	0.8790***	-0.2474***	0.991

Table 6.16: GJR-GARCH for the Bearish periods

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\alpha}+\hat{\beta}+0.5\hat{\gamma}$
Bear Period Jan 2018-Dec 2018				
Bitcoin	0.1484*	0.8072***	0.0271	0.969
Ethereum	0.0328*	0.9337***	0.06488	0.999
Gold	0.7760***	0.5101***	-0.9890***	0.791
JSE	0.0502***	1.0000***	-0.1023***	0.999
Bear Period Jul 2019-Feb 2020				
Bitcoin	0.0079	1.0000***	-0.0179	0.999
Ethereum	0.0947	0.5906***	0.2763	0.823
Gold	0.0000	1.0000***	-0.0035	0.998
JSE	0.0000	1.0000***	-0.0028	0.998

6.11 Conclusion

Stylised facts in financial data have been well documented since the early 19th century, with researchers such as Mandelbrot (1972) and Black (1976) being some of the pioneers. The stylised facts are model-free and seek to assist in the creation of better models for the data. While the financial data used to build up these stylised facts are mainly stock return data, there is no guarantee that these will be present in cryptocurrency. This paper sought to identify the presence of these stylised facts on cryptocurrency or the lack thereof. The study also made use of Gold and JSE/FTSE for comparison. The data was also split into periods of Bullish and Bearish cryptocurrency seasons. The study focused mainly on two cryptocurrencies that are, Bitcoin and Ethereum. The study reveald that, similar to stock return data, the cryptocurrency returns are heavy-tailed, negatively skewed, and possess volatility clustering. A notable difference is the extent of volatility, where the cryptocurrency is more volatile than both Gold and stock returns. The Hurst exponent also showed that there is a higher persistence in the cryptocurrency than one observed in Gold and stock returns. For the entire period, the gamma parameter for the cryptocurrencies was negative, indicating the inverse leverage effect, contrary to the normal leverage effect witnessed on stock returns. The bull period had negative gamma for all assets. The bear period had negative gamma for Gold and JSE but positive for cryptocurrency except for Bitcoin in the first period. The existence of inverse leverage effect is not new as it was also reported in studies such as those by Zhang et al. (2021) and Huang et al. (2022). An explanation for this inverse leverage effect may be that investors use cryptocurrency as a safe haven during market turmoil.

These results can be considered by investors in cryptocurrency when allocating assets in their portfolios. Analysts can also use them when choosing models to use or developing new models to model cryptocurrencies. Another way to use these results is to consider if the cryptocurrency is in the Bullish or Bearish season and make decisions accordingly. Future studies could assess whether cryptocurrency can be used as a hedge in the South African market.

6.12 Chapter summary

The purpose of this chapter was to investigate the stylised facts that exist in cryptocurrency. The study also compared the results with those of the JSE stock returns. This is because cryptocurrency is a different asset class with stock data. The known financial properties are from stock data. The results as shown in the conclusion, show that the two asset classes have similar properties. These characteristics however differ on intensity. Cryptocurrency was more volatile than the JSE market and it also possessed inverse leverage.

This section gives an idea of the models that can be fit for the data. Models should be able to capture the above stylised facts. In the next chapter the study considers modelling the volatility using the GARCH-type models and the development of hybrid GARCH models.

7. MODELING CRYPTOCURRENCY RISKS USING HYBRID GARCH MODELS

This chapter focuses on finding the best GARCH type models that can capture the stylised effects of cryptocurrencies. The best models will then be further modified to allow for extreme value distribution in the conditional variance. This will be novel in the area of cryptocurrency hence the resulting models will be a hybrid GARCH models. Once the hybrid models are formulated, then their ability in quantifying risk using Value at Risk (VaR) is examined. The results of the VaR will be backtested using the Kuipiec test.

7.1 *Introduction*

Amid the 2008 global recession, a new type of digital currency known as cryptocurrency was born. This innovation was a paperless form of currency that would soon change the dynamics of online payments by introducing a new player in the market. Despite a slow start, cryptocurrency quickly took off, with many new digital coins entering the market. In January 2021 VanDenburgh and Daniels (2021) reported that the market cap of all cryptocurrencies reached 1 trillion dollars for the first time. This record would be followed by Bitcoin in February 2022, being the first cryptocurrency and the fastest asset to reach a market cap of 1 trillion dollars. This growth would be a sign to many that cryptocurrency was now a significant player in finance.

The initial cryptocurrency, Bitcoin, was introduced by Nakamoto (2008). He argued a need for a new payment system that was free from institutional control and did not rely on trust but required proof of work to complete transactions. This new system would be based on blockchain technology. The number of coins created would be publicly known, allowing price discovery by the law of supply and demand by matching buyers and sellers according to a price that both sides find acceptable. Since the creation of Bitcoin, over 10,000 different coins have been made, with their total market capitalisation reaching a record 2.8 trillion dollars as of November 2021, a 600% increase from the 390 billion dollar market capitalisation in November 2020, as reported by cryptocurrency tracker CoinGecko (n.d.). Despite the growth in the number of cryptocurrencies, Bitcoin remains the dominant coin. According to CoinMarketCap (n.d.), as of November 2021, Bitcoin accounted for around 41% of the total market capitalisation. In contrast, the second dominant coin, Ethereum, accounted for just below half the market capitalisation of Bitcoin, accounting for approximately 19% of the total market capitalisation, leaving the rest of the coins combined accounting for only 39%.

As evident by the high increase in market capitalisation, cryptocurrency has become one of the go-to investments for both individual and corporate investors. With this growth also came a growing interest in modelling cryptocurrency. One of the well-known properties of cryptocurrency is its high volatility. This volatility is primarily due to the speculative nature of cryptocurrency. Studies such as those by Catania and Grassi (2021), Zhang et al. (2018) and Hu et al. (2019) give evidence of the highly volatile nature of cryptocurrency. Due to its volatile nature, it is a risky investment. Therefore, there is a need to be able to develop models that can help evaluate the risk. A typical financial measure of risk is the Value at Risk (VaR). This measure has been used in past studies that we will discuss here.

Kasse et al. (2021) analysed the VaR using the Asymmetric Power Autoregressive Conditional Heteroscedasticity (APARCH). The study found the ARIMA (6,1,1) and APARCH (1,1) as the best models to capture the value at risk. The models used in this study have limitations because they do not account for some of the stylised facts, such as volatility clustering and leverage effects. Other studies used the GARCH type models, which are well known to capture most of the sytlised facts in most financial data. For example, Chu et al. (2017) used different GARCH type models to study 7 cryptocurrencies, and the results showed that the best models to measure the VaR were the IGARCH and GJR-GARCH models. Using similar models, Ngunyi et al. (2019) studied eight cryptocurrencies: Bitcoin, Ethereum, Ripple, Stellar, Litecoin, Monero, Dash, and NEM. They made use of asymmetric and symmetric GARCH models under different error distributions. The heavy-tailed error distributions made the best VaR forecast, particularly the skewed t-Student and the skewed-GED. This agrees with the well-known fact that financial data has heavy tails and, as such, cryptocurrency is not spared.

Due to having heavy tails in financial data, some researchers have used the Extreme Value Theorem (EVT) to model stock returns, which has been shown to improve the VaR estimates. For example, in a study using 12 world stock indices, Omar et al. (2020) used the EVT theorem to model the market risk of 12 world stock indices. They used the Peaks Over Threshold model to fit the Generalised Pareto Distribution (GPD). The results showed that the use of EVT helped improve the forecasting of risk. Chinhamu et al. (2015) also use EVT to model tail-related risk measures, such as VaR and ES, for the gold market. The findings were that EVT, in particular, GPD, was more appropriate to describe the conditional excess distributions of a heteroscedastic gold log return series and provided adequate estimations for VaR and ES. Likewise, Edem and Ndengo (2021) use EVT, in particular the Generalised Extreme Value Distribution (GEVD) and the GPD, to model the Value at Risk (VaR) and the Expected Shortfalls (ES) for the Bank of Kigali stock. The findings showed that GPD was a better fit for the tail of the data, giving more accurate risk measures than other distributions. For cryptocurrency, limited studies are using the extreme value theorem. This can be explained by the fact that cryptocurrency is new and is still growing. Gkillas and Katsiampa (2018) used the extreme value theorem to model Bitcoin, Ethereum, Ripple, Bitcoin Cash, and Litecoin. The study was done by fitting a GPD to the marginal distribution of the returns of each cryptocurrency, using the peaks-over-threshold method to extract extremes. Afterwards, the risk was assessed using value at risk methods. The revelation was that Bitcoin Cash was the riskiest, while Bitcoin and Litecoin were the least risky assets. Islam and Das (2021) applied the extreme value theorem to Bitcoin returns. They concluded that the GEV model using the BM approach to the Bitcoin returns data was better than the GPD model using the POT approach. In another study, Hussain et al. (2021) used the GPD to model extreme daily returns in the Bitcoin market from 2017 to 2019. The returns were directly applied to the GPD, with the results showing that EVT provides a better way to comprehend tail returns.

Reading through the literature, the authors noticed that past studies focused on applying the standard GARCH type models combined with mainly the student t, skewed student t and the general error distribution. The authors are also aware that models provide a better fit when they capture the different stylized facts that the data may possess. The stylized facts of cryptocurrency are similar to those of stock returns, but cryptocurrency volatility is more severe and persistent than stock returns Kaseke et al. (2021). This difference motivated the current study to use the Extreme Value theorem in estimating the risk associated with cryptocurrency. The study by Hussain et al. (2021) which employed EVT, did not use any model but applied the GPD to returns directly. This direct application neglects the critical aspect of modelling, which requires models to capture stylized facts such as volatility clustering, leverage effects, and autocorrelations. The study also did not further estimate the risk associated with cryptocurrency. This study seeks to improve the current knowledge base by using the extreme value theorem to develop hybrid GARCH models that use the GDP and the GEVD. The results from the hybrid GARCH models will be used to estimate the value at risk. Three cryptocurrencies will be used; Bitcoin, Ethereum, and Dogecoin return data. Our study contributes to the literature by adding the hybrid GARCH models as alternatives to model cryptocurrency risk. It also uses updated data that may capture any new properties of cryptocurrency.

7.2 Study objectives

Examining past literature revealed that all the studies that used GARCH type models to model the VaR in cryptocurrency used the common error distributions: the student t, skewed student t and the general error distribution. A search for the application of the Extreme Value Theorem shows only one study that applied the General Pareto distribution. This study by Gkillas and Katsiampa (2018) applied the GPD directly to the returns before extracting the extremes for VaR analysis. No cases of studies of cryptocurrency using GARCH-type models combined with the Extreme Value Theorem were found. It is, therefore, the aim of this paper to fill this gap by developing a hybrid GARCH model. Two steps will be used to achieve this. First, the GARCH type models are applied to the cryptocurrency returns data. The best model is selected amongst the GARCH-type models. Secondly, the residuals are extracted from the selected model and are used to fit the EVT distributions. Introducing these models will broaden the spectrum of models that can be applied in estimating the VAR of cryptocurrencies. Being in Africa, which as per Tumblr and Team (2022), saw the cryptocurrency market grow by over 1200% by value in 2021, hence making Africa the third fastest-growing cryptocurrency economy. This knowledge is crucial to assist African Investors and the world make data-driven investment choices.

7.3 Data and methodology

Three cryptocurrencies were selected for this study. These are, namely, Bitcoin, Ethereum, and Dogecoin. The selections are because Bitcoin and Ethereum are the two most prominent cryptocurrencies, accounting for 60% of the market capitalisation as per CoinMarketCap (n.d.) data. Another reason is that the market movements usually follow the direction of Bitcoin, as in studies such as by Katsiampa et al. (2019), Moratis (2021) and Kyriazis (2019). The study includes Dogecoin because it is one of the most popular meme coins. Due to its lower price, it was one of the biggest gainers in cost and the most popular amongst entry-level cryptocurrency users. The results from these three will then be generalised to the whole market. The data was obtained as daily closing prices Pt from *Financial News and Stock quotes* (n.d.). The sample period considered is from 4 June 2017 to 27 May 2021, resulting in 1454 observations. The period is considered based on data availability as the coins were not created during the same period. Dogecoin only came into existence years after the other two. The closing price P_t was converted to daily log returns r_t, and then multiplied by 100 to make them percentage daily log returns. The calculations are as shown below:

$$r_t = \ln(P_t/P_{t-1}) \times 100. \tag{7.1}$$

The initial step before analysing the data is running diagnostic tests that check for stationarity, autocorrelations, and heteroscedasticity/ARCH effects. If the data is not stationary, transformations will be applied to make it stationary. Stationarity means that the statistical properties of the series do not change over time, making predictions easier. The presence of ARCH effects is required as it warrants us to move to the next stage, which is fitting the GARCH type models. These are ideally designed to capture the ARCH effects. Three GARCH type models are used: the Generalised Autoregressive Conditional Heteroskedasticity (GARCH) of Bollerslev (1986), the Exponential GARCH model by Nelson (1991) and the GJR-GARCH model by Glosten et al. (1993).

The GARCH was developed as an improvement of the ARCH model of Engle (1982) which required many parameters to capture the volatility of the time series. The GARCH, in turn, proved to be parsimonious, requiring, in most cases, the GARCH(1,1) to be sufficient to capture the volatility process. The EGARCH added the ability to capture the asymmetric effect and the volatility clustering, while the GJR can capture the leverage effects by capturing the positive and negative returns separately. All the GARCH type models are in two parts: the conditional mean equation $\mu_t = E(r_t|Ft-1$, which captures the data as a function of the error term and other variables; the second part is the conditional variance equation $\sigma_t^2 = Var(r_t|Ft-1 = E[(r_t-\mu)^2|Ft-1]$, which determines the conditional variance evolution of the error term from the mean equation as a function of past conditional variances and lagged errors. In most cases, the mean equation is captured by the ARMA(P,Q) model. The difference between the models lies in the way the conditional variance is captured. The conditional variance equations for each model are shown below:
For the GARCH model:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i {\varepsilon^2}_{t-i} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \tag{7.2}$$

where σ_t^2 is the volatility process, ω a constant, α_i and β_j are the ARCH and GARCH parameters respectively. We require that $\alpha_i \geq 0$, $\beta_j \geq 0$, and $0 \leq \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j \leq 1$.

For the EGARCH model

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{s} \frac{|a_{t-i}| + \gamma_i \, a_{t-i}}{\sigma_{t-1}} + \sum_{j=1}^{m} \beta_j \ln(\sigma_{t-1}^2), \tag{7.3}$$

where the error term is $\epsilon_{t-1} = a_{t-i}/\sigma_{t-1}$. A positive a_{t-i} results in $\alpha_i(1 +$

γ_i)$|\epsilon_{t-1}|$ contribution to the log volatility while a negative a_{t-i} results in α_i(1 - γ_i)$|\epsilon_{t-1}|$. γ is the leverage effect parameter of a_{t-i}.

For the GJR-GARCH

$$r_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = \omega + \alpha^2 \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma S_{t-1} \epsilon_{t-1}^2, \tag{7.4}$$

where, S_{t-1} is an indicator variable such:

$$S_{t-1} = \begin{cases} 1, & \text{if } \epsilon_t \leq 0, \\ 0, & \text{if } \epsilon_t > 0. \end{cases}$$

γ is the asymmetric parameter and is the constant while α_i and β_j are the ARCH and GARCH parameters, respectively. The indicator function S_{t-t} takes the value of one in the case of $\epsilon_t < 0$ and 0 otherwise. If the gamma is positive and statistically significant there is a negative asymmetric volatility response.

After the GARCH type models, the extreme value theorem is employed. The section that follows discusses the EVT in detail

7.4 *Extreme value theorem*

Smith (1986) defines the extreme value theorem as the branch of statistics concerned with the extremes of the distributions. It seeks to model the tail distribution of data. EVT has seen an increase in use in the financial sector because extreme events in finance can be catastrophic for investors. Two main methods exist to model the extreme value theorem: the Block maxima method and the Peaks over the threshold method. The block maxima approach is implemented by dividing the data into blocks of equal length, such as months, weeks, or years. From these blocks, the maximum values of each block are analysed. For the POT method, a threshold is selected from the data, and all the data points above the threshold are considered for analysis. The distributions used for this study employ these approaches separately. The GEVD uses the Block maxima method, while the GPD uses the POT approach. We summarise the distributions and their fittings below.

7.4.1 *Generalised extreme value distribution*

The Generalised extreme value distribution was proposed by Jenkinson (1955). It sought to take into account all the different types of extreme value distribution of Gnedenko (1943) (1943), which are

- The Gumbel family known as Type 1 where $\xi = 0$, hence a CDF become

$$F(x) = \exp[-\exp(-x)], for \infty < x < \infty. \tag{7.5}$$

- The Fr'echet family known as Type 2 where $\xi > 0$, hence the CDF is

$$F(x) = \begin{cases} exp[-(1+\xi x)]^{-\frac{1}{\xi}} & \text{if } x > -\frac{1}{\xi}, \\ 0 & \text{if } otherwise. \end{cases} \tag{7.6}$$

- The Weibull family known as Type 3 where $\xi < 0$, hence CDF is

$$F(x) = \begin{cases} exp[-(1+\xi x)]^{-\frac{1}{\xi}} & \text{if } x < -\frac{1}{\xi}, \\ 1 & \text{if } otherwise. \end{cases} \tag{7.7}$$

The tail distribution determines the maximum limiting distribution F (x). The Gumbel family has a right tail that declines exponentially, while for the Fr'echet family, it declines by a power function, and it is finite for the Weibull.

7.4.1.1 Block maxima

The block maxima method is used to estimate the GEVD parameters. The method works by blocking the data into equal but none overlapping blocks and then applying the EVT to each block to obtain a block maxima M_{ni}. In general the block distributions are not known but for large enough n, the distribution of the normalised block maximas $y_{n,i} = (M_{n,i} - \mu)/\sigma$ converge asymptotically to a generalised extreme value distribution.
The collection of these normalised maximas gives a new data set from which the unknown parameters, are estimated. The estimation can be done using various methods, including the maximum likelihood.

7.4.2 Generalized pareto distribution

The GPD makes use of the POT method. It is defined by two parameters which are the shape parameter ξ and the scale parameter β. Using the notation of Tsay (2005), the GDP is defined as follows:

$$F_{\xi,\beta}(x) = \begin{cases} 1 - [1 + \frac{\xi x}{\beta}]^{-1/\xi} & \text{if } \xi \neq 0, \\ 1 - \exp[-x/\beta] & \text{if } \xi = 0. \end{cases} \tag{7.8}$$

where $\beta > 0$, $x \geq 0$ when $\xi \geq 0$, and $0 \leq x \leq -\beta/\xi$ when $\xi < 0$.

7.4.2.1 Peaks-over-threshold

The GPD employs the POT method, which considers values that exceed a certain threshold. Given the threshold μ , the excess is given by those returns r_t that are greater than, μ i.e. all $r_t > \mu$. The exceedance, hence, is given by $x = r_t - \mu$. These data points are, in turn, modelled separately from those that do not exceed the threshold. The excess distribution function is then defined as follows:

$$F_\mu(x-\mu) = \frac{F_x - F_\mu}{1 - F_\mu}. \tag{7.9}$$

By Coles et al. (2001), shows that $F_\mu(x-\mu)$ is approximately a generalized Pareto family i.e.

$$F(x) = 1 - \frac{N_\mu}{n}(1 + \frac{\hat{\xi}}{\hat{\beta}(x-\mu)}). \tag{7.10}$$

7.4.2.2 Threshold choice

The choice of the threshold value is critical as it significantly impacts the analysis. The threshold value has to be chosen so that it is not too high and also not too low. This is because if it is low, then one is likely to get biased results; on the other hand, if it is too high, then it means only a few data points will be available for estimating the shape and scale parameters, leading to high variance estimates. This is emphasised by various authors, such as Chinhamu and Chifurira (2019) and Ilupeju (2016). In this paper, the Laeken package, which calculates the threshold using the van Kerms rule of thumb of Van Kerm (2007) is used. Apart from that the mean excess plot is also used for graphically identifying the threshold level.

7.5 Value-at-Risk (VaR)

Investors must be able to quantify the risks that they are likely to face in their investments. This allows them to take positions that they can bear and have capital reserves for continuation in the worst market conditions. One measure of the risk is the value at risk. It is a measure that measures the amount one is likely to lose. For example, a 5% Value at Risk is the amount one can expect to lose in the best 5% of the worst-case scenarios. This value is the 5% quantile of the distribution. Correct estimation of VaR is essential because wrong calculations can have severe implications for the financial status of a company or a financial institution. This is because VaR does not consider the size of losses on the remaining days (i.e. the other 5% for 95% VaR). The loss

may be high enough to lead to liquidation on such days. Overestimation, on the other hand, may result in more capital requirements than necessary.
VaR is defined as the p^{th} quantile of a given distribution F. This means that VaR_p is the p^{th} quantile of the distribution F. This is defined as $VaR_p = F^{-1}(1-p)$, where F^{-1} is the inverse of F known as the quantile function and $0 < p < 1$.

Using GEVD and GPD we can also model and approximate the VaR. Using the notation by Tsay (2005), the GPD the VaR for a small upper tail probability p, is estimated as follows:

$$VaR_q = \gamma - \frac{\psi(\gamma)}{\xi}\{1 - [\frac{T}{N_\gamma}(1-q)]^{-\xi}\}, \tag{7.11}$$

where, γ is the threshold parameter, T is the sample size, N_γ is the number of exceedances, and $\psi(\gamma)$ and ξ are the scale and shape parameters of the GPD distribution.
For the GEVD,

$$VaR_q = \begin{cases} \mu - \frac{\sigma}{\xi}\{1 - [-n\ln(1-p)]^{-\xi}\} & \text{if } \xi \neq 0, \\ \mu - \sigma[-n\ln(1-p)] & \text{if } \xi = 0. \end{cases} \tag{7.12}$$

where n is the subperiod length and μ, σ and ξ are the maximum likelihood estimates of the GEVD.

7.6 Empirical analysis

In order to have an understanding of the data, the return series of the three cryptocurrencies are plotted against time. The resulting plots are shown in Figure 7.1 below. It can be observed that the data fluctuates around the value zero, suggesting some form of stationarity in the mean. For most periods, the returns of Bitcoin and Ethereum are bound by -20% and 20%, with a few occasions having spikes above. On the other hand, the Dogecoin returns seem to be spikier, with the bounds mainly being around -50% to 50%. This means that Dogecoin is a riskier asset compared to the other two. All three series' spikes suggest that the log-returns are not variance stationary. Also, periods of high volatility are clustered together, and those of low volatility are also clustered together. This indicates the presence of volatility clustering, which is a stylised fact of financial time series. Dogecoin had one day that had an abnormal increase of over 6000% on the 2nd of January 2020 and reverted to the average price range on the next day. This was over 150 times the average price. To avoid bias, it was treated as an outlier. The observation was removed and replaced with the average of the previous day and the next day.

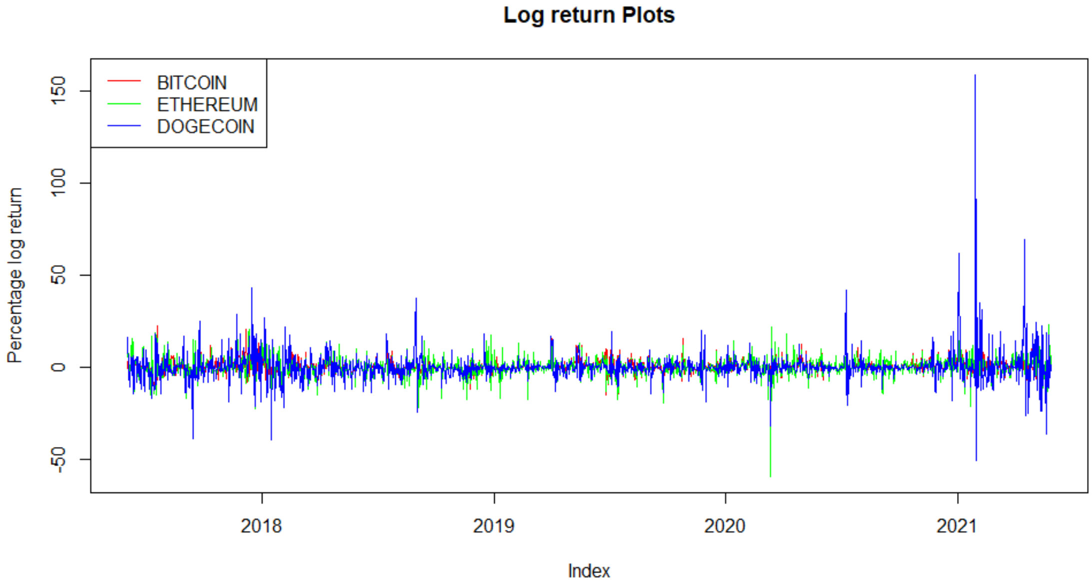

Figure 7.1: Log return plots from 4 June 2017 to 27 May 2021

Using graphical analysis, the data seems to be stationary. However, we also perform formal statistical tests. For this purpose, three different tests were used; these were the Augmented Dickey-Fuller test by Dickey and Fuller (1981), the Phillips-Perron (PP) test by Phillips and Perron (1988), and the Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) test. The tests are presented in Table 7.1. The ADF test and the PP test have a null hypothesis that the series is not stationary. All three cryptocurrencies gave p-values that were less than 0.05; therefore the null hypothesis was rejected for the alternative of stationarity. For the KPSS test, the null hypothesis is that the series has stationarity. The null hypothesis was not rejected based on p-values that were not significant. This means that all the tests agreed that the series were all stationary.

Table 7.1: Tests for stationarity

Cryptocurrency	ADF test	PP test	KPSS test
Bitcoin	-10.163***	-1667.7***	0.12765
Ethereum	-10.713***	-1719.2***	0.33921
Dogecoin	-10.501***	-1525.1***	0.42726*

Note: *,**,*** is statistical significance at the 0.05, 0.01 and 0.001 critical level respectively.

The data was further interrogated using numerical tests and measurements. Table 7.2 below shows the summarised descriptive statistics and tests for normality tests. The minimum percentage daily return for all data series shows that the biggest losses were greater than 40%, which is -49% for Bitcoin,-58.9% for Ethereum, and -50.2% for Dogecoin. The maximum daily returns are 22.8% for Bitcoin, 23.1% for Ethereum, and 158.4% for Dogecoin. These values are quite high for standard stable investment returns, confirming the presence of extreme events on cryptocurrency returns. The average daily returns are 0.19%, 0.18%, and 0.32% for Bitcoin, Ethereum, and Dogecoin, respectively. The standard deviations are approximately 4.4% and 5.6% for Bitcoin and Ethereum, respectively. On the other hand, the standard deviation for Dogecoin is 8.6%, which is almost twice that of the other two. This suggests that Dogecoin is more volatile as compared to the other cryptocurrencies. Bitcoin and Ethereum are negatively skewed, while Dogecoin is positively skewed. This points to more extreme events on the negative side for Bitcoin and Ethereum but more extreme positive events for Dogecoin. For all three data series, the excess kurtosis is positive, with Dogecoin almost eight times that of others. These results coincide with the general results of financial time series data, which is known to be skewed, have high kurtosis, and have volatility clustering.

Table 7.2: Descriptive statistics

Cryptocurrency	Bitcoin	Ethereum	Dogecoin
Minimum	-49.7278	-58.9639	-50.2321
Maximum	22.7602	23.0772	158.3848
Mean	0.1886	0.1767	0.3213
Median	0.1633	0.1399	-0.0612
Standard deviation	4.3932	5.6645	8.6123
Skewness	-0.9050	-0.9565	5.0577
Kurtosis	13.5524	10.3800	85.0086
Jarque Bera	11364***	6772.8***	445265***
Ljung-Box	29.701***	51.263***	27.552*****

Note: *,**,*** is statistical significance at the 0.05, 0.01 and 0.001 critical level respectively. Standard errors are in parenthesis

The normality assumption is rejected using the Jarque Bera Test, as evident by the significant p-values shown in Table 2. A visual test was also done by using the Q-Q plot shown in Figure 7.2 below. The Q-Q plots show that the tails deviate from the standard normal line for all three data series. This suggests that the tails are heavier or thicker than the tails of a standard normal distribution. Apart from the descriptive statistics, we perform the test for serial autocorrelation, also referred to as a test for ARCH effects. To do this, we use the Ljung-Box test of Ljung and Box (1978). From Table 7.2, the tests were all significant, indicating the presence of ARCH effects. This shows the heteroscedasticity nature of the returns, which justifies using GARCH type models. The Ljung-Box test on the squared returns was all significant, indicating the presence of ARCH effects. Evidence of serial correlation in these return series contrasts with the assertion of informational efficiency in the EMH of Fama (1970).

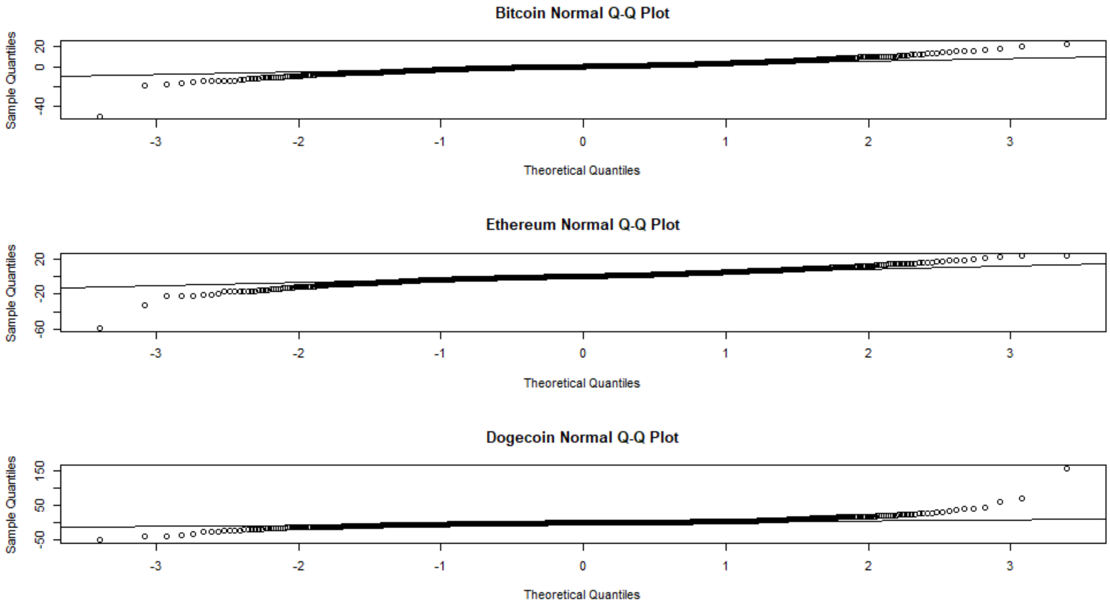

Figure 7.2: QQ plots

7.6.1 GARCH type models

In this section three GARCH type models were fit, which are the GARCH, EGARCH and GJR-GARCH to each of the cryptocurrencies. These models were fit with the normal distribution governing the innovations. The maximum likelihood calculated estimates are presented in Tables 7.3, 7.4 and 7.5 below.

Table 7.3: ML parameter estimates of the GARCH-type models with normal innovations for Bitcoin

Parameter estimate	GARCH (1,1)	EGARCH (1,1)	GJR-GARCH (1,1)
$\hat{\mu}$	0.2499 (0.1018)*	0.1263 (0.1050)*	0.1705 (0.1023)*
$\hat{\omega}$	1.1794 (0.2436)****	0.2323 (0.0722)****	1.3537 (0.2993)****
$\hat{\alpha}$	0.1010 (0.0196)****	-0.0588 (0.0168)****	0.0665 (0.0162)****
$\hat{\beta}$	0.8399 (0.0237)****	0.9271 (0.0239)****	0.8212 (0.0275)****
$\hat{\gamma}$		0.184000 (0.036799)***	0.109465 (0.0313)****
AIC	5.6982	5.6939	5.6863
BIC	5.7127	5.7121	5.7045

Note: *,**,*** is statistical significance at the 0.05, 0.01 and 0.001 critical level respectively. Standard errors are in parenthesis

From Table 7.3, the mean parameter is significant at 5% for all three cryptocurrencies while the other parameters are also significant at 1%. Based on Both the AIC and BIC values, the GJR-GARCH is the best fit for the Bitcoin return data.

Table 7.4: ML parameter estimates of the GARCH-type models with normal innovations for Ethereum

Parameter estimate	GARCH (1,1)	EGARCH (1,1)	GJR-GARCH (1,1)
$\hat{\mu}$	0.1817 (0.1322)	0.1476 (0.1346)	0.1390 (0.1340)
$\hat{\omega}$	1.7722 (0.4507) ***	0.2147 (0.0226) ***	2.0081 (0.5191) ***
$\hat{\alpha}$	0.1047 (0.0176) ***	-0.0325 (0.0144) ***	0.0840 (0.0186) ***
$\hat{\beta}$	0.8478 (0.0247) ***	0.9422 (0.0070) ***	0.8372 (0.0271) ***
$\hat{\gamma}$		0.1923 (0.0231) ***	0.0467 (0.0263) *
AIC	6.2040	6.2040	6.2028
BIC	6.2185	6.2222	6.2210

Note: *,**,*** is statistical significance at the 0.05, 0.01 and 0.001 critical level respectively. Standard errors are in parenthesis

Table 7.4 shows the results for the maximum likelihood parameter estimates for GARCH-type model with normal innovations fitted to the Ethereum returns. It can be seen that the the mean parameter is not significant under all three models. The other paramors are all significant at 1% level. The GJR-GARCH is the best model as seen by the lower AIC and BIC values.

Table 7.5: ML parameter estimates of the GARCH-type models with normal innovations for Dogecoin

Parameter estimate	GARCH (1,1)	EGARCH (1,1)	GJR-GARCH (1,1)
μ	0.0543 (0.1373)	0.4127 (0.1140) ***	0.1364 (0.1320)
$\hat{\omega}$	0.5594 (0.1018) ***	0.0606 (0.0032) ***	0.4535 (0.0762) ***
$\hat{\alpha}$	0.0749 (0.0064) ***	0.1117 (0.0080) ***	0.11707 (0.0103) ***
$\hat{\beta}$	0.9241 (0.0057) ***	0.9905 (0.000006) ***	0.9355 (0.0047) ***
$\hat{\gamma}$		0.0920 (0.0039) ***	-0.1072 (0.0111) ***
AIC	6.6736	6.6815	6.6163
BIC	6.6881	6.6997	6.6345

Note: *,**,*** is statistical significance at the 0.05, 0.01 and 0.001 critical level respectively. Standard errors are in parenthesis

Table 7.5 shows the maximum likelihood parameter estimates for GARCH-type models with normal innovations fitted to the Dogecoin returns. Only the EGARCH model had a significant mean parameter. The other parameters were significant at 1% for all models. As it was for Bitcoin and Ethereum, the GJR-GARCH was the best fitting model. Overall, of all three, the GJR-GARCH was the best fit. This suggests that the leverage effect parameter plays a crucial role in modelling cryptocurrency, as also seen by the significance. The next step is to examine the model adequacy for the GJR-GARCH for all three cryptocurrencies. First, we consider the normality assumption using the Q-Q plot of the standardised residuals shown in Figure 7.3 below. A significant deviation from the normal line is observed on the ends, suggesting heavy tails on the data. Statistical tests are performed to confirm the departure from normality further using the JarqueBera and Shapiro-Wilk tests. Both had p-values < 0.01 under all three cryptocurrencies. Further, the excess kurtosis for the residuals was 12.773842, 6.816993, and 34.462656 for Bitcoin, Ethereum, and Dogecoin, respectively. Armed with the statistical tests and visual evidence, we conclude that cryptocurrencies require heavy tail distributions to capture their heavy tails adequately. To deal with this, the GJR-GARCH is modified to create a hybrid model, as discussed in the following section.

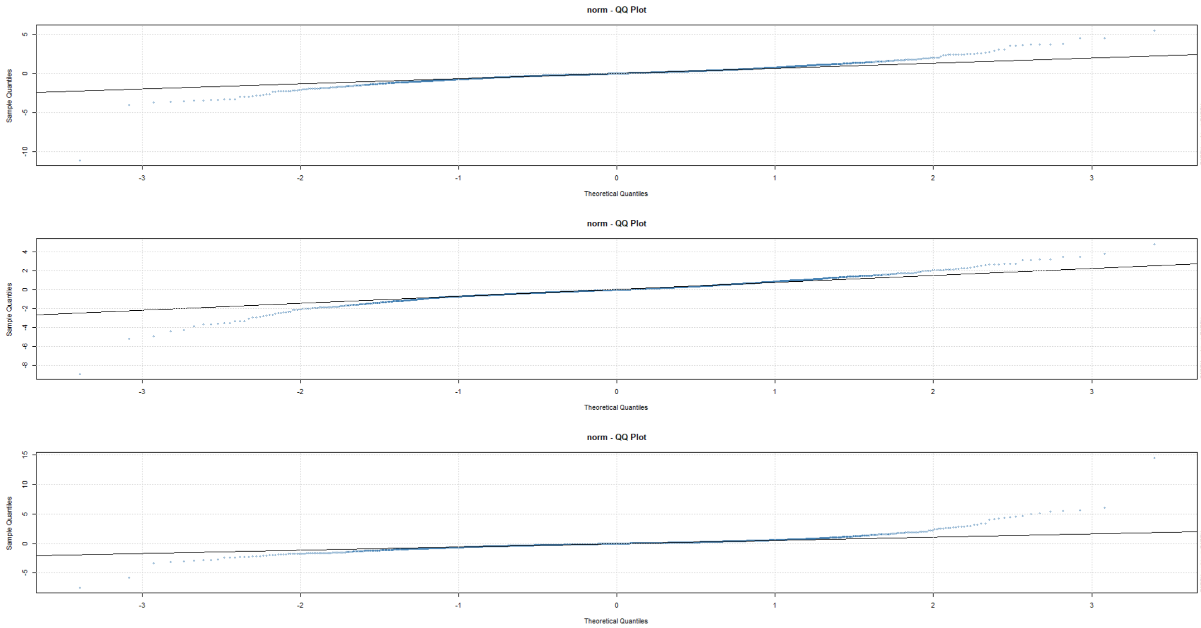

Figure 7.3: Normal Q-Q plot of standardized residuals of Bitcoin, Ethereum and Dogecoin GJR-GARCH model

7.7 Hybrid models

To cater for the heavy tails, we developed a hybrid GJR-GARCH model. This we achieve by combining the GJR-GARCH model from the previous section with heavy tail distributions. We make use of the GPD and the GEVD. To do this, we extract the standardised residuals from the GJR-GARCH and fit them to the GPD and the GEVD distributions. Before fitting our distribution to the standardised residuals, we check if the data is independent and identically distributed (i.i.d), as it is a standard requirement before fitting data to distributions. We use the Bartels rank test for i.i.d. random variables, which has the null hypothesis that the series consists of i.i.d. random variables. The test returned p-values greater than 0.1 for all the three cryptocurrency residuals. Therefore, all our series are i.i.d. We also ran the BDS test for i.i.d random variables, and the test confirms that the series are i.i.d. We also check for heteroscedasticity in the standardised residuals using the Ljung box test. The results showed that there was no heteroscedasticity.

7.7.1 Fitting the generalised pareto distribution (GPD)

We fit the GDP separately for the upper tail (gains) and the lower tail (losses). To obtain the estimates for the lower tail, we need to obtain a separate threshold. We achieve this by multiplying the standardised residuals e_t by -1 giving $-e_t$.

We then calculate the threshold values for the upper tail and lower tail based on e_t and $-e_t$. To obtain the threshold, we use the paretoScale function of the Laeken package in R. This function estimates the scale parameter for the pareto distribution from the given data. The mean excess plot was also used to give a visual identification of the threshold.

7.7.2 Bitcoin

Using the paretoScale function, we obtain a threshold value of 1.811689 for the positive residuals and a threshold of 1.900744 for the negative innovation. The mean excess plot shown in Figure 7.4 has the dotted confidence bounds covering the mean excess as required. The threshold is estimated to be around 1.5 to 2 in both plots. This confirms the value we obtained using the paretoScale function.

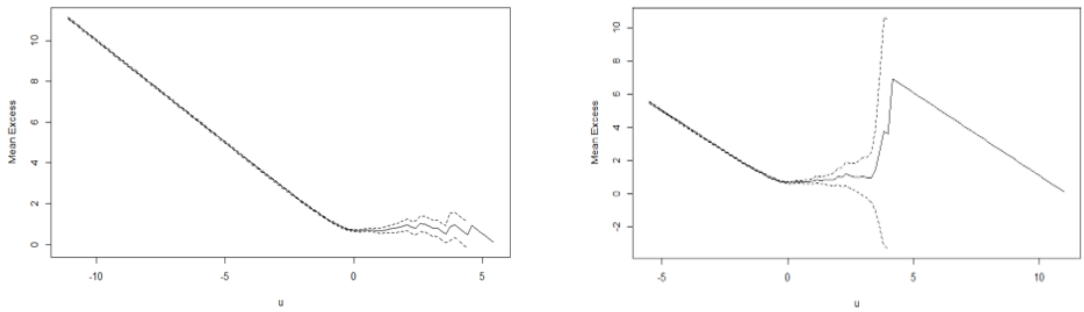

Figure 7.4: Bitcoin Mean excess plot for negative and positive residuals

We fitted the GPD with the thresholds obtained above, and the obtained maximum likelihood estimates are shown in Table 7.6 below. The corresponding standard errors are given in brackets. For the upper tail, the shape parameter $\xi = -0.058$ and the scale parameter ν=0.929, while for the lower tail, ξ=0.238 and ν =0.6951.

Table 7.6: Bitcoin GDP model estimates

f	EVT threshold $\hat{\tau}$	Observations in tail	tail	shape parameter $\hat{\xi}$	scale parameter $\hat{\nu}$
Zt	1.8117	43		-0.0580(0.181)	0.9287 (0.2186)
- Zt	1.9007	43		0.2379(0.1699)	0.6951(0.1567)

Note: Standard errors are in parenthesis

Figure 7.5 shows the model diagnostic plots for the GDP model for the upper tail (Left pane) and the lower tail (Right pane). The probability plots and the QQ plots for both losses and gains do not diverge much from the line, suggesting that our standardised residuals follow the GDP model for both the lower and upper tail. The return level plots also suggest estimating the VaR using the GDP. We, therefore, conclude that the GDP is an improvement to the model and can be used to model Bitcoin.

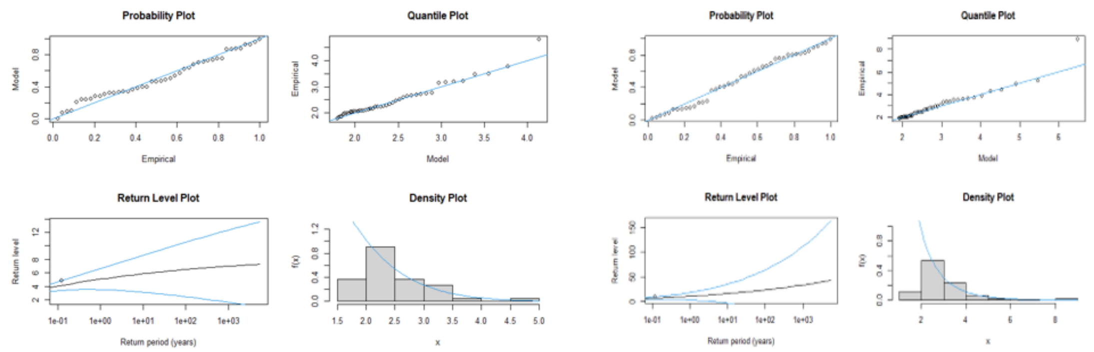

Figure 7.5: Bitcoin GDP model diagnostic plots

7.7.3 Ethereum

The threshold values for Ethereum were found to be 1.777042 for the positive standard residuals and 1.886594 for negative standard residuals. These values are confirmed by the mean excess plots shown in Figure 7.6.

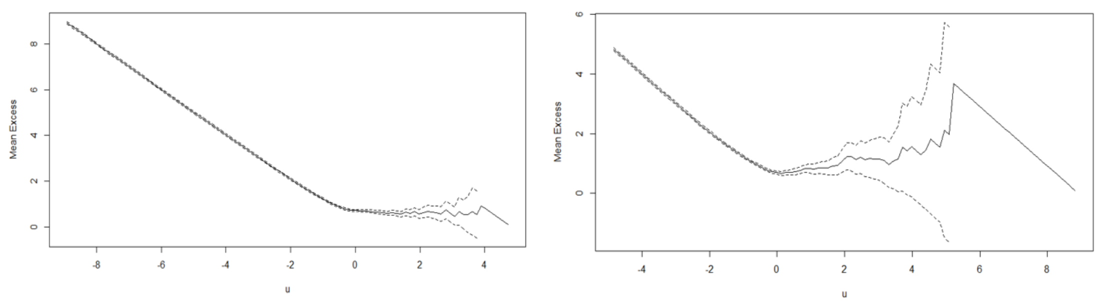

Figure 7.6: Ethereum Mean excess plot for negative and positive residuals

The maximum likelihood estimates and their standard errors are reported in Table 7.7. For the upper tail, ξ= -0.109 and ν =0.757 while for the lower tail, ξ=0.1314 and ν =0.937.

Table 7.7: Ethereum GDP model estimates

f	EVT threshold $\hat{\tau}$	Observations in tail	tail	shape parameter $\hat{\xi}$	scale parameter $\hat{\nu}$
Zt	1.7770	44		-0.1094(0.1320)	0.7572(0.1512)
- Zt	1.8866	43		0.1314(0.1749)	0.9373(0.2167)

Note: Standard errors are in parenthesis

Figure 7.7 shows the model diagnostic plots for the GDP model for the upper tail (left pane) and the lower tail (right pane). Similar to the Bitcoin standardised residuals, the Ethereum standardised residuals show that they fit the GDP model well, as evident from the QQ plots and the probability plots. The return level plots also suggest estimating the VaR using the GDP.

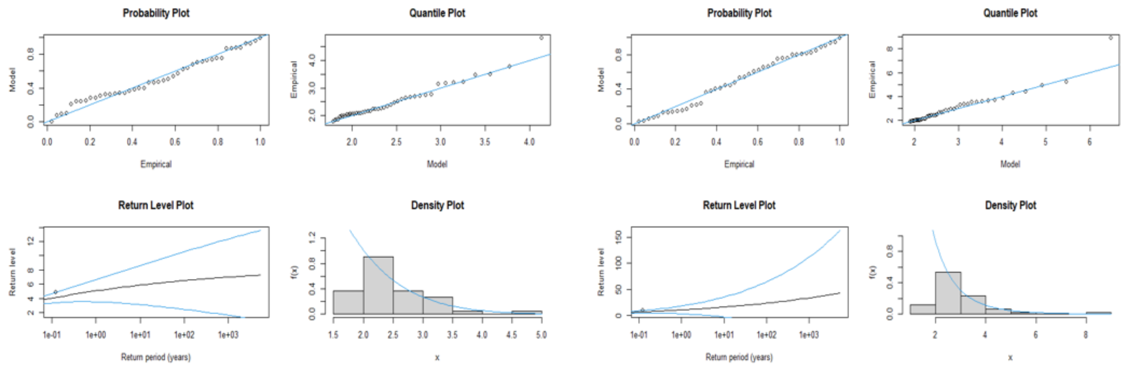

Figure 7.7: Bitcoin GDP model diagnostic plots

7.7.4 Dogecoin

The threshold values for Dogecoin were found to be 1.905238 for the positive standard residuals and 1.644386 for negative standard residuals. These values are confirmed by the mean excess plots shown in Figure 7.8.

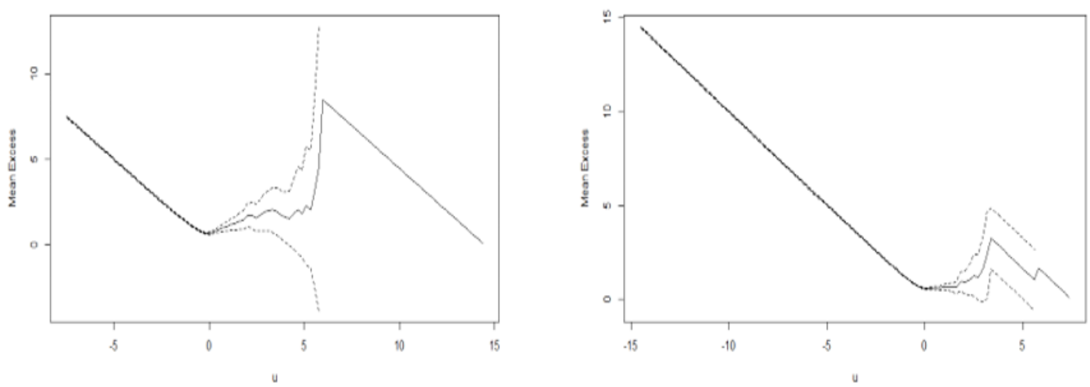

Figure 7.8: Dogecoin Mean excess plot for negative and positive residuals

The maximum likelihood estimates and their standard errors are reported in Table 7.8. For the upper tail, ξ= 0.209 and ν =1.236 while for the lower tail, ξ=0.478 and ν =0.3697.

Table 7.8: Ethereum GDP model estimates

f	EVT threshold $\hat{\tau}$	Observations in tail	tail	shape parameter $\hat{\xi}$	scale parameter $\hat{\upsilon}$
Zt	1.9052	43		0.2088(0.1810)	1.2362(0.2886)
- Zt	1.6444	43		0.4781(0.2392)	0.3697(0.1013)

Note: Standard errors are in parenthesis

Based on the diagnostic plots shown in Figure 7.9, which show both positive and negative residuals not deviating from the QQ and probability plot lines significantly, we conclude that the GDP is a good fit for estimating the Dogecoin. The return plots also justify the use of the GDP for VaR estimations.

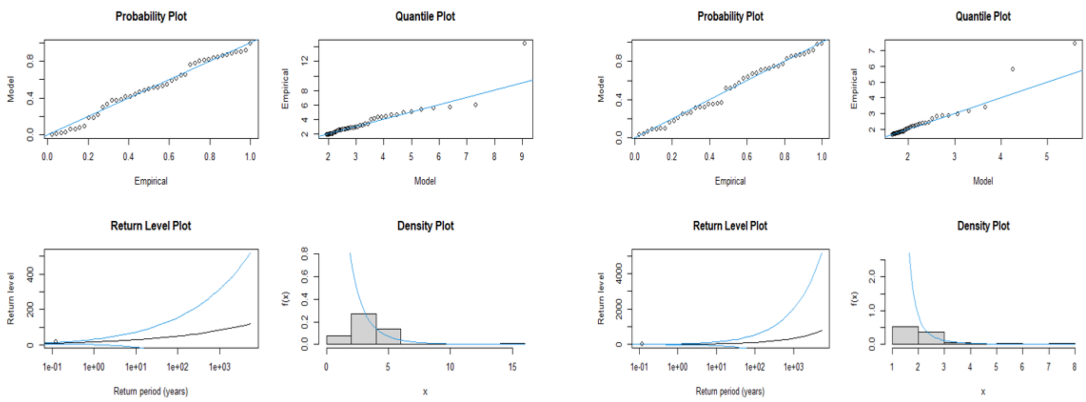

Figure 7.9: Dogecoin GDP model diagnostic plots

7.7.5 Fitting the generalized extreme value distribution (GEV)

The GEV is applied to both the positive and negative standardized residuals of the GJR-GARCH. The block size used was 7. The model's maximum likelihood parameter estimates and standard errors (in brackets) are shown in Table 7.9 below.

Table 7.9: ML Parameter estimates for the GEV

	$\hat{\xi}$	$\hat{\sigma}$	$\hat{\mu}$	AD test
Bitcoin				
Zt	0.1069	0.6184	0.83599	0.3415
	(0.0574)	(0.0373)	(0.0490)	(0.9039)
- Zt	0.2015	0.5672	0.8226	0.1716
	(0.0564)	(0.0355)	(0.0446)	(0.9963)
Ethereum				
Zt	0.0041	0.5981	0.9121	0.3732
	(0.0490	(0.0337)	(0.0465)	(0.8747)
- Zt	0.2132	0.5778	0.8493	0.6573
	(0.0540	(0.0359)	(0.0450)	(0.5954)
Dogecoin				
Zt	0.4338	0.5135	0.6141	0.2009
	(0.069978	(0.0377)	(0.0377)	(0.9902)
- Zt	0.2226	0.4542074	0.6736	0.36564
	(0.0620	(0.0293)	(0.03623)	(0.8819)

Note: Standard errors are in parenthesis

The last column in Table 7.8 shows the results of the Anderson-Darling Goodness of fit test for each model. The p-value for each asset is > 0.05 hence, the null hypothesis is retained. This means that the GEVD is a good fit for all the data. The diagnostic plots are shown in Figures 7.10, 7.11, and 7.12 for Bitcoin, Ethereum, and Dogecoin, respectively.

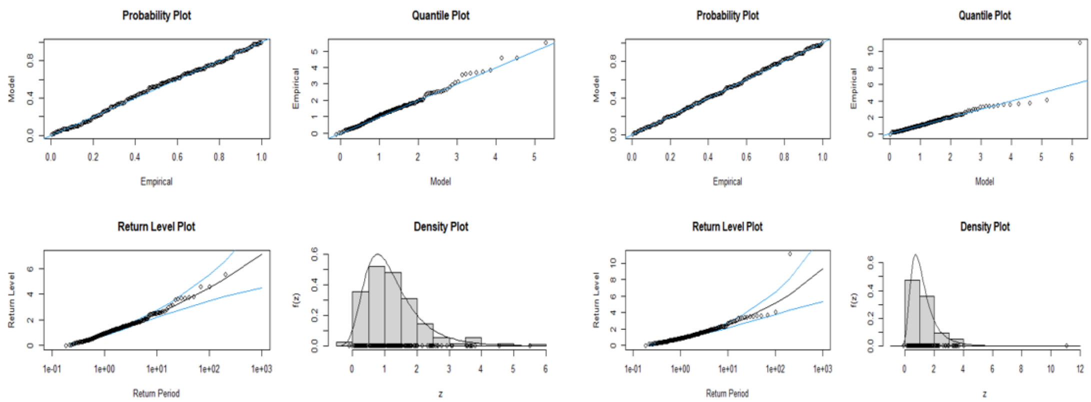

Figure 7.10: Bitcoin GEVD model diagnostic plots

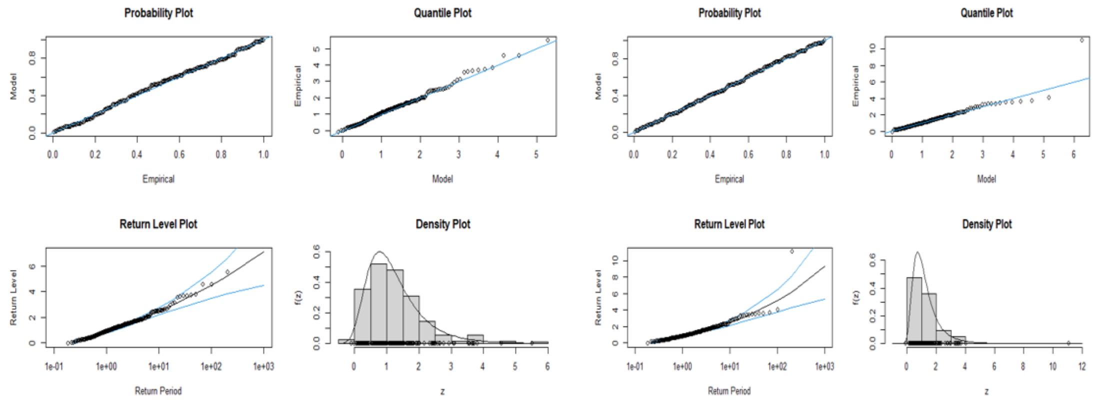

Figure 7.11: Ethereum GEVD model diagnostic plots

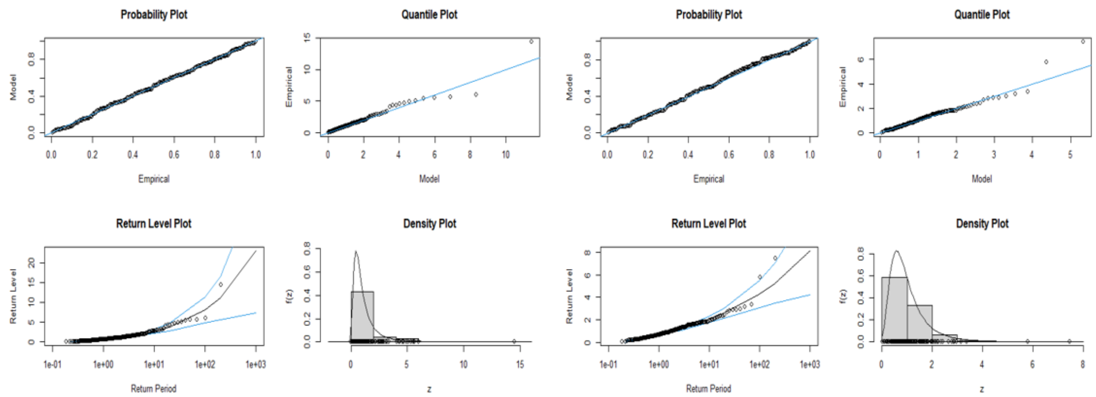

Figure 7.12: Dogecoin GEVD model diagnostic plots

7.8 *Estimating Value at Risk (VaR)*

In the previous section, the results showed that both the GPD and the GEVD are adequate in modelling the three cryptocurrencies. In this section, the two distributions will be compared in terms of their ability to capture the Value at Risk (VaR).

Table 7.10 shows the quantiles for the gains and losses of the positive and negative returns, respectively. The gains and losses are represented at 90%, 95%, and 99% for each asset for both GDP and the GEVD. For all assets, the largest gains and losses are recorded using the GEVD model at all quantiles.

Table 7.10: Value at Risk Estimates

Distribution	VaR estimates Positive returns			VaR estimates Negative returns		
	90%	95%	99%	90%	95%	99%
BITCOIN						
GPD	0.66222	1.33849	2.80765	1.16559	1.55760	2.76058
GEVD	2.40906	2.99750	4.50972	2.43741	3.12878	5.11956
ETHEREUM						
GPD	0.81015	1.38618	2.56667	0.8314375	1.41020	2.97882
GEVD	2.26440	2.69960	3.68987	2.51774	3.24390	5.36472
DOGECOIN						
GPD	0.59752	1.31589	3.44530	1.30305	1.47272	2.16964
GEVD	2.57243	3.72405	8.13820	2.00044	2.58561	4.31429

The VaR was back-tested using the Kupiec test. The p-values of the in-sample Kupiec test are summarised in table 7.11. The p-values for the 99% VaR using the GDP are greater than 0.05 for all assets. As a result, the null hypothesis cannot be rejected, whereas, for the GEVD, all but the 95% value for Dogecoin have p-values less than 0.05. The results tell us that the GDP model is the best model to capture the VaR for the three cryptocurrencies.

Table 7.11: Backtesting of VaR results

Distribution	p-values for Kupiec test for positive returns			p-values for Kupiec test for negative returns		
	90%	95%	99%	90%	95%	99%
BITCOIN						
GPD	0.001	0.00264	0.67937	0.00794	0.83732	0.26162
GEVD	0.001	0.001	0.001	0.001	0.001	0.001
ETHEREUM						
GPD	0.001	0.001	0.90399	0.001	0.04726	0.90399
GEVD	0.001	0.001	0.001	0.001	0.001	0.001
DOGECOIN						
GPD	0.001	0.11937	0.88612	0.001	0.23276	0.80125
GEVD	0.001	0.001	0.32959	0.001	0.001	0.001

7.9 Conclusion

The focus of this chapter was to examine the suitability of GARCH type models combined with heavy tail distribution in the modelling of cryptocurrency VaR. Three cryptocurrencies were used: Bitcoin, Ethereum, and Dogecoin. The GARCH type models were initially run, and they managed to capture most of the stylised facts. The best model was the GJR-GARCH. This result agrees with the results of Chu et al. (2017), and Ngunyi et al. (2019). However, the GJR-GARCH assumed the normal error distribution and could not capture the heavy tails. We then developed hybrid GARCH models by including the heavy tail distributions to solve this problem. The GPD and the GEVD distributions

were applied to the standardised residuals of the GJR-GARCH model. The results showed that both distributions improved the fit. Using results from the new hybrid models, the VaR was estimated under the GDP and the GEVD. After backtesting the results of the VaR estimates obtained under all three cryptocurrencies, the GPD outperformed the GEVD in capturing the VaR. The conclusion is that the GJR-GARCH with GDP innovations can be used as an alternative model to calculate the VaR.
The finding is helpful for investors as they can accurately estimate their possible risk and allocate capital accordingly. In further research, comparisons can be made with the results of other heavy-tailed distributions.

7.10 Chapter summary

This chapter introduced a new hybrid GARCH model to model cryptocurrency. The hybrid GARCH model was created by combining the traditional GARCH-type models with the Extreme Value theorem. The results showed that the new hybrid model adequately quantifies the VaR.

8. A COMPARATIVE ANALYSIS OF THE NATURE OF VOLATILITY IN CRYPTOCURRENCY AND THE JSE MARKET

This chapter compares the volatility of cryptocurrency and that of the JSE market. It considers the effect of structural breaks in the data and how they may affect the estimates. This is important as studies have shown that the presence of structural breaks bring about a bias in the estimates. This implies that making decisions for data with structural breaks using models that do not account for the structural breaks may result in losses for the investors. We therefore test for the presence of structural breaks in the data and then we employ GARCH models that do not account for the structural breaks and compare them with GARCH models that account for structural breaks.

8.1 Introduction

Of all the stylised facts that come with financial data, volatility, which is the fluctuation of asset return prices, is the most crucial. This is because volatility determines the riskiness of an asset. For investors, assets with high volatility signify more risk and, therefore, the likelihood of losing their investment. However, a higher probability of higher returns is also associated with high volatility. Therefore, volatility can be good or bad for the investor, depending on whether the investor is risk-averse or risk-neutral. For institutions that deal with client investments, assets with high volatility are a no-go area as they pose a risk of losing shareholder money. Likewise, the average risk-averse investor tends to avoid volatile assets. Instead, they seek investments in more stable markets with slow growth over time. A lack of proper understanding of volatility and its risk may lead to liquidation. Such extreme cases give rise to the need to understand volatility properly.

Stock markets tend to be stable over time, with periods of high returns followed by other periods of high returns and equally low returns followed by periods of low returns. These stretches of similar magnitude give rise to what is termed "volatility clustering." Investors usually take advantage of this clustering to put up their trading strategies. For example, an entry into low volatility can be taken as a signal for the coming low volatility season; hence, traders are prepared for it. However, once in a while, the market goes off pattern, causing devastating effects on investors' investments. A recent example is the 2008 global crisis. Before the global crisis, the markets were relatively stable, with returns on investments increasing ever so slightly. However, once the crisis struck in 2008, a massive shift in the return on investment occurred. Huge losses and, in some cases, even liquidation were ensured. This downside of volatility explains why it is the most studied phenomenon in stock markets. During the global crisis, the first idea of cryptocurrency was published by Nakamoto (2008). Nakamoto (2008) introduced Bitcoin, a new form of currency that would be free from institutional control and based on blockchain computer technology. Over a decade later, Bitcoin and other cryptocurrencies created after it have become one of the most talked-about investment areas. While we do not go through the details of the creation of Bitcoin in this chapter, comparisons will be made on cryptocurrency returns with stock market returns.

Even though cryptocurrency was introduced as an alternative currency, it is now being used as an asset. This is highlighted in studies such as those of Baur et al. (2018) and Glaser et al. (2014), which show that people are now approaching cryptocurrency as an asset rather than an alternative means of exchange. This is not surprising given the rapid increase of Bitcoin from being free on launch, reaching 10 cents in 2010, and attaining an all-time high of over $60000 in 2021. This deviation from an alternative currency to an asset has been subject to many studies which compare the characteristics of cryptocurrency with other financial assets such as stocks, gold, and even the currency itself. Despite the rapid growth in the value of cryptocurrencies, the growth has been marred by high volatility. In some instances, they were losing over 80% of their current value. Such incidents include the recent value loss of Bitcoin from an all-time high of over $60 000 in 2021 to reaching $18 000 in June 2022. In the same period, Ethereum, the second-largest cryptocurrency by market cap, fell from an all-time high of almost $5000 to just over $900. This tendency to be volatile leads to the study of the nature of cryptocurrency volatility being a priority so that investors have a better understanding of how to invest amid the turmoil.

In studying volatility, the focus is on its characteristics as it is not directly observable. The most studied features of volatility are mean reversion, volatility persistence, and volatility asymmetry. Volatility is said to be persistent if today's return affects the unconditional variance of many periods in the future. Mean reversion refers to the tendency of volatility to return to some average after swinging up and down. Asymmetry refers to the different impacts of positive and negative shocks. These characteristics are essential indicators for investors when determining how to profit from their investments, whether through day trading or long-term holding. One fundamental aspect of investing is understanding the difference in the volatility aspects of different financial assets. While volatility characteristics are the same across financial assets, there is a difference in the level and frequency of such characteristics. For example, given a market shock, the market's reaction speed, i.e. mean reversion, to its expected value is different for each asset. Furthermore, the reaction to similar shocks may differ. This difference can be noted in the reaction of cryptocurrency and the stock market when there is a market shock. For example, interest changes can quickly impact stocks' prices but not affect the cryptocurrency market. As a result, investors can shift some of their funds away from troubled stocks and into cryptocurrency until stability returns.

8.2 Literature review

The high volatility of cryptocurrency has been the major talk in the financial markets, with many investors voicing their worries about the level of volatility. Numerous studies have reiterated the highly volatile nature of cryptocurrency. The high volatility has been attributed to factors such as speculation, lack of regulation, bad news amongst others. The explanation for the unusually high volatility found in cryptocurrencies is not as simple as it seems. Hence, various reasons have been brought forward. In their research, Corbet et al. (2018) found that the introduction of Bitcoin futures resulted in the destabilisation of the Bitcoin market, fueling the increase in volatility. The study also revealed that using Bitcoin futures as a hedging tool was ineffective and that Bitcoin futures did not affect the nature of Bitcoin as a speculative asset rather than a currency. Akyildirim et al. (2020) applied the DCC GARCH models to study the relationship between cryptocurrencies such as Bitcoin and LTC and the United States and European financial markets as measured by the VIX and VSTOXX, respectively. The study revealed that cryptocurrency volatility increased during periods when investor fear was high. The GARCH models also revealed strong correlations with volatility products such as the VIX and VSTOXX.

Similarly, the stock market volatility, which studies have shown to be below that of cryptocurrency, has different causes. However, for stocks, the leading causes are policies made by governments. These policies tend to alter the investors' perspectives and their expected profit margins. Despite the policies being made with market stabilisation in mind, Smith Jr (1988) argues that the markets would be better off with less institutional interference. This point of self-regulatory is the main reason for the creation of cryptocurrency. Based on the observations so far, such decentralisation has not lessened volatility. Another reason for the high volatility and the leverage effect can be derived from Duffee (2002) who argued that balance-sheet effects are a potential source of asymmetric volatility. The study shows that a drop in the value of the stock increases financial leverage, which makes the stock riskier and increases its volatility. In contrast, an increase in price causes the opposite. Cryptocurrency has shown in its entire existence that the value tends to drop suddenly, which, following the reason given by Duffee (2002) increases the financial leverage, fueling the volatility.

The GARCH-type models have been the go-to to determine the half-life and volatility. The most seminal models are the GARCH, EGARCH, and the GJR-GARCH. For most cases, these models use the student t, skewed student t and the general error distributions. Examples of such studies include those of Muguto and Muzindutsi (2022) who used the GARCH, GJR-GARCH and EGARCH to quantify the half-life of volatility in the BRICS markets and the G7 markets. The same models were used by John et al. (2019) who studied the half-life of Bitcoin. The models helped determine the half-life in these studies, but they also show that a slight change in persistence significantly impacts the half-life value. In a similar study, Ghoddusi et al. (2020) uses the same GARCH models but accounts for structural breaks in the data. Dummy variables accounted for the structural breaks. The results showed that the presence of structural breaks leads to overestimated persistence and hence an overestimated half-life. This finding is coherent with the study of Lamoureux and Lastrapes (1990), where the GARCH model was applied to stock-return data and yielded a high measure of persistence close to 1. This high persistence they discovered was attributed to the presence of deterministic shifts in the unconditional variance, i.e., the structural breaks. Using the US stock market indices, the same conclusion is reached by Chatzikonstanti (2017) in the stock market. Besides using dummy variables, the study split the data at breakpoints, and the change in persistence was recorded per segment. Overall, these studies show that accounting for structural breaks is crucial to avoid bias in results.
With regards to mean reversion and momentum (persistence), Zaremba et al. (2021) and Pavlov (2022) explain that for traders, mean reversion and persistence, if notable, can be exploited for profit-making. These traders believe in buying the upward momentum and selling the downward momentum. However, for this to occur, the momentum and mean reversion should be known. In particular, Zaremba et al. (2021) found a powerful one-day reversal in cryptocurrencies. Cryptocurrencies with low returns on the previous day strongly outperform those with high last day's returns. The study also shows that 2% of the biggest coins exhibit momentum rather than reversal. The remaining 98% exhibit means reversion. This difference, the study argues, is due primarily to the liquidity effects, whereby small cryptocurrencies are less liquid. Hence, they suffer from demand and supply shocks that liquidity cannot absorb. With regards to stocks Cubbin et al. (2006), studies the mean reversion in the JSE market and found evidence of mean reversion with portfolios of low price-to-earning ratio (P/E) shares significantly outperforming portfolios of high P/E shares.

When it comes to price discovery in cryptocurrency, using Bitcoin in their study Corbet et al. (2018) discovered that price discovery was by uninformed investors

taking part in the spot market. This conclusion was based on the information leadership share, which had 97% of the information affecting Bitcoin prices reflected in the spot market, while the remaining 3% reflected in the futures market. This finding contradicts the widely known fact that futures markets tend to influence price discovery.
In his study, Kwon (2020), compared the tail behaviour of cryptocurrency (Bitcoin) to that of the US dollar, gold, and the stock market index. The study using the conditional autoregressive value at risk found similarities in the tail behaviours of Bitcoin and the dollar and also the stock market concerning contemporaneous correlation. However, the correlation was negative, suggesting that Bitcoin could be used as a hedge for the dollar and the stock market in a portfolio. The study further reveals through single predictive regression that the tail of stock market return is associated with the risk premium on Bitcoin's return. When the tail risk of the stock market return is high, investors' demand for Bitcoin rises.

Bouri, Azzi and Dyhrberg (2017) while studying the relationship between price returns and volatility changes in the Bitcoin market using asymmetric GARCH models, they found that Bitcoin had inverse asymmetric volatility, meaning that the shocks to the returns were positively correlated to the volatility. They discovered that the reverse property vanished after the crush period. The takeaway was that before the crash, Bitcoin could have been used as a safe haven, but not after. The same study used S&P500 returns for the same periods before and after the 2013 crash. The results showed that negative return shocks to US equities lead to increased volatility, which differs from Bitcoin market results. Furthermore, the study compared a portfolio of 50% Bitcoin and 50% S&P500 to a portfolio with 100% S&P500. The demonstration results showed that adding Bitcoin reduced the portfolio risk for both periods, but mainly before the crash when the inverse asymmetry was found. In a similar study, López-Cabarcos et al. (2019)used the GARCH and EGARCH models to examine the effects of Bitcoin volatility on S&P500 VIX returns and investor sentiment. The results showed evidence of higher volatility in bitcoin as compared to S&P500. The volatility of Bitcoin was unstable during speculative periods. Another study by Dasman (2021), which used t-tests and F tests to compare the mean returns and variance of Bitcoin, the Indonesia Composite Index, and gold, found that Bitcoin was more volatile than gold and the stock market. The average return of Bitcoin was significantly higher, while the stock returns and gold were not significantly different from each other.

From the reviewed literature, it is clear that most studies look at stock and cryptocurrency volatilities separately. Of those that studied both simultaneously, the main focus was on the developed markets, or they focused on the spillover between cryptocurrency and stocks. The results obtained from these studies are then generalised for the developing markets. Studies show that volatility is higher in cryptocurrencies than in stocks. However, considering that the state of the developing markets is more dynamic, generally underdeveloped, and most likely inefficient compared to the developed markets, it is justified that the findings may be different. Further, developing markets are still a go-to investment area for investors seeking higher returns due to the inefficiency they exhibit. This consistent attempt to profit from inefficiency may result in unusual volatility in developing markets.

From the information presented by the reviewed literature may not apply in general to all markets. This is a gap that this study seeks to explore and fill. The common feature of the study with others is that it will use cryptocurrency returns and GARCH-type models but will differ in the sample period as more information was available at the time of our research compared to the reviewed studies. This additional data also sets the study apart because it presents an opportunity to learn more about cryptocurrency properties that may not have been revealed in the past. The study also brings into account the effect of the presence of structural breaks.

8.2.1 Data and methodology

This study uses the daily log-returns of the Bitcoin, Ethereum, and FTSE/JSE 40 index. Bitcoin and Ethereum were chosen as the representations of cryptocurrencies because they are the two biggest by market cap. The JSE market is represented by the FTSE/JSE 40 because it takes the top 40 SA companies listed on the JSE, making it a fair representation of the entire market. All the data was retrieved as daily closing prices from *Financial News and Stock quotes* (n.d.). The daily returns are calculated from the daily closing prices. The sample period runs from 18 September 2017 to 27 May 2021, with 1348 observations per asset. We chose this period to allow for uniformity in the data series based on the availability of the data. The daily prices P_t were converted to returns using the following formula

$$R_t = \frac{P_{t+1} - P_{t-1}}{P_{t-1}} \tag{8.1}$$

The simple returns R_t are converted to log returns. The reason for the conversion is that log returns have statistical properties that are more tractable. They can be used with many statistical theories, such as the need for normalisation (Quigley and Ramsey 2008). The log-returns are obtained as follows:

$$r_t = \ln(P_t/P_{t-1}) * 100 \tag{8.2}$$

The log-returns are multiplied by 100 so as to work with percentage returns rather than raw returns, which will have many decimal places.

Statistic	Bitcoin	Ethereum	JSE
nobs	1348	1348	1348
Minimum	-39.1816	-44.5472	-9.9229
Maximum	22.5512	25.9572	9.4798
Mean	0.2674	0.3254	0.0306
Median	0.1622	0.1400	0.0739
Variance	17.9959	28.8176	1.5220
Stdev	4.2422	5.3682	1.2337
Skewness	-0.2777	-0.2886	-0.3023
Kurtosis	7.8431	6.1876	10.1323

Table 8.1: Descriptive statistics of daily log-returns Bitcoin, Ethereum, Dogecoin and JSE

As seen in Table 8.1, the average daily returns were highest in cryptocurrency, with 0.27% for Bitcoin and 0.33% for Ethereum, while the JSE market had 0.03% which is nine times less than the daily return of Bitcoin. The standard deviations indicate higher volatility in cryptocurrency, with Bitcoin having 4.2 and 5.4 for Ethereum, compared to 1.2 for the JSE. These results are not surprising as they confirm the findings of many studies, such as those by Bouri, Molnár, Azzi, Roubaud and Hagfors (2017) which showed that cryptocurrency has higher returns and a higher standard deviation than stocks. All three assets exhibited negative skewness, indicating heavier left tails compared to the upper tail. This result suggests that the lower returns are more probable than the higher returns. Negative skewness in the Bitcoin and Ethereum return series is also observed by Baur and Dimpfl (2018), and Kim et al. (2021). Another observation is that the highest and lowest returns of the JSE are almost symmetrical, i.e., a maximum of 9.48% and a minimum of -9.92%. This observation contradicts the two cryptocurrencies with higher absolute values on the negative side, as seen by the maximums of 22.55%, 25.96% and minimums of -39.18, -44.55% for Bitcoin and Ethereum, respectively. An important observation compared with other studies is that the skewness and other descriptives are sample period dependent. Periods concentrating on the bull period and those with more of the bear period will have different results. For example Brauneis and Mestel (2018) and Uzonwanne (2021) used sample periods that had negative skewness. Stationarity was checked using the Argumented Dicky Fuller test. This test is essential because stationarity is a model assumption for the models used. Stationary data means there is consistency in series properties, making the model results more reliable, while non-stationary data results in non-consistent and biased results. From the results in Table 8.2, the null hypothesis of lack of stationarity is rejected, concluding that our data is stationary.

Table 8.2: Stationarity tests for the returns

	Test statistic	1%	5%	10%	Conclusion
Panel A daily returns					
Bitcoin	-24.8279	-2.58	-1.95	-1.62	Stationary
Ethereum	-24.5003	-2.58	-1.95	-1.62	Stationary
JSE	-25.1035	-2.58	-1.95	-1.62	Stationary

Apart from the statistics, we also considered the log returns' visual plots, which are presented in Figure 8.1. These plots strongly indicate the presence of heteroscedasticity and volatility clustering for all the return series. The returns of Bitcoin and Ethereum fluctuated around -10 and 10, while those of the JSE fluctuated around -5 and 5. This shows that the two cryptocurrencies have higher returns and losses than the JSE. Periods of random extreme shocks are also visible. As expected, these extreme shocks are similar to those in cryptocurrency. For example, there was a significant negative shock for Bitcoin and Ethereum around July 2019, but none on the JSE.

Similarly, there was a substantial negative shock in the JSE market at the end of 2019, followed by a substantial positive shock. These shocks do not coincide with any of the shocks in the cryptocurrency plots. According to López-Cabarcos et al. (2019), this could indicate that investors were fleeing the stock market during the turmoil.

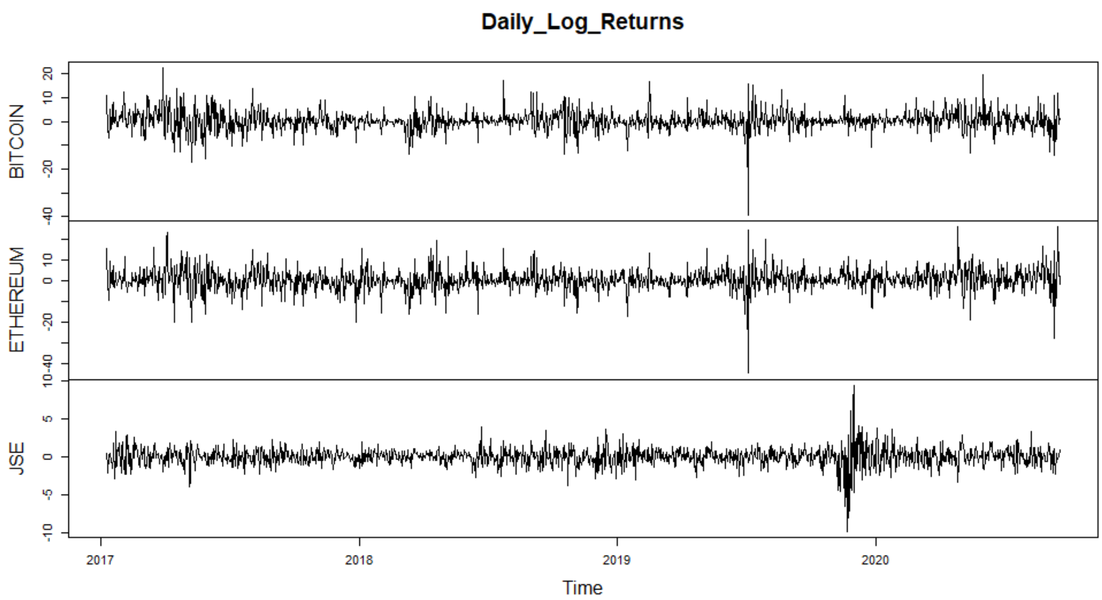

Figure 8.1: Daily log returns

In general, for financial data, serial dependence in a stock return series r_t tends to be weak or does not exist. This is the case with the ACF plots shown in Figure 8.2, where the ACF plots generally have weak low-order correlations. We test the presence using the Portmanteau test, and the results confirm that the low lag correlations are significant. LeBaron (1992) mentions that serial autocorrelations can be introduced into the series during the final stock price index and return calculation. However, this is more common in high-frequency data. The low order autocorrelations also indicate that the markets are not efficient according to the efficient market hypothesis. The squared residuals will have significant autocorrelations. However, these are more pronounced in JSE returns and less in cryptocurrency returns. These significant squared returns show that there are ARCH effects present in the data. This conclusion is cemented by performing the ARCH effects test using the ARCH LM test. The results all had a p-value of less than 0.05 for all three assets. Therefore, we rejected the null hypothesis of no ARCH effects.

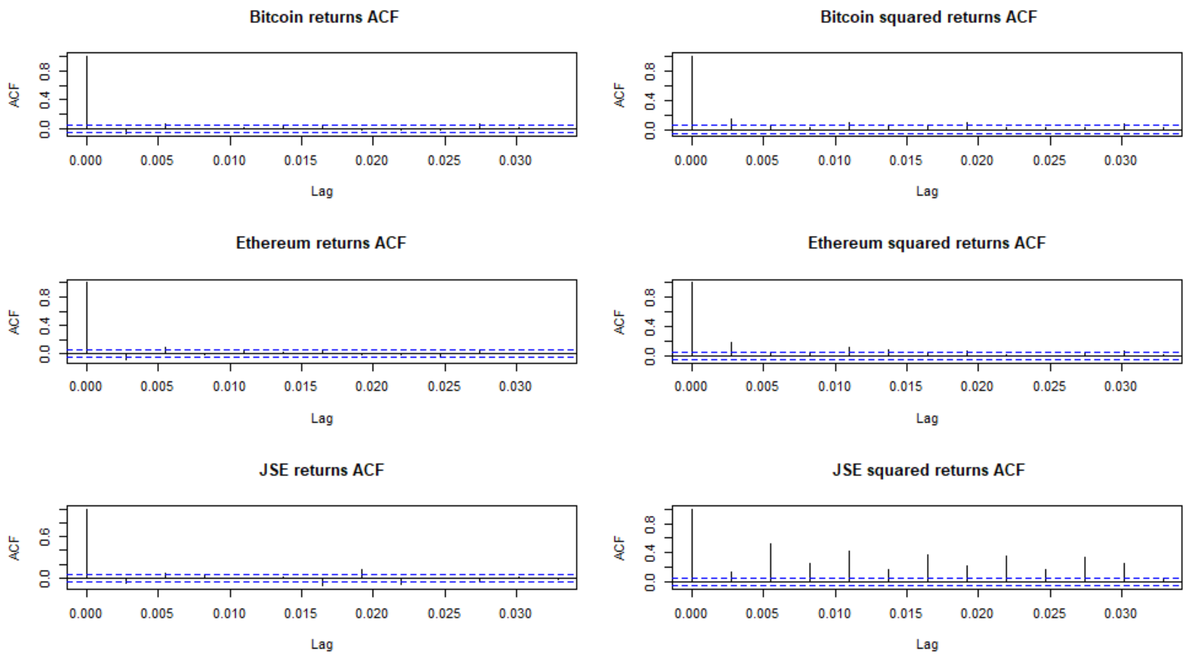

Figure 8.2: ACF plots of returns and squared returns

The QQ plots in Figure 8.3 below show that for all the three assets, the normal assumption is not a good fit for the data, as evident by the data deviating from the line at the ends. This deviation indicates tails that are heavier than the normal distribution, i.e., leptokurtic. To affirm the findings from the QQ plots, we also carried out the Jarque-Bera (JB) test for normality. The results of the test had p values < 0.05. Hence, the normality for log return is rejected.

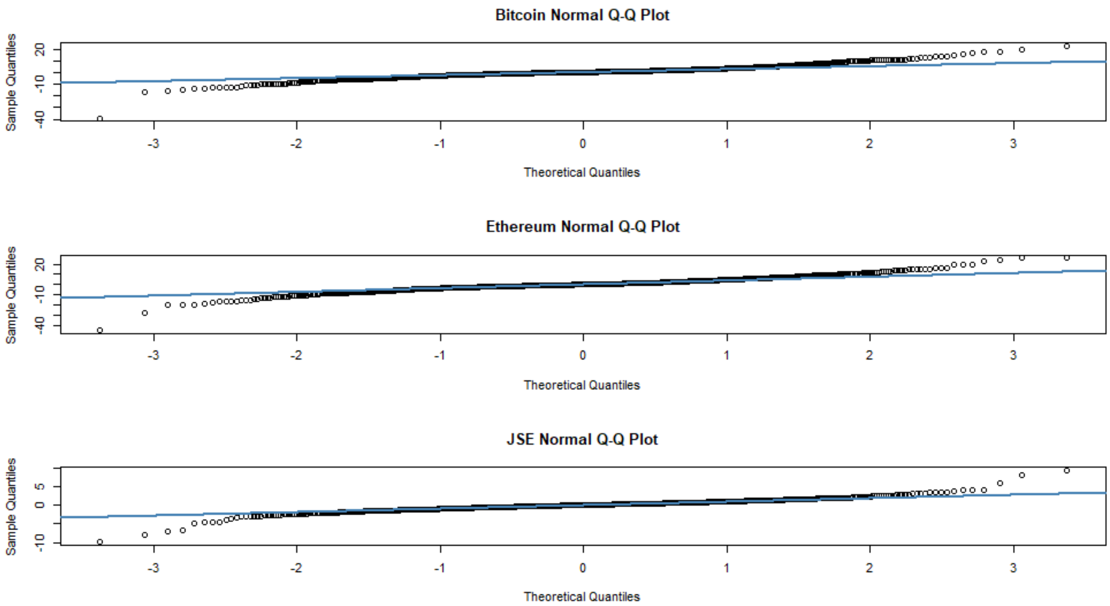

Figure 8.3: QQ plots under the normal distribution assumption

The presence of ARCH effects means we can use the GARCH models, which are the go-to models for heteroscedasticity. Three main GARCH models will be used in this study. These are; the GARCH model of Bollerslev (1986), the Exponential GARCH (EGARCH) model of Nelson (1991), and the GJR-GARCH of Glosten et al. (1993).

The GARCH model was developed to improve on the shortcomings of the ARCH model, of Engle (1982) which required many parameters to capture the volatility process adequately. The GARCH model, on the other hand, reduces the number of parameters required to the point where GARCH (1,1) is sufficient in most cases. This parsimonious model is achieved because, in the GARCH model, the conditional variance is modelled to depend on the past squared residuals and past conditional variance. Despite the modifications made to the GARCH model, it still did not capture some features found in other time series, such as the leverage effects. All the GARCH models follow a similar framework which consists of two equations; the first being the mean model, which captures the conditional mean of the process. The second is the conditional variance equation. Given r_t to be the log return of an asset at time index t, then the mean equation is given by:

$$r_t = \mu_t + a_t, \quad \mu_t = \sum_{i=1}^{p} \phi_1 y_{t-i} - \sum_{i=1}^{q} \theta_i a_{t-i}, \qquad y_t = r_t - \phi_0 - \sum_{i=1}^{k} \beta_i x_{it}, \tag{8.3}$$

where y_{t-i} the return series without the effect of the explanatory variables, a_t is the shock or innovation of an asset return at time t. The mean equation for all

GARCH models is the same, but they differ in how the conditional variance σ_t^2 evolves over time. The conditional variances for the three models to be used are as follows;

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon^2{}_{t-i} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2. \tag{8.4}$$

$$\ln(\sigma_t^2) = \omega + \sum_{i=1}^{s} \frac{|a_{t-i}| + \gamma_i \, a_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^{m} \beta_j \ln(\sigma_{t-i}^2). \tag{8.5}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{s} (\alpha^2 + \gamma_i S_{t-i}) \epsilon_{t-i}^2 + \beta \sigma_{t-i}^2. \tag{8.6}$$

The GARCH, EGARCH, and GJR-GARCH conditional variance equations are represented by equations 4,5 and 6. ω is a constant, β captures the past conditional variance effects on the current volatility, α captures the past shock effects on the current conditional variance. We perform model adequacy tests to get the final lags required per model, which means modelling lower lags until we get insignificantly higher lags. The GARCH model assumes that the data has no serial autocorrelations. This assumption means that, if present, the autocorrelations must first be removed. The removal is achieved by fitting an AR(p) model with the lag p determined by the significant lags from an ACF plot.

For the standard GARCH model, the sufficient conditions to ensure positivity and stationarity of the conditional variance are that, $\omega > 0, \alpha \geq 0, \beta \geq 0$ and that $\alpha + \beta < 1$. The persistence of volatility is measured by $\alpha + \beta$. To ensure covariance stationarity, then $\alpha + \beta < 1$. Despite its success, the standard GARCH model has its limitations. While it captures stylised facts such as persistence and mean-reversion, it retains the weakness of the ARCH model of responding equally to both positive and negative shocks.

The GJR-GARCH improves the GARCH model by allowing the asymmetric effect to be modelled. This asymmetry is captured by an additional parameter, γ, which indicates the presence or absence of the leverage effect. If $\gamma = 0$, then there is no leverage effect. However, if $\gamma > 0$, negative shocks will increase the volatility more than how positive shocks would. The opposite is true for $\gamma < 0$, which means positive shocks increase the volatility more than negative shocks. An indicator function S_{t-i} is used to capture the asymmetry. This indicator takes the value of one in the case of a negative shock and zero in the case of a positive shock. Hence a negative shock contributes $\alpha_i + \gamma_i \epsilon_{t-i}^2$ which is higher than $\alpha_i \epsilon_{t-i}^2$ which is the contribution from a positive shock (Tsay (2014)). The sufficient conditions to ensure positivity and stationarity of the conditional variance are that, $\omega > 0, \alpha \geq 0, \beta \geq 0, \gamma \geq 0$. To ensure covariance stationarity, we require $\alpha + \beta + \frac{1}{2}\gamma < 1$ (Caporin and Costola (2019)).

The EGARCH also improves the GARCH by considering the asymmetry and leverage effects. It differs from the GJR-GARCH in that there is no need for additional constraints to avoid violating the none negativity conditions since the conditional variance is modelled using the natural log, making the variance positive by construction. However, $\omega > 0$ and $\alpha + \beta < 1$ should still hold. γ is the leverage effect parameter of a_{t-i}. If γ is 0, then there is no leverage effect. Otherwise, if γ is negative, negative shocks will increase the volatility more than positive shocks. The opposite is true for a positive γ. The persistence for the EGARCH model is given by $\sum_{i=1}^{p} \beta_i$. If the persistence parameter is less than 1, the return series will exhibit mean reversion. However, if the persistence parameter is equal to 1, then the series follows the random walk (Gbenro and Moussa 2019).

Another concept closely related to persistence is the mean reversion. The mean reversion of volatility is defined by Engle and Patton (2007) as the average level of volatility to which volatility will eventually return. This means that, despite any wild swings, the volatility will remain at its average level. In other words, this means that, in the long run, current volatility will not affect future volatility. To get the average number of periods the volatility takes to revert to its mean, we employ the half-life by Engle and Patton (2007). They defined the half-life as the time it takes the volatility to move halfway back towards its unconditional mean. The formula is given as follows:

$$\ell = \frac{\ln(0.5)}{\ln(Volatility\ persistence)} \tag{8.7}$$

where ℓ is the half life and the Volatility persistence is the volatility persistence of the model used. In the case of the GARCH, then the $\ell = \frac{\ln(0.5)}{\ln(\alpha+\beta)}$, for the EGARCH, $\ell = \frac{\ln(0.5)}{\sum \beta}$ (Mert and Demireli 2020), and for the GJR-GARCH $\ell = \frac{\ln(0.5)}{\ln(\alpha+\beta+\frac{1}{2}\gamma)}$.

Another issue with GARCH models is that they do not account for structural breaks in the data. Studies such as those by Diebold (1986) and Lamoureux and Lastrapes (1990), have shown that structural breaks in time series data should not be ignored as they have an impact on the volatility measures. These studies show that persistence, in particular, is often over-estimated when structural breaks are not taken into account. Using such biased results would have ripple effects on decisions derived from the results. With this in mind, in this study, breakpoints were detected using the Pruned Exact Linear Time (PELT) method from the Changepoint package in R developed by Killick and Eckley (2014). These change points are then included in the GARCH type models in the mean and variance equations as dummy variables. The results are then compared to those of models without the structural breaks.

The three GARCH models were fitted to each asset under three error distributions, the student t, the skewed student t, and the generalised error distribution, to find the best fit model for each asset. The normal distribution was not included since the preliminary analysis of the data showed that it was leptokurtic and diverted from the normal distribution. Hence, heavy-tailed distributions were preferred to capture the tails. These heavy-tailed distributions are ideal as they can capture the rare events in finance, such as random crashes and booms that yield extreme values. In the literature, the choice of the error distribution has varied for stock return data, but the most used is student t, skewed student t and the general error distribution. For cryptocurrency, the same distributions have proven to be efficient in capturing the tails. The choice of the error term distribution usually does not change the results when using the three distributions. However, since with new data, new patterns may emerge, we employ all three distributions. The best model will be chosen using two information criteria: the AIC and the BIC.

8.3 Results and discussions

Three models were fit for each of the assets. The best model was selected using two information criteria, the AIC and BIC values, as shown in Table 8.3 below. Based on the information criterion, the best models were the EGARCH(1,1) with Student t errors for Bitcoin, the GARCH (1,1) with GED for Ethereum and the GJR-GARCH(1,1) with GED for JSE.

Table 8.3: Model selection

	ERROR	BITCOIN			ETHEREUM			JSE		
CRITERION		GARCH	EGARCH	GJR-GARCH	GARCH	EGARCH	GJR-GARCH	GARCH	EGARCH	GJR-GARCH
AIC	STD	5.3575	**5.3452**	5.3580	5.9407	5.9402	5.9421	2.9279	2.9090	2.9045
BIC			**5.3761**	5.3889	5.9678	5.9711	5.9730	2.9549	2.9399	2.9354
AIC	SSTD	5.3584	5.3464	5.3589	5.9398	5.9390	5.9412	2.9235	2.9011	2.8973
BIC			5.3812	5.3936	5.9709	5.9738	5.9759	2.9544	2.9358	2.9321
AIC	GED	5.3614	5.3580	5.3628	**5.9316**	5.9322	5.9331	2.9282	2.9092	**2.9043**
BIC		5.3885	5.3889	5.3937	**5.9587**	5.9631	5.9640	2.9552	2.9401	**2.9352**

Note: All models are fit with an AR(2)component which is not shown due to space constraints.

The resulting model parameters of the three final models are presented in Table 8.4 below. The two cryptocurrencies had positive and significant AR terms, showing autocorrelation in the log returns. The presence is a sign that future returns of the cryptocurrencies can be explained by their past values and that there is a general mean reversion. For the JSE, the AR terms were not significant. The AR terms in the JSE were only added in the model to allow a more balanced model comparison between the cryptocurrency and the JSE market. Another implication of significant serial autocorrelations is that this is evidence against the weak form of market efficiency. This inefficiency enables investors who can employ technical strategies to profit by setting up positions based on the information.

The effect of past shocks on returns, as shown by the parameter α was positive for both Bitcoin and Ethereum. However, it was only significant for Bitcoin, indicating that for Bitcoin, positive past shocks could affect the future returns and volatility of Bitcoin. This finding also points to Bitcoin being an inefficient market, which agrees with the suggestions of significant autocorrelations. For JSE, α was rounded off to zero in the model. The β was significant for all three assets, suggesting that the past volatilities can explain the future volatility. This finding is not surprising as it supports the volatility clustering identified on the return plots and is also a known stylised fact of financial data.

For Bitcoin, the EGARCH model had a positive γ, indicating an inverse leverage effect; this means that positive returns increased volatility more than negative ones. This finding agrees with the findings by Huang et al. (2022) who also find inverse leverage effect in both Bitcoin and Ethereum. In another study, Zhang et al. (2021) discovered that the inverse leverage effect exists in the short run before and after the introduction of Bitcoin futures trading; however, the inverse leverage effect changes to the usual leverage effect in the long run. The JSE model had a positive and significant asymmetry parameter. Therefore means that negative shocks increase volatility as compared to positive shocks. This effect shows that the market is more leveraged and therefore more risky, so the volatility should increase (Bauwens et al. 2012).

Table 8.4: Maximum likelihood estimates for the selected models

Asset	Bitcoin	Ethereum	JSE
Model	AR(2) + EGARCH	AR(2) + GARCH	AR(2)+ GJR-GARCH
$\hat{\mu}$	0.1409**	0.0836*	0.0303
AR1	-0.0765**	-0.1097***	0.0089
AR2	0.0548***	0.0468***	-0.0168
$\hat{\omega}$	0.0090**	1.2331**	0.0326***
$\hat{\alpha}$	0.0125	0.0841***	0
$\hat{\beta}$	0.9971***	0.8758***	0.8987***
$\hat{\gamma}$	0.2023***		0.1448***
$\hat{\eta}$	2.7907***	0.9696***	1.5901***
Persistence	1.0095	0.9599	0.8987
Half-Life	235.7388	16.9637	23.5884

Note: *,**,*** is statistical significance at the 0.05, 0.01 and 0.001 critical level respectively.

Model diagnostics are carried out to check the model's adequacy in capturing volatility. The first diagnostic is to examine the models' standardised residuals. Using the QQ plots in Figure 8.4, we examined if the data fit the distribution assumptions reasonably. The plots show that the heavy-tailed error distributions fit the data reasonably well. The heavier tails of the std error used for Bitcoin and the GED error used for Ethereum and JSE are an improvement from the normal error, which failed to capture the heavy tails. At the ends, a few points deviate from the line, indicating the presence of some rather extreme events. This result is not surprising as financial data is known to produce extreme data points due to random and extreme market shocks. The ACF plots of the residuals are also considered to check if the autocorrelations are removed to give white noise residuals. The ACF plots in Figure 8.4 show that there is one significant low autocorrelation for all three assets. Despite the minor lag, the models describe the conditional mean adequately. Different GARCH and error combinations were tried, but the lower-order autocorrelation remained. This indicates that we may need other models or a change in the error type to be able to remove the significant autocorrelation. A suggestion would be to try extreme value theory distributions.

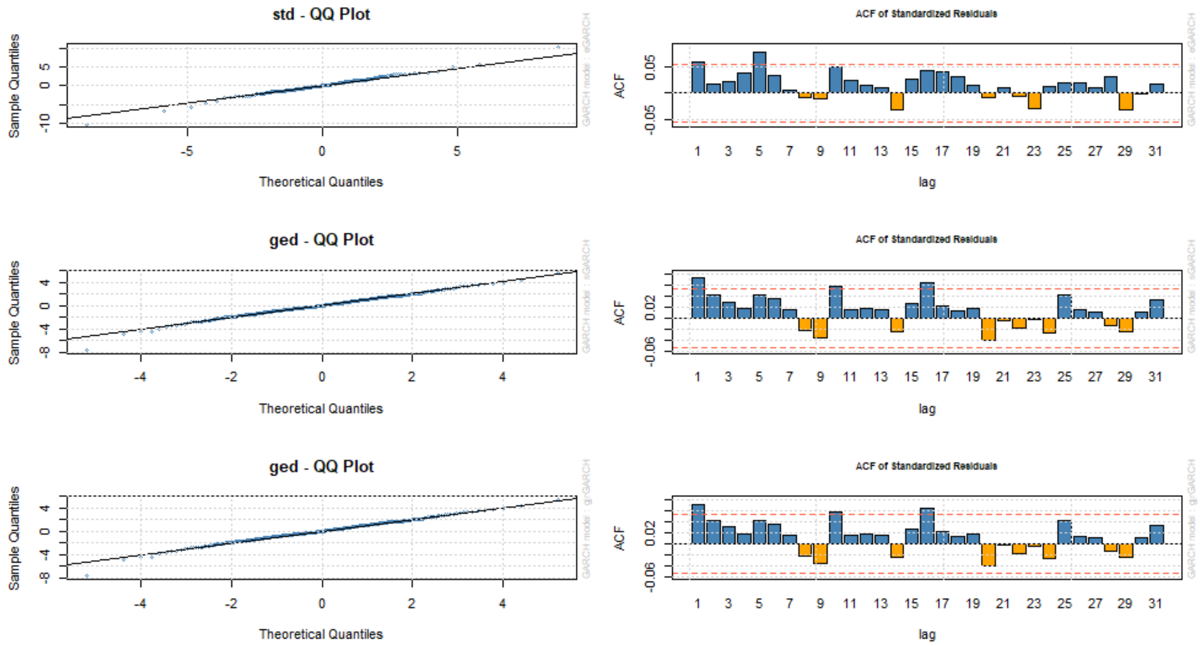

Figure 8.4: QQ plots and ACF plots for the standardised residuals: First row is Bitcoin, Second row is Ethereum and at the bottom is JSE

The conditional volatility plots are presented in Figure 8.5 below. It is visible that the volatility is higher in the two cryptocurrencies as compared to the JSE. The similarities in the volatility plots for cryptocurrencies are striking. The scale range is similar between the two plots, with a maximum value of just over 10, and periods of high volatility are the same. For example, the volatility is slightly higher around Jan-Feb 2018 for both, coinciding with the 2018 crypto-crush after the 2017 boom. Another spike is witnessed in the cryptocurrencies around February/March 2020, which is not observed in the JSE market. This period coincides with the beginning of the COVID-19 pandemic. In June and July 2020, there was high volatility in the JSE market, but the levels in cryptocurrency were normal. This observation is consistent with other studies using the developed markets, such as one by Mariana et al. (2021). Using Bitcoin and Ethereum, they observed concurrent high volatility in the cryptocurrency, which was not observed in the stocks. The same occurred when volatility was high in the stocks. It was not the same in the two cryptocurrencies. Another explanation of the difference in the volatility magnitude is the safe-haven effect explained by Mariana et al. (2021), where investors move funds to cryptocurrency in times of turmoil in the stock market. Therefore, the differing volatility periods suggest using cryptocurrency in a portfolio with JSE as a hedge.

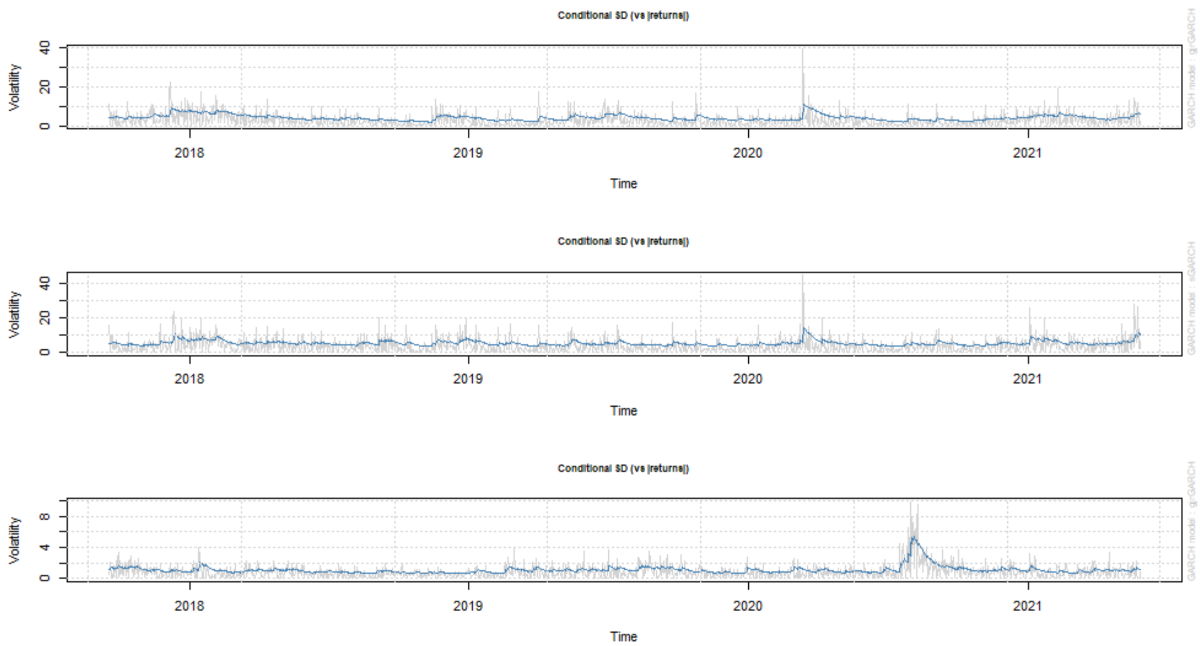

Figure 8.5: Volatility plots with Bitcoin on top followed by Ethereum and lastly JSE

8.3.1 Structural breaks

Table 8.55 reports the results the breakpoints identified from the change point package using the PELT method and a penalty of 45. For Bitcoin and Ethereum, 4 breakpoints each were identified. As expected the occurrence of these breakpoints were in similar position (not exact) for three of the four identified breakpoints. This is not surprising as both are in the same market and have a tendency to face similar market shocks. For the JSE, 2 breakpoints were identified, and these did not coincide with the cryptocurrency ones. These structural breaks will be accounted for in the GARCH-type models using dummy variables.

Table 8.5: Breakpoints identified in the return series

Asset	Number of breaks	Position of breaks	Break point dates
Bitcoin	4	207,906,918,1185	2018/04/12, 2020/03/11, 2020/03/23, 2020/12/15
Ethereum	4	485,902,914,1202	2019/01/15, 2020/03/07, 2020/03/19, 2021/01/01
JSE	2	1033,1069	2020/07/16, 2020/08/21

A visual representation of the identified structural breaks is given in Figure 8.6 below.

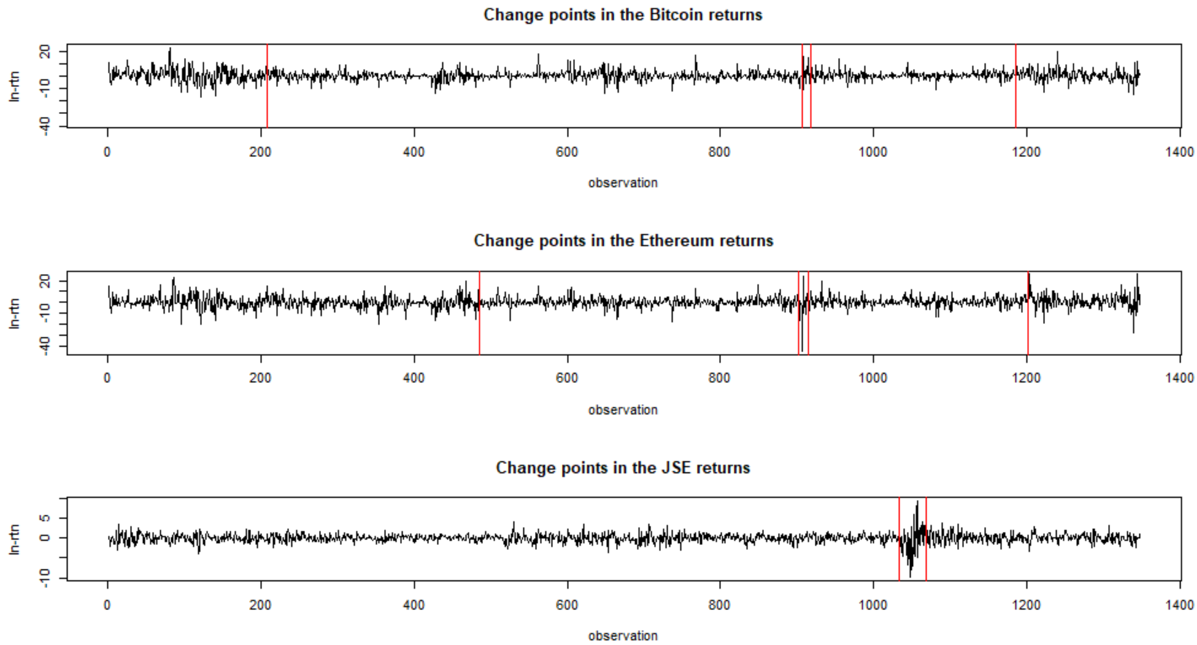

Figure 8.6: Identified structural breaks using the PELT method:

From Table 8.6 the structural breaks are only significant in the mean equation of the Ethereum and the JSE model. For the variance equation, the break point parameter is estimated as 0 for Ethereum and JSE returns but is a positive value for Bitcoin which is however statistically insignificant. A surprising result is the difference in results between Bitcoin and Ethereum. We further interrogated this bizarre result by running the Bitcoin under GJR-GARCH and GARCH model and the results had a significant break under the mean equation. This revelation shows that the significance of the structural break variable is sensitive to the model used. We therefore focus on the implication of the persistence rather than the significance of the parameter itself.

Table 8.6: Maximum likelihood estimates for models with structural breaks

Asset	Bitcoin	Ethereum	JSE
Model	AR(2) + GARCH	AR(2) + GARCH	AR(2)+ GJR-GARCH
$\hat{\mu}$	0.1375**	0.08335***	0.0313
AR1	-0.0770***	-0.1085***	0.0064
AR2	0.0541**	0.0480***	-0.0161
S.Break	2.6780	-2.4896***	-1.0035***
$\hat{\omega}$	0.01207***	1.2457 **	0.0316***
$\hat{\alpha}$	0.0110	0.0850 ***	0
$\hat{\beta}$	0.9957***	0.8745 ***	0.9024 ***
$\hat{\gamma}$	0.2013***		0.1387 ***
S.Break	0.3210	0	0
$\hat{\eta}$	2.8248***	0.9668 ***	1.5703***
Persistence	0.9957	0.9594	0.9718
Half-Life	161.9447	16.7400	24.2010

Note: *,**,*** is statistical significance at the 0.05, 0.01 and 0.001 critical level respectively.

Regarding the model's parameters, the same conclusions are attained from the model with structural breaks. For the cryptocurrency, the AR parameters remained significant, whilst for the JSE, the non-significance is retained. The mean and the alpha parameter of the JSE are insignificant, which agrees with the results of Muguto and Muzindutsi (2022) who also used the structural breaks, which did not change these parameters. Their research revealed a significant structural break for the variance equation. This structural break parameter was 0.0003, which was close to our zero estimate. Critical for our study is the effect of structural breaks on the persistence of volatility and, therefore, the half-life. There is an abundance of research that shows that ignoring the presence of structural breaks causes persistence over estimation in most cases. One such study is by Lamoureux and Lastrapes (1990) who analysed daily stock-return data and a Monte Carlo simulation experiment to confirm the hypothesis that GARCH measures of persistence in variance are sensitive to structural changes. Our results show that for cryptocurrency, this is true, in particular for Bitcoin, whose half-life went from 235 to 162 days. Surprisingly for Ethereum, there was a slight increase in persistence, which did not change the half-life. For JSE, the persistence also increased and led to a 1-day increase in the half-life.

Figure 8.7 are the model diagnostic plots for the models with structural breaks. The QQ plots show that the error distributions are a good fit. The ACF plots still have 2 small overlapping lags for the two cryptocurrencies but shows no significant lag for the JSE residuals.

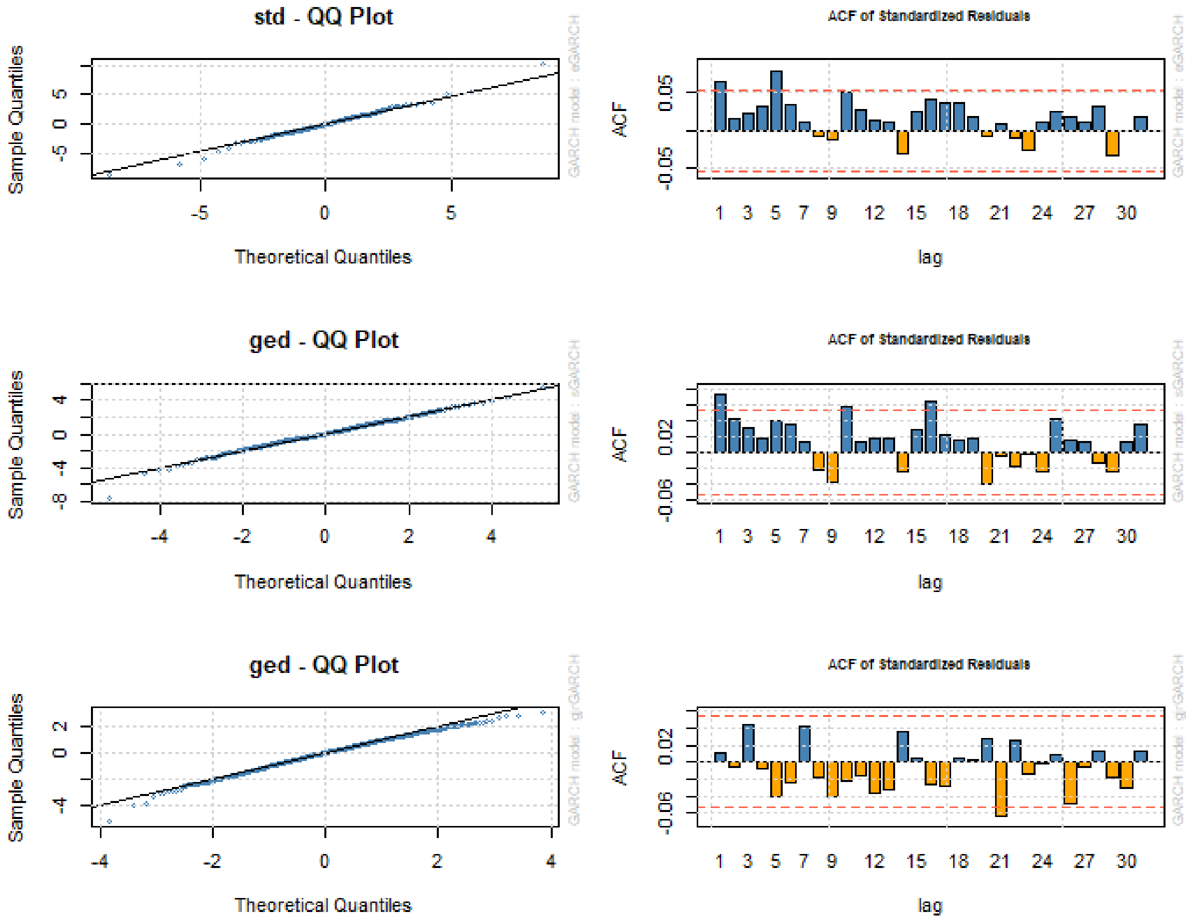

Figure 8.7: QQ plots and ACF plots for residuals for models with structural breaks: First row is Bitcon, Second row is Ethereum and at the bottom is JSE

The results of this study are of significant importance to investors, especially those investing in the JSE market, who need to make informed decisions on how to fully utilise cryptocurrency in their portfolios. Firstly, the findings reveal how the volatility in the cryptocurrency market is always higher than that of the JSE market. This finding warns risk-averse investors to be wary of investing in cryptocurrency. However, risk-tolerant investors can invest with the knowledge of the persistence patterns and hence can limit potential losses by exiting positions based on the persistence and half-life knowledge. Persistence is lower in the JSE market than in the cryptocurrency market, a phenomenon also observed in the developed markets. This observation implies that investors in the JSE can take advantage of the long memory in the Bitcoin market to have risk-adjusted positions in their portfolios. However, it should be noted that the same does not hold with Ethereum, as it had lower persistence than that observed in Bitcoin. This Ethereum result is surprising as, generally, Ethereum moves in tandem with Bitcoin. Such a revelation also serves as a warning for

investors not to assume that all cryptocurrencies behave like Bitcoin.
With these results, portfolio allocation can be made with the know-how of the behaviour of cryptocurrency. Secondly, we observed the effect of structural breaks on the persistence and, subsequently, the half-life. There is a need to use models that cater for structural breaks whenever breaks are detected.
Thirdly, there is evidence of an inverse leverage effect in Bitcoin, unlike the leverage effect that is observed in the JSE. Running the EGARCH for Ethereum also shows that there is inverse leverage. Therefore, investors must be wary of the effect of the positive shocks on increasing volatility rather than the negative shocks. Lastly, based on market inefficiency in cryptocurrencies, especially Bitcoin, some investors who use technical analysis can take advantage of the inefficiency to make abnormal profits. For example, investors, once they identify the structure of the correlations and the level of persistence, can use this to improve their forecasts and, consequently, the potential profitability of their trading strategy. Such findings may indicate the need for some form of cryptocurrency market regulation. The inefficiency also means that traditional asset pricing models such as the CAPM and APT will be inadequate to model returns in these markets (Muguto and Muzindutsi 2022).

8.4 Conclusion

Because different markets have different behaviours, any findings in one market may not apply to the others. These differences arise because of different capitalisations and also different policies. Therefore, in line with this argument, this study compared cryptocurrency behaviour with that of the JSE market. Similar studies have been conducted, but they have focused on developed economies, leaving African economies with generalised conclusions. Such sweeping generalisations may prove fatal for local investors. The study, therefore, fills the gap in two ways, making a comparison with a developing market and also accounting for structural breaks in the comparisons.

The study used GARCH-type models in the absence and presence of structural breaks. The results showed that cryptocurrency possesses most features associated with financial data, such as fat tails, volatility clustering, asymmetry, and persistency. However, it is the magnitude of these features that differs. Similar to developed markets, cryptocurrency volatility is higher than that of the JSE stock market. For the risk-averse investor, they are better off investing in stocks than in cryptocurrency. The presence of structural breaks influenced persistence. However, we noticed that the significance was affected by the model used. A slight difference in persistence has a considerable impact on the half-life. This observation itself casts doubt on the reliability of the half-life measures from GARCH models, as changing models may considerably affect the half-life. Investors or researchers must be wary of these issues before relying on the first result they see. Proper model selection is of paramount importance. The study also revealed that cryptocurrency markets are not efficient. This evidence was more substantial for Bitcoin than for Ethereum. Inefficiency in the market allows for market manipulation as investors can take up positions based on this inefficiency. Observations of the inefficiency of the cryptocurrency also raise the debate on the need for regulation in the market. This idea, however, remains controversial as the goal of cryptocurrency is, in itself, to be free from institutional control. On the issue of inefficiency, the study also revealed that contrary to popular belief that developing markets are inefficient, the JSE exhibited efficiency; this is also found in the study by Muguto and Muzindutsi (2022).

Another important takeaway is that the high volatility periods of cryptocurrency and the JSE market do not coincide. This finding means that investors in the JSE market can look into moving funds to cryptocurrency during the turmoil in the JSE market. In essence, Bitcoin and Ethereum act as a safe haven. Such conclusions were also reached by researchers in European and Asian markets, such as Bouri, Molnár, Azzi, Roubaud and Hagfors (2017), Selmi et al. (2018) and Mariana et al. (2021). However, according to Bouri, Molnár, Azzi, Roubaud and Hagfors (2017), safe-haven properties differ across markets. Other researchers also share such sentiments. Moreover, it has been shown that Bitcoin's volatility behaves differently across time. López-Cabarcos et al. (2019) puts it this way, when stock markets are volatile, Bitcoin can be used as a safe haven; when stock markets are stable, Bitcoin becomes appealing to speculative investors.

Overall, having an understanding of the nature of the volatility of different assets is essential for investors. In this case, we provide volatility comparisons between the JSE stock market and the two biggest cryptocurrencies, Bitcoin and Ethereum. This information, we believe, can assist African investors in having information calculated from our local markets and not relying on generalisations made from developed markets.

Limitations of this paper are that we only considered the JSE as a representative of developing markets, yet there are many other developing markets. The developing markets are all at different stages; therefore, our results may be valid only for the JSE market. Future studies should consider examining the volatility features using the extreme value theorem to try and improve the capture of extreme events. Models that naturally incorporate structural breaks, such as regime-switching models, are also an area to look into to deal with the structural shifts. Another area to consider is combining machine learning and time series models. Deep learning methods which extend Artificial Neural Network models have capacity to handle high dimensional time series data including non-linearities.

8.5 Chapter summary

This chapter showed that the presence of structural breaks in cryptocurrency and that the structural breaks have an effect on the estimates. It also showed that for JSE the structural breaks were not significant in the sample period used. While the introduction of the dummy variable helps in making better estimates, it also considers the period as whole and gives an aggregated result. A better way would be to model the sections that are found to have similar characteristics on their own. One way to do this is to use regime changing models. These models allow the data to reveal regimes of similar characteristics and models those regimes separately. In the next chapter the regime switching models are considered in modeling the cryptocurrency.

9. ESTIMATING THE VOLATILITY IN CRYPTOCURRENCY AND THE JSE USING MARKOV SWITCHING GARCH MODELS

9.1 Introduction

Financial time series modelling has been a subject of constant research, with researchers constantly looking for more robust models, particularly in the wake of the financial crisis, which often led to failures of the current models. The traditional models used in time series, such as the AR and MA models, have proven inadequate in modelling financial data. This inadequacy is because financial data does not adhere to the assumptions such as constant variance and normality that these traditional models assume. Financial data is characterised by none constant variance, volatility clustering, and leverage effects, all of which cannot be captured by the traditional models.

Volatility is one of the most critical aspects that need to be modelled in finance. It is essential because it reflects the riskiness of an asset, with higher volatility indicating higher risk and, therefore, a higher probability of making losses (Muguto and Muzindutsi 2022). Investors are keen to know the risk level associated with their investments and need to know how to forecast the risk. This need was intensified by the global financial crisis in 2008 as investors lost out on their investments, and some were totally liquidated.

In an attempt to solve the issue of modelling financial time series data, Engle (1982) introduced the autoregressive conditional heteroscedasticity (ARCH) model, which would later earn him a Nobel prize in economics. The ARCH model revolutionised how time-series data with time-varying volatility would be modelled. His student Bollerslev (1986) later improved this model. Despite the improvements, the GARCH model still had weaknesses, especially the capturing of asymmetric and leverage effects. To circumvent this issue, various scholars modified the GARCH model. Nelson (1991) developed the EGARCH, and Glosten et al. (1993) developed the GJR-GARCH model to capture the volatility asymmetry and the leverage effect. Together these models have been applied in numerous studies proving to be useful tools of trade in finance and economics.

With the coming of cryptocurrency, a new class of financial data was created.

Just like any other financial asset, there was a sudden need to understand the characteristics of this new asset class. Studies such as those of Kaseke et al. (2021), Drożdż et al. (2018) have given evidence that cryptocurrency has similar stylised facts as those of stock return data. However, the magnitude of these stylised facts tends to differ. The most distinguishable stylised facts are the existence of an inverse leverage effect in some cryptocurrencies and that cryptocurrencies are generally way more volatile than stock returns.
Various studies using different models have been carried out in the study of cryptocurrency. Bouoiyour and Selmi (2015) used the standard GARCH model to examine the Bitcoin market and found the presence of inverse asymmetry to positive and negative shocks. Katsiampa (2017) makes a comparative study on the GARCH type models, i.e., GARCH, EGARCH, TGARCH, Asymmetric Power ARCH (APARCH), Component GARCH (CGARCH) and Asymmetric Component GARCH (ACGARCH). The study found the AR-CGARCH to be the best volatility model. This result they suggestively attribute to the model having both a short-run and a long-run component of conditional variance. Bouri, Azzi and Dyhrberg (2017) focus on the asymmetric GARCH models to compare the US VIX and the Bitcoin volatility. The results showed an inverse relationship between the US VIX and the volatility of Bitcoin. Similarly Baur and Dimpfl (2018) compares the 20 largest cryptocurrencies using asymmetric GARCH models. In this study, they find an inverse leverage effect in cryptocurrency, a result also found by Kaseke et al. (2021).

One critical aspect of volatility modelling is the ability of models to capture all the stylised facts about the data. Not accounting for all the stylised facts leads to biased estimates and, therefore, wrong forecasts, which may be detrimental to investment decisions. Lamoureux and Lastrapes (1990) found the presence of structural breaks to be critical in time series modelling. The study showed that volatility persistence is overestimated when structural breaks are not taken into account. This notion is shared by Thies and Molnár (2018) who detected structural breaks in Bitcoin using a Bayesian change point approach. They split the data into segments as detected by the structural change points. Afterwards, they calculated the average returns, showing that some periods had positive and negative returns despite the combined period showing positive returns. These studies and others echo the need for models that account for these structural breaks in their structure.

The issue of structural breaks in time series data is not new or unique to cryptocurrency but occurs across the divide. Due to the negative impact of the structural breaks in model performances, researchers have been working on fixing the deficiencies of the available models for decades. One early known model that seeks to account for structural changes is by Quandt (1958). He developed a model for the identification of switching points of linear regression. This model motivated the creation of the Markov switching regime model of Hamilton (1989). The Markov switching model used different regimes to characterise changes in the parameters of an autoregressive process, making each regime model periods of homogenous statistical properties, therefore eliminating the bias of structural breaks. Various other regime-switching models were created after that, and we consider some studies that used them.

Studies by Runfang et al. (2017) and Zhang et al. (2015) compare the performance of a single regime and the two regime GARCH type models in forecasting crude oil market volatility. Both studies show that the two-regime models improved in-sample forecasts. Oseifuah and Korkpoe (2019) and Shiferaw (2018) apply the MSGARCH models to the JSE stock return data. In both studies, they used the Bayesian approach. The results all showed that the regime-switching model outperformed the single regime model. The regimes differed in the unconditional volatility level and the persistence of the volatility. This finding showed that the regimes are decided on the volatility of the periods. Therefore, under two regime models, the regimes can be classified as low-volatility or high-volatility.

The regime-switching models have also been applied in cryptocurrency. Figà-Talamanca et al. (2021) employs Autoregressive Hidden Markov chain regime models, whereas Chappell et al. (2019) employs Hidden Markov chain to investigate the volatility of cryptocurrencies such as Bitcoin, Ethereum, and Litecoin, among others. In both studies the regime switching models outperformed the standard single regime models. MSGARCH models have also been applied in studies such as those by Ardia, Bluteau and Rüede (2019) and Maciel (2021). These two studies also indicated that the regime-switching models are better than single regime models.
In this chapter, the focus was on modelling the regime changes of two cryptocurrencies and the JSE. As per the literature review presented, it is not the first study to use regime-switching models in cryptocurrency and the JSE. However, this study contributes to the available literature. First, cryptocurrency is relatively new, and we believe it has yet to reveal its characteristics in terms of stylised facts. As more users enter the space and investors learn more, the characteristics will likely settle into a specific known range. The study considers these two asset classes simultaneously, with data covering the same range. This allows us to make a more reliable comparative study than comparing different studies done at different time intervals. Moreover, this study compares with the JSE, a developing market. There have been comparison studies such as by Suda and Spiteri (2019) and Caferra and Vidal-Tomás (2021), but these considered the developed markets. This means the results retained are more applicable to developed markets. Applying them to developing markets without research could be disastrous, as these markets are at various stages of development. In

addition, unlike previous studies, which used the Markov chain Monte Carlo (MCMC) method, we use the maximum likelihood (ML) method to estimate the MSGARCH.

9.2 Methodology

We follow the structure of Ardia, Bluteau, Boudt, Catania and Trottier (2019). By letting r_t be the log returns of the series at time t. Here we also assume that r_t has been demeaned and has no serial correlations. This means that

$$\begin{aligned} E(r_t) &= 0 \; \forall t, \\ Cov(r_t, r_s) &= 0 \; \forall s, t. \end{aligned}$$

This means that the series must have the mean subtracted, and if it has serial correlations, these must be removed either by using the AR model as suggested by the ACF plots and then modelling the residuals
The MSGARCH model allows for regime-switching in the conditional variance only. The general form of the MSGARCH model is given by

$$r_t|(s_t = k, F_{t-1}) \sim D(0, h_{k,t}, \xi_k), \tag{9.1}$$

where F_{t-1} is the information set available at time $t-1$. $D(0, h_{k,t}, \xi_k)$ is a continuous distribution with mean zero conditional variance $h_{k,t}$ and ξ_k is a vector of shape parameters. The latent variable s_t takes on values { 1,2,...,k} which represent the none overlapping k separate regimes. Therefore for a 2 state regime model, k will be 2. This means that the regimes are piecewise models describing the time varying conditional variances specific to that regime.

To implement the model, the MSGARCH package uses two approaches for the state variable dynamics s_t. These are the Markov-switching GARCH model of Haas et al. (2004*b*) which is a first-order homogeneous Markov chain, and the multinomial distribution of Haas et al. (2004*a*) which characterizes the mixture of GARCH models.

The latent variable s_t is governed by an un-observed first-order ergodic homogeneous Markov chain with a $K \times K$ transition probability matrix A

$$A = \begin{bmatrix} p_{11} & \cdots & p_{1,K} \\ \vdots & \ddots & \\ p_{K1} & & p_{KK} \end{bmatrix},$$

where $P(S_t = j|S_{t-1} = i) = p_{ij}$ is the transition probability from state i at time t-1 to state j at time t. By rules of probability, it suffices that $0 < p_{ij} < 1 \;\; \forall i, j \in$

$1, .., K$ and the sum of all transition probabilities for each regime should sum up to one i.e. $\sum_{j=1}^{k} p_{ij} \quad \forall i \in 1, .., K$.

The Haas et al. (2004*a*) approach is also known as the mixture of GARCH. Here the state s_t is assumed to be sampled independently over time from a multinomial distribution with a vector of probabilities $\boldsymbol{\omega} = (\omega_1, ..., \omega_K)^T$, that is $P[s_t = k] = \omega_k$. Each of the K members of the mixture has a GARCH-type model defined for it.

For the Haas et al. (2004*b*) approach, we assume that the conditional variance $h_{k,t}$ for r_t follows a GARCH-type model. This means that for each regime s_t, $h_{k,t}$ is a function of past returns r_{t-1} and past variance $h_{k,t-1}$ and the regime specific parameter set θ_k as follows:

$$h_{k,t} \equiv h(r_{t-1}, h_{k,t-1}, \theta_k). \tag{9.2}$$

When K is 1, we have a single regime model equivalent to the standard GARCH type model. The GARCH type models considered here are the standard GARCH (Bollerslev 1986), EGARCH (Nelson 1991) and the GJR-GARCH (Glosten et al. 1993).

GARCH model

$$h_{k,t} = \omega_k + \alpha_k r_{t-1}^2 + \beta_k h_{k,t-1}. \tag{9.3}$$

To ensure positivity we require the following to hold, $\omega_k > 0, \alpha_k \geq 0, \beta_k \geq 0$. For the covariance stationarity assumption it must hold that $\alpha_k + \beta_k < 1$.

EGARCH model

$$\ln(h_{k,t}) = \omega_k + \alpha_k(|\eta_{k,t-1}| - E[|\eta_{k,t-1}|] + \alpha_k \eta_{k,t-1} + \beta_k \ln(h_{k,t-1}), \tag{9.4}$$

where the expectation $E[|\eta_{k,t-1}|]$ is with respect to the distribution conditional on regime k. η_k captures the leverage effect specific to regime k. Since $\ln(h_{k,t})$ is modelled, positivity is automatically ensured on $h_{k,t}$ even if the parameters are negative. However to maintain the covariance-stationarity in each regime we assume that $\beta_k < 1$.

GJR-GARCH model

$$h_{k,t} = \omega_k + (\alpha_k + \eta_k \boldsymbol{I}\{r_{t-1} < 0\}) r_{t-1}^2 + \beta_k h_{k,t-1}, \tag{9.5}$$

where $k = 1, .., K$, $\boldsymbol{I}$ is a indicator function that takes the value of one if the condition $r_{t-1} < 0$ holds and zero otherwise. This then means that the parameter η_k controls the extent of asymmetry response to shocks. The necessary conditions for positive conditional variance are $\omega_k,\ \alpha_k > 0$, $\eta_k \geq 0$, $\beta_k \geq 0$. For regime stationarity, we impose that $\alpha_k + \kappa_k \eta_k + \beta_k < 1$. Where $\kappa_k = P[r_t < 0 | s_t = k, F_{t-1}]$.

The models will be fit under three conditional distributions for the innovation terms. These are the Student-t (std), the skewed Student-t (sstd) and the general error (ged) distribution. The pdf's for the conditional distributions are given by:

Student t distribution

The probability density function (PDF) of the standardized Student-t distribution is given by:

$$f(\epsilon, \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{(\nu-2)\pi\Gamma(\frac{\nu}{2})}} \left(1 + \frac{\eta^2}{\nu - 2}\right)^{-\frac{\nu+1}{2}}, \tag{9.6}$$

where Γ is the Gamma function and ν is the degrees of freedom. To ensure the existence of the second order moment we require $\nu > 2$.

Skewed Student t distribution
Developed by Lambert and Laurent (2001) by applying Fernández and Steel (1998) method to the student t distribution. The pdf is as follows:

$$g(\epsilon|\xi, \nu) = \begin{cases} \frac{2}{\xi + \frac{1}{\xi}} \varrho f[\xi(\varrho\epsilon + \bar{\omega})|\nu] & \text{if } \epsilon < -\bar{\omega}\ /\varrho, \\ \frac{2}{\xi + \frac{1}{\xi}} \varrho f[\xi(\varrho\epsilon + \bar{\omega})/\xi|\nu] & \text{if } \epsilon \geq -\bar{\omega}/\varrho. \end{cases} \tag{9.7}$$

where f is the pdf of the standardized student t distribution, ν is the degrees of freedom and ξ is the skewness parameter. ϱ and $\bar{\omega}$ are defined as:

$$\varrho^2 = (\xi^2 + \frac{1}{\xi^2} - 1) - \bar{\omega}^2, \quad \bar{\omega} = \frac{\Gamma[(\nu-1)/2]\sqrt{\nu-2}}{\sqrt{\pi}\Gamma(\nu/2)}(\xi - \frac{1}{\xi}).$$

GED distribution
The PDF of the standardized generalized error distribution (GED) is given by:

$$f(\epsilon, \nu) = \frac{\nu \exp^{-\frac{1}{2}|\epsilon/\lambda|^{\nu}}}{\lambda 2^{(1+1/\nu)}\Gamma(1/\nu)}, \quad \lambda \equiv \left(\frac{\Gamma(1/\nu)}{4^{1/\nu}\Gamma(3/\nu)}\right)^{\frac{1}{2}}. \tag{9.8}$$

The GED can model the error with both positive and negative kurtosis unlike the student t which can only model kurtosis greater than 3.

9.3 Data

For the analysis, the daily log-returns of the Bitcoin, Ethereum, and FTSE/JSE 40 index were employed. Bitcoin and Ethereum will represent the cryptocurrency while the FTSE/JSE 40 will be a representative of the JSE market. The JSE market is represented by the FTSE/JSE 40 because it takes the top 40 SA companies listed on the JSE, making it a fair representation of the entire market. All the data was retrieved as daily closing prices from *Financial News and Stock quotes* (n.d.). The daily returns are calculated from the daily closing prices. The sample period considered is from 18 September 2017 to 27 May 2021, consisting of 1348 observations per asset. The sample period was chosen to cater for the availability of Ethereum which was only created in the end of 2015. To obtain the log returns, the daily prices P_t were converted to log returns using the following formula:

$$r_t = \ln \frac{P_t}{P_{t-1}} \tag{9.9}$$

Table 9.1 reports the daily log returns summary statistics for the three asset classes. All the assets exhibit negative skewness, and positive excess kurtosis, a clear sign of departure from normality. The average daily returns are highest for cryptocurrency, where Bitcoin and Ethereum have an average of 0.27% and 0.33%, respectively. On the other hand, the JSE had an average daily return of 0.03% which is almost ten times less than the cryptocurrency returns. The high average is, as expected, accompanied by high daily returns for cryptocurrency, with Bitcoin having the maximum daily return of 4.2% and 5.4% for Ethereum. The JSE had a maximum of 1.2%. The corresponding minimums were 22.5%, 26% and 9.5% for Bitcoin, Ethereum, and JSE, respectively. As usual, higher returns are accompanied by higher risks, as is the case with Bitcoin and Ethereum, which have a daily standard deviation of 4.24 and 5.37, respectively, compared to the JSE with 1.23.

Table 9.1: Summary statistics of daily log-returns Bitcoin, Ethereum,and JSE

Statistic	Bitcoin	Ethereum	JSE
nobs	1348	1348	1348
Minimum	-39.1816	-44.5472	-9.9229
Maximum	22.5512	25.9572	9.4798
Mean	0.2674	0.3254	0.0306
Median	0.1622	0.1400	0.0739
Variance	17.9959	28.8176	1.5220
Stdev	4.2422	5.3682	1.2337
Skewness	-0.2777	-0.2886	-0.3023
Kurtosis	7.8431	6.1876	10.1323

Figure 9.1 below shows the log return plots for the three assets. From the plots, volatility clustering is clearly visible for all three. We also notice that cryptocurrencies are more volatile, as evidenced by more extreme movements bound by -20 and 20 versus the JSE, which is bound by -5 and 5.

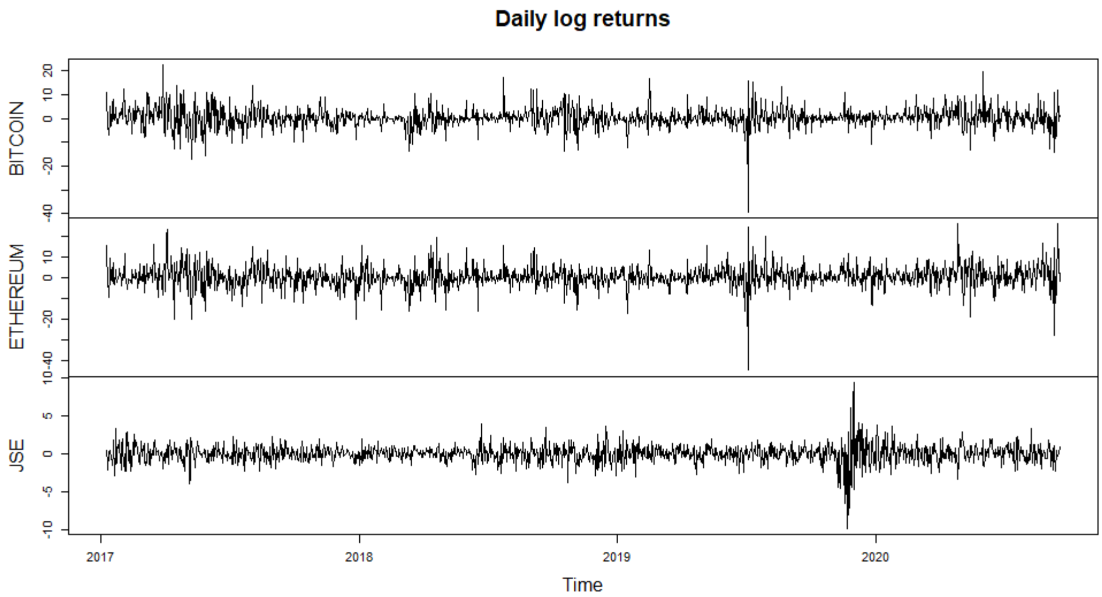

Figure 9.1: Daily log return plots for Bitcoin, Ethereum and the JSE/FTSE40 for the period 18 September 2017 to 27 May 2021

9.4 Empirical results

Each asset is fitted to 18 models (54 for all assets combined) that are a combination of GARCH, EGARCH, and GJR-GARCH; under three error distributions are the student t, skewed student t, and the general error.
We evaluate the goodness of fit using the AIC values. The AIC compares the best model amongst the fitted and does not necessarily mean that the model is the correct one for the data.

The results are shown in Table 9.2 below. The best models for the two cryptocurrencies were the two regime models under the skewed student t error distribution. The best model for Bitcoin was an EGARCH, and for Ethereum, it was a GJR-GARCH. The JSE results show that the single regime outperformed the two regime model. This means that the gain in the complexity of the two regime model was not enough to outdo the single regime. This result contradicts the findings of Oseifuah and Korkpoe (2019), who discovered that the two regime models outperformed the single regime models in the JSE. However, the differences may be attributed to the different sample periods. The other studies covered the financial crisis, which may have included a significant change in structure. Due to the need to use a comparable period with cryptocurrency, which only became available after the financial crisis, our data does not include the global crisis.

Table 9.2: AIC values for GARCH-type models.

Model	Student t	Skewed student t	General error
BITCOIN			
Single Regime			
GARCH	7206.4398	7205.808	7253.1951
GJR	7216.7994	7207.8184	7255.0664
EGARCH	7189.5184	7188.6881	7213.648
Two Regime			
GARCH	7186.2513	7186.2259	7182.7945
GJR	7190.117	7190.1425	7188.2179
EGARCH	7176.5127	**7176.2846**	7187.2558
ETHEREUM			
Single Regime			
GARCH	7994.6869	7990.8368	7992.7339
GJR	7996.6939	7992.8376	7994.754
EGARCH	7993.5634	7989.2379	7993.1422
Two regime			
GARCH	8011.9027	8001.5724	7997.5266
GJR	8014.7133	**7988.8338**	8001.3074
EGARCH	8030.4886	8009.8559	8000.4702
JSE			
Single Regime			
GARCH	3958.4436	3948.9646	3959.706
GJR	3932.9159	**3919.7507**	3933.0486
EGARCH	3942.8035	3929.8371	3943.5503
Two regime			
GARCH	3964.4407	3958.1746	3962.0955
GJR	3946.9145	3930.6321	3940.3376
EGARCH	3939.8998	3937.4962	3943.8598

We consider the three best models for further analysis for each asset. The parameter estimates are shown on Table 9.3 below:

Table 9.3: Maximum likelihood estimates for the selected models, V1 and V2 are the calculated unconditional volatilities

Parameter	Bitcoin	Ethereum	JSE
Regime 1			
$\hat{\omega}_1$	-0.0071**	0.5013*	0.0313***
$\hat{\alpha}_1$	0.16 ***	0.0981*	0
$\hat{\eta}_1$	0.022 *	0.0002	0.1378***
$\hat{\beta}_1$	0.9966 ***	0.8728***	0.9036***
$\hat{\nu}_1$	3.4383 ***	3.1611***	12.4971***
$\hat{\xi}_1$	1.0512***	1.1472***	0.863***
V_1	16.3	77.7	
Regime 2			
$\hat{\omega}_2$	3.7247	45.5077***	
$\hat{\alpha}_2$	-0.4484**	0.0003	
$\hat{\eta}_2$	-0.5584	0.1051	
$\hat{\beta}_2$	-0.0586	0.0001	
$\hat{\nu}_2$	88.3268***	81.5447***	
$\hat{\xi}_2$	1.3413***	0.8182***	
P_{11}	0.903***	0.7797***	
P_{21}	0.7144***	0.6949***	
V_2	119.9	133	

Note: *,**,*** is statistical significance at the 0.1, 0.05 and 0.01 critical level respectively. Standard errors are in parenthesis

For Ethereum under the GJR-GARCH, we observe that the two regimes are heterogeneous. The results show that the leverage parameter η_k is positive and significant under both regimes. This result means past negative returns increase volatility more than past positive returns. The reaction is weaker in regime one i.e. $\eta_1 = 0.0002$, than in regime 2, $\eta 2 = 0.1051$. Using 365 uninterrupted trading days since cryptocurrency is traded throughout the year, we get the unconditional volatilities of 71.7 for the first and 133.0 for the second regime. The persistence is different for the two regimes, with regime one being highly persistent at 0.971 and regime two not persistent at 0.0529. In short, we can classify regime one as "calm market conditions" characterised by low volatility, high persistence, and weak reaction to past negative returns. On the other hand, regime two can be classified as "turbulent market conditions" characterised by high volatility, a strong reaction to negative returns, and low persistence. The posterior probabilities $P_{1,1}$ and $P_{2,2}$, show that both regimes are persistent, with regime 1 being slightly more persistent than 2. The stable state probabilities were 0.7593 for regime one and 0.2407 for regime two, indicating a higher probability of being in regime one than in two.

For Bitcoin, the two regimes are also heterogeneous. The second regime parameters are negative. Nevertheless, for the EGARCH model, parameters can be negative because the model estimates $\ln(h_{k,t})$ which means that $h_{k,t}$ will be positive regardless of the parameter sign. In this regard, we continue with the model. For regime one, η_1 is positive. This positiveness indicates the reverse/inverse leverage effect. Here positive returns increased volatility more than negative ones. However, the reactions are not strong, given that the estimate is 0.022. Under regime 2, the η_2 is negative and of greater magnitude than under regime one $\eta 1$. This means there is a normal leverage effect where past negative returns increase volatility more than past positive returns. Volatility is also higher in the first regime and barely persistent in the second regime. The stable start probabilities were 0.8804 for regime one and 0.1196 for regime two. This means that most of the time, the market is calm. Summarising the regimes can similarly classify the regimes as we did for Ethereum, where the first regime is the calm market conditions and the second is the turbulent market conditions.

Overall, regimes one and two are associated with low and high unconditional volatilities for cryptocurrencies. This finding is consistent with the findings of Maciel (2021). Based on the stable state probabilities, we can also conclude that the cryptocurrency markets are more likely to be calm than turbulent. Once in these states, the results also show from the posterior probabilities $P_{1,1}$ and $P_{2,2}$ that the market persists longer in the calm market, but the turbulent period is short-lived. The tail probabilities do not have a similar pattern, but the values do not differ much across the regimes.

The single regime model of the JSE returns shows a high volatility persistence of 0.9725. We also note that the leverage effect is significant and positive for the JSE, indicating the existence of volatility asymmetry. Therefore, in the JSE market, negative returns will increase volatility more than positive returns.
Figure 9.2 shows the smoothed probabilities, $P[s_t = k|F_T]$, for the two regimes for both Bitcoin and Ethereum. The persistence of regime one is, as shown by the model parameters, higher than that of regime two, as indicated by the high probabilities. The volatility plots also show high volatility, that is, most of the time, above 50%. The volatility plots are similar but differ in intensity.

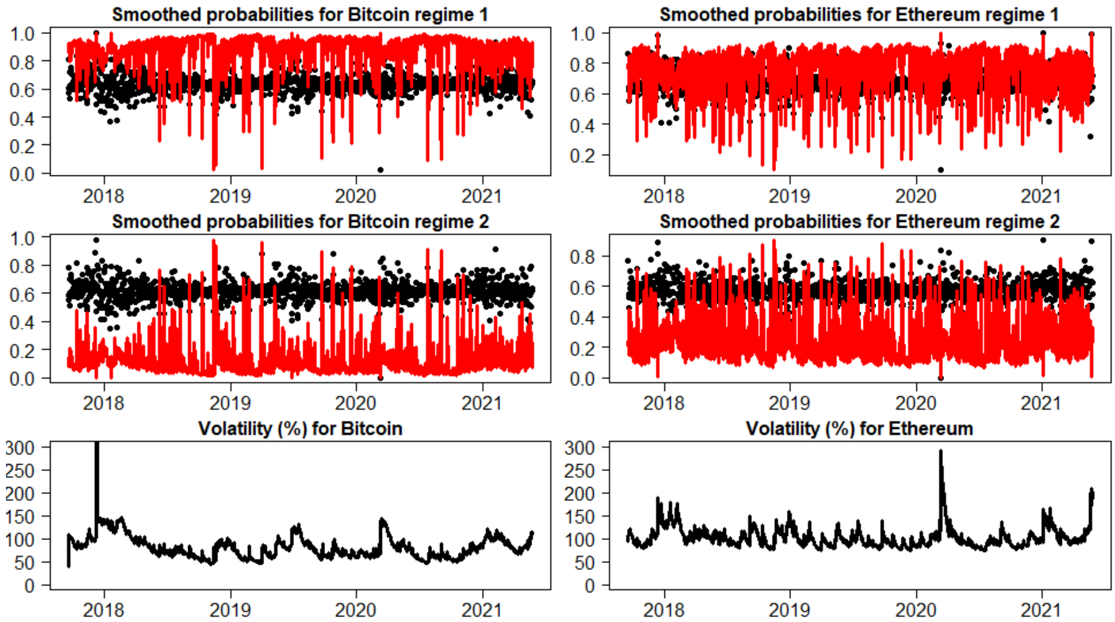

Figure 9.2: Left: Smoothed probabilities for the two-regime model for Bitcoin and the volatility plot. Right: Smoothed probabilities for the two-regime model for Ethereum and the volatility plot.

9.5 Discussion and Conclusion

There is an increasing need for reliable models when estimating or forecasting the volatility of financial assets. One of the fastest-growing investment areas and also one of the least studied is cryptocurrency. The current literature focuses on the traditional GARCH-type models. While these models have been proven adequate in stock data under certain conditions, the need for improvement in cryptocurrency has been proven. Studies such as those by Lamoureux and Lastrapes (1990) and Diebold (1986) have shown that structural breaks affect the traditional GARCH model estimates. This chapter attempts to find a suitable model to describe the volatility dynamics in cryptocurrency while also comparing it with the JSE by using Markov switching GARCH models.

We include the JSE stock returns because we want to make a comparative study between the two asset classes. The literature consists of studies such as those by Suda and Spiteri (2019), İçellioğlu and Öner (2019) and Fakhfekh et al. (2021) which focus on comparing cryptocurrency with developed markets. This means the results are more specific to those markets, and applying them in developing markets may lead to losses for investors.

The results showed that the two regime models outperformed the single regime models for the two cryptocurrencies. This finding is consistent with the findings of Ardia, Bluteau and Rüede (2019) who also found the two regime models outperform the single regime models for Bitcoin. Similar findings are made by Figà-Talamanca et al. (2021), Maciel (2021) and Mensi et al. (2019). However, this was not the case with the JSE stock data, where the single regime model outperformed the two regime models. This contradicts the findings of Oseifuah and Korkpoe (2019) and Makatjane and Molefe (2020) who found the single regime less effective. However, we are quick to note that we used a period that starts in 2017, whilst the studies mentioned above-considered sample periods that cover the global financial crisis of 2008. The global crisis may have played a role as there was a significant shift in the market, while there were no significant shifts in our period. This finding also highlights that market shifts are less frequent in the stock markets than in cryptocurrency. Historical evidence shows that there have been multiple cryptocurrency market crashes due to the speculative nature of the market.

Regarding market dynamics, the results show that cryptocurrency is more volatile than stock market data. These findings agree with the findings from developed markets. For our sample period, we also identify that for both Bitcoin and Ethereum, regime one is the calm market conditions, and it is highly persistent. In contrast, the high volatility regime is less persistent. This agrees with the findings of Chkili (2021). There is also a higher probability of being in regime one for both cryptocurrencies, as evident from the stable state probabilities. The cryptocurrency also exhibited inverse leverage, a result also found Huang et al. (2022), Zhang et al. (2021) and Kaseke et al. (2021). On the other hand, the JSE had the normal leverage effect, confirming the results of Kaseke et al. (2021) and Muguto and Muzindutsi (2022). This finding suggests the use of cryptocurrency as a safe haven for investors in the JSE.

The implications of these results for investors are as follows: Due to observed differences in volatility levels and persistence across regimes, investors ought to pay attention to market conditions before they apply their trading strategies. Here they must focus on whether the cryptocurrency is in a turbulent or calm period and note the persistence level for that period, especially for futures trading. Those wanting to hedge their JSE portfolios with cryptocurrency have to account for the inverse leverage effect as a strategic viewpoint.

Future studies may focus on using artificial intelligence and allowing the data to drive the search for the best model via supervised learning algorithms. Another area is applying multivariate time series models to identify how the assets are co-dependent, allowing more accurate forecasts and better portfolio allocation strategies.

9.6 Chapter summary

This chapter focused on investigating the effect of accounting for regimes in the GARCH models. This was motivated by the fact that market characteristics change depending on the prevailing conditions. Hence, accounting for the regimes may help separate data int homogenous regimes and improve model estimates. The results confirmed that two regime models outperform the single regime models.

10. INVESTIGATING THE SPILLOVER EFFECTS BETWEEN CRYPTOCURRENCY AND JOHANNESBURG STOCK EXCHANGE

This chapter introduces the multivariate models. The focus is on modelling the codependence that may occur between cryptocurrencies and between cryptocurrency and the JSE market. The co-dependance is refereed to as the spillover effects. This because the in finance due to globalisation, when one influential market receives a significant enough shock(usually negative) the effetcs of the shocks tend to spill into the other markets. The purpose therefore of this chapter is to see which cryptocurrency has such an effect on others an also to see if shocks in JSE affects the cryptocurrency markets and vice-versa.

10.1 Introduction

Since its introduction in 2008 in a concept paper by Nakamoto (2008) and its eventual creation in 2009, Bitcoin has become one of the most sought-after investment options. The early days after its creation saw most people label it a joke and a scam. However, despite all this, Bitcoin has grown in value from virtually 0 dollars on launch to reaching one dollar for the first time in February 2012 before reaching its all-time high of 68 564.40 dollars in November 2021 *Bitcoin* (n.d.), earning itself the name "digital gold" in the process. While it was the first cryptocurrency, other alternative cryptocurrencies have also been created. The number of cryptocurrencies has since skyrocketed, with the number of coins listed on CoinMarketCap (n.d.) being over 10 000 as of May 2021, with many others not being listed on the website. The emergence of other alternative coins resulted in a reduction in the market dominance of Bitcoin, which, however, still commands the market with a market share of nearly 50%. The combined all-time high of the market value of the cryptocurrency market rose from around 578 billion dollars in November 2020 to a staggering 3 trillion dollars in November 2021. This represented over a fivefold growth in market value. Indeed, with such monetary value, cryptocurrency could no longer be ignored as a financial product. Like any other investment asset, investors need to know the characteristics and behaviour of cryptocurrency. There is a need to understand the risk associated with the cryptocurrency market and how it can play a part in portfolio allocation.

Various studies have been carried out on the cryptocurrency market, with most studies focusing on modelling the volatility of the major cryptocurrencies such as Bitcoin and Ethereum. These studies, such as those by Bouri, Molnár, Azzi, Roubaud and Hagfors (2017), Al-Yahyaee et al. (2020) and Drożdż et al. (2018) show that cryptocurrencies are more volatile as compared to the stock market, making them a risky investment. However, studies have also shown that Bitcoin and other coins can play a pivotal role in portfolio allocation as a diversifier, as shown in the studies by Bouri, Molnár, Azzi, Roubaud and Hagfors (2017) and that of Guesmi et al. (2019). The volatile nature of Bitcoin also revealed another common occurrence in the cryptocurrency market: that the fall in the price of Bitcoin would result in others losing their value. The same would be observed if Bitcoin increased in value. These occurrences pointed towards spillover effects amongst the cryptocurrencies. Investors and risk managers have long emphasised the importance of understanding spillovers across financial markets. Past studies have shown that different financial markets are not entirely secluded but do face some form of spillover effects from one another. As expected, the most commonly studied spillovers have been between stock market data from different parts of the world. This is because they have been around for a long time and are also the most invested in. The presence of spillovers in the stock markets is identified in studies such as ones by Arouri et al. (2011) and Raputsoane (2008).

With the insurgence of cryptocurrencies as a go-to investment alternative, like stock markets, the cryptocurrency markets are not expected to be immune to volatility spillovers. These markets also exist within a world economy and rely on fiat currencies such as the US dollar. As a result, the spillovers can also be among the cryptocurrencies and country stock markets. The knowledge of return and volatility spillovers across different cryptocurrencies is of great importance to investors and speculators as this may help in portfolio adjustments. For example, suppose it is understood that one cryptocurrency coin has a spillover effect on another. In that case, it allows portfolio managers to adjust their portfolio allocations to mitigate the spillover effects. This is a common practise for investors as they want to hedge their risk. A lack of a well-diversified portfolio will result in investors or portfolio managers exposing themselves to higher risk than they would with a well-calculated portfolio. Furthermore, volatility spillover information is vital in other financial applications, such as option pricing, value at risk, and hedging, as they rely on estimating conditional volatility. It is hence the goal of this study to extend the limited literature on dynamic connectedness and integration in cryptocurrency markets and the equity markets, in this case, with the JSE.

From the reviewed literature presented in the section below and numerous others not mentioned, and to the best of our knowledge, there seem to be no studies on the spillover effects between cryptocurrencies and the Johannesburg Stock Exchange (JSE). As the largest stock exchange in Africa, we see fit to fill this gap by investigating the spillovers between cryptocurrencies and the JSE. We consider Bitcoin, Ethereum, Dogecoin, and JSE stock return data. The study seeks to address two main objectives. Firstly, it investigates the presence of unidirectional transmission or bidirectional transmission spillover effects amongst the chosen cryptocurrencies. It also considers the direction of the spillovers from cryptocurrencies and the JSE. Secondly, the study examines the dynamic correlations between cryptocurrencies and the JSE. The volatility spillovers are captured using the DCC-GARCH model. We chose this model as it makes use of dynamic correlations. The study will contribute by adding knowledge about the financial contagion between cryptocurrencies and the South African JSE market. The rest of the paper will be organised as follows: the literature review, the introduction of the data and the methodology; the application; and the conclusion.

10.1.1 Literature Review

While there have been numerous studies of the spillover effects of stock/equity markets, there have been fewer studies on the spillover effects amongst cryptocurrencies. Of those existing, the focus is mainly on the volatility spillover effects between the cryptocurrencies and Europe and Asian stock markets. As literature shows, the go-to models for volatility modelling are the GARCH-type models. These models have also been the most used to detect spillover effects. Yang and Zhou (2017) use structural Vector Autoregressive models to study the spillover of the U.S. Treasury bond market and global markets for stocks and commodities. Their key finding was that the U.S. stock market is at the centre of the international volatility spillover network; hence, investors must consider the spillover when creating portfolios. Other studies used the GARCH type models, as seen in the study by Mikhaylov (2018) where he employed the FIGARCH model to examine the feedback between the exchange rate and the stock market. The study showed that the volatility spillover effect is observed in one direction: from the exchange rate to the stock market. Using another GARCH type model Another class of models that have been the most used are the ones employing time-varying correlations. These were used in studies, such as one by Yousaf et al. (2020) who used the BEKK-GARCH model to show that there is a unidirectional return transmission from Mexico to the U.S. stock market during the global financial crisis. This study also revealed that there was bidirectional volatility transmission between the U.S. and the stock markets of Chile and Mexico during the global financial crisis. Khalfaoui et al. (2019) also used a time-varying correlation model, the DCC-GARCH model, to examine the spillover effects between oil-importing countries like the USA and the oil-exporting countries like Saudi Arabia and Russia. The results showed that oil-importing countries were affected by lagged oil price shocks, and there is less evidence of interdependence between stock markets for both oil-importing and oil-exporting countries.

The study of cryptocurrency relationships is not the first, as there have been other studies. Literature abounds with studies mainly about the relationship between Bitcoin and other cryptocurrencies. This is not surprising as Bitcoin was the first cryptocurrency and also has the biggest market share by capitalisation. Examples of such studies include Ji, Bouri, Lau and Roubaud (2019) whose study of the spillover of Bitcoin, Ethereum, and Litecoin showed that Litecoin is at the centre of returns and volatility connectedness, sharing with Bitcoin the dominant transmitting role for total return and volatility spillovers. Likewise, Luu Duc Huynh (2019) makes use of the VAR and SVAR (Structural Vector Autoregressive Model) Granger causality and Student's t Copulas to estimate the spillover effects between Bitcoin and Ethereum. The study revealed that Ethereum is likely to be an independent coin, unlike Bitcoin, which tends to have a sensitive recipient, influenced by the other coins.

Apart from the spillovers between the cryptocurrencies themselves, there have been studies of the spillover effects between cryptocurrencies and other markets. Ji, Bouri, Roubaud and Kristoufek (2019) employ the transfer entropy approach, where they use the transfer entropy to capture information on the spillover direction between cryptocurrencies and energy commodities. The study revealed a weak connection between cryptocurrencies and energy commodities. Similarly, Symitsi and Chalvatzis (2018) using VAR-GARCH, find evidence of short-term return spillovers from energy and technology stock indices to Bitcoin. At the same time, Bitcoin has long-run volatility effects on fossil fuel and clean energy stocks. Another study, Huynh et al. (2020) studies the interconnectedness among 14 cryptocurrencies and their linkages with gold prices by employing transfer entropy. The study showed that the cryptocurrencies intended to be tied to the U.S. dollar are likely to be volatile by sending and receiving shocks. They conclude, using transfer entropy, that gold is still a good investment for investors looking to hedge in the cryptocurrency market. However, in so doing, it is crucial to differentiate between different cryptocurrencies.

A few studies consider the volatility spillovers between cryptocurrencies and stock markets. Kristoufek (2015) investigated the drivers of Bitcoin, both technical and speculative. Using the wavelength approach, they found that the Bitcoin and Chinese markets have strong positive correlations practically at all scales and during the entire examined period. Zhang and He (2021) using the MSV model with dynamic correlation and Granger causality to study the spillover effect between bitcoin, gold, crude oil, and major stock markets. They found that Bitcoin is more independent and less affected by the volatility of major stocks and the gold market volatility. On the other hand, gold has a one-way spillover effect from stock market volatility. This finding suggested using Bitcoin as a component of a high-risk asset allocation. Another study by Qarni et al. (2019) used the VAR and NVAR models and found that there was low-level contagion between Bitcoin and the U.S. market, but at high frequency, there was more connectedness between the two. This suggested rapid processing of the information between the markets.

10.2 Data and methodology

In this section, we will discuss the data and methodology used in our paper. We first discuss the data.

Data

This study uses the daily log-returns of the Bitcoin, Ethereum, Dogecoin and FTSE/JSE 40 index. We chose Bitcoin and Ethereum because these are the two biggest cryptocurrencies by market share. We include Dogecoin because it was one of the fastest-growing cryptocurrencies in popularity and the percentage growth in price. The daily returns are calculated from the closing prices obtained from za.investing.com. The sample period runs from 18 September 2017 to 27 May 2021. We chose this period to allow for uniformity in the data series based on the availability of the data. Given the price P_t, we calculate the return as follows:

$$R_t = \frac{P_{t+1} - P_{t-1}}{P_{t-1}}. \tag{10.1}$$

then the log returns are given by

$$r_t = \ln(R_t / R_{t-1)}. \tag{10.2}$$

The choice of using the log returns is because they offer us a scale-free unit of measure and have tractable statistical properties.

Figure 10.1 shows the return series of the four different assets. It can be seen that the returns fluctuate around zero. However, all three cryptocurrencies are spikier than the JSE returns, indicating that they have higher volatility than the JSE returns. This finding is similar to other studies where cryptocurrency returns were found to be more volatile than stocks. An example is the study Chaim and Laurini (2018) where they used Bitcoin and the SP500 daily log-returns. The results showed that Bitcoin was more volatile than the SP500. Our data further reveals that Dogecoin has the biggest spikes, especially in 2021, when there was a massive increase in Dogecoin purchases. The bigger spikes coincide with the boom of Dogecoin after it got the endorsement of high-profile people, with Elon Musk being at the forefront.

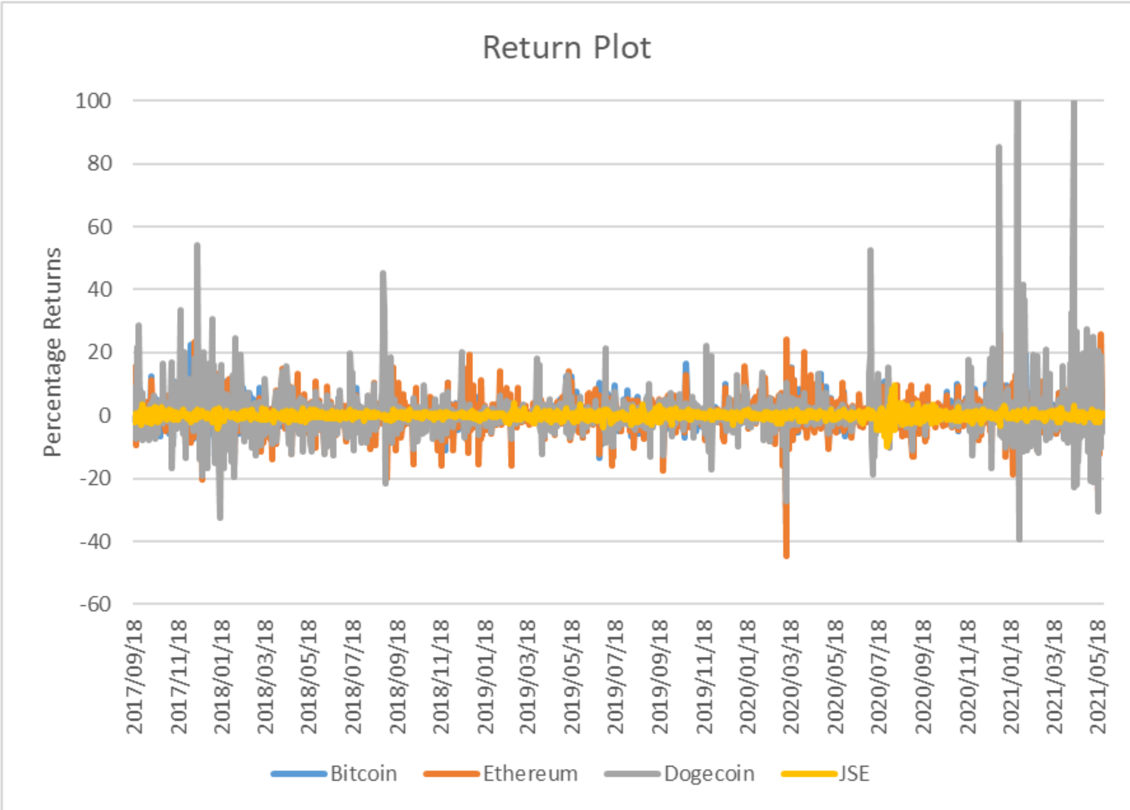

Figure 10.1: Return series of Bitcoin, Ethereum, Dogecoin and JSE

Figure 10.2 below shows the market dominance of Bitcoin from 2013 to 2021. The graph shows the decline in the overall dominance of Bitcoin as other cryptocurrencies began gaining more of the market share. The inclusion of Dogecoin was made due to its rapid rise in demand, which saw it briefly overtake Bitcoin as the most sought after coin, as revealed by Google trends.

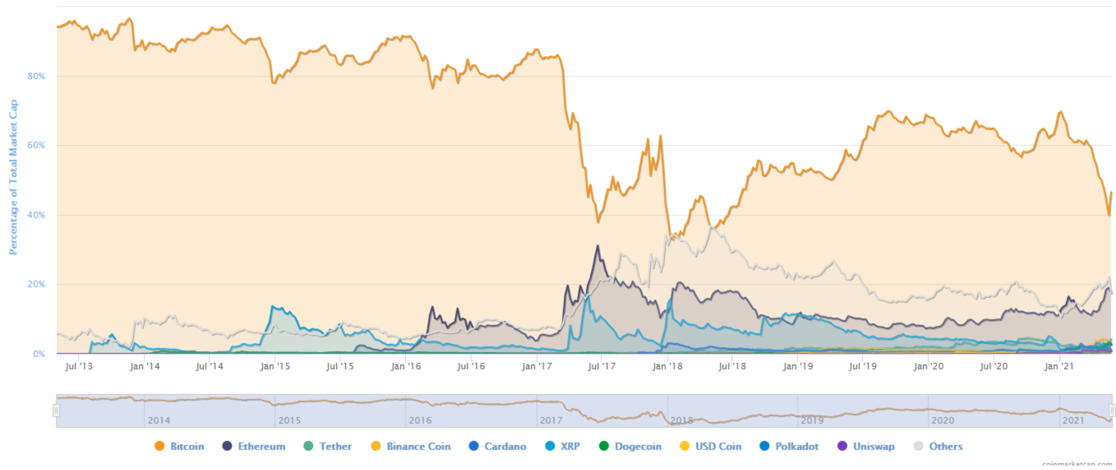

Figure 10.2: Market capitalisation of cryptocurrency

The descriptive statistics presented in Table 10.1 below further cement what the graphs showed. The minimum percentage returns observed are on Dogecoin with a minimum of -98%, followed by Ethereum and Bitcoin with -44.5% and -39.2%, respectively. These are all high compared to the JSE, with the minimum as -9.9% loss. The maximum returns followed a similar pattern; however, Dogecoin had a surprising 6191% daily increase. For the JSE, the minimum and the maximum percentage return are similar, as can be seen by a maximum return of 9.5%. The average returns are also higher on cryptocurrencies, with the highest percentage return observed in Bitcoin, with an average daily return of 5.4%. This is way above all other returns, as the second-highest was Ethereum with 0.33% and Bitcoin with 0.27%. The JSE is the least, with an average daily return of 0.03%. The accompanying standard deviations show that the higher returns are accompanied by high standard deviations, i.e., more volatile, implying more risk. The skewness and kurtosis suggest that the data is not normally distributed. Dogecoin deviated more from the other return series.

Table 10.1: Descriptive statistics of daily log-returns Bitcoin, Ethereum, Dogecoin and JSE

Statistic	Bitcoin	Ethereum	Dogecoin	JSE
nobs	1348	1348	1348	1348
Minimum	-39.1816	-44.5472	-98.4328	-9.9229
Maximum	22.5512	25.9572	6191.5110	9.4798
Mean	0.2674	0.3254	5.4303	0.0306
Median	0.1622	0.1400	-0.0369	0.0739
Variance	17.9959	28.8176	28615.5298	1.5220
Stdev	4.2422	5.3682	169.1613	1.2337
Skewness	-0.2777	-0.2886	36.2874	-0.3023
Kurtosis	7.8431	6.1876	1323.7077	10.1323

10.3 Methodology

10.3.1 Stationarity

For most time series models, we require stationary data. This is because stationarity guarantees consistency in data properties leading to more accurate estimates. We make use of the Augmented Dickey-Fuller test (ADF) by Dickey and Fuller (1981) and Phillips and Perron test (PP) of Phillips and Perron (1988) unit root tests, and the Zivot and Andrews test Zivot and Andrews (2002) for Structural Break of Stationarity to our variables.

10.3.2 Pearson correlation

The Pearson correlation is the most common method to measure the relationship between two variables in tandem. It measures the strength and direction of the relationship between two variables. The correlation coefficient is represented as follows:

$$\bar{\rho} = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i^n (x_i - \bar{x})^2}\sqrt{\sum_i^n (y_i - \bar{y})^2}}, \tag{10.3}$$

where x and y are the two variables whose relationship is to be measured with $\bar{x}$ and $\bar{y}$ being the means of x and y respectively.

10.3.3 Vector Autoregressive model (VAR)

The Vector Autoregressive model (VAR) is a multivariate model. It is one of the simplest multivariate models.
A VAR(p) is defined as follows,

$$Z_t = \phi_0 + \sum_i^p \phi_i Z_{t-i} + \epsilon_t, \tag{10.4}$$

where $Z_t = (Z_{1t}; ::::; Z_{kt}; ::::; Z_{Kt})$ is a set of K endogenous variables, ϕ_i are $(K \star K)$ matrices of coefficients and ϵ_t is a K-dimensional process with $E(\epsilon_t) = 0$ and time invariant positive definite covariance matrix $E(\epsilon_t \epsilon_t^T) = \sum_{\epsilon_t}$ (whitenoise).

The lag order for the VAR(p) model is selected using different Information criterion. These are:

$AIC(\ell) = \ln |\hat{\sum}_{a,\ell}| + \frac{2}{T}\ell k^2.$

$BIC(\ell) = \ln |\hat{\sum}_{a,\ell}| + \frac{\ln(T)}{T}\ell k^2.$

$HQ(\ell) = \ln|\hat{\sum}_{a,\ell}| + \frac{2\ln[\ln(T)]}{T}\ell k^2.$

where T is the sample size, $\hat{\sum}_{a,\ell}$ is the Maximum likelihood estimate of $\sum_a$ the covariance matrix.

AIC is the Akaike information criterion proposed in Akaike (1974), BIC is the Schwarz Bayesian Information Criterion of Schwarz et al. (1978), and HQ is the Hannan-Quinn criterion as proposed by Quinn (1980). The lag selected by the minimum of the criteria is used.
The advantage of the VAR model is that it can be used to investigate causality between time series variables. We have two types of causality, as stated by Granger (1969).

- Granger causality, where we say X Granger causes Y when the future values of Y are predicted more accurately if we incorporated information gained from past values of X.
- Instantaneous Granger causality, where X instantaneously Granger causes Y when the future values are predicted more accurately when we add future values of X to past values of X in the prediction.

We will use the Granger causality to investigate the presence or absence of any unidirectional or bi-directional causality between the cryptocurrencies and the JSE index.

10.3.4 ARCH effects

Before applying GARCH models, we test for heteroscedasticity in the return residuals using the Lagrange Multiplier (LM) test suggested by Engle (1982). The presence of ARCH effects justifies the use of GARCH models.

10.3.5 Dynamic conditional correlation models (DCC-GARCH)

Engle (2002) proposed a new model that allows for correlations to change over time. Given $\boldsymbol{\Sigma_t}$ =$[\sigma_{ij,t}]$ which is the volatility matrix of a_t given $F_{t.1}$, the information available at time $t-1$. Then the DCC model is defined as

$$\rho_t = D_t^{-1}\boldsymbol{\Sigma_t}D_t^{-1}. \qquad (10.5)$$

D_t: n $\times$ n, diagonal matrix of the time varying volatilities i.e. $diag\{\sigma_{11,t}^{1/2}....\sigma_{kk,t}^{1/2}\}$. The DCC model starts by modeling the volatility series $\{\sigma_{ii,t}\}$ using univariate GARCH models.

By letting $\gamma_t = (\gamma_1, \ . \ . \ . \ , \gamma_k)'$ be the vector of marginally standardized innovations with $\gamma_{it} = a_{it}\sqrt{\sigma_{ii,t}}$. Then, ρ_t is the conditional correlation matrix of the standardized disturbances γ_t.
The dynamic dependence of ρ_t are then modeled in two ways. The first type being the DCC by Engle which is defined as follows:

$$\rho_t = J_t Q_t J_t, \tag{10.6}$$
$$Q_t = (1-\theta_1-\theta_2)\bar{Q} + \theta_1 Q_{t-1} + \theta_2 \gamma_{t-1}, \tag{10.7}$$

where $\bar{Q}$ is the unconditional covariance matrix of the standardized innovation vector γ_t and θ_1 and θ_2 are non-negative real numbers satisfying the condition that $0 < \theta_1 + \theta_2 < 1$. The θ_1 and θ_2 govern the dynamic dependence in the equation above.

Tse and Tsui (2002) proposed the second type of DCC models which is defined as:

$$\rho_t = (1-\theta_1-\theta_2)\bar{\rho}_t + \theta_1\rho_{t-1} + \theta_2\psi_{t-1}, \tag{10.8}$$

where $\bar{\rho}_t$ is the unconditional correlation matrix of η_t. θ_i are non-negative real numbers satisfying additional constraint $0 < \theta_1 + \theta_2 < 1$, and $\psi t - 1$ is a n $\times$ n matrix whose elements are functions of the lagged observations of γ_t.
Tsay (2013) explains that both models start with the unconditional covariance matrix of γ_t. However, they differ in the way local information at time $t-1$ is used. The DCC model of Engle (2002) uses γ_t only so that Q_t must be re-normalized at each time index t. However, the DCC model of Tse and Tsui (2002) uses local correlations to update the conditional correlation matrices.

10.4 Empirical findings

10.4.1 Stationarity

Table 10.2 shows the results of the stationarity tests. We find that all variables are stationary at the 1% significance level, i.e. no unit root is present. Our variables are therefore integrated at I(0). Hence, we can carry out our models that require stationary data.

Table 10.2: Test of stationary.

Asset	Augmented Dickey-Fuller	Phillips Perron	Zivot Andrews
Bitcoin	-24.8279***	-38.9583***	-39.606***
Ethereum	-24.5003***	-39.7465***	-40.4337***
Dogecoin	-25.1035***	-39.1591***	-39.5977***
JSE	-26.0968***	-37.2346***	-37.3257***

10.4.2 Correlation

As seen in Table 10.3, Bitcoin and Ethereum have a strong positive correlation, which is significant at the 1% significance level. It shows that these two coins have linear dependence. The correlation between Dogecoin and each of Bitcoin and Ethereum is weak, positive, and insignificant. This can be explained by the difference in the nature of the coins. Ethereum and Bitcoin are utility coins, meaning they have a practical use in cryptocurrency. On the other hand, Dogecoin is a meme coin that was initially created as a joke before gaining popularity. The JSE has a weak negative correlation with all coins. This is not surprising as it is a collection of stocks and not a cryptocurrency. These results suggest a spillover effect between Bitcoin and Ethereum, where an increase in one may lead to an increase in the other. The spillovers can also be present within the other assets, but we have to use other statistical tests to confirm any spillovers.

	Bitcoin	Ethereum	Dogecoin	JSE
Bitcoin	1			
Ethereum	0.7476***	1		
Dogecoin	0.0216	0.0212	1	
JSE	-0.0145	-0.0364	-0.0025	1

Table 10.3: Pearson correlation matrix for the sample period Jul 2, 2016 to April 31, 2021

10.5 Causality using the VAR model

For all of the model specifications under consideration, the AIC (n), HQ (n), the SC (n), and the FPE (n) criterion all suggest an optimal lag of k = 1. We then fitted a 5-variable VAR model with a lag of 1. The results of the VAR then allowed us to study the dynamic interactions and the causalities among the assets using the Granger Causality. The results in Table 10.4 show contagion effects among the two top cryptocurrencies, Bitcoin and Ethereum. Bitcoin has an instantaneous cause on Ethereum but does not granger cause it. On the other hand, Ethereum has both instantaneous and Granger causality on Bitcoin. There is no contagion effect from Dogecoin and JSE on Bitcoin or Ethereum, and the reverse from Bitcoin or Ethereum to JSE or Dogecoin. The VAR model and its Granger causality have limitations in determining the dynamic dependency. Lütkepohl (2013) explains that the lack of a Granger-causal relationship from one group of variables to the remaining variables cannot necessarily be interpreted as a lack of a cause and effect relationship as a VAR representation characterises a joint distribution of sets of random variables. More assumptions are required for cause and effect relationships than those offered by the VAR model. Hence, we consider other ways to measure the spillovers.

Table 10.4: Granger causality test results

Model variables	Cause	Ftest	Granger cause	Chi square	Instantaneous cause
Bitcoin and JSE	Bitcoin	0.0906	No	0.3818	No
	JSE	0.3555	No	0.3818	No
Bitcoin Ethereum	Bitcoin	1.0917	No	481.09	Yes
	Ethereum	9.7554	Yes	481.09	Yes
Bitcoin and Dogecoin	Bitcoin	3.4312	No	0.8126	No
	Dogecoin	0.07332	No	0.8126	No

Table 10.5 shows the causality of each asset in the system, which includes all the variables: Bitcoin, Ethereum, JSE, and Dogecoin. The results show that despite most people saying Bitcoin causes the effect in all coins, evidence points to Ethereum being the one leading to the overall cause. Bitcoin has instantaneous cause on the system but does not Granger cause the system. On the other hand, Ethereum has both Granger and instantaneous causality effects. The JSE and Dogecoin do not pose any form of causality to the system. This result is not surprising given that the JSE is a South African market that is not influential in cryptocurrency. Dogecoin is also a less dominant coin whose price shifts do not have a significant effect in the cryptocurrency space.

Table 10.5: System Causality

Cause	Ftest	Granger cause	Chi square	Instantaneous cause
Bitcoin	0.94708	No	482.82	Yes
Ethereum	7.825	Yes	483.1	Yes
JSE	0.96775	No	2.876	No
Dogecoin	0.4008	No	1.3963	No

10.6 Dynamic conditional correlation model

The study employs a two-asset Dynamic Conditional Correlation model to investigate the existence of spillovers from Bitcoin to each asset separately. We make use of an autoregressive with one lag (AR1) to capture the autocorrelation of the residuals as suggested above by the AIC(n), HQ(n), SC(n) and the FPE(n). During the estimation of the univariate GARCH, we make use of GARCH (1,1) with the student t error distribution to capture the fat tails in the data. We run the DCC models by comparing the Bitcoin return series against another one at a time instead of all the return series simultaneously. This we do to limit biased estimates and also to avoid higher dimensions, which are tricky to base conclusions on.

Table 10.6 below shows the two asset DCC-GARCH results. The jointdcca measures the impact of past shocks on the current conditional correlation, i.e., it measures the short-run volatility spillover. The jointdccb measures the impact of past correlation on the current conditional correlation, i.e., it measures the long-run volatility spillovers. Based on the output, we see a significant positive short-run spillover from Bitcoin to Ethereum. The jointdccb is also significant. Hence, there is also a significant positive spillover from Bitcoin to Ethereum. The previous section, using the VAR model, had shown instantaneous causality. There is no evidence for short- and long-run volatility spillovers from Bitcoin to Dogecoin. This finding is consistent with the findings we made using the VAR model. For the Bitcoin and JSE models, the dcca and dccb are both insignificant; hence, there is no short-run and long-run spillover effect from Bitcoin into the JSE market. This finding is also consistent with the outcome from the VAR model. However, there are short and long-run volatility spillovers from Dogecoin to Ethereum.

Table 10.6: DCC model estimates

Bitcoin and Ethereum			Bitcoin and Dogecoin		
Variable	Estimate	Pr(> $\|t\|$)	Variable	Estimate	Pr(>$\|t\|$)
Bitcoin.mu	0.1570	0.0120	Bitcoin.mu	0.1570	0.0053
Bitcoin.ar1	-0.0734	0.0026	Bitcoin.ar1	-0.0734	0.0006
Bitcoin.omega	0.1204	0.4676	Bitcoin.omega	0.1204	0.4770
Bitcoin.alpha1	0.0736	<0.001	Bitcoin.alpha1	0.0736	<0.001
Bitcoin.beta1	0.9254	<0.001	Bitcoin.beta1	0.9254	<0.001
Bitcoin.shape	3.2857	<0.001	Bitcoin.shape	3.2857	<0.001
Ethereum.mu	0.1497	0.1143	Dogecoin.mu	-0.1170	0.0090
Ethereum.ar1	-0.1149	<0.001	Dogecoin.ar1	-0.1321	0.0021
Ethereum.omega	1.5932	0.0276	Dogecoin.omega	11.2585	0.5939
Ethereum.alpha1	0.1254	0.0007	Dogecoin.alpha1	0.7979	0.0000
Ethereum.beta1	0.8608	<0.001	Dogecoin.beta1	0.2011	0.2802
Ethereum.shape	3.0398	<0.001	Dogecoin.shape	2.6846	0.0180
Jointdcca1	0.0731	0.0005	Jointdcca1	0.0000	1.0000
Jointdccb1	0.8988	<0.001	Jointdccb1	0.9262	0.9864

Bitcoin and JSE			Dogecoin and JSE		
Bitcoin.mu	0.1570	0.0109	Dogecoin.mu	-0.1170	0.0812
Bitcoin.ar1	-0.0734	0.0025	Dogecoin.ar1	-0.1321	<0.001
Bitcoin.omega	0.1204	0.3979	Dogecoin.omega	11.2585	0.0457
Bitcoin.alpha1	0.0736	<0.001	Dogecoin.alpha1	0.7979	<0.001
Bitcoin.beta1	0.9254	<0.001	Dogecoin.beta1	0.2011	0.2437
Bitcoin.shape	3.2857	<0.001	Dogecoin.shape	2.6846	<0.001
JSE.mu	0.0584	0.0243	JSE.mu	0.0584	0.0242
JSE.ar1	0.0011	0.9687	JSE.ar1	0.0011	0.9687
JSE.omega	0.0400	0.0205	JSE.omega	0.0400	0.0204
JSE.alpha1	0.0889	0.0003	JSE.alpha1	0.0889	0.0003
JSE.beta1	0.8792	<0.001	JSE.beta1	0.8792	<0.001
JSE.shape	8.3653	<0.001	JSE.shape	8.3653	<0.001
Jointdcca1	0.0581	0.1316	Jointdcca1	0.0144	<0.001
Jointdccb1	0.2130	0.2854	Jointdccb1	0.8155	<0.001

The correlation plots of the assets are shown in Figure 10.3 below. We notice strong positive correlations between Ethereum and Bitcoin that range from 0.6 to 0.85 across the time horizon. This is not surprising as both are cryptocurrencies known to have utility in the cryptocurrency space; hence they tend to follow each other. The correlations between Bitcoin and Dogecoin are also positive throughout the sample period. However, these correlations are weekly positive, ranging between 0.01 and 0.06. The positive correlation can be explained by the fact that these are both cryptocurrencies. However, the weak nature can be explained by the fact that Dogecoin is less of a utility coin but instead is meme coin based on a Shiba Inu dog. Despite being a meme coin, Dogecoin received endorsement from prominent people like Elon Musk. Its affordability for low-income investors resulted in its rise in the cryptocurrency space. This growth can be seen by the increase in the correlation with Bitcoin after 2020. The correlations between Bitcoin and JSE returns range from a negative -0.1

to a positive 0.06 across the sample period. This result is not surprising as these two are in different financial asset classes. It shows that the behaviour of the cryptocurrency is different from that of the JSE one. The correlations between Ethereum and Dogecoin show that the correlations are similar to the ones observed between Dogecoin and Bitcoin. The similarity is in the graph shape and also the magnitude. This is not surprising as we have established a strong relationship between Bitcoin and Ethereum. However, the spillovers from Dogecoin to JSE are highly significant for both the short and long run, meaning that these two have financial contagion. This raises questions about the common but unverified claim that some cryptocurrency users refer to Dogecoin as the new cash while Bitcoin is the new Gold.

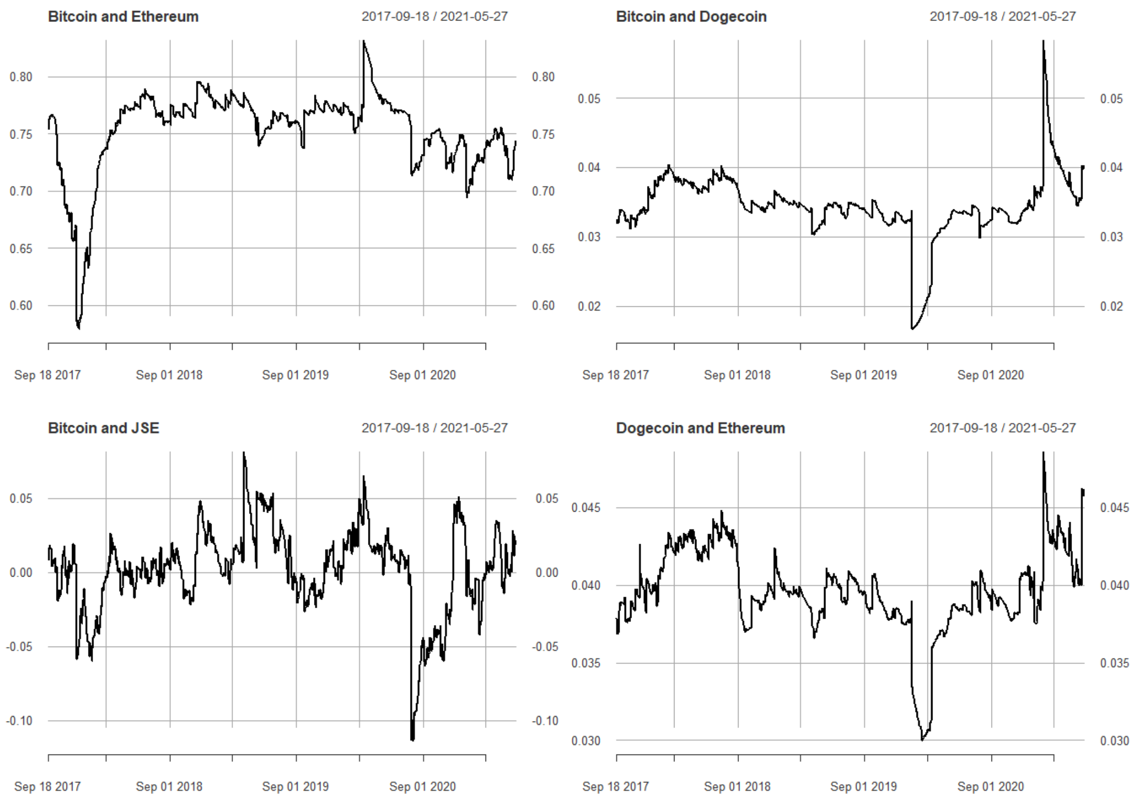

Figure 10.3: Dynamic Conditional Correlation plots from 2017-09-18 to 2021

10.7 *Further discussion and conclusion*

Whilst there have been studies on the spillover effects of Bitcoin and other cryptocurrencies and Bitcoin with European or American markets, there has been no focus on the South African market. This study considered the spillover between Bitcoin, Ethereum, and Dogecoin and added JSE/FTSE 40 return data. As expected, the findings showed that Bitcoin has a high positive correlation with Ethereum. This is because these are the top two cryptocurrencies by market capitalisation, and both have high utility in the crypto space. Our study showed that the two cryptocurrencies have similar dynamic correlations with both Dogecoin and the JSE. In terms of spillover effects using both the VAR model and the DCC-GARCH model, we find evidence of short-run spillover effects from Bitcoin to Ethereum. The VAR model, however, does not show long-run spillovers from Bitcoin to Ethereum. The DCC-GARCH shows there are long-run spillovers. Unlike the VAR model, which is time-invariant, the DCC-GARCH model takes into account dynamic correlations and volatility presence, making it a more useful model. We thus conclude that there is a financial contagion between the two cryptocurrencies. This finding is no surprise as the two have high correlations, as evident in the Pearson correlation and the dynamic correlation plots. Ideally, investors or portfolio managers seeking to diversify and reduce portfolio risks cannot use these two to offset the risk of the other. On the other hand, there are no short-run or long-run spillover effects between Bitcoin and JSE, as evident from both the VAR and DCC-GARCH models. This suggests that there is an opportunity for investors in South African stocks to use Bitcoin as a hedge. The result is similar to the findings by Dyhrberg (2016) except that the study considered FTSE stock rather than JSE stock data. Interestingly, there are no spillover effects from Bitcoin to Dogecoin in both the long and short run. This opens up an opportunity for those who want to create a portfolio with cryptocurrencies that offset each other's risk.

Overall, the study allows portfolio managers to look at cryptocurrency as a hedge in the South African market. Future studies can look at the spillover behaviour during different market times, such as during a financial crisis and when not in a financial crisis, to further see the possibility of cryptocurrency as a safe haven for investors in the South African market during times of financial crisis. One may also consider the market behaviour during and before the COVID-19 period.

10.8 Chapter summary

This chapter investigated the spillovers that occur amongst cryptocurrencies. It also investigated the spillovers between the cryptocurrencies and the JSE. The results were used to evaluate the possibilities of hedging. It also looked at possibilities of using cryptocurrency as a safe haven for South African market investors.

11. A DISCUSSION ON FACTORS AFFECTING THE GROWTH AND ADOPTION OF CRYPTOCURRENCY IN AFRICA AND THE WORLD

This chapter discusses factors that affect the mass adoption of cryptocurrency. There are no models used for this section except a review of different sources that discuss cryptocurrency growth and hindrances to growth.

11.1 Introduction

At the end of the twentieth century, a new type of currency known as cryptocurrency emerged. The first cryptocurrency was Bitcoin, introduced in 2008 in a concept paper by a faceless character named Satoshi (Nakamoto 2008). The paper introduced Bitcoin as a new payment system based on blockchain technology. The idea was to have a decentralised currency free from institutional control. This new currency would also remove the need for a third party to facilitate transactions and be trustless.

The first Bitcoin was minted in the subsequent year, with Satoshi as the first beneficiary. The initial days saw little to no demand for Bitcoin, with a few tech enthusiasts participating. Coins were given for free in universities, but due to their low value, most students exchanged the coins for meals at the cafeteria. One of the most famous sales is when a man sold 10 000 in exchange for pizza, marking one of the most iconic regrettable moves. Despite the slow start, Bitcoin began gaining traction, increasing in value. Other alternative coins began to emerge to compete with Bitcoin, with Ethereum eventually becoming the second most valuable coin by market capitalisation. Undoubtedly, there has been an increasing trend in the use of cryptocurrency based on metrics such as volume traded, total market value, number of users, and other measures. According to crypto.com, the global crypto population increased by 178% in 2021, rising from 106 million in January to 295 million in December. Regarding the adoption rate, there was a 37.5% growth in the second half of 2021 compared to a 33.3% growth at the same time in 2020. These figures suggest a year-on-year adoption growth. While these figures show positive growth, a lot more growth is required for cryptocurrency to be considered a competitor to the traditional fiat system. Hinzen et al. (2022) explains how Bitcoin is used by more people than fiat currencies of some developed economies. However, the adoption rate of Bitcoin is low in those countries compared to their fiat currencies, which have a near 100% adoption rate within these countries' borders. He illustrates this with New Zealand as an example, where New Zealand has a 5 million population, compared to over 30 million Bitcoin users. Despite this, New Zealand has an approximately 100% adoption rate in New Zealand, while Bitcoin, the purported world currency, has less than 1% in the world. From this example, cryptocurrency is in its infancy regarding mass adoption.

There have been several studies that have looked at cryptocurrency adoption drivers and hindrances. The studies have been conducted via surveys and some through analysing market data. We review some below.
Alzahrani and Daim (2019) investigates cryptocurrency user intentions, and the results indicate that users evaluate and make decisions primarily from economic and social perspectives. The top four criteria influencing cryptocurrency adoption are investment opportunity, subjective norms, business acceptance, privacy, and global attention. The study by Nseke et al. (2018) reveals that in Africa, the main factors for the use of cryptocurrency were the low transaction cost, low level of entry, and quickness of processing transactions. The study also shows drawbacks due to strong volatility, lack of user-friendliness, and its usage in crime. Using the theory of planned behaviour (TPB), Mazambani and Mutambara (2019) evaluate the lagging of South Africa in cryptocurrency adoption. The findings indicated that attitude and perceived behavioural control positively impacted the intention to adopt cryptocurrency. On the other hand, most deemed cryptocurrency too technical; hence, they were not keen on adopting it as a transaction mechanism. Similar findings are arrived at by Walton and Johnston (2018), who studied the South African market and found that perceived benefit, usefulness, ease of use, and trust-related risk indirectly affect the intention to use Bitcoin. The barriers to adoption consisted of the complex nature of Bitcoin and its high degree of volatility.

In another study by Saiedi et al. (2021), the results revealed that bitcoin adoption is driven by perceived failings of traditional financial institutions and systems. The study also shows that the unwillingness of merchants to accept cryptocurrency as a form of payment hinders adoption. Similarly, Connolly and Kick (2015) studied bitcoin adoption, and the findings show that organisation adoption is more important than consumer adoption because consumers cannot use cryptocurrency if organisations do not accept it as payment. Wood et al. (2017) examine the factor associated with adopting Bitcoin as a means of financial exchange. By employing structural equation modelling and the partial least squares method on data from a survey of 121 global cryptocurrency community members, they found that relative advantage and ease of use positively impacted the intention to use Bitcoin. The study also showed that compatibility had a significant positive impact.

Looking at the literature, most of the research focuses on the subjective views of the surveyed individuals. The research focuses more on the use of cryptocurrency and the factors driving its use. Most articles are based on developed countries and fewer on developing countries. Focusing on developing countries means the factors considered are generally universal factors and neglect issues of affordability, which are an issue in developing countries. This paper investigates and discusses issues that have seen cryptocurrency fail to make inroads in Africa and the world.

11.2 Methodology

This chapter is a discussion paper. The goal is to review different sources on the adoption of cryptocurrency. The data source will include published research articles, website publications, and general public comments on social media. We also consider known cases of structural failures and other issues that have hit the cryptocurrency space. Factors seen fit for discussion as determined by the frequency with which these reasons are mentioned in research and the news will be included.

11.3 Factors hindering cryptocurrency adoption

Cryptocurrency growth has been steady, with a more significant increase occurring in 2020-2021 during the COVID-19 pandemic. This growth may be because most stock markets were closed, so people wanted new investment areas.
The most significant factor that has hindered Africa's development can be linked to poverty. Africa is home to some of the world's poorest countries by economic measures such as GDP. It is also poverty that is one of the main reasons why cryptocurrency adoption is lagging in Africa. We will discuss some of the issues that arise from poverty.

11.3.1 Poverty

Cryptocurrency use is linked to massive infrastructure. For example, one needs high-end computers to earn cryptocurrency by participating in mining. One needs the internet, a smartphone, or a computer to log into an exchange. These requirements mean that the majority of Africans are already excluded. Kirkman et al. (2002) mention that lack of technological support obstructs digital technology adoption at the national level. This technological support comes at a cost that most African countries cannot afford.

11.3.1.1 lack of digital literacy

Another issue arising from poverty is the lack of digital literacy. *Digital Literacy in Africa* (2022) defines digital literacy as the ability to use information and communication technologies to find, evaluate, create, and communicate information, requiring both cognitive and technical skills. Africans are lagging behind the global average in digital literacy. This lagging behind is tied to most Africans' being from developing countries. Without any know-how to use the internet, it becomes impossible for one to be able to trade or use cryptocurrency as a payment form. The lack of digital literacy goes hand in hand with the lack of internet access and the devices themselves. *Digital Literacy in Africa* (2022) states that digital literacy is no longer an option but has become a fundamental requirement for people's daily lives.

11.3.1.2 Internet access

Linked to digital literacy is Internet access. According to the internet usage statistics by (*World Internet Users Statistics and 2022 World Population Stats* n.d.), Africa has the world's worst internet penetration, way less than the global average. In the first quarter of 2020, the internet penetration in Africa was 39.3%, while the world average was 58.8%. This figure would improve in quarter 1 of 2021 to 42.3% for Africa and 65.3% for the world average. This represents a 3% for Africa against a 7.5% increase on the world average. As per the first quarter of 2022, African usage was 43.1% compared to the world average of 66.2%, an increase of 0.8% and 0.9%, respectively. This means that the core requirement to use cryptocurrency is already not accessible to over half of the African population. Furthermore, some of those who have access to it have limited access due to costs.

11.3.1.3 The cost of data in Africa

Africa has some of the world's most expensive data. Accompanied by the low income in Africa, it means that most people can not afford it. With the high data cost, any technology based on internet access will generally not spread as fast in remote African areas as it would in other regions of developed countries.

11.3.1.4 High rural population

In most of Africa, infrastructure such as electricity and the internet is mainly available in urban areas, with rural areas lacking basic services. Electricity and water availability are already issues, with the internet being worse off. According to *Rural population, percent in Africa* (n.d.) the rural population in Africa in 2021 was 52.13%. The highest value was in Burundi: 85.94% and the lowest was in Gabon: 9.58%. Unlike some developed countries, African rural areas tend to suffer from a lack of development.

11.3.2 Fear of abuse by criminals

One of the major concerns raised by governments and other national regulators is that cryptocurrency can be used as a way of transferring illegal proceeds made by criminals. The claim is that one can easily buy into cryptocurrency and siphon money out of countries with the responsible reserve banks having no control.

11.3.3 Decentralisation of cryptocurrencies

One of the factors used by cryptocurrency promoters is that it is decentralised. Decentralisation means that there is no institutional control over the asset. However, there is an argument that this may not exactly be the case. In cryptocurrency, decentralisation is established in the fact that for the creation or mining of a block, many participants are involved, and this is done in such a way that there is no way that an individual can influence the operations. For example, for Bitcoin, which uses a proof of works concept, the amount of computational power required to verify blocks is achieved by different miners pooling together resources to form a mining pool. Anyone across the world can be part of the mining of the blocks, hence decentralising it.

11.3.4 Too many cryptocurrencies

The number of cryptocurrencies available is on a constantly rising trend. The number of coins listed on CoinMarketCap (n.d.) as of June 2022 was more than 20 000. All the coins come promising to offer a different aspect. The high number of different coins makes it hard for users, especially new ones, to know which coins to buy and use.

11.3.5 Limited usability

To perform a transaction, one must have a public and private key. This means that a person requiring someone to perform a transaction on their behalf must give them their key, which can risk their privacy and lead to an unsolicited transaction.

11.3.6 Environmental concerns

Cryptocurrency mining requires enormous amounts of energy. According to Rauchs et al. (2021) Bitcoin uses on average 84.87(lower limit 38.84- the upper limit of 139.50) terawatts of power per annum. This figure is more than most of the world's countries use per annum. The production of the energy required for mining also means extra atmospheric pollution. According to Stoll et al. (2019), Bitcoin mining emits 22.0 to 22.9 megatons of carbon dioxide into the atmosphere annually, making cryptocurrency a significant contributor to global air pollution and climate change.

11.3.7 High volatility

In general, financial assets are known to be volatile. This volatility is such a defined phenomenon that it has been developed into a traded index. The volatility of cryptocurrencies is even higher than that of stock market returns. Cryptocurrency is punctuated with sudden price increases and decreases that, to some extent, can make investors super wealthy or suddenly lose most of their investments. The volatility is mentioned in academic papers such as that of Kaseke et al. (2021)

11.3.8 Too much supply

One of the characteristics of money is scarcity. In cryptocurrency, the number of coins is usually stated before coins are minted. A few coins, such as Dogecoin, have an unlimited supply, meaning they will continue to be minted infinitely at a given rate. This tends to weaken the coin as it goes against the scarcity principle.

On the other hand, there are coins such as Bitcoin, which have a low supply compared to others. This low supply creates value for the coin. For example, bitcoin has a supply of 21 million coins, and Ethereum has a total supply of 121,421,087. These low numbers create demand, hence a push in the prices. Such coins can be attractive to investors. Graf (2014) writes that "Bitcoin has brought authentic rival scarcity into the realm of digital goods. This is not the artificially imposed, legally constructed scarcity of intellectual property legislation. The Bitcoin protocol has set up a type of scarcity that is inherent to and inseparable from the nature of the digital good"
This property has not been achieved by all coins, especially the recent ones where some creators have made coins with supplies in the quadrillions. The coins will start cheap and hence affordable. The problem is that they sell dreams to the public. Calculations show that if the price reaches a dollar for coins with a billion supply, then the market cap for that coin will be 1 billion; likewise, for

those with a trillion supply, the market cap will be a trillion. If the coins with a 10 trillion supply reach $0.1, the total market cap for that coin is 1 trillion United States dollars. This value is unrealistic given that the total cryptocurrency market cap ranges from 1.7 to 2 trillion dollars, with an all-time high of over 3 trillion dollars in November 2021. The creators of coins with trillions and quadrillions of coins are selling dreams to investors by telling them the coins will grow to a cent or a dollar. The founders therefore are more likely to be selling dreams that do not exist. Because of this, some creators burn their coin supply to allow price increases by creating scarcity. However, not all creators do so. An example of burnt tokens is Shiba Inu, which has burned and saw a steady gain in token value. In some cases, even 50% may not impact coins with a high supply. For example, coins in the trillions and quadrillions burning 50% would still leave the supply in the trillions and quadrillions.

11.3.9 Existence of whales

According to Academy (n.d.) a whale is an individual or organisation that holds a large amount of a particular cryptocurrency. A whale may also be defined as a person with enough coins or tokens to cause a significant impact on the market prices, either by buying or selling large amounts. Such holders tend to have enough power to move markets. For most coins, a few individuals own the majority of the coins. This means that, instead of community control, those individuals are essentially the guardians of the coins. Figures 11.1 and 11.2 show the charts for the ownership of Dogecoin and Bitcoin, respectively, as obtained from *Bitcoin, Ethereum, Dogecoin, Litecoin stats* (n.d.). For Dogecoin, one address owns 21.56% of the coins in circulation. One hundred thirty-seven accounts hold 70.67% of the total supply. For Bitcoin, five addresses own 4.05%. Two thousand ninety-five addresses hold 41.6% of the total. While this is better than the Dogecoin case, big moves by these holders can still manipulate the coins. However, it states that most big holders often avoid conventional trading and, instead, they buy and sell coins off the exchange books, in what is known as Over the Counter (OTC) trading. In the case of Dogecoin, Elon Musk, affectionately known as the Doge Father for his influence in the Doge community, raised concern about how these large holdings could hinder the growth of Dogecoin. He offered to exchange the Dogecoins for real money so that he could put the Dogecoins up for sale in the market. Despite the plea, the whales did not sell their holdings.

Dogecoin distribution

Balance, DOGE	Addresses	% Addresses (Total)	Coins	USD	% Coins (Total)
(0 - 0.1)	459776	10.63% (100%)	6,718 DOGE	$415.09	0% (100%)
[0.1 - 1)	153130	3.54% (89.37%)	59,486 DOGE	$3,676	0% (100%)
[1 - 10)	1122918	25.95% (85.84%)	3,694,563 DOGE	$228,293	0% (100%)
[10 - 100)	816451	18.87% (59.88%)	35,171,249 DOGE	$2,173,286	0.03% (100%)
[100 - 1,000)	1024742	23.68% (41.01%)	376,059,338 DOGE	$23,237,293	0.28% (99.97%)
[1,000 - 10,000)	538229	12.44% (17.33%)	1,671,851,517 DOGE	$103,306,315	1.24% (99.69%)
[10,000 - 100,000)	176896	4.09% (4.89%)	4,923,531,503 DOGE	$304,232,697	3.64% (98.46%)
[100,000 - 1,000,000)	30349	0.7% (0.8%)	8,135,807,318 DOGE	$502,724,233	6.02% (94.82%)
[1,000,000 - 10,000,000)	3757	0.09% (0.1%)	9,500,659,103 DOGE	$587,060,555	7.03% (88.8%)
[10,000,000 - 100,000,000)	565	0.01% (0.02%)	15,012,161,340 DOGE	$927,624,881	11.1% (81.78%)
[100,000,000 - 1,000,000,000)	120	0% (0%)	25,073,998,825 DOGE	$1,549,361,524	18.54% (70.67%)
[1,000,000,000 - 10,000,000,000)	16	0% (0%)	41,344,359,341 DOGE	$2,554,732,497	30.57% (52.13%)
[10,000,000,000 - 100,000,000,000)	1	0% (0%)	29,161,944,038 DOGE	$1,801,962,040	21.56% (21.56%)

History

Figure 11.1: Token ownership of Dogecoin as of July 2022 :Image source bitinfocharts.com

Bitcoin distribution

Balance, BTC	Addresses	% Addresses (Total)	Coins	USD	% Coins (Total)
(0 - 0.00001)	3258222	7.61% (100%)	15.57 BTC	$311,715	0% (100%)
[0.00001 - 0.0001)	7942807	18.56% (92.39%)	343.56 BTC	$6,876,465	0% (100%)
[0.0001 - 0.001)	10543378	24.64% (73.82%)	4,074 BTC	$81,539,735	0.02% (100%)
[0.001 - 0.01)	10567836	24.7% (49.18%)	40,109 BTC	$802,780,664	0.21% (99.98%)
[0.01 - 0.1)	6773560	15.83% (24.49%)	221,663 BTC	$4,436,604,560	1.16% (99.77%)
[0.1 - 1)	2820178	6.59% (8.66%)	873,570 BTC	$17,484,544,164	4.58% (98.61%)
[1 - 10)	734195	1.72% (2.06%)	1,851,608 BTC	$37,060,025,970	9.7% (94.03%)
[10 - 100)	133476	0.31% (0.35%)	4,300,031 BTC	$86,065,348,042	22.53% (84.33%)
[100 - 1,000)	13684	0.03% (0.04%)	3,854,418 BTC	$77,146,374,734	20.19% (61.8%)
[1,000 - 10,000)	2101	0% (0.01%)	5,060,841 BTC	$101,293,000,263	26.51% (41.61%)
[10,000 - 100,000)	89	0% (0%)	2,102,386 BTC	$42,079,359,481	11.01% (15.1%)
[100,000 - 1,000,000)	5	0% (0%)	779,548 BTC	$15,602,692,586	4.08% (4.08%)

History

Figure 11.2: Token ownership of Bitcoin as of July 2022 :Image source bitinfocharts.com

11.3.10 Rug pull scams

One of the disturbing occurrences in cryptocurrency of late is the issue of scams. People have been duping people with their hard-earned investments, only for projects to be dropped midway. The most common scam is what is known as a "rug pull".

According to coinmarketcap.com, (CoinMarketCap n.d.) a rug pull is a malicious manoeuvre in the cryptocurrency industry whereby crypto developers abandon a project and run away with investors' funds. This move is typical in the decentralised finance (Defi) ecosystem, which refers to financial applications utilising blockchain technology to allow digital transactions between multiple parties, such as borrowing and lending. These scammers usually target the decentralised exchanges (DEXs) as they allow users to list tokens for free without any audit. This is not the case with centralised exchanges, which audit the project before listing the token. The scammers list the tokens on DEX and use leading tokens such as Ethereum as trading pairs. They then promote the token using various social media platforms, such as YouTube and Telegram. The tokens are listed on the DEX and paired with a leading cryptocurrency like Ethereum. They then promote the token using various methods, such as creating Telegram groups and social media accounts, especially on Twitter. Once a significant number of investors invest in the token, the creators withdraw everything from the liquidity pool, driving the coin's price to zero.

Austin (2022) states three types of rug pull exist. These are liquidity theft, limiting selling orders and dumping.

In liquidity theft, token developers withdraw all the coins from the liquidity pool. The liquidity pool is a collection of the cryptocurrency backing assets secured by a smart contract for use in exchanges, loans, and other trades. The liquidity pool, in essence, is the backing of the token's value. Removal of this pool means that the value the investors had put in is gone, driving the token price to zero. The unsuspecting investors lose while the creators make vast amounts of money.

Under limit selling orders, fraudulent orders are put on investors to deceive them. The developers code the tokens so only they can sell them, forcing investors to be buyers at prices they set. Once the developers have reached their price target, they sell their holdings, obtaining the paired currency. The sudden sell drives the token price down, leaving the investors with worthless coins. An excellent example of a limiting selling order is the Squid Token scam of 2021. The creators took advantage of the hype of the Netflix show called the Squid Game and created a new token called the Squid. The token started selling at $0.01, and in less than a week, the price soared to over $2000. As the price soared, investors realised they could not sell their tokens. The creators then sold their holdings, and the price plummeted by more than 99% leading to enormous losses for the investors.

Dumping, also known as a pump-and-dump scheme, is a type of fraud characterised by the offenders hoarding a particular token over time and then causing an artificial increase in price through spreading misinformation (Pumping) before selling off their stock to unsuspecting buyers at the inflated price. Due to the dumping, the price will fall, leaving the late buyers to incur a loss. Forms of pumping include spreading lies about prospective projects linked to a particular coin using influential accounts on social media, especially on Twitter, where the main account of the coin spreads misinformation. One scam known in the cryptocurrency world is the telegram and discord groups where the offenders advertise that they are trying to beat whales at their game by joining in numbers to pump a coin they will announce at a particular time. The promoters ask members to invite more people and promise them huge returns. Once the time hits, the group administrators post the coin that everyone must buy so that the price goes up. The unsuspecting individuals all rush to buy, causing a surge in the price, and within minutes the offenders dump their lot, leaving the victims nursing their wounds as the price plummets.

11.3.11 Government effect

As stated before, one of the causes of Bitcoin's value decline is the policies of governments. One country that has shaken the cryptocurrency market like no other is China. Its position may explain this as the fastest growing economy, and its population size affects the overall market. *Coindesk* (n.d.) report cites the first ban in China occurred in 2013 when the government banned banks from handling cryptocurrency. This call was emphasised in 2021, when the cryptocurrency was starting to boom. The government also banned the mining of Bitcoin, shutting down mining rigs. The latest ban, at the time of writing, was on the 24th of September 2021, when the Chinese central bank declared all transactions involving Bitcoin and other virtual currencies illegal and moved to block any use of the virtual currencies. This announcement saw the price of Bitcoin fall by around 9% to around $40,000. At the same time, Ethereum fell from around $3100 to around 2800. This price drop spread across all the cryptocurrencies, losing significant value hours after the campaign to block the use of unofficial digital money. The reason is that the government says the state financial system may be exposed indirectly to the risks in the cryptocurrency market. However, there are some positive notations in Elsavado, where the country officially announced bitcoin as legal tender. The Elsavado government offered citizens a $30 incentive for citizens who downloaded and installed the national bitcoin wallet app called chivo. Despite the government endorsements, most people remain reluctant to invest in Bitcoin, with most citing the issue of complexity and fear of the volatile nature of cryptocurrency. *CBC News* (2021) reported that three face-to-face public opinion surveys performed in 2020 showed that most Salvadorans disagreed with the government's decision to make Bitcoin a legal currency.

11.3.12 Lack of regulation

The lack of regulation is linked to the government's effect. Because most local financial institutions and governments have not regulated cryptocurrency, most people fear participating. It is human nature that we prefer regulation as it gives a sense of safety and trust. In Africa, there have been mixed reactions to cryptocurrency. Jazeera (2022) reports that the Central African Republic has adopted bitcoin as its official currency, with the government passing a bill governing the use of cryptocurrency. Countries such as Zimbabwe and Nigeria, on the other hand, have had their central banks prohibit local banks from dealing with cryptocurrencies.

11.3.13 Influential individuals

Like in any investment area, some individuals command a lot of respect and followership, so their opinions significantly impact the markets. For example, Elon Musk, the world's richest man as of July 2022, according to *Real time billionaires* (n.d.) can move the markets via his tweets. Researchers have studied this phenomenon as Huynh (2022) who studied 10,850 tweets by Musk and found that the tone of the world's wealthiest person can drive the Bitcoin market, having a Granger causal relation with returns. However, they did not find evidence that the tweets caused Bitcoin volatility. The same conclusion is reached by Ante (2021) who, using an event study approach on a sample of 47 cryptocurrency-related Twitter events, identified significant positive abnormal returns and trading volume following such events. This study also showed that, on average, price effects are only significant for Dogecoin-related Tweets but not for Bitcoin.

11.3.14 Limited liquidity

In its simplest form, liquidity is defined as the ease of converting cryptocurrency into cash and whether this can be achieved without a negative impact on the cryptocurrency price. While the major cryptocurrencies such as Bitcoin and Ethereum are liquid, there have been some problems with the new coins. Some users will offer to sell their holdings at the current reflective price or lower, but they will still not get takers. Users would prefer a currency they can easily trade for cash or use for exchange instantly with no liquidity issues.

11.3.15 Shilling

Shilling is defined as the active promotion or endorsement of a particular cryptocurrency to generate hype and increase the number of investors. The more people invest, the more the demand for the coin increases, causing the price to increase. The problem with shilling, however, is that some people are only in it to lure unsuspecting individuals into investing, and once the price is high, the huge holders will dump their coins, leading to a price fall. The common methods of shilling include using influential individuals to promote a coin. Most people do so without research on the coin or the founders but do it for the money. Another way of shilling is through promising social media users coins if they change their profile pictures and headers to promote a particular coin. The danger is that the victims of scam shilling lose their faith in cryptocurrency and label it as a scam.

11.3.16 Hacking

Another reason used by cryptocurrency promoters in the early days was that cryptocurrency is hackproof. This meant that no one could hack into the system to manipulate the code for personal gain or to destroy the networks altogether. A look at the history shows that this is not the case, as there have been cases of hacking and stealing funds from exchanges and directly on the networks. Cryptocurrencies like Bitcoin use ledgers that have multiple copies across different miners. For one to hack, they must have access to over half of those in what is termed the 51% attack. This shows that, in theory, it is possible, although improbable in reality. Exchanges, however, can be hacked, as was the case with Binance, where hackers stole over $40 million in bitcoins in 2019, according to *Coindesk* (n.d.).

11.4 Discussion

Throughout the reviewed literature and the information from published news from news sites and blogs, there is a general agreement that issues are hindering mass adoption. The most cited reason affecting adoption is the high volatility nature of cryptocurrency. Most people liken investing in cryptocurrencies to gambling. While inflation affects the fiat currency, the price volatility rate in cryptocurrency leaves much to be desired. This high volatility also explains the equating of Bitcoin to Gold. The price of Gold is not as volatile and has managed to be on an upward trend for decades. However, the high volatility of cryptocurrency also comes with opportunities for high rewards. For Africa, the main hindrance has been seen to be poverty. Most Africans are battling to feed their citizens and do not consider technological infrastructure a priority. This means that people do not have internet access and therefore lag in the main requirement to use cryptocurrency. The other major factor is the lack of knowledge about cryptocurrency. Fiat currency is tangible and is also known to be backed by Gold and issued by the state. These factors give consumers a form of trust in the currency, unlike cryptocurrency, which has no regulation and is backed by cryptography that a layman in the street does not easily understand. The main reason is the need for access to the internet and technology for transactions compared to traditional fiat currency. This reason alone disqualifies most of the African population and the world at large. The high volatility and the random attacks by hackers also pose a massive threat to the general public, which is not keen on trading but keeps their coins as they would with fiat. The lack of regulation leaves the public prone to rug puling with no consequences and no hope of any recovery. As for investors, cryptocurrency trading strategies based on the fundamentals of trading and technical analysis may fail. This is because different factors can move cryptocurrency. One of the possible strategies to survive in this market is to follow social media sentiments. This method is supported by López-Cabarcos et al. (2021) whose study results suggest that Bitcoin could act as a safe haven and that Bitcoin investors could consider sentiment about the stock market from social networks rather than market volatility when designing their investment strategies. This finding suggests that Bitcoin investors are more "technological" and therefore pay more attention to the information from these media. Domjan et al. (2021) states that in the world and especially in developing countries, blockchain technology can solve trust issues and provide affordable transfers.

Conclusion

In conclusion, cryptocurrency is indeed a new and novel means of payment and exchange. However, we do not see it as a mainstream currency accepted by the majority soon. The high volatility nature of cryptocurrency leaves a lot to be desired. It takes away the peace that comes with the fiat currency of most countries. While fiat currencies can be affected by inflation, they do so gradually, whereas cryptocurrency can lose more than half their value overnight. The lack of public education hinders growth and fast adoption because the general populace does not understand cryptocurrency and its workings. Keegan (n.d.) argues that it is not if global adoption of cryptocurrency will happen, but when it will happen. This assertion may seem logical given the current trend in growth. However, given the evidence of constant scams, and high volatility swings, the author believes the "when" is not anytime soon unless there is a proper clean-up in cryptocurrency. Otherwise, traditional money is here to stay, at least for the foreseeable future. This study is of significance as it highlights the issues that the promoters of cryptocurrency should look at if it is going to be an alternative to the current financial system. While cryptocurrency seeks to avoid institutional control, responsible governments have a duty to protect citizens from being taken advantage of. Therefore, this study also highlights the issues that governments and financial institutions can require to be addressed before allowing the unrestricted use of cryptocurrency. The study's limitations include the fact that it was based on a review of published articles and opinion pieces. We gathered the reasons from social interactions and our views. We did not perform any statistical tests. Therefore, researchers can conduct surveys for future studies by administering questionnaires to people. The findings can then be derived from the surveys by analysing the responses qualitatively and quantitatively. This will help to reduce subjective bias.

11.5 Chapter summary

The chapter focused on issues pertaining to cryptocurrency adoption. It gave a list of areas that the responsible authorities can look at to enhance the adoption of cryptocurrency by the general population.

12. DISCUSSION AND CONCLUSION

The arrival of cryptocurrency created an alternative market for investors. However, cryptocurrency studies are still limited compared to other financial assets such as stock returns and exchange rates. The studies are even fewer when it comes to the context of examining cryptocurrency in the context of developing economies. While the world has become a global village, it still stands that studies carried out on developed markets may not hold in developing markets. This is because these two markets are at different stages. For example, when it comes to the market efficiency of Fama (1998), developed markets tend to be more efficient, with any new information absorbed quickly into the price. On the other hand, developing markets tend to be inefficient, opening up opportunities for technically oriented investors to take advantage of the inefficiency in setting up their trading strategies.

In current literature, most studies focus on modelling cryptocurrency using the well-known financial models mainly used for stock returns without any adjustments. While these models work, they are generally made to capture the stylised facts of the stock return data. Cryptocurrency is classified as financial data, but within financial data, the stylised facts vary in magnitude and frequency. With this in mind, we believe care has to be taken when modelling cryptocurrency.

In this study, we used five data sets: the daily log returns of Bitcoin, Ethereum, Dogecoin, FTSE/JSE 40, and Gold. Gold and Dogecoin were only included in one article each, while the rest were in every article. The first focus of this study was the investigation of the stylised facts of cryptocurrency, which we presented along with those observed in the JSE stock market. The results revealed that the stylised facts known to be present in financial data are also present in cryptocurrency. However, the magnitude and frequency of these stylised facts differ. The major finding regarding stylised facts was that cryptocurrency is more volatile than stock markets. This finding agrees with many studies, including those of Mariana et al. (2021). Another finding was the presence of the inverse leverage effect in cryptocurrency. This is supported by the findings of Zhang et al. (2021), Huang et al. (2022) and Ardia, Bluteau and Rüede (2019). Armed with the knowledge of the stylised facts, the study moved on to modelling the data.
One of the main stylised facts was the non-normality distribution of the data.

Fitting the GARCH models with normal distributions failed to capture the heavy-tailed nature of the innovations. This result led to the creation of the hybrid GARCH models. These were made by combining the ordinary GARCH models with heavy-tailed distributions. The distributions used were the GEVD and the GPD. The results showed that the GPD improved the GARCH models and the VaR estimates. The next chapter compared the volatility nature of the cryptocurrencies with the FTSE/JSE40 index returns. Structural breaks were accounted for using dummy variables in the GARCH-type models to further the comparisons. The results showed that the cryptocurrency market is inefficient and could be easily manipulated by alert investors. The study also showed that not accounting for structural breaks leads to overestimating volatility persistence.

Guided by the structural breaks effect, we introduced the Markov switching GARCH models to examine if having multiple regimes can improve the GARCH model results. The results showed that for cryptocurrency two regime models outperformed single regime models. A finding consistent with those of other studies such as by Ardia, Bluteau and Rüede (2019), Figà-Talamanca et al. (2021) and Maciel (2021). The regimes were characterised by their different volatility and persistent levels. The first regime we classified as the "calm market conditions" characterised by low volatility, high persistence, and a weak reaction to past negative returns. Regime two we classified as "turbulent market conditions", characterised by high volatility, lower persistence and a strong reaction to negative returns. The same classification is reached for regimes one and two by Chkili (2021). Further interrogation using Markov switching GARCH models revealed that the leverage effect might vary across regimes. This result we also observed using GARCH models that accounted for structural changes. The JSE, on the other hand, had a normal leverage effect, as also found by Muguto and Muzindutsi (2022). These findings are important to investors as they not only know that there is an inverse leverage effect, but also, if they formulate trading strategies, they must take note of the prevailing market conditions.

From the GARCH-type models and also the Markov switching models, we observed that cryptocurrency volatility is highly persistent. The same is observed in the JSE, but more for cryptocurrency. This finding points out the market inefficiency of these markets. The cryptocurrency returns also had significant serial autocorrelation, further cementing the finding of inefficiency in the market. The JSE did not exhibit significant serial correlations, suggesting market efficiency, a finding also made by Muguto and Muzindutsi (2022). The inefficiency of cryptocurrency means investors can take advantage and set up trading strategies. Another finding on the regime switching model was that in the case of persistence, the high volatility regime is less persistent than the low volatility regime. This means investors should account for this before taking up their positions, especially those who do options trading.

Regarding volatility spillovers, there was evidence of unidirectional and bidirectional causality in some assets. Using the VAR and the DCC model, there was evidence of financial contagion between Bitcoin and Ethereum. There was no evidence of financial contagion between Bitcoin and Dogecoin using both models. This result is unsurprising as the correlation between the two assets was very weak at 0.026. Compared to the JSE, both the VAR and DCC models did not show the presence of any market contagion. This result was not surprising as the JSE market is very small, yet the world market determines cryptocurrency prices. The lack of financial contagion allows JSE investors to use cryptocurrency, especially Bitcoin and Ethereum, as a hedge. Overall, this study has provided new insights into the stylised facts of cryptocurrency and how they compare to stock returns, the most studied financial data.The study has also helped find the ideal models to use for cryptocurrency. These models also help make better Value at Risk estimates. These findings mean that investors are better positioned when making risk assessments for their investment strategies. Alert investors can manipulate the markets for personal gains based on the inefficiency of the cryptocurrency markets. On the other hand, policymakers, governments, and financial bodies may see this as a call for regulation in the cryptocurrency markets.

The limitations of this study were mainly twofold. The study used three cryptocurrencies to represent a market with thousands of coins. This may mean that our results will likely not apply to the market as a whole. The second limitation is that we wanted to compare with developing markets to get findings not based on the developed markets, but the study used only the JSE market. The result may not necessarily hold for other developing markets. However, we believe these results are better for use by other developing markets than using developed markets' results. Another limitation is that the study did not exhaust all GARCH-type and multivariate models that capture spillovers. However, there are many models; therefore, using all of them may be unrealistic, although adding more would be adventagous.

For future studies, more models can be used in the comparisons. IGARCH, GARCH-M, and TGARCH can be used for univariate volatility models, while CCC GARCH, GO GARCH, and EMWA can be used for multivariate cases. Another area to consider is the use of machine learning. Models such as the Support Vector Machine (SVM), a supervised machine learning algorithm, can be applied to the GARCH model. Artificial Neural Network (ANN) models for deep modelling can also be used for our data. The ANN takes in the data, processes it in multiple layers, and improves with each run. These machine learning models have an advantage as they can handle high-dimensional time series data and work with non-linearity.

BIBLIOGRAPHY

Academy, B. (n.d.), 'Whale'.
URL: https://academy.binance.com/en/glossary/whale

Agbeyegbe, T. D. (1994), 'Some stylised facts about the jamaica stock market', *Social and Economic Studies* pp. 143–156.

Akaike, H. (1974), A new look at the statistical model identification, *in* 'Selected Papers of Hirotugu Akaike', Springer, pp. 215–222.

Akyildirim, E., Corbet, S., Lucey, B., Sensoy, A. and Yarovaya, L. (2020), 'The relationship between implied volatility and cryptocurrency returns', *Finance Research Letters* **33**, 101212.

Al-Yahyaee, K. H., Mensi, W., Ko, H.-U., Yoon, S.-M. and Kang, S. H. (2020), 'Why cryptocurrency markets are inefficient: The impact of liquidity and volatility', *The North American Journal of Economics and Finance* **52**, 101168.

Alexander, S. S. (1961), 'Price movements in speculative markets: Trends or random walks', *Industrial Management Review (pre-1986)* **2**(2), 7.

Alzahrani, S. and Daim, T. U. (2019), Evaluation of the cryptocurrency adoption decision using hierarchical decision modeling (hdm), *in* '2019 Portland International Conference on Management of Engineering and Technology (PICMET)', IEEE, pp. 1–7.

Ante, L. (2021), 'How elon musk's twitter activity moves cryptocurrency markets', *Available at SSRN 3778844* .

Ardia, D., Bluteau, K., Boudt, K., Catania, L. and Trottier, D.-A. (2019), 'Markov-switching garch models in r: The msgarch package', *Journal of Statistical Software, Articles* **91**(4), 1–38.
URL: https://www.jstatsoft.org/v091/i04

Ardia, D., Bluteau, K. and Rüede, M. (2019), 'Regime changes in bitcoin garch volatility dynamics', *Finance Research Letters* **29**, 266–271.

Arias-Oliva, M., Pelegrín-Borondo, J. and Matías-Clavero, G. (2019), 'Variables influencing cryptocurrency use: a technology acceptance model in spain', *Frontiers in Psychology* **10**, 475.

Arouri, M. E. H., Lahiani, A. and Nguyen, D. K. (2011), 'Return and volatility transmission between world oil prices and stock markets of the gcc countries', *Economic Modelling* **28**(4), 1815–1825.

Auger, I. E. and Lawrence, C. E. (1989), 'Algorithms for the optimal identification of segment neighborhoods', *Bulletin of mathematical biology* **51**(1), 39–54.

Austin, C. (2022), 'Defi: Top 10 list of crypto rug pulls'.
URL: https://coinscreed.com/

Baur, D. G. and Dimpfl, T. (2018), 'Asymmetric volatility in cryptocurrencies', *Economics Letters* **173**, 148–151.

Baur, D. G., Hong, K. and Lee, A. D. (2018), 'Bitcoin: Medium of exchange or speculative assets?', *Journal of International Financial Markets, Institutions and Money* **54**, 177–189.

Bauwens, L., Hafner, C. M. and Laurent, S. (2012), *Handbook of volatility models and their applications*, Vol. 3, John Wiley and Sons.

Bitcoin (n.d.).
URL: https://www.coindesk.com/price/bitcoin/

Bitcoin, Ethereum, Dogecoin, Litecoin stats (n.d.).
URL: https://bitinfocharts.com/

Black, F. (1976), Studies of stock market changes, proceedings of the american statistical association, *in* 'Proceedings of the 1976 American Statistical Association, Business and Economical Statistics Section (American Statistical Association, Alexandria, VA)'.

Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroskedasticity', *Journal of econometrics* **31**(3), 307–327.

Bouoiyour, J. and Selmi, R. (2015), 'Bitcoin price: Is it really that new round of volatility can be on way?'.

Bouri, E., Azzi, G. and Dyhrberg, A. H. (2017), 'On the return-volatility relationship in the bitcoin market around the price crash of 2013', *Economics* **11**(1).

Bouri, E., Molnár, P., Azzi, G., Roubaud, D. and Hagfors, L. I. (2017), 'On the hedge and safe haven properties of bitcoin: Is it really more than a diversifier?', *Finance Research Letters* **20**, 192–198.

Brauneis, A. and Mestel, R. (2018), 'Price discovery of cryptocurrencies: Bitcoin and beyond', *Economics Letters* **165**, 58–61.

Caferra, R. and Vidal-Tomás, D. (2021), 'Who raised from the abyss? a comparison between cryptocurrency and stock market dynamics during the covid-19 pandemic', *Finance Research Letters* **43**, 101954.

Campbell, J. Y., Campbell, J. W., Lo, A. W.-C., MacKinlay, A. C., Champbell, J. J., LO, A. A., Lo, A. W. et al. (1997), *The econometrics of financial markets*, princeton University press.

Caporin, M. and Costola, M. (2019), 'Asymmetry and leverage in garch models: A news impact curve perspective', *Applied Economics* **51**(31), 3345–3364.

Caporin, M. and McAleer, M. (2010), 'Ranking multivariate garch models by problem dimension', *Available at SSRN 1601236* .

Catania, L. and Grassi, S. (2021), 'Forecasting cryptocurrency volatility', *International Journal of Forecasting* .

CBC News (2021).
URL: https://www.cbc.ca/news/world/el-salvador-bitcoin-1.6166977

Cebrián-Hernández, Á. and Jiménez-Rodríguez, E. (2021), 'Modeling of the bitcoin volatility through key financial environment variables: an application of conditional correlation mgarch models', *Mathematics* **9**(3), 267.

Chaim, P. and Laurini, M. P. (2018), 'Volatility and return jumps in bitcoin', *Economics Letters* **173**, 158–163.

Chambers, J. M., Cleveland, W. S., Kleiner, B. and Tukey, P. A. (2018), *Graphical methods for data analysis*, Chapman and Hall/CRC.

Chappell, D. R. et al. (2019), 'Regime heteroskedasticity in bitcoin: A comparison of markov switching models', *Journal of Applied Economic Sciences (JAES)* **14**(65), 621–641.

Chatzikonstanti, V. (2017), 'Breaks and outliers when modelling the volatility of the us stock market', *Applied Economics* **49**(46), 4704–4717.

Chinhamu, K. and Chifurira, R. (2019), 'Evaluating south africa's market risk using asymmetric power auto-regressive conditional heteroscedastic model under heavy-tailed distributions', *Journal of Economic and Financial Sciences* **12**(1), 1–11.

Chinhamu, K., Huang, C.-K., Huang, C.-S. and Chikobvu, D. (2015), 'Extreme risk, value-at-risk and expected shortfall in the gold market', *International Business and Economics Research Journal (IBER)* **14**(1), 107–122.

Chkili, W. (2021), 'Modeling bitcoin price volatility: long memory vs markov switching', *Eurasian Economic Review* **11**(3), 433–448.

Choudhry, M. (2013), *An introduction to value-at-risk*, John Wiley and Sons.

Chu, J., Chan, S., Nadarajah, S. and Osterrieder, J. (2017), 'Garch modelling of cryptocurrencies', *Journal of Risk and Financial Management* **10**(4), 17.

Coindesk (n.d.).
URL: https://www.coindesk.com/

CoinGecko (n.d.), 'Cryptocurrency prices, charts, and crypto market cap'.
URL: https://www.coingecko.com/

CoinMarketCap (n.d.), 'Cryptocurrency prices, charts and market capitalizations'.
URL: https://coinmarketcap.com/

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

Connolly, A. and Kick, A. (2015), 'What differentiates early organization adopters of bitcoin from non-adopters?'.

Cont, R. (2001), 'Empirical properties of asset returns: stylized facts and statistical issues'.

Corbet, S., Lucey, B., Peat, M. and Vigne, S. (2018), 'Bitcoin futures-what use are they?', *Economics Letters* **172**, 23–27.

Costa, N., Silva, C. and Ferreira, P. (2019), 'Long-range behaviour and correlation in dfa and dcca analysis of cryptocurrencies', *International Journal of Financial Studies* **7**(3), 51.

Cubbin, E., Eidne, M., Firer, C. and Gilbert, E. (2006), 'Mean reversion on the jse', *Investment Analysts Journal* **35**(63), 39–47.

Dasman, S. (2021), 'Analysis of return and risk of cryptocurrency bitcoin asset as investment instrument', *Accounting and Finance Innovations* p. 51.

De Vries, C. G. and Leuven, K. (1994), 'Stylized facts of nominal exchange rate returns'.

Delatte, A.-L. and Lopez, C. (2013), 'Commodity and equity markets: Some stylized facts from a copula approach', *Journal of Banking and Finance* **37**(12), 5346–5356.

Dickey, D. A. and Fuller, W. A. (1981), 'Likelihood ratio statistics for autoregressive time series with a unit root', *Econometrica: journal of the Econometric Society* pp. 1057–1072.

Diebold, F. X. (1986), 'Comment on modelling the persistence of conditional variance', *Econometric Reviews* **5**(1), 51–56.

Digital Literacy in Africa (2022).
URL: https://www.do4africa.org/en/digital-literacy-in-africa/

Domjan, P., Serkin, G., Thomas, B. and Toshack, J. (2021), Introducing blockchain: TomorrowâĂŹs railroads, *in* 'Chain Reaction', Springer, pp. 1–8.

Dowd, K. (2007), *Measuring market risk*, John Wiley and Sons.

Drożdż, S., GÈľbarowski, R., Minati, L., OświÈľcimka, P. and Watorek, M. (2018), 'Bitcoin market route to maturity? evidence from return fluctuations, temporal correlations and multiscaling effects', *Chaos: An Interdisciplinary Journal of Nonlinear Science* **28**(7), 071101.

Duffee, G. R. (2002), 'Balance sheet explanations for asymmetric volatility', *University of California at Berkeley working paper* .

Dufour, J.-M. and Roy, R. (1985), 'Some robust exact results on sample autocorrelations and tests of randomness', *Journal of Econometrics* **29**(3), 257–273.

Dyhrberg, A. H. (2016), 'Bitcoin, gold and the dollar–a garch volatility analysis', *Finance Research Letters* **16**, 85–92.

Edem, K. D. and Ndengo, M. (2021), 'Modeling bank of kigali stock risks in rwanda stock exchange using extreme value distribution', *Journal of Financial Risk Management* **10**(3), 225–240.

Edwards, A. W. and Cavalli-Sforza, L. L. (1965), 'A method for cluster analysis', *Biometrics* pp. 362–375.

Engle, R. (2002), 'Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models', *Journal of Business and Economic Statistics* **20**(3), 339–350.

Engle, R. F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation', *Econometrica* **50**(4), 987–1007.
URL: http://www.jstor.org/stable/1912773

Engle, R. F. and Bollerslev, T. (1986), 'Modelling the persistence of conditional variances', *Econometric reviews* **5**(1), 1–50.

Engle, R. F. and Patton, A. J. (2007), What good is a volatility model?, *in* 'Forecasting volatility in the financial markets', Elsevier, pp. 47–63.

Erdman, C. and Emerson, J. W. (2008), 'bcp: an r package for performing a bayesian analysis of change point problems', *Journal of statistical Software* **23**, 1–13.

Fakhfekh, M., Jeribi, A., Ghorbel, A. and Hachicha, N. (2021), 'Hedging stock market prices with wti, gold, vix and cryptocurrencies: A comparison between dcc, adcc and go-garch models', *International Journal of Emerging Markets* .

Fama, E. F. (1965), 'The behavior of stock-market prices', *The journal of Business* **38**(1), 34–105.

Fama, E. F. (1970), 'Efficient capital markets: A review of theory and empirical work', *The journal of Finance* **25**(2), 383–417.

Fama, E. F. (1998), 'Market efficiency, long-term returns, and behavioral finance', *Journal of financial economics* **49**(3), 283–306.

Fernández, C. and Steel, M. F. (1998), 'On bayesian modeling of fat tails and skewness', *Journal of the american statistical association* **93**(441), 359–371.

Figà-Talamanca, G., Focardi, S. and Patacca, M. (2021), 'Regime switches and commonalities of the cryptocurrencies asset class', *The North American Journal of Economics and Finance* **57**, 101425.

Financial News and Stock quotes (n.d.).
URL: https://za.investing.com/

Foka Nzaha, S., Kala Kamdjoug, J. R. and Ndassi Teutio, A. O. (2022), Exploring the drivers of cryptocurrency adoption intent in africa, *in* 'World Conference on Information Systems and Technologies', Springer, pp. 394–403.

Frankovic, J., Liu, B. and Suardi, S. (2021), 'On spillover effects between cryptocurrency-linked stocks and the cryptocurrency market: evidence from australia', *Global Finance Journal* p. 100642.

Gbenro, N. and Moussa, R. K. (2019), 'Asymmetric mean reversion in low liquid markets: Evidence from brvm', *Journal of Risk and Financial Management* **12**(1), 38.

Ghoddusi, H., Morovati, M. and Rafizadeh, N. (2020), 'Asymmetric bitcoin volatility under structural breaks', *Available at SSRN 3589781* .

Gkillas, K. and Katsiampa, P. (2018), 'An application of extreme value theory to cryptocurrencies', *Economics Letters* **164**, 109–111.

Glaser, F., Zimmermann, K., Haferkorn, M., Weber, M. C. and Siering, M. (2014), 'Bitcoin-asset or currency? revealing users' hidden intentions', *Revealing Users' Hidden Intentions (April 15, 2014). ECIS* .

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

Gnedenko, B. (1943), 'Sur la distribution limite du terme maximum d'une serie aleatoire', *Annals of mathematics* pp. 423–453.

Gopane, T. J. (2021), An assessment of inter-market volatility and shock dynamics of bitcoin as a digital asset, *in* 'International Conference on Digital Economy', Springer, pp. 104–117.

Graf, K. S. (2014), 'Revisiting conceptions of commodity and scarcity in light of bitcoin'.

Granger, C. W. (1969), 'Investigating causal relations by econometric models and cross-spectral methods', *Econometrica: journal of the Econometric Society* pp. 424–438.

Guesmi, K., Saadi, S., Abid, I. and Ftiti, Z. (2019), 'Portfolio diversification with virtual currency: Evidence from bitcoin', *International Review of Financial Analysis* **63**, 431–437.

Haas, M., Mittnik, S. and Paolella, M. S. (2004*a*), 'Mixed normal conditional heteroskedasticity', *Journal of financial Econometrics* **2**(2), 211–250.

Haas, M., Mittnik, S. and Paolella, M. S. (2004*b*), 'A new approach to markov-switching garch models', *Journal of financial Econometrics* **2**(4), 493–530.

Hamilton, J. D. (1989), 'A new approach to the economic analysis of nonstationary time series and the business cycle', *Econometrica: Journal of the Econometric Society* pp. 357–384.

Hamilton, J. D. and Susmel, R. (1994), 'Autoregressive conditional heteroskedasticity and changes in regime', *Journal of econometrics* **64**(1-2), 307–333.

Hannan, E. J. and Quinn, B. G. (1979), 'The determination of the order of an autoregression', *Journal of the Royal Statistical Society: Series B (Methodological)* **41**(2), 190–195.

Herranz, E. (2017), 'Unit root tests', *Wiley Interdisciplinary Reviews: Computational Statistics* **9**(3), e1396.

Hinzen, F. J., John, K. and Saleh, F. (2022), 'BitcoinâĂŹs limited adoption problem', *Journal of Financial Economics* **144**(2), 347–369.

Houthakker, H. S. (1961), 'Systematic and random elements in short-term price movements', *The American Economic Review* **51**(2), 164–172.

Hu, A. S., Parlour, C. A. and Rajan, U. (2019), 'Cryptocurrencies: Stylized facts on a new investible instrument', *Financial Management* **48**(4), 1049–1068.

Huang, J.-Z., Ni, J. and Xu, L. (2022), 'Leverage effect in cryptocurrency markets', *Pacific-Basin Finance Journal* p. 101773.

Hurst, H. E. (1951), 'Long-term storage capacity of reservoirs', *Transactions of the American society of civil engineers* **116**(1), 770–799.

Hussain, S. I., Masseran, N., Ruza, N. and Safari, M. A. M. (2021), Predicting extreme returns of bitcoin: Extreme value theory approach, *in* 'Journal of Physics: Conference Series', Vol. 1988, IOP Publishing, p. 012091.

Huynh, T. L. D. (2022), 'When elon musk changes his tone, does bitcoin adjust its tune?', *Computational Economics* pp. 1–23.

Huynh, T. L. D., Nasir, M. A., Vo, X. V. and Nguyen, T. T. (2020), '"small things matter most": The spillover effects in the cryptocurrency market and gold as a silver bullet', *The North American Journal of Economics and Finance* **54**, 101277.

İçellioğlu, C. Ş. and Öner, S. (2019), 'An investigation on the volatility of cryptocurrencies by means of heterogeneous panel data analysis', *Procedia Computer Science* **158**, 913–920.

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

Islam, M. T. and Das, K. P. (2021), 'Predicting bitcoin return using extreme value theory', *American Journal of Mathematical and Management Sciences* **40**(2), 177–187.
URL: https://doi.org/10.1080/01966324.2021.1950086

Jackson, B., Scargle, J. D., Barnes, D., Arabhi, S., Alt, A., Gioumousis, P., Gwin, E., Sangtrakulcharoen, P., Tan, L. and Tsai, T. T. (2005), 'An algorithm for optimal partitioning of data on an interval', *IEEE Signal Processing Letters* **12**(2), 105–108.

Jankeeparsad, R. W. and Tewari, D. D. (2022), 'Bitcoin: An exploratory study investigating adoption in south africa', *Interdisciplinary Journal of Information, Knowledge, and Management* **17**, 191–214.

Jarque, C. M. and Bera, A. K. (1987), 'A test for normality of observations and regression residuals', *International Statistical Review/Revue Internationale de Statistique* pp. 163–172.

Jazeera, A. (2022), 'Bitcoin is now official currency in central african republic'.
URL: https://www.aljazeera.com

Jenkinson, A. F. (1955), 'The frequency distribution of the annual maximum (or minimum) values of meteorological elements', *Quarterly Journal of the Royal Meteorological Society* **81**(348), 158–171.

Ji, Q., Bouri, E., Lau, C. K. M. and Roubaud, D. (2019), 'Dynamic connectedness and integration in cryptocurrency markets', *International Review of Financial Analysis* **63**, 257–272.

Ji, Q., Bouri, E., Roubaud, D. and Kristoufek, L. (2019), 'Information interdependence among energy, cryptocurrency and major commodity markets', *Energy Economics* **81**, 1042–1055.

John, A., Logubayom, A. I. and Nero, R. (2019), 'Half-life volatility measure of the returns of some cryptocurrencies', *Journal of Financial Risk Management* **8**(1), 15–28.

JP Morgan, B. (1995), *'Riskmetrics technical manual'*, JP Morgan Bank. New York.

Karlsson, L. (2002), 'Garch modelling: Theoretical survey, model implementation and robustness analysis', *Unpublished master's thesis, Kungl Tekniska Högskolan* .

Kaseke, F. (2015), Modelling volatility in stock exchange data: a case study of three Johannesburg Stock Exchange (JSE) companies., PhD thesis, University of Kwazulu-Natal.

Kaseke, F., Ramroop, S. and Mhwambi, H. (2021), 'A comparison of the stylised facts of bitcoin, ethereum and the jse stock returns', *African Finance Journal* **23**(2), 50–64.

Kasse, I., Mariani, A., Utari, S. and Didiharyono, D. (2021), 'Investment risk analysis on bitcoin with applied of var-aparch model', *JTAM (Jurnal Teori dan Aplikasi Matematika)* **5**(1), 1–13.

Katsiampa, P. (2017), 'Volatility estimation for bitcoin: A comparison of garch models', *Economics letters* **158**, 3–6.

Katsiampa, P., Corbet, S. and Lucey, B. (2019), 'Volatility spillover effects in leading cryptocurrencies: A bekk-mgarch analysis', *Finance Research Letters* **29**, 68–74.

Keegan, F. (n.d.), 'Crypto mass adoption: A matter of when, not if'.
URL: `https://www.nasdaq.com/articles/crypto-mass-adoption%3A-a-matter-of-when-not-if`

Kendall, M. G. and Hill, A. B. (1953), 'The analysis of economic time-series-part i: Prices', *Journal of the Royal Statistical Society. Series A (General)* **116**(1), 11–34.

Khalfaoui, R., Sarwar, S. and Tiwari, A. K. (2019), 'Analysing volatility spillover between the oil market and the stock market in oil-importing and oil-exporting countries: Implications on portfolio management', *Resources Policy* **62**, 22–32.

Killick, R. and Eckley, I. (2014), 'changepoint: An r package for changepoint analysis', *Journal of statistical software* **58**(3), 1–19.

Killick, R., Fearnhead, P. and Eckley, I. A. (2012), 'Optimal detection of changepoints with a linear computational cost', *Journal of the American Statistical Association* **107**(500), 1590–1598.

Kim, J.-M., Jun, C. and Lee, J. (2021), 'Forecasting the volatility of the cryptocurrency market by garch and stochastic volatility', *Mathematics* **9**(14), 1614.

Kirkman, G. S., Osorio, C. A. and Sachs, J. D. (2002), 'The networked readiness index: Measuring the preparedness of nations for the networked world', *Korea* **4**, 20.

Kristoufek, L. (2015), 'What are the main drivers of the bitcoin price? evidence from wavelet coherence analysis', *PloS one* **10**(4), e0123923.

Kupiec, P. H. et al. (1995), *Techniques for verifying the accuracy of risk measurement models*, Vol. 95, Division of Research and Statistics, Division of Monetary Affairs, Federal.

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

Kwon, J. H. (2020), 'Tail behavior of bitcoin, the dollar, gold and the stock market index', *Journal of International Financial Markets, Institutions and Money* **67**, 101202.

Kyriazis, N. A. (2019), 'A survey on empirical findings about spillovers in cryptocurrency markets', *Journal of Risk and Financial Management* **12**(4), 170.

Lahmiri, S., Bekiros, S. and Salvi, A. (2018), 'Long-range memory, distributional variation and randomness of bitcoin volatility', *Chaos, Solitons and Fractals* **107**, 43–48.

Lambert, P. and Laurent, S. (2001), Modelling financial time series using garch-type models with a skewed student distribution for the innovations, Technical report, Catholic University of Leuven.

Lamoureux, C. G. and Lastrapes, W. D. (1990), 'Persistence in variance, structural change, and the garch model', *Journal of Business and Economic Statistics* **8**(2), 225–234.

LeBaron, B. (1992), 'Some relations between volatility and serial correlations in stock market returns', *Journal of Business* pp. 199–219.

Ling, S. and Li, W.-K. (1997), 'Diagnostic checking of nonlinear multivariate time series with multivariate arch errors', *Journal of Time Series Analysis* **18**(5), 447–464.

Ljung, G. M. and Box, G. E. (1978), 'On a measure of lack of fit in time series models', *Biometrika* **65**(2), 297–303.

López-Cabarcos, M. Á., Pérez-Pico, A. M., Piñeiro-Chousa, J. and Šević, A. (2021), 'Bitcoin volatility, stock market and investor sentiment. are they connected?', *Finance Research Letters* **38**, 101399.

López-Cabarcos, M., Pérez-Pico, A., Pineiro-Chousa, J. and Ševic, A. (2019), 'Bitcoin volatility, stock market and investor sentiment', *Are they connected* .

Lütkepohl, H. (2013), *Introduction to multiple time series analysis*, Springer Science and Business Media.

Luu Duc Huynh, T. (2019), 'Spillover risks on cryptocurrency markets: A look from var-svar granger causality and student's t copulas', *Journal of Risk and Financial Management* **12**(2), 52.

Maciel, L. (2021), 'Cryptocurrencies value-at-risk and expected shortfall: Do regime-switching volatility models improve forecasting?', *International Journal of Finance and Economics* **26**(3), 4840–4855.

Makatjane, K. D. and Molefe, E. K. (2020), 'Predicting regime shifts in johannesburg stock exchange all-share index (jse-alsi): A markov-switching approach', *Eurasian Journal of Economics and Finance* **8**(2), 95–103.

Manahov, V. and Hudson, R. (2013), 'Herd behaviour experimental testing in laboratory artificial stock market settings. behavioural foundations of stylised facts of financial returns', *Physica A: Statistical Mechanics and its Applications* **392**(19), 4351–4372.

Mandelbrot, B. (1972), 'Certain speculative prices"(1963)', *The Journal of Business* **45**(4), 542–543.

Mandelbrot, B. B. (1997), The variation of certain speculative prices, *in* 'Fractals and scaling in finance', Springer, pp. 371–418.

Mandelbrot, B. B. and Van Ness, J. W. (1968), 'Fractional brownian motions, fractional noises and applications', *SIAM review* **10**(4), 422–437.

Mariana, C. D., Ekaputra, I. A. and Husodo, Z. A. (2021), 'Are bitcoin and ethereum safe-havens for stocks during the covid-19 pandemic?', *Finance research letters* **38**, 101798.

Mazambani, L. and Mutambara, E. (2019), 'Predicting fintech innovation adoption in south africa: the case of cryptocurrency', *African Journal of Economic and Management Studies* .

Mensi, W., Al-Yahyaee, K. H. and Kang, S. H. (2019), 'Structural breaks and double long memory of cryptocurrency prices: A comparative analysis from bitcoin and ethereum', *Finance Research Letters* **29**, 222–230.

Mert, U. and Demireli, E. (2020), 'Asymmetric garch-type and half-life volatility modelling of usd/kzt exchange rate returns', *Eurasian Research Journal* **2**(2), 7–18.

Mikhaylov, A. Y. (2018), 'Volatility spillover effect between stock and exchange rate in oil exporting countries', *International Journal of Energy Economics and Policy* .

Mills, F. C. (1927), Introduction to" the behavior of prices", *in* 'The Behavior of Prices', NBER, pp. 31–36.

Mitchell, W. C. (1938), *The making and using of index numbers*, US Government Printing Office Washington DC.

Mitra, S. K. (2012), 'Is hurst exponent value useful in forecasting financial time series', *Asian social science* **8**(8), 111–120.

Moratis, G. (2021), 'Quantifying the spillover effect in the cryptocurrency market', *Finance Research Letters* **38**, 101534.

Muguto, L. and Muzindutsi, P.-F. (2022), 'A comparative analysis of the nature of stock return volatility in brics and g7 markets', *Journal of Risk and Financial Management* **15**(2), 85.

Nakamoto, S. (2008), 'Bitcoin: A peer-to-peer electronic cash system', *Decentralized Business Review* p. 21260.

Nelson, D. B. (1991), 'Conditional heteroskedasticity in asset returns: A new approach', *Econometrica: Journal of the Econometric Society* pp. 347–370.

Ngunyi, A., Mundia, S. and Omari, C. (2019), 'Modelling volatility dynamics of cryptocurrencies using garch models'.

Nseke, P. et al. (2018), 'How crypto-currency can decrypt the global digital divide: bitcoins a means for african emergence', *International Journal of Innovation and Economic Development* **3**(6), 61–70.

Omar, C., Mundia, S. and Ngina, I. (2020), 'Forecasting value-at-risk of financial markets under the global pandemic of covid-19 using conditional extreme value theory'.

Orskaug, E. (2009), Multivariate dcc-garch model:-with various error distributions, Master's thesis, Institutt for matematiske fag.

Oseifuah, E. K. and Korkpoe, C. H. (2019), 'A markov regime switching approach to estimating the volatility of johannesburg stock exchange (jse) returns', *Investment management and financial innovations* **16**(1), 215–225.

Pavlov, A. (2022), 'Momentum vs mean reversion in crypto'.
URL: https://medium.com

Peters, E. E. (1994), *Fractal market analysis: applying chaos theory to investment and economics*, Vol. 24, John Wiley and Sons.

Peters, E. E. (1996), *Chaos and order in the capital markets: a new view of cycles, prices, and market volatility*, John Wiley and Sons.

Phillips, P. C. and Perron, P. (1988), 'Testing for a unit root in time series regression', *Biometrika* **75**(2), 335–346.

Qarni, M. O., Gulzar, S., Fatima, S. T., Khan, M. J. and Shafi, K. (2019), 'Intermarkets volatility spillover in us bitcoin and financial markets', *Journal of Business Economics and Management* **20**(4), 694–714.

Quandt, R. E. (1958), 'The estimation of the parameters of a linear regression system obeying two separate regimes', *Journal of the american statistical association* **53**(284), 873–880.

Quigley, L. and Ramsey, D. (2008), 'Statistical analysis of the log returns of financial assets', *Financial mathematic, University of Limerick* **32**.

Quinn, B. G. (1980), 'Order determination for a multivariate autoregression', *Journal of the Royal Statistical Society: Series B (Methodological)* **42**(2), 182–185.

Raputsoane, L. (2008), *Exchange rate volatility spillovers and the South African currency*, South African Reserve Bank.

Rauchs, M., Blandin, A., Dek, A. and Wu, Y. (2021), 'Cambridge bitcoin electricity consumption index', *Retrieved from Cambridge Center for Alternative Finance website: https://www. cbeci. org/cbeci/comparisons* .

Real time billionaires (n.d.).
URL: https://www.forbes.com/real-time-billionaires

Regnault, J. (1863), *Calcul des chances et philosophie de la bourse*, Pilloy.

Runfang, Y., Jiangze, D. and Xiaotao, L. (2017), 'Improved forecast ability of oil market volatility based on combined markov switching and garch-class model', *Procedia computer science* **122**, 415–422.

Rural population, percent in Africa (n.d.).
URL: https://www.theglobaleconomy.com/rankings/rural_population_percent/Africa.

Saiedi, E., Broström, A. and Ruiz, F. (2021), 'Global drivers of cryptocurrency infrastructure adoption', *Small Business Economics* **57**(1), 353–406.

Schwarz, G. et al. (1978), 'Estimating the dimension of a model', *The annals of statistics* **6**(2), 461–464.

Scott, A. J. and Knott, M. (1974), 'A cluster analysis method for grouping means in the analysis of variance', *Biometrics* pp. 507–512.

Selmi, R., Mensi, W., Hammoudeh, S. and Bouoiyour, J. (2018), 'Is bitcoin a hedge, a safe haven or a diversifier for oil price movements? a comparison with gold', *Energy Economics* **74**, 787–801.

Sen, A. and Srivastava, M. S. (1975), 'On tests for detecting change in mean', *The Annals of statistics* pp. 98–108.

Shahzad, S. J. H., Bouri, E., Roubaud, D., Kristoufek, L. and Lucey, B. (2019), 'Is bitcoin a better safe-haven investment than gold and commodities?', *International Review of Financial Analysis* **63**, 322–330.

Shanaev, S. and Ghimire, B. (2021), 'A fitting return to fitting returns: Cryptocurrency distributions revisited', *Available at SSRN 3847351* .

Shapiro, S. S. and Wilk, M. B. (1965), 'An analysis of variance test for normality (complete samples)', *Biometrika* **52**(3/4), 591–611.

Sheppard, K. (2013), 'Financial econometrics notesï£¡', University of Oxford.

Shiferaw, Y. (2018), The bayesian ms-garch model and value-at-risk in south african agricultural commodity price markets, Technical report, University of Johannesburg.

Smith Jr, C. W. (1988), 'Market volatility: Causes and consequences', *Cornell L. Rev.* **74**, 953.

Smith, R. L. (1986), 'Extreme value theory based on the r largest annual events', *Journal of Hydrology* **86**(1-2), 27–43.

Stavroyiannis, S. (2018), 'A note on the nelson-cao inequality constraints in the gjr-garch model: is there a leverage effect?', *International Journal of Economics and Business Research* **16**(4), 442–452.

Stoll, C., Klaaßen, L. and Gallersdörfer, U. (2019), 'The carbon footprint of bitcoin', *Joule* **3**(7), 1647–1661.

Suda, D. and Spiteri, L. (2019), Comparing market phase features for cryptocurrency and benchmark stock index using hmm and hsmm filtering, *in* 'International Conference on Business Information Systems', Springer, pp. 195–207.

Symitsi, E. and Chalvatzis, K. J. (2018), 'Return, volatility and shock spillovers of bitcoin with energy and technology companies', *Economics Letters* **170**, 127–130.

Takaishi, T. (2018), 'Statistical properties and multifractality of bitcoin', *Physica A: statistical mechanics and its applications* **506**, 507–519.

Thies, S. and Molnár, P. (2018), 'Bayesian change point analysis of bitcoin returns', *Finance Research Letters* **27**, 223–227.

Thompson, S. (2011), 'The stylised facts of stock price movements', *The New Zealand Review of Economics and Finance* **1**.

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

Tsay, R. S. (2013), *Multivariate time series analysis: with R and financial applications*, John Wiley and Sons.

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

Tse, Y. K. and Tsui, A. K. C. (2002), 'A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations', *Journal of Business and Economic Statistics* **20**(3), 351–362.

Tumblr, A. C. T. and Team, A. C. (2022), 'P2p platforms, remittances, and savings needs power africa's grassroots cryptocurrency adoption'.
URL: https://blog.chainalysis.com/reports/africas-grassroots-cryptocurrency-adoption/

Uzonwanne, G. (2021), 'Volatility and return spillovers between stock markets and cryptocurrencies', *The Quarterly Review of Economics and Finance* **82**, 30–36.

Van Kerm, P. (2007), 'Extreme incomes and the estimation of poverty and inequality indicators from eu-silc'.

VanDenburgh, W. M. and Daniels, R. B. (2021), 'Pragmatic realities of bitcoin and crypto-investing', *The CPA Journal* **91**(6/7), 36–41.

Vincent, O. and Evans, O. (2019), 'Can cryptocurrency, mobile phones, and internet herald sustainable financial sector development in emerging markets?', *Journal of Transnational Management* **24**(3), 259–279.

Walton, A. J. and Johnston, K. A. (2018), 'Exploring perceptions of bitcoin adoption: The south african virtual community perspective', *Interdisciplinary Journal of Information, Knowledge, and Management* **13**, 165.

Wood, J., Jang, H., Lenskiy, A. and Khan, G. F. (2017), 'The diffusion and adoption of bitcoin: a practical survey for business', *International Business Management* **11**, 1278–1288.

World Internet Users Statistics and 2022 World Population Stats (n.d.).
URL: https://www.internetworldstats.com/stats.htm

Yang, Z. and Zhou, Y. (2017), 'Quantitative easing and volatility spillovers across countries and asset classes', *Management Science* **63**(2), 333–354.

Yousaf, I., Ali, S. and Wong, W.-K. (2020), 'Return and volatility transmission between world-leading and latin american stock markets: Portfolio implications', *Journal of Risk and Financial Management* **13**(7), 148.

Zaremba, A., Bilgin, M. H., Long, H., Mercik, A. and Szczygielski, J. J. (2021), 'Up or down? short-term reversal, momentum, and liquidity effects in cryptocurrency markets', *International Review of Financial Analysis* **78**, 101908.

Zhang, C., Ma, H., ARKORFUL, G. B. and Peng, Z. (2021), 'The impacts of futures introduction on spot market volatility: Evidence from the bitcoin market', *Available at SSRN 3903735* .

Zhang, J. and He, Q.-z. (2021), 'Dynamic cross-market volatility spillover based on msv model: Evidence from bitcoin, gold, crude oil, and stock markets', *Complexity* **2021**.

Zhang, W., Wang, P., Li, X. and Shen, D. (2018), 'Some stylized facts of the cryptocurrency market', *Applied Economics* **50**(55), 5950–5965.

Zhang, Y., Chan, S., Chu, J. and Sulieman, H. (2020), 'On the market efficiency and liquidity of high-frequency cryptocurrencies in a bull and bear market', *Journal of Risk and Financial Management* **13**(1), 8.

Zhang, Y.-J., Yao, T. and He, L.-Y. (2015), 'Forecasting crude oil market volatility: can the regime switching garch model beat the single-regime garch models?', *arXiv preprint arXiv:1512.01676* .

Zivot, E. and Andrews, D. W. K. (2002), 'Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis', *Journal of business and economic statistics* **20**(1), 25–44.

PUBLICATIONS AND SUBMISSIONS

Available online at

Sabinet

African Finance Journal Vol. 23, Issue 2 (2021) 50-64

A Comparison of the Stylised Facts of Bitcoin, Ethereum and the JSE Stock Returns

Forbes Kaseke [a, *], Shaun Ramroop [a], Henry Mhwambi [a]

[a] *School of Mathematics, Statistics and Computer Science, University of Kwazulu-Natal (PMB)*

ABSTRACT

This study investigates the stylised facts of cryptocurrency using Bitcoin and Ethereum and comparing them with those of Gold and FTSE/JSE 40 index. We also divide our data to bearish and bullish periods of cryptocurrency. Various known properties of financial data are investigated on the four assets. The evidence shows that cryptocurrencies possess similar stylised facts with the Gold and JSE but has some minor differences such as the volatility level. Examining the bearish and bullish periods separately did not reveal much change on the stylised facts.

JEL Classification: C10, C18

Keywords: Bitcoin, Ethereum Gold, volatility, GARCH

1. Introduction

Financial data are different and exist under different markets. Decades of research have shown that for some assets there exist statistical properties that are very similar although varying to some degree. The homogenous properties are mostly seen on assets within the same class or within the same markets. These types of properties are known as stylised facts. While stylised facts are not the same, they help develop a general understanding of assets, which in turn aid in in model development and hence decision-making. For any data, an adequate model depends on the unobservable data generating process. Whilst the process is unobservable, we can observe the patterns within the data. It is this pattern that the stylised facts reveal.

In financial data, the unique characteristics were first reported early in the 20th century by various scholars. Mitchell (1938) was the first to report the existence of fat tails in financial data, this was also reported by Mills (1927)) and other scholars such as Alexander (1961). Kendall and Hill (1953) and Houthakker (1961) were among the first to give empirical evidence that prices were non-stationary. This would lead to most models requiring transformed data prices such as the returns, which were in turn stationary. Cont (2001) discuss the various stylised facts that exist in financial markets and further explain how these invalidate the use of traditional time series models. The perfect model would require to take into account these various stylised facts. Thompson (2011) explains that if a model provides a reasonable approximation of the actual data generating process for a stock price, then the fitted model should exhibit the stylized facts. In essence, the stylised facts further act as a model check.

In 2008, Nakamoto (2008) proposed a new form of a financial asset that is known as cryptocurrency. This new form of currency was to be paperless and an online virtual currency based on computer blockchain technology. The paper argued that this peer to peer transaction was purely computer-based requiring proof of works to prevent double-spending which had been a problem with other online payments. In 2009 the actual creation of the first cryptocurrency Bitcoin occurred. Since then, many other cryptocurrencies have been created. Among the new cryptocurrencies is Ethereum, which is currently the second-biggest cryptocurrency by market value as per coinmarketcap.com data.

*Corresponding author.
Email address: forbeskaseke@gmail.com (F. Kaseke)

Investment Management and Financial Innovations, Volume 19, Issue 4, 2022

BUSINESS PERSPECTIVES

LLC "CPC "Business Perspectives"
Hryhorii Skovoroda lane, 10,
Sumy, 40022, Ukraine

www.businessperspectives.org

Forbes Kaseke (South Africa), **Shaun Ramroop** (South Africa),
Henry Mwambi (South Africa)

A COMPARATIVE ANALYSIS OF THE VOLATILITY NATURE OF CRYPTOCURRENCY AND JSE MARKET

Abstract

Despite the rapid growth of developing markets, aided by globalization, comparative studies of cryptocurrency and stock market volatility have focused on the developed markets and neglected developing ones. In this regard, this study compares cryptocurrency volatility with that of the Johannesburg Stock Exchange (JSE), a developing market. GARCH-type models are applied to daily log returns of Bitcoin, Ethereum, and the FTSE/JSE 40 in two ways. Firstly, the models are applied directly; secondly, structural breaks are tested and accounted for in the models. The sample period was from September 18, 2017, to May 27, 2021. The results show higher volatility and higher volatility persistence in cryptocurrency than in the JSE market. They also show that persistence is overestimated for cryptocurrencies when structural breaks are not accounted for. The opposite was true for the JSE.

Moreover, the two cryptocurrencies were found to have close to identical volatility plots that differ from that of the JSE. High volatility periods of cryptocurrency also did not coincide with that of JSE and those of JSE did not coincide with the cryptocurrency ones. There is also evidence of an inverse leverage effect in cryptocurrency, which opposes the normal leverage effect of the JSE market.

Received on: 7th of July, 2022
Accepted on: 16th of August, 2022
Published on: 13th of October, 2022

Forbes Kaseke, Ph.D. in Statistics, College of Agriculture, Engineering and Science, School of Mathematics, Statistics and Computer Science, University of Kwazulu-Natal, South Africa. (Corresponding author)

Shaun Ramroop, Professor, College of Agriculture, Engineering and Science, School of Mathematics, Statistics and Computer Science, University of Kwazulu-Natal, South Africa.

Henry Mwambi, Professor, College of Agriculture, Engineering and Science, School of Mathematics, Statistics and Computer Science, University of Kwazulu-Natal, South Africa.

Keywords cryptocurrency, volatility, Bitcoin, persistence, asymmetry, structural breaks

JEL Classification C22, C58, G11, G15

INTRODUCTION

Volatility is the most crucial of all the stylized facts that come with financial data. This is because it determines the riskiness of an asset. Higher volatility signifies more risk and a higher likelihood of making losses.

Nakamoto (2008) suggested the first idea of cryptocurrency. He introduced Bitcoin, a new form of currency based on blockchain computer technology and free from institutional control. Over a decade later, Bitcoin and other cryptocurrencies created afterward are considered alternative investment options. While the initial purpose was to create an alternative currency, cryptocurrencies are now considered an asset. Baur et al. (2018) and Glaser et al. (2014) highlight the classification of cryptocurrency as an asset. This revelation is not surprising given the rapid increase of Bitcoin from being given for free on launch, reaching 10 cents in 2010, and attaining an all-time high of over $60,000 in 2021. However, despite the rapid growth in the value of cryptocurrencies, the growth has been marred by high volatility. There is, therefore, a need to study this volatility so investors can better understand how to invest amid the turmoil.

Conflict of interest statement:
Author(s) reported no conflict of interest

Modeling Cryptocurrency Risk Using Hybrid GARCH Models

Abstract

The last decade has seen the birth and growth of a new form of digital currency known as cryptocurrency. This has been seen as a new alternative area of investment, but the adoption has been met with critics due to the high volatility, which poses a high risk to investors. As a new asset, less is known about the properties of cryptocurrency. This in turn makes models used to quantify the risk an ongoing research topic. Published literature is dominated by GARCH type models which make use of the three common error distributions, that is, the student t, the skewed student t, and the generalised error distribution. This paper seeks to add to the knowledge of the models that can be used to quantify the risk by introducing a hybrid GARCH model. The hybrid GARCH model is developed by combining traditional GARCH type models with extreme value distributions, with the main focus being the Generalised Pareto Distribution and the Generalised Extreme Value Distribution. Risk is then quantified using VaR. The VaR is backtested using the Kuipiec test. Using Bitcoin, Ethereum, and Dogecoin, the results show that the GJR-GARCH model with GPD innovations is effective in estimating the VaR of cryptocurrency.

Keywords

Cryptocurrency, volatility, GPD, Extreme value theory.

Introduction

Amid the 2008 global recession, a new type of digital currency known as cryptocurrency was born. This innovation was a paperless form of currency that would soon change the dynamics of online payments by introducing a new player in the market. Despite a slow start, cryptocurrency quickly took off, with many new digital coins entering the market. In January 2021 VanDenburgh and Daniels (2021) reported that the market cap of all cryptocurrencies reached 1 trillion dollars for the first time. This record would be followed by Bitcoin in February 2022, being the first cryptocurrency and the fastest asset to reach a market cap of 1 trillion dollars. This growth would be a sign to many that cryptocurrency was now a significant player in finance.

Estimating the volatility in Cryptocurrency and the JSE using Markov switching GARCH models

Abstract

Past studies have shown that structural breaks affect the model parameters and any inferences made from them. This paper tests the presence of regimes in the volatility dynamics of cryptocurrency and the JSE through the Markov-switching GARCH (MSGARCH) models. We compare single-regime GARCH models to two-regime GARCH models. The sample period is from September 18th, 2017, to May 27th, 2021. The results showed that the two regime models for the sample period outperformed the cryptocurrency's single regime model. However, for the JSE, there was no added benefit in using the two regime models. The study findings are helpful to investors as they highlight the need to know the current market state before laying out their trading strategy.

Introduction

Financial time series modelling has been a subject of constant research, with researchers constantly looking for more robust models, particularly in the wake of the financial crisis, which often led to failures of the current models. The traditional models used in time series, such as the AR and MA models, have proven inadequate in modelling financial data. This inadequacy is because financial data does not adhere to the assumptions such as constant variance and normality that these traditional models assume. Financial data is characterised by none constant variance, volatility clustering, and leverage effects, all of which cannot be captured by the traditional models.

Volatility is one of the most critical aspects that need to be modelled in finance. It is essential because it reflects the riskiness of an asset, with higher volatility indicating higher risk and, therefore, a higher probability of making losses (Muguto and Muzindutsi 2022). Investors are keen to know the risk level associated with their investments and need to know how to forecast the risk. This need was intensified by the global financial crisis in 2008 as investors lost out on their investments, and some were totally liquidated.

In an attempt to solve the issue of modelling financial time series data, Engle (1982) introduced the autoregressive conditional heteroscedasticity (ARCH) model, which would later earn him a Noble prize in economics. The ARCH model revolutionised how time-series data with time-varying volatility would

Investigating the spillover effects between cryptocurrency and Johannesburg Stock Exchange

Forbes Kaseke, Shaun Ramroop, Henry Mhwambi
School of Mathematics, Statistics and Computer Science, University of Kwazulu-Natal (PMB)

Abstract

We investigate the spillovers effects from Bitcoin to Ethereum, Dogecoin, and the JSE40 index. Using the VAR model and the DCC-GARCH model we investigate the short run and long run spillover effects from Bitcoin to the other cryptocurrencies and the JSE. Our findings indicate that Bitcoin has significant spillover effects on Ethereum but not on Dogecoin and JSE. There also exist strong positive conditional correlations between Bitcoin and Ethereum. The conditional correlations fluctuate from negative to a weak positive conditional correlation between Bitcoin and JSE across the considered period.

Keywords

Bitcoin, volatility, DCC-GARCH, Spillovers.

Introduction

Since its introduction in 2008 in a concept paper by Nakamoto (2008) and its eventual creation in 2009, Bitcoin has become one of the most sought-after investment options. The early days after its creation saw most people label it a joke and a scam. However, despite all this, Bitcoin has grown in value from virtually 0 dollars on launch to reaching one dollar for the first time in February 2012 before reaching its all-time high of 68 564.40 dollars in November 2021 *Bitcoin* (n.d.), earning itself the name "digital gold" in the process. While it was the first cryptocurrency, other alternative cryptocurrencies have also been created. The number of cryptocurrencies has since skyrocketed, with the number of coins listed on CoinMarketCap (n.d.) being over 10 000 as of May 2021, with many others not being listed on the website. The emergence of other alternative coins resulted in a reduction in the market dominance of Bitcoin, which, however, still commands the market with a market share of nearly 50%.

The combined all-time high of the market value of the cryptocurrency market rose from around 578 billion dollars in November 2020 to a staggering 3 trillion

A discussion on factors affecting the growth and adoption of cryptocurrency in Africa and the world

Forbes Kaseke, Shaun Ramroop, Henry Mwambi
School of Mathematics, Statistics and Computer Science, University of Kwazulu-Natal (PMB)

Abstract

There has been growing concern about the viability of cryptocurrency as an alternative form of payment. The main concern is the lack of mass adoption across the divide. While there is a general increase in the users, the market penetration remains low for a currency aimed to be an alternative to the fiat system. This paper aims to discuss the factors hindering the mass adoption of cryptocurrency, with the main focus being Africa. We review the literature on the adoption studies and the reported problems associated with cryptocurrency. We also discuss factors associated with technology and the market area of cryptocurrency. In our conclusion, we show that one of the most significant barriers to mass adoption in Africa is poverty and the volatile nature of cryptocurrency.

Keywords

Bitcoin, volatility, Africa.

Introduction

The end of the twentieth century saw the emergence of a new form of currency known as cryptocurrency. The first cryptocurrency was Bitcoin, introduced in 2008 in a concept paper by a faceless character named Satoshi (Nakamoto 2008). The paper introduced Bitcoin as a new payment system based on blockchain technology. The idea was to have a decentralised currency free from institutional control. This new currency would also remove the need for a third party to facilitate transactions and be trustless.

The first Bitcoin was minted in the subsequent year, with Satoshi as the first beneficiary. The initial days saw little to no demand for Bitcoin, with a few tech enthusiasts participating. Coins were given for free in universities, but due to their low value, most students exchanged the coins for meals at the cafeteria. One of the famous sales is when a man sold 10 000 in exchange