UNIVERSITY OF NAIROBI

SCHOOL OF MATHEMATICS

Testing Weak Form of Market Efficiency of Exchange Traded Funds at NSE Market

Cavin Oyugi Ongere

Reg. No: I56/12130/2018

A Masters Project Submitted in Partial Fulfillment of the Requirements for the Award of the Degree of Masters of Science in Actuarial Science in the School of Mathematics, University of Nairobi, Kenya.

August, 2020

Declaration and Approval

This project research is my original work and has not been presented for a degree in any other university or institution of learning for examination.

Cavin Oyugi Ongere

Reg.No. 156/12130/2018

Signature: .... Date ....

Approval

This project research has been has been under our supervision and has our approval for submission.

Prof. Philip Ngare

School of Mathematics,

University of Nairobi,

Nairobi, Kenya.

Signature .... Date: ....

Acknowledgment

Gratitude, Thanksgiving and Appreciation is all that fills my heart with the completion of this Master's Degree Project. All Glory to the Almighty God for guiding me through. My sincere appreciation goes to Prof. Philip Ngare for his unwavering support, guidance and encouragement without which, the completion of this project would not have been easy. Your attention to detail will forever be my guide in my professional life. I am forever indebted to you. I would like to thank my parents as well, whose ceaseless encouragements have seen me this far. Words may not be enough to express my gratitude for you. The project, would not have also been a great success without the emotional support of my family; my dearest wife Stacey Awaa, my little princess Adriana and my little knight Benson, all whose ostensible love and support has continued to inspire me. To friends and classmates from whom endless pieces of encouragement, prayers and support flowed, thank you. I may not mention all but I acknowledge that the success of this project is as a result of all your collective and collaborative support. I say thank You. I wish to be there for all of you as you have been to me. For the pieces of advice, for the encouragements, for the spiritual and emotional support; Gratitude and Thanks!!

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Acronyms

MAPFE Mean Absolute Percentage Forecast Error

i.i.d. Independent and identically distributed

M.L.E. Maximum Likelihood Estimation

EMH Efficient Market Hypothesis

NSE Nairobi Stock Exchange

NSE-20 Nairobi Stock Exchange 20 Share Index

DFA Detrended Fluctuation Analysis

H_0 Null Hypothesis

H_0 Alternative Hypothesis

NASI NSE ALL Share Index

RMSEA Root Mean Square Error of Approximation

ARFIMA Autoregressive fractionally integrated moving average

ETF Exchange Traded Fund

Abstract

Market efficiency is defined as a case when the prices in a market reflects the information, which is currently available within the market, or otherwise, the market is said to be inefficient whenever the prices of a financial security in not reflected by the information available to all local and international investors who are trading in the securities market. Having a better understanding of the market efficiency when trading an Exchange Traded Fund of any given set of securities in an exchange market is extremely vital for any prospective investor who need to make sound investment decisions as well as market predictions. When trading in a market with few traders who likes dominating the market through insider trading, it is more likely to experience securities exchange market without confidence of investors thus depicting weak form of the efficient market efficiency. Nairobi Securities Exchange market is important in the economy especially for those companies that are looking forward to capital or startups in the Kenyan Market from a global perspective. While testing of the weak form of efficient market hypothesis or EMH of the Nairobi stock exchange (NSE) is done through daily as well as weekly securities index data from NSE 20 share index over the period, 2nd February 2002 to May 2nd 2019. The research study applies the use of secondary NSE data that was derived from Nairobi Stock Exchange market website. This research has deviated from the normal and conventional linear approach to test market efficiency and use of using unit roots to test serial correlation. The daily returns in aspect to skewness as well as kurtosis was found to be non-normal. Similar demonstrations resulted from the Kolmogorov Smirnov test. From the results, null hypothesis of the normality was not rejected. In this research, there is the use of fractional integration thus utilization of ARFIMA to test long term memory and even the traditional unit root test is incorporated to compare both results thus giving a perfect conclusion on whether NSE stock market is definitely weak form efficient. Moreover, NSE-20 share Index stocks are used to make an Exchange Traded Fund that is priced and forecasted that is important for investors looking forward to make investments at the NSE Market in Kenya due to its mimicking ability. Ultimately, the forecasted values of ETF is done on the trendlines similar to the NSE-20 share Index trends, which investors to make informed financial decisions when buying any securities traded in NSE market.

Chapter 1

Introduction

1.1 Background of the Study

A securities market efficiency studies is an essential concept in financial modeling, especially in terms of a comprehensive of the working of the capital markets as well as in their performance together with development of any country's economy progress for the investors see[Cao et al., 2018]. In past couple of years, the efficient market hypothesis (EMH) has been a huge topic for many researchers in financial due to the significant implications when determining on whether an investor will buy securities of a given company or not at the same time enabling the investors to understand the trends within the market prior to making their investment decisions.

[Fama, 1965] did a definition of the efficient market as one where the existing prices fully does reflect all kinds of information available that investors should know when making decisions. It means that any security market that is an efficient can only be achieved when the prices do represent the absolute true values of the securities at the same time in turn, the scarce savings can be automatically allocated to all resourceful investments thus making way that benefits to investors looking for ways to enhance the economic activities in a country. [Copeland et al., 2005]. Fama in his publication also suggested three models that can be used when testing market efficiency were namely the sub-martingale model, the random walk as well as fair game models, which are important in making an excellent investment decision on the stock market thus leading to profits from the selected portfolio of companies.

For instance, the fall in securities prices started this year mainly reflecting concern and uncertainty over the global spread of Covid-19 thus having a huge effect on the market economy see [Odhiambo et al., 2020a]. This has led to the experienced securities prices falling by over 15% across many stocks worldwide. Any investor always believes that the securities prices in the market will have a reflection of a real economic activities in the specific country.

When an investor have information about the fundamentals of a securities exchange market in terms of its market efficiency, they will be capable of making an informed investment decision on whether it is viable to buy a certain type of security or not depending on the value of a specific index that trades in the securities market[Osoro and Jagongo, 2013] for the investors.

1.1.1 Stock Market

Stock market can always be described as any form of structured market used for purchasing or/and vending financial instruments such as bonds, shares, swaps, options and commodities. A stock market exchange is a specific location where all purchasers and sellers conglomerate to make a tradoff of securities of different characteristics. A company listed by stock exchange market must satisfy the laid down trading requirements see[Kairu, 1976]. Stock markets facilities free transfer of shares between companies and investors thus mobilize people's savings and directing them to growth of the economy. The efficiency of any form of emerging market has been so vital to investors as regulatory reforms are made and barriers are removed for internationally equity investments.

In Kenya, the dealing or trading in securities started in 1920's whenever the market was in a rudimentary stage for the colonist investors who inviting foreigners to make tradings on the floor of trade. However, there was none of a structured market with rules and regulations that would govern the activities of securities brokerage. Trading that took place was dependent on the gentleman's agreements that lacked standard commissions which were charged on clients who were indebted to honor the terms and conditions of contractual commitments when making good delivery at the same time settling relevant incurred costs during the period.

In 1980s when the Kenyan Government discovered the significance of designing and implementing policy reforms that would easily improve manageable economic conditions both with a both efficient and effective way after development of the infrastructure for trading thus allowing for many investors who would wish to trade within the securities market at the same time increasing its liquidity.

1.1.2 Kenya Stock Market

Kenya became independent from British colonial powers in 1963, however it became a republic in 1964 that allows it to open market for those international traders who were looking forward to invest in the country. It took the country many more years to introduce a democratic system that came in the year 1991 that was having an impact on the kind of trade activities taking place in the Nairobi Securities Exchange due to worst political environment for investors. The above phenomena was attained after many years of turbulence and pressure foreign countries. Currently Kenya is among the best performing economies in Eastern Africa with the highest GDP when compared to all other neighboring countries. In the year 1964, NSE-20 share index was introduced at the same time it was the main index in the country's security market. However it is only after 1993 that NSE-20 share index started performing so well as a result of relaxed taxation, less control on foreign investments and exchange controls. Unfortunately, political instability has been a key causative of high market volatility mostly during the general elections.

In many cases, for instance, the Nairobi Stock Exchange (NSE) uses indices to determine the trends in the market such as Nairobi All Share Index as well as NSE-20 share Index among many others that are being used as a market index to determine whether the economy of the country is doing well or not. Its measure is an overall indicator of the market performance see[Kirui et al., 2014]. The index always incorporates all the tradings in the number of shares of the day from those companies that are perceived to perform well thus making them lucrative for the investors for investment purposes. It has attention that is therefore important on the determining the trends on the overall market capitalization and the price movements of the selected securities that are being traded on the market.

While in Kenya, thee stock market is determined using three stock indices like NSE-20 Index, the NSE All Share Index as well as MCSE Share Index. The NSE-20 Index is the most commonly used since it incorporates 20 companies cutting across all sectors in the economy. This index has always been a great importance in the world markets NSE being one of these growing markets see[Mumo et al., 2017]. The index has helped the world market in the analysis and portfolio management. Therefore, the index value is used when determining the performance of a security market and the institutions as well as the individuals can get to know how the market is performing and their investments in general see [Odhiambo et al., 2020b].

The companies that formed the NSE-20 share index are given as follows Saini Ltd, WPP Scangroup Ltd, Nation Media Group, Bamburi Cement Ltd, Kenya Power & Lighting Ltd, Kenol Kobil Ltd, British American Investments Co. Ltd, Kengen Ltd, CIC Insurance Group, Kenya Commercial Bank, Diamond Trust Bank Ltd, The Cooperative Bank, Barclays Bank Ltd, CFC Stanbic Holdings Ltd, Equity Bank Ltd, East African Breweries Ltd, Athi River Mining, British American Tobacco Kenya Ltd, Safaricom Ltd and Centum Investments Ltd. The 20 share index is dominated by the financial sector, all other sectors of the economy are represented.

All the companies in NSE are categorized into different forms. The forms includes Banking, agricultural,Construction and Allied, telecommunication among others. Table 1.1 shows all categories traded in NSE:

No.	Sector
1.	Insurance
2.	Exchange Traded Funds
3.	Agricultural
4.	Commercial and Services
5.	Real Estate Investment Trust
6.	Telecommunication
7.	Energy and Petroleum
8.	Investment
9.	Banking
10.	Construction and Allied
11.	Automobiles and Accessories

Table 1.1: Key Sectors of Trade at the Nairobi Securities Exchange Market

1.1.3 Capital Market Efficiency

The characteristics of efficient market security market includes the ability securities prices responding accurately and instantaneously to new information as the information in the market will always determine the trends of the prices because of the forces of demand as well as supply within the market [Rösch et al., 2017]. The trading rules makes it impossible to use simulation experiments to make superior returns in an efficient market. In addition, the insider information cannot be used in making super returns as insiders are not permitted to trade in their companies and both private and public information adjust quickly to the price of the security as some investors may take advantage of it to make super-normal profits from the information that is not available readily to participants of the public when making their tradings.

It is key to note that there should be randomness within the changes of the expected prices of securities when trading in the market at the same time the changes in the mean returns are brought about by changes in interest rates as well as risk levels of individuals. According to [Fama, 1965], EMH can be categorized depending on the available information thus helping in making sound investment decisions during trade on the Nairobi Securities Exchange market.

1.2 Forms of Efficient Market Hypothesis

1.2.1 A Weak form of an Efficient Market Hypothesis

In the weak form of an efficient market hypothesis exist when the market price of an investment security incorporates all the information that is present in the historical prices of that particular investment. Should the weak form of EMH holds, then trading rules that are based only in the historical data about the price of the security should not be capable of generating higher investment returns for an investor. The above trading rule form that use historical data of NSE to predict the present or future prices that is referred to as technical or chartist analysis.

1.2.2 A Semi Strong form of an EMH

In this form, the market price of a security incorporates all the publicly available information. For market that is in the semi strong form of efficiency, fundamental analysis of the publicly all the available information should not be capable of generating higher expected return from investment. Fundamental analysis uses information of the issuer of the security such as profitability, liquidity, turnover and level of gearing together with other economic conditions such as inflation and real interest rates to establish the real value of security whether it is cheap or dear.

1.2.3 A Strong form of an EMH

In this form of the EMH, the market prices of the security incorporates both the privately and publicly all available information about a given stork or security. The private information is also referred to as the insider information. Strong of EMH holds if the insider information cannot be used in beating the traders in the market. From Strong form of EMH, many investors can make huge profits in terms of the privately available information that can be used when making the trades that would enable them make huge profits.

1.3 Exchange Traded Funds

An exchange-traded fund (ETF) is defined as an investment fund that is traded on the securities exchange market such as securities. An ETF holds many assets like stocks, bonds, commodities, and often operates within a mechanism of arbitrage designed to maintain it trading close to par value within the securities market called net asset value, with low levels of volatility existing between the trading period. Most of the ETFs often track the value of an index, like a stock index like NSE-20 Share Index or bond index see[Bae and Kim, 2020]. Many ETFs possess attractive investments features that makes them the best options for investors due to their low costs, high tax efficiency, as well as security or stock-like features when trading in the market.

ETF trader often buy or sell the respective ETFs directly from authorized dealers who may happen to be large broker-dealers capable of drafting agreements that leads to creation of units, comprising of large blocks of thousands of ETF stocks that are to be traded in the securities market. The authorized participants also may wish about making investments in the ETF shares especially for the long period of investment commonly act as market makers whenever they are on the open market as well as their capability of exchanging the created units using their underlying securities thus offering ETF shares liquidity while helping in making sure that the intraday market price do approximate the exact value of the net asset value traded on the underlying assets[Wang and Xu, 2019].

An ETF has a feature of combining the valuation feature of a unit investment trust or mutual fund that can be bought or sold when the market closes after the trading day for calculation of the net asset value, which is allowing traders to buy and sell the products during the period of investment.

An ETF just like any type of mutual fund, it can own assets such as bonds, gold bars, and stocks among others before dividing ownership into securities often held by the shareholders of the specific stocks that makes the company. In most of the times, the shareholders are authorized to a given amount profit in terms of shares or can just get either interest or dividends, after liquidation of the fund.

ETFs are much similar in several ways to traditional mutual funds, however, they can be either bought or sold on a daily such as stocks on a given stock exchange like NSE through an agency or dealer of stocks [Gastineau, 2008]. The capability of purchasing and redeeming creation units often provides an ETF that an arbitrage mechanism, which aims at minimizing the potential deviation in between the existing market price as well as net asset value of the respective ETF stocks. Current ETFs do have transparent portfolios, thus making institutional investors to understand the contents of portfolio assets to assemble when purchasing a creation unit, which will trade in the market.

1.4 Statement of the Problem

Long time ago investors could buy shares or sell them the same day. The above could pose a bigger threat as investors could assess the trend of the securities prices before predicting future prices thus getting chances of making arbitrage profits. The above was countered by the new policy where once shares were bought then they could only be sold after a fortnight. The main aim of the study is about testing the weak form of market efficiency in NSE by assessing whether a person can use the prices of the securities to predict future prices.

Firstly, the securities exchange market has been playing and will continue playing important virtues for the citizens who want to develop culture of thrifting and saving. Furthermore, the securities exchange market is assisting in the transfer of savings of investment in the productive enterprises as an alternative to keep their savings and make the money grow. When an investor makes the savings, he or she will invest the money into the economy that offers an opportunity to create wealth that ultimately make the economy grows.

Thirdly, any form of robust securities exchange market should assist in the efficient and rational allocation of capital that has been a scarce resource for those looking to make sound investment in the Kenyan economy. Moreover, a stock market should promote very high accounting standards of resource management as well as transparency during business management of the stock traded. This has a meaning that a stock exchange is an avenue of improving the finance access of different types of users through the provision of flexibility for trading customization. Lastly, a stock exchange market needs to offer all its investors with one efficient mechanism thus enhancing liquidity of their investments in market during securities trading.

1.5 Hypothesis of the Study

During the research, it is important to make the first hypothesis that the null hypothesis is the stock prices in the Nairobi Stock market do follow a Gaussian distribution and an alternative hypothesis that the stock prices in the Nairobi Stock market does not follow a Gaussian distribution. The second null hypothesis is that the stock prices are random during the study period against the alternative hypothesis of the stock prices are not random in this study period. This is important since modeling of the indices will be used when making forecasting of the future trends for investors.

Another key principle of the weak form of the EMH is in the randomness of securities prices thus making it improbable when finding price patterns as a way of taking advantage of movement of the individual prices. To be more specific, daily securities price fluctuations can sometimes be over entirely independent from each other; thus, making an assumption that price momentum in the market does not exist, which may be difficult during trading. Moreover, past earnings in terms of growth does not always forecast current or even the growth in future especially in terms of stock earnings.

1.6 Objectives of the study

1.6.1 General Objective

The general objective is this study is to test the weak form of market hypothesis of Exchange Traded Funds at Nairobi Securities Exchange market using NSE-20 share Index.

1.6.2 Specific Objectives

The Specific objectives of this study are:

1. To Test the weak form of market hypothesis of NSE-20 share Index at NSE
2. To Determine how the NSE-20 Share Index expresses its long memory
3. To Predict the future prices of NSE-20 Share securities at NSE

1.7 Significance of the Study

As a way of testing the weak form of an efficient market hypothesis that assists investors during the selection of portfolio and enables investors to have equal chances making profit from their investments. Market Efficiency Hypothesis has always been an essential concept for many investors looking forward when holding internationally portfolios in a diversified way. In addition, with increased investments movement globally across international boundaries thanks to the high levels of integration experienced in many economies of the world. Having a understanding of efficiency as an investor of the emerging markets can offer an individual a competitive advantage thus gaining more returns on investment within a shorter time period.

The Nairobi Securities Exchange Market offer investors both locally and internationally an opportunity to earn proceeds from any preferred security in terms of a company where they would like holding a stake through acquisition of the company. For example, an investor looking for ownership can purchase as many as possible shares that will offer them an opportunity to hold shareholding certificate thus participating in decision makings of the company. In addition, many developing economies in the world in terms of securities markets are now starting to get momentum that is reliable as well as profitable chance for the investors.

For any type of an investor looking for a ownership through the stock market there has only been a win or even lose position when doing trading in a given market. Thus, the efficient security market necessity is vital because in any form of an inefficient market, an investor may decide to generate abnormal profits by generating abnormal losses from poor decisions that is different from rest of market investors. When doing test for a given weak form of an efficient market hypothesis, this offers information that will help an investors in the portfolio selection while making sure that all investors get a similar chance of making profit from sale of ownership.

1.8 Limitations of the study

Testing of the weak form of an efficient market hypothesis helps many investors during the process of selection of portfolio and enables investors to have equal chances making profit from their investments. However, the concept of martingale that dictates that the presently available information on the securities prices have a greater impact on the prediction of future prices making historical data on a stock less significant in predicting the potential value of returns of the future. This makes the use of semi-strong as well as strong forms of the efficient market hypothesis have an advantage in testing the liquidity of the market especially for those investors who are looking for short term tradings on short term securities.

Weak form of efficiency also does not consider the concept of technical analysis as accurate or asserts that in some cases fundamental analysis may be flawed during analysis. It is therefore very difficult, as per weak form efficiency for an investor to outperform the market when making the short-term financial decisions. For instance, if an investor decides on type of efficiency, they will have to believe on the existence of no point in having an excellent financial expert or an experienced portfolio manager to help in making financial analysis. In the same way, investors may be advocating for weak form of efficiency with an assumption of randomly picking an investment from a portfolio of stock traded in a securities exchange market that will offer the same returns.

Chapter 2

Literature Review

The philosophy of an efficient market is mainly concerned with assessment of whether prices of securities at any particular time do reflect the information that is available for the public investors[Fama, 1965]. There has always been a natural mechanism through which the price competition among financial markets make the prices to converge to an efficient state. The convergence to efficient state is also caused by exploitation of arbitrage opportunities by the operators. Thus, with the market operators taking advantage of price differential, the above forces will always push the prices of the securities to their expected values. Thus, profit chances are also eliminated as the markets tend towards optimum price, which has a meaning that the specific market is therefore efficient.

The convergence mechanism offers information on how the market receives about the new information that would change on the prices. The speed of accessing the new information is faster due to advancement in technology and free accessibility of new information at costless or affordable price [Arouri et al., 2010]. Though empirical study has been a subject of huge stock market, the same is never said to be used in emerging markets [Jefferis and Smith, 2004]. Nevertheless, in the recent times new empirical methods that have been developed has provided a path through which analysis of efficiency can be done in both developed together with emerging markets [Shaker, 2013] in their growth capacity.

According to [Antoniou et al., 1997], conventional methods for testing of efficiency have just been developed to test markets that show high levels of liquidity, well-educated investors with top quality information from few other institutional impediments that has existed before. On the other extreme emerging markets have low liquidity, considerable volatility and investors are less informed as the access to information is unreliable. Thus according to [Arouri et al., 2010] emerging markets are less efficient as compared to developed market due to many market imperfections. That is the reason why latest studies on the emerging markets concentrate on weak form of market hypothesis while the studies done in the markets of developed countries on the study of all forms of the efficient market hypothesis.

Many factors often affects the inefficiency of the emerging market and they include discontinuous and less frequent trading, poor quality information, and inappropriate accounting regulations, restriction in capital flow and discrimination in taxation [Arouri et al., 2010]. Many researchers for instance, [Kisaka et al., 2008] has published on the response of emerging securities markets in the African setup while proposing many reforms implemented in the revitalizing process that captured efficiency of the market as well as volatility during 1988 year to December 1999, mainly for the Nairobi Stock Exchange market.

The implemented reforms included enactment of regulatory framework, automation of the trading process and relaxing trade restrictions on many foreign investors. From the investigation reports, the above researchers established that markets that utilizes modern technology, have tight regulatory systems and removed trade barriers for foreign investors have greater efficiency and reports lower market volatility. Thus, research showed that reforms reduce volatility while increasing market efficiency. According to the research done by[Mlambo et al., 2003], most of the securities markets around the world are working daily to ensure that they enhance market efficiency by passing of information at the same time making the information as present as possible to the large number of investors and introducing electronic trading process. The above reforms makes participants to have equal chances in accessing vital trading information.

Most researchers that have been conducted deals a lot on the work during testing of the weak form of the NSE. In the year 2007, [Mlambo and Biekpe, 2007] had studied the weak form of market efficiency in many of the African markets. They used variables such as daily closing security prices of stocks and even volume traded of the stocks for Kenyan market from Jan 1997 to May 2002. In their study, they used run tests together with serial correlation tests to test the dependence between the stock prices. They concluded that Kenyan market was found to be more efficient in the weak form as significant number precedented by the random walk.

[Green et al., 2002] did a paper that was investigating revitalization process enhancing stock market micro structure of NSE. They also tested the market efficiency. Data used in the investigation was the monthly data form Nairobi Stock exchange by looking at NSE-20 share index from the year January 1870 to December 1999. They used unit root test as well as serial correlation test. The researchers concluded that NSE is not efficient in the weak form see[Simons and Laryea, 2005] In addition, [Al-Khazali et al., 2007]did a research paper to ascertain that index of stock market price s did follow random walk process.

The research was done in the following countries Egypt, Nigeria, Mauritius, Morocco, Botswana and Kenya. The historical data used were from NSE 20 share index form the 3rd week of January to the final week of August 1998. In the methodology, they used multiple variance ratio test of the Chow and Denning to test independence of stock prices. The analyzed results showed that hypothesis of the random walk has been rejected the stock prices were auto-correlated. [Dickinson and Muragu, 1994], Investigating if behavior of Kenyan market was consistent with the weak form of EMH.

The paper used data from 30 top traded equity securities that are listed in the Kenya stock exchange market from 1979 to the year 1988. In their investigation , [Dickinson and Muragu, 1994] used serial correlation tests and even the run tests. The conclusion on that the Kenyan stock market is not efficient in the weak form. [Chesire, 2014] investigated weak form of market hypothesis by use of daily stock prices of Kengen for the time period 17th may 2006 to even December 2009 and also with Kenya power and Lighting from 2nd January 2002 to 31st December 2009. [Chesire, 2014] did use a serial correlation test, Run tests and Durbin Watson tests. The results showed that the NSE is not always efficient in the weak form. However the researchers did not tell us the reason for the choice of the Kengen and Kenya Power and Lightning companies. [Gil-Alana et al., 2015] has tested the inefficiency with long memory, anomalies and persistence of the NSE 20 share using date from 2001 to 2009. The above researchers used the concept of fractional integration, they concluded that there is evidence of long memory. [Vitali and Mollah, 2010] also carried a research using unit root, autocorrelation, variance ratio as well as runs test on the daily indices of Kenya with other six African countries between time periods 1999 to the year 2009. The results showed rejection of the random walk hypothesis for Kenya. Nearly all the above named researchers investigated EMH of NSE exchange through the application of unit root tests or tests that are related to random walk hypothesis [Ajao and Osayuwu, 2012]. In a deeper analysis , unit root case poses restrictions, , Thus this paper considers more flexible and general approach to study NSE 20 share index other than being restricted to $1(0)/I(1)$ dichotomy. Thus, in this research paper a possibility of fractional integration with "d" as a fraction is applied.

In regard to [Adelman et al., 1965], a high number of sum total economic time series take a similar shape as the spectral density that rises vertically as the frequency heads to zero. Nevertheless, continued differencing the given data may result to overdifferencing when the respective frequency is at zero. The aggregation causes of fractional integration when an AR is aggregated when it heterogeneous. [Gil-Alana and Robinson, 1997] this is because fractional process is applied in explaining the dynamics of monetary and economic time series. Through his paper, [Sowell, 1992] had modeled long-run behavior of stocks using the fractional ARIMA model, which makes it for those investors who would want to make investment opportunities in the Kenyan market based on the trends of the market that follows ARIMA model.

Chapter 3

Methodology

3.1 Introduction

Market efficiency hypothesis is made that the prices of securities are random at the same time that distribution is Gaussian distributed. This means that it is important in testing on whether the market exhibit the characteristics of being efficient[Dias et al., 2020]. The main advantages of this Gaussian distribution is that it has only two measures, which is mean and variance that will describe all the distribution. In addition, the distribution makes basic assumptions when modeled as underlying CAPM (capital asset pricing model). The histogram of all prices is computed with the curve for normal distributions fitted with an aim of ascertaining whether the respective distribution of all price values do fit the Gaussian distribution.

A Gaussian distribution, which is not symmetric has in many cases, or more often has had a tail that ends in the distribution known as skewed. In addition, when the tail is running towards much larger values, then it is safe to say that it skewed positively to the right and when the tail is towards those smaller values, the distribution is then negatively skewed towards the left. Kurtosis is a statistical measure that indicate the extent to which, for the standard deviation, recordings cluster do around a given central points[Drożdż et al., 2007]. If in any case in a distribution cluster more compared to those in the Gaussian distribution, the distribution is termed as a leptokurtic. On the other hand when cases cluster far less when compared in the Gaussian distribution, the distribution is known as platokurtic. The values for kurtosis as well as skewness are zero if all observed distributions then it is called a Gaussian distribution..

3.2 Nairobi Stock Exchange 20 Share Index

All indices in any market always shows how strong a stock market is for the investors who may be having interest to make investments whether locally or internationally. When an investor understand on how these indices works, he or she can use the information to look at the different securities that trades in the Nairobi Securities Exchange market before making a decision on what to do when deciding to make an investment. On the other hand, an AMP will arise whenever an investor has more information on the kind of investments to make especially when the market are down from the effects of Covid-19 pandemic that is now being experienced in the world.

The Table 3.1 below shows some of the commonly used NSE stock indices in the market.

Stock Index	Trading Price in Kshs.
NSE -20 Share Index	1,957.54
NSE-25 Share Index	3,298.34
NSE All Share Index	142.54
FTSE ASEA Pan African Index	176.39
FTSE NSE Kenya 15 Index	188.38
FTSE NSE Kenya 25 Index	923.62
FTSE NSE Kenya Govt. Bond Index	96.45

Table 3.1: All NSE Securities Indices as at 30th June 2020

Remark 1. The NSE 20 Share Index price is the most commonly used index since it captures the best performing companies in the NSE thus giving ideas to savvy investors in the market to learn when deciding on prudent financial decisions on where to put their money. In addition, it has a long historical data that is vital to modeling and prediction trends of the Nairobi Securities Exchange market for all stakeholders.

3.3 Mathematical Modeling of NSE-20 Share Index

It is vital to model NSE-20 index as follows with respect with time, t as follows:

$$\pi_t = \alpha^T Z_t + e_t \tag{3.3.1}$$

where $t = 1, 2, 3, ...$ and Z_t refers to a vector of deterministic terms that could join the line where $Z_t = 1$ while e_t error term is represented by that follows the form process below;

$$(1 - L)^c y_t = u_t \tag{3.3.2}$$

where $t = 1, 2, 3, ...$

In which $c \in \mathbb{R}$ is at a point when the equation is zero, which means that covariance stationary process of a spectral density function is greater than zero at the same time bounded at the zero frequency giving room for weak autocorrelation of the ARMA.

$$M(L)u_t = (L)U_t \tag{3.3.3}$$

In this case ML is an AR polynomial while (L) is an MA the error term becomes the white noise. $C = 0$ is when $I(0)$ representation; $c = 1$ refers to the unit root. This was established through the US series of macroeconomics which showed that only one that was not exposed to unit roots and not the common trend-stationary processes.

Setting c in equation (3.3.2) to be a fraction and possibly have the values greater than one while $(1-L)^c$ can be expressed through binomial expansion where for all $c \in \mathbb{R}$.

$$(1-L)^c = \sum_{k=0}^{\infty} (-1)^j L^j = \sum_{k=0}^{\infty} \begin{pmatrix} c \\ j \end{pmatrix} (-1)^j L^j = 1 - cL + \frac{c(c-1)}{2} L^2 \tag{3.3.4}$$

Therefore, it is easy to define the following equation 3.3.4 as follows;

$$(1-L)^c y_t = y_t - c y_{t-1} + \frac{c(c-1)}{2} e_{t-1} \tag{3.3.5}$$

In the equations (3.3.5), c serves an essential, it shows the degree of reliance of the series. When the c value is high the association level is also high. For instance, in equation (3.3.2) where $c > 0$ expresses the, “long memory” that is how its decay of its autocorrelation takes place and that the density function is bounded at origin.

3.4 Hypothesis Testing

Standard unit root formula is applied thereafter an estimation of the fractional differencing estimate by applying whittle function in all frequency domain. According to [Robinson, 1994] of the testing procedures drought about are applied within estimating if the actual value of c in $I(c)$ models. Lastly, Lagrange multiplier methods should be shown as the perfect one in the fractional integration. It is applied in testing for the null hypothesis $H_0 : c = c$ for numbers in equation 3.3.1 and 3.3.2 and various types of disturbances.

A confidence interval for true values can be constructed since it follows the normal distribution. There are additional parametric methods that can be employed like the maximum likelihood estimation (MLE) this is applied alongside the Detrended Fluctuations Analysis.

An example of $ARFIMA(p,c,q)$model can be given in the form:

$$V(L)(1-L)^c(e_t - r - s_t) = L(W_t) \tag{3.4.1}$$

where $W_t \sim iid(0,\sigma^2)$

In this case $V(L)$ is an AR of order p polynomial while (L) is an MA of order q. When the log NSE-20 is not stationary when performing the initial differencing, regression comes first. MLE is applied in determining parameters in equation (3.3.4) with an assumption of the error term being normally distributed. As a way of making a good choice, correct model several tests have to be done for a perfect choice.

3.5 Detrended Fluctuation Analysis (DFA) of NSE-20 Share Index

For it to be easy during investigation of the existence of long memory, there is need to check if trend exists. In case of the existence of trend in the NSE data, the Hurst did an analysis of rescaled range as well as other non-trending methods would easily give spurious results after the analysis. *DFA* is well developed robust method that has been used widely in ascertaining the scaling behavior of NSE noisy data with all the available diverse trends that exists during modeling.

For purposes of a record, Let $[Y(i), I = 1, 2, .., N]$ where the value of N is denoted as the record length. The procedure of *DFA* has the following 4 steps as;

Step One

There is a need to determine the profile of $Y(i)$

$Y(i) =$ sum from ($k = 1$ to i) of Yk–Y ,where Y is the mean of the record.

Step Two

The second step is to divide the profile $Y(i)$ into all $Ns = N/s$ boxes . in each of the boxes, we will fit all integrated time series by the use of polynomial function , $P_y(i)$ normally defined as the local trend. For the order of *DFA*, n order of polynomial function is then applied to the estimation of fitting. We will subtract the local trend in every box to obtain the detrended function of $Y_s(i)$.

Where ; $Ys(i) = Y(i)$–$P_y(i)$

Step Three

In each and every box of the size s, it is easy to determine the root mean square (RMS) fluctuation of $F(s)$ where $F(s) =$ square root ($1/Ns$ sum of $(Ys^2(i))$ from $i = 1$ to Ns

Step Four

The fourth step is to repeat the procedure for different box sizes, with different scales, In case there is a power-law relation in between values of $F(s)$ and s. Then

$$F(s) \sim\sim S \tag{3.5.1}$$

This is indicating the presence of the scaling property with the parameter (h), known as the scaling exponent or the fluctuation exponent is representing the correlation proportion of the NSE data. When looking at the correlation exponent Z_s , determined from the autocorrelation function, which is a similar estimation for the value of $F(s)$ is given by;

$$F(s) \sim\sim S^{1-\frac{z}{2}} \tag{3.5.2}$$

By comparing equation (3.5.1) and (3.5.2), we will be able to obtain;

$$Z = 2 - 2h \text{ for } 0 < Z < 1$$

The above is a brief certification of the relationship between Z and h . The analysis is done using R statistical programming language.

In the use of DFA method, a straight line in the log-log graph indicate the presence of statistical affinity. Often, the scaling exponent has been defined as the slope of the straight line fitting to the log-log graph. In fact, this exponent generalization is the Hurst exponent.

3.6 Forecasting of NSE-20 Share Index Prices Using Normal Distribution

A lognormal) distribution commonly known as log-normal distribution that is a continuous PDF of a specific random variable where natural logarithms of the individual random variables are normally distributed as [Rosenkrantz, 2003]. This means that when you take the natural logarithms of the random variables, the random variables will normally distributed having a mean of θ and variance of σ^2. Therefore, if a random variable X with log-normally distributed, then $Y = ln(X)$ follows a normal distribution. On the contrary side, if Y has a normal or Gaussian distribution, then the value of Z distribution of the exponential function , $X = exp(Z)$, has a log-normal distribution. The feature of a log normal distribution that makes it necessary when modeling the NSE-20 Share Index only takes positive values while taking care of the possibility of extreme events during trading in the financial markets.

Let Z to be defined as a standard Gaussian distribution with mean of 0 and variance of 1. This means that distribution of a random variable

$$Y = e^{\theta+\sigma Z} \tag{3.6.1}$$

where Y has a log-normal distribution with parameters mean of θ and variance of σ^2. Let the probability density function be denoted by $f(y)$, then then;

$$f(y) = \frac{d}{dy}(P[Y \leq y] = \frac{d}{dy}P[lnY = lny] = \frac{d}{dy}[\Phi(\frac{lny-\theta}{\sigma}] \tag{3.6.2}$$

$$f(y) = \zeta(\frac{lny-\theta}{\sigma}) * \frac{d}{dy}(\frac{lny-\theta}{\sigma})$$

$$f(y) = \zeta(\frac{lny-\theta}{\sigma}) * \frac{1}{dy}$$

$$f(y) = \zeta(\frac{lny-\theta}{\sigma}) * \frac{1}{dy}$$

$$f(y) = \frac{1}{y} * \frac{1}{\sigma\sqrt{2\pi}} exp\{(\frac{lny-\theta}{\sigma})\} \tag{3.6.3}$$

In terms of the cumulative density function,CDF, the log-normal distribution will be;

$$F(Y) = \int_{ally} f(y)dy$$

$$F(Y) = \Phi(\frac{lny-\theta}{\sigma}) \tag{3.6.4}$$

From equation (3.6.3), the mean as well as variance of the log-normal distribution will be given by;

$$mean = [E(y)] = e^{(\theta+\frac{\sigma^2}{2})} \tag{3.6.5}$$

$$variance = [E(y)]^2(e^{\sigma^2} - 1) \tag{3.6.6}$$

3.6.1 Return Modeling

The series of the returns for NSE-20 share Index can be defined as;

$$R_t = \frac{lnS_{t+1}}{lnS_t} \tag{3.6.7}$$

$$R_t = logS_{t+1} - lnS_t$$

where R_t is the return on a NSE-20 share Index at time t with the NSE-20 share Index at time $t+1$ for S_{t+1} and t for S_t respectively during the time of trade. It is important to note that the Return on NSE-20 share Index is modeled as log-normal distribution with the mean as well as variance of returns given above.

3.6.2 Non-normality Assumption Test

For purposes of testing for normality, we will assume that NSE- 20 share index Prices follows a Gaussian distribution with mean of θ and variance of σ^2. The pdf will be given by;

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} exp\{-\frac{1}{2}(\frac{lny-\theta}{\sigma})^2\} \tag{3.6.8}$$

The distribution of NSE-20 share Index, which has been derived from the NSE market, are assumed not to follow a normal distribution to take into consideration the presence of extreme events such the existing presence of Covid-19 that massively affects the behavioral of investors. Using the procedure of [Antoniou et al., 2004], which as assumption that all data available at NSE is not normal. We will obtain the MLEs of the normal distribution as well as their corresponding standard errors. The results are shown below;

Parameters	Estimates	Standard Error
θ	3888.886	0.0189152
σ	809.5395	0.0134072

Table 3.2: Maximum Likelihoods Estimates of the Normal Distribution

From the above Table 3.2, the estimation of the parameters are as follows; $\theta = 3888.886$ and $\sigma = 809.5395$ with the standard error of 0.0189152 and 0.0134072 respectively. While these are parameters for the normal distribution, we can use the same parameters to obtain the parameters of the log-normal distribution as in equation (0.5) and (0.6) respectively.

$$e^{(\theta+\frac{\sigma^2}{2})} = 3888.886 \tag{3.6.9}$$

$$[E(y)]^2(e^{\sigma^2} - 1) = 655354.3 \tag{3.6.10}$$

Solving equation 3.6.9 and 3.6.10 simultaneously;

$$(e^{\sigma^2} - 1) = \frac{655354.3}{3888.886^2} = 0.0433342$$

$$e^{\sigma^2} = 1.0433342 \Longrightarrow \sigma^2 = ln1.0433342$$

$$\sigma = \sqrt{0.04242} \Longrightarrow \sigma = 0.206 \tag{3.6.11}$$

Replacing back the value of σ to equation 3.6.9 to obtain the value of θ,

$$\theta = ln(3888.886) - \frac{0.206}{2}$$

$$\theta = 8.256 \tag{3.6.12}$$

From equation (3.6.11) and (3.6.12), we can derive the value of NSE-20 share Index follows a given log-normal distribution that has parameters of $\theta = 8.256$ and $\sigma =$ 0.206) i.e. $Y \sim LG(\theta = 8.256, \sigma = 0.206)$.

Chapter 4

Data Analysis

4.1 Introduction

It is essential to check for normality of data. Using the *qqplot* function in R statistical language of NSE-20 share Index data, it is easier when determining on whether the sets of data given forms a Gaussian distribution before setting of data can be used.

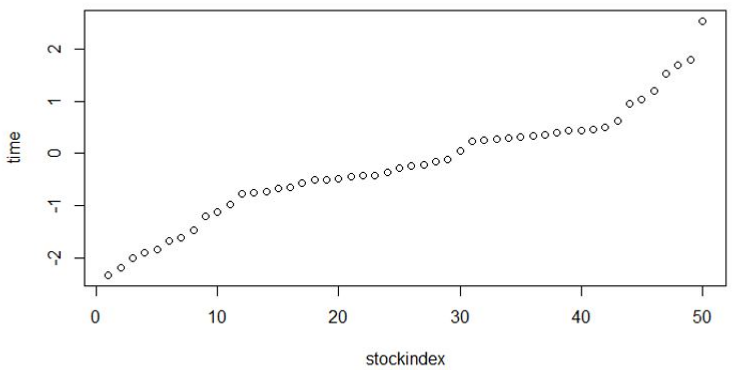

Figure 4.1.1: Test for Normality of NSE-20 Share Index

From Figure 4.1.1, it is easy to tell that the given data on NSE-20 share Index follows a Gaussian distribution that is important in modeling especially when looking for ways of forecasting for making investment decisions.

In addition, we can Test for Normality as follows;

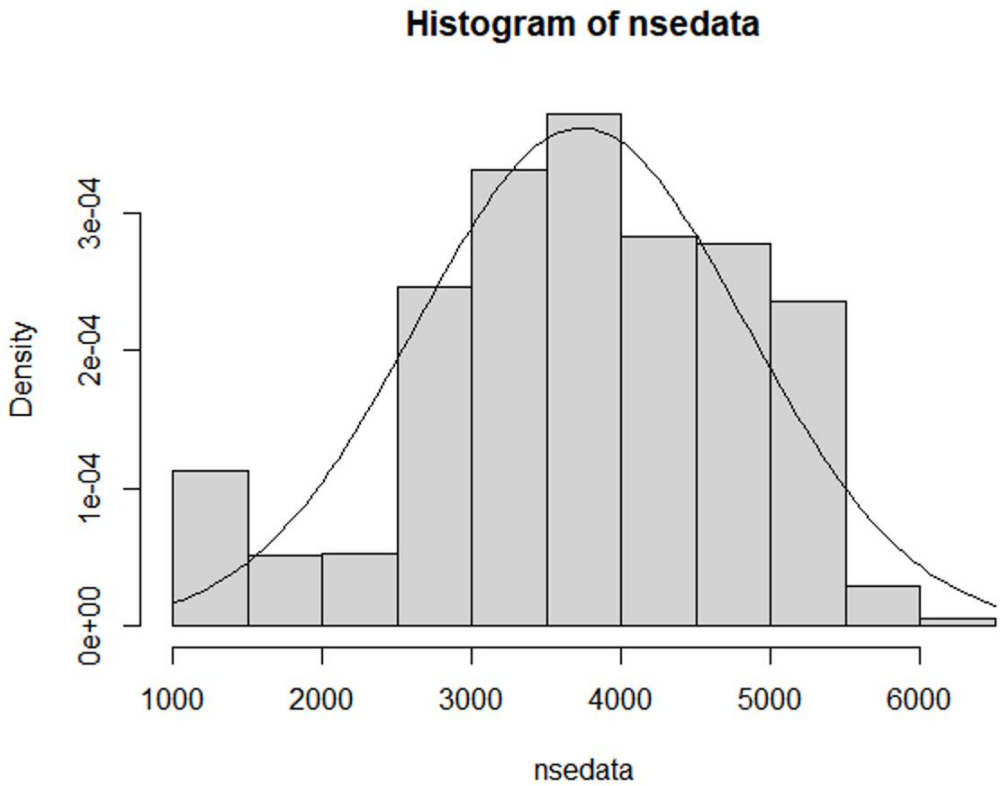

Figure 4.1.2: Testing for Normality of Log-Normal NSE Data

From Table 4.1.2, it is easy to note that getting the log Normal values of the NSE-20 Share Index Data becomes Normal, which makes it easy when modeling.

4.2 Data Analytics

4.2.1 Empirical Analysis

The past researchers have been testing existence of unit root alone. However, less research is done on testing the long memory. As a way of fully investigating the existence within the weak form of the efficient market hypothesis, this research will be testing both unit root test as well as long memory. NSE 20 share trend

The NSE trends from 2002 to the year 2019 has been depicted in the figure 4.2.1 below;

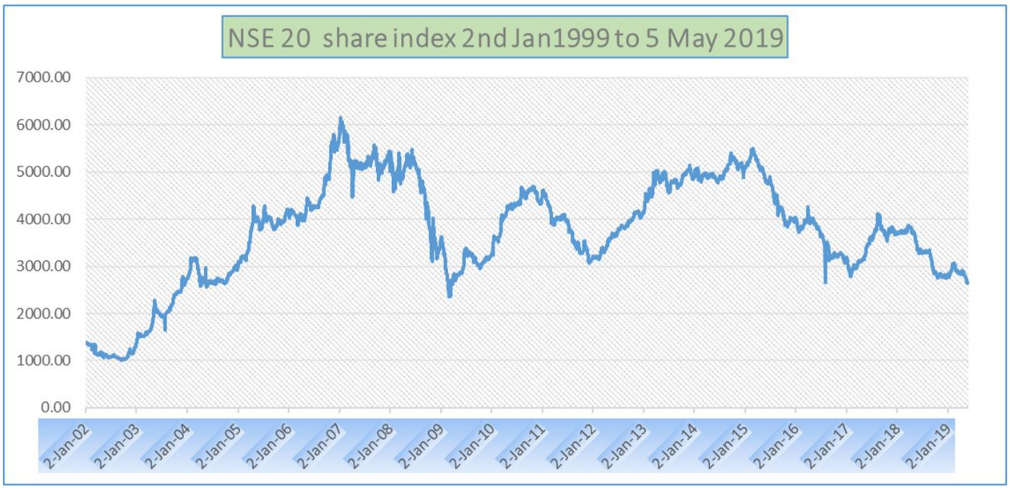

Figure 4.2.1: Trends of NSE-20 Share Index With Time

4.2.2 Stationarity of NSE-20 Share Index Data

Two periods determined as pre-crisis and crisis was included in this research. Before 2007/08 PEV and AFTER Stationarity tests/ unit root tests Non-stationary data, as a statistical rule, are unprecedented and cannot be predicted or modeled. Note that the first lag value is statistically important, whilst partial auto-correlations for all other lags are not statistically important. From the suggestions of a possible $AR(1)$ model for the NSE data, it is illustrated as below:

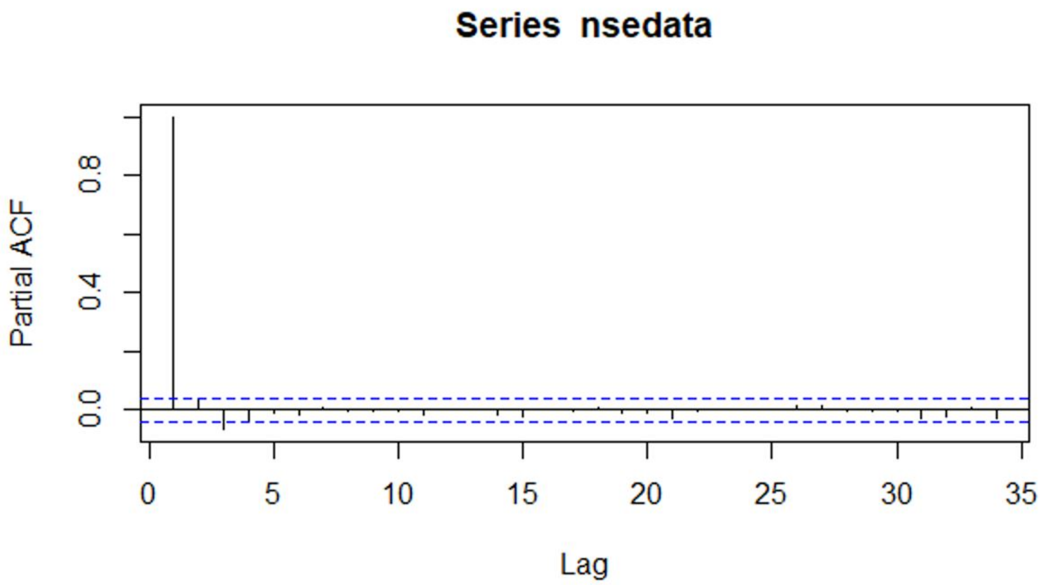

Figure 4.2.2: Series of the NSE-20 Share Index Data

The Stationarity data is shown in the figure 4.2.3 below as follows;

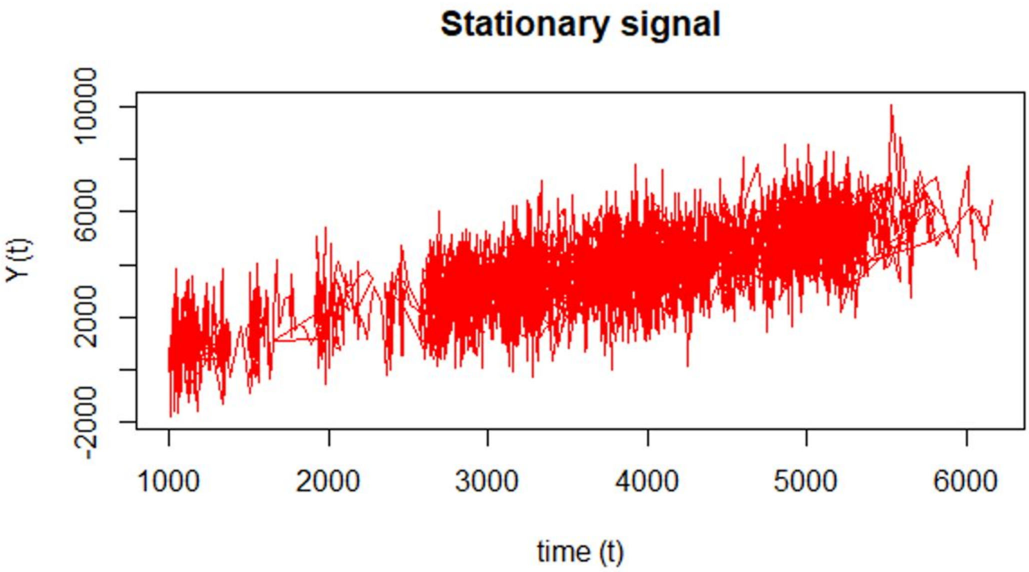

Figure 4.2.3: Stationarity of NSE-20 Share Index Data

4.2.3 Unit root tests

In this unit root test, the paper test for stationarity.

Data Type:	NSE Data	
Dickey-Fuller = -1.5944	Lag order of 3	p-value = 0.7501
Alternative hypothesis:	Stationary	
We will fail reject H_0 at 5%	hence there is unit root	

Table 4.1: Augmented Dickey-Fuller Test

We fail to reject H_0 at 5% since we do not have enough confidence at a level of 5% to make such statistical claims thus concluding that the NSE data exhibits the property of stationarity.

Phillips-Perron	Unit Root Test	
Data Type:	NSE Data	
Dickey-Fuller Z(alpha)= -4.3894	Truncation lag parameter of 10	p-value = 0.8649
Alternative hypothesis:	Stationary	

Table 4.2: Phillips-Perron Unit Root Test

From the Table 4.2, we can conclude that the NSE Data is stationary.

KPSS Test	for	Level Stationarity
Type of Data:	NSE Data	
KPSS Level = 7.8814,	Truncation lag parameter = 10	p-value = 0.01

Table 4.3: KPSS Test for Level Stationarity

The Test of type DF-GLS are shown in the Table 4.4 below as follows;

lm(formula data	= dfgls.form,	= data.dfgls)		
Residuals:				
Min	1Q	Median	3Q	Max
-840.40	-13.02	0.31	14.45	887.28
Coefficients:	Estimate	Std. Error	t value	$Pr(>\|t\|)$
yd.lag	-0.0001291	0.0002383	-0.542	0.588
yd.diff.lag1	0.0698108	0.0152783	4.569	$5.03e^{-06}$
yd.diff.lag2	0.1079643	0.0153079	7.053	$2.03e^{-12}$
yd.diff.lag3	0.0319363	0.0153081	2.086	0.037
yd.diff.lag4	-0.0145806	0.0152789	-0.954	0.340
Residual	standard error:	40.13 on 4283	degrees of freedom	
Multiple R-squared:	0.01956,	Adjusted R-squared:	0.01842	
F-statistic:	17.09 on 5 and	4283 DF,	p-value: $< 2.2e^{-16}$	
Value	of test-statistic	is: -0.5418		
Critical values	of DF-GLS are:	1pct=-2.57	5pct=, -1.94	10pct=-1.62

Table 4.4: Test of type DF-GLS

Remark 2. The test summary demonstrates that the test statistic is ranging from -1.2 to -1.2. Approximately 10% of critical values for the DF-GLS testing range from -2.57 to 2.57, which is not the suitable critical value when dealing with ADF test in case an intercept as well as time trend are all included within Dickey-Fuller regression. From the asymptotic distributions that demonstrates the testing statistics also unique when looking at all critical values used in the experiment. .

lm(formula data	= dfgls.form,	= data.dfgls)		
Residuals:				
Min	1Q	Median	3Q	Max
-840.40	-13.02	0.31	14.45	887.28
Coefficients:	Estimate	Std. Error	t value	$Pr(>\|t\|)$
yd.lag	-0.0001291	0.0002383	-0.542	0.588
yd.diff.lag1	0.0698108	0.0152783	4.569	$5.03e^{-06}$
yd.diff.lag2	0.1079643	0.0153079	7.053	$2.03e^{-12}$
yd.diff.lag3	0.0319363	0.0153081	2.086	0.037
yd.diff.lag4	-0.0145806	0.0152789	-0.954	0.340
Residual	standard error:	40.13 on 4283	degrees of freedom	
Multiple R-squared:	0.01956,	Adjusted R-squared:	0.01842	
F-statistic:	17.09 on 5 and	4283 DF,	p-value: $< 2.2e^{-16}$	
Value	of test-statistic	is: -0.5418		
Critical values	of DF-GLS are:	1pct=-2.57	5pct=, -1.94	10pct=-1.62

Table 4.5: Zivot-Andrews Unit Root Test

Since the test is left-sided from Table 4.5, it is improbable when rejecting the null hypothesis stating that the NSE-20 Share Index data has the property of stationary after applying the DF-GLS test.

Box-Ljung test
data: nse1
X-squared = 103384, df = 25, p-value less than $2.2e^{-16}$

Table 4.6: Box-Ljung test

The p-value of the NSE data is 2.2e-16, which means that there is non-zero autocorrelations evidence in the in-sample forecast errors defined at lags 1-20. When speaking quantitatively, one should apply the use of built-in test when testing stationariy. In addition, the use of Ljung-Box test helps in examining whether there is a significant evidence for all non-zero correlations at the stated lags (1 to 25 shown in the Figure 4.2.4 below). This comes with null hypothesis of independence in a state time series especially for a non-stationary signal that has lower p-value.

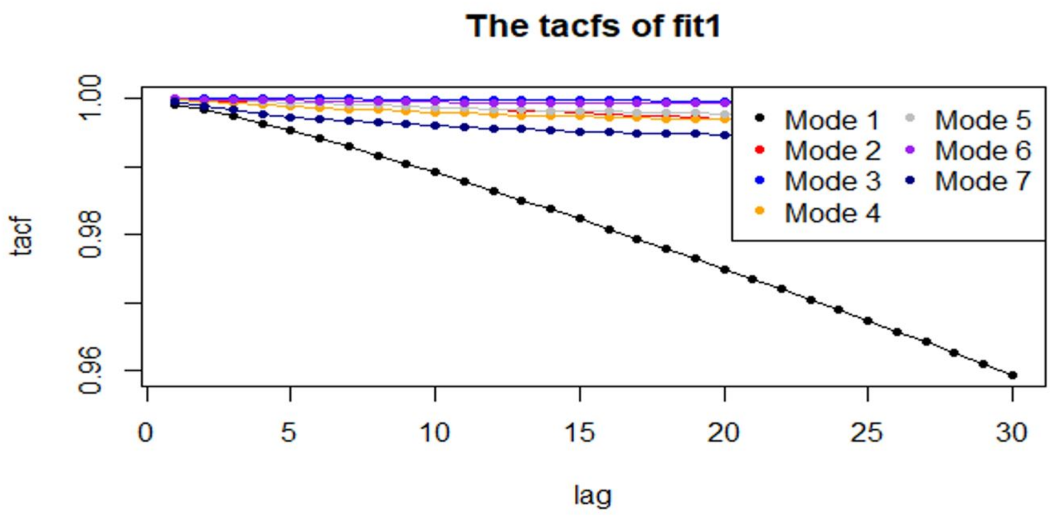

Figure 4.2.4: The Tacts of Fit1

The Tacts of Fit1 shows the trends during the period of forecasting for the NSE-20 share index that is helpful for decision makers looking for prudent financial decisions.

		Arfima	Analysis		
bHat	alphaHat	HHat	dHat	LL	convergence
0.010	0.010	0.995	0.495	-11135.286	0.000

Table 4.7: Arfima analysis

Remark 3. The list with components: bHat transformed optimal parameters alphaHat estimate of alpha HHat estimate of H dHat estimate of d phiHat estimate of phi thetaHat estimates of the theta will be optimized value of the Whittle estimated log-likelihood of $d = 0.495$.Thus, the value of $d > 0$ meaning there is long memory in the NSE 20 share index data.

4.2.4 DFA empirical Tests

Test for long memory empirical approximation of the confidence intervals linked to the Detrended Fluctuation Analysis (DFA)

	Detrended	fluctuation	analysis	for	NSE	Data			
H estimate:	0.5803184								
Domain:	Time								
Statistic:	RMSE								
Length of series:	2606								
Block detrending model:	x ~ 1 + t								
Block overlap fraction :	0								
Scale ratio:	2								
Scale	4.000	8.000	16.000	32.000	64.00	128.00	256.00	512.00	1024.00
RMSE	28.132	41.617	61.927	89.869	142.49	201.58	342.87	504.21	627.68

Table 4.8: Detrended fluctuation analysis for nsedata

From the Detrended Fluctuation Analysis (DFA) test, the NSE data fits well on the predicted values during the investment period.

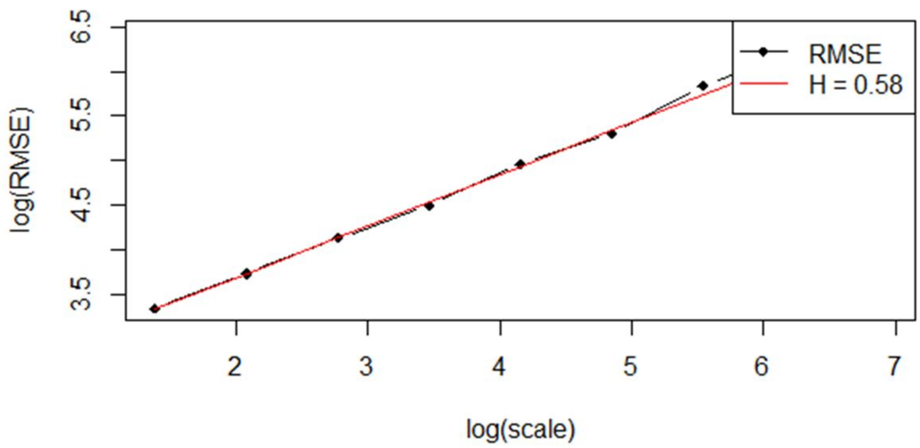

Figure 4.2.5: Root Mean Square Error Vs Time

The Root Mean Square Error (RMSE) commonly defined as the absolute standard deviation of the residuals (commonly known as prediction errors). The residuals are defined as measure of how far apart data points are from the regression line or how spread out the respective residuals are. This means that the measure will tell you on how congested the NSE data is around its best fit of line.

The Hurst exponent of H [3] has been used in some fields as well as its value that is linked to specified characteristic of a given independent stochastic process. The H value is also bounded between (0,1). If H is always equal to 0.5, which is the independent stochastic process does not demonstrate a long-term memory especially if H is greater than 0 .5 of the series is persistent at the same time the process is also linked by the trend reinforcing memory. Alternatively, if H is less than 0 .5 then the series is said to be anti-persistent.

4.3 Interpretation of Results

During fitting the NSE-20 Share Index Data, it is calculated using the estimated parameters as follows;

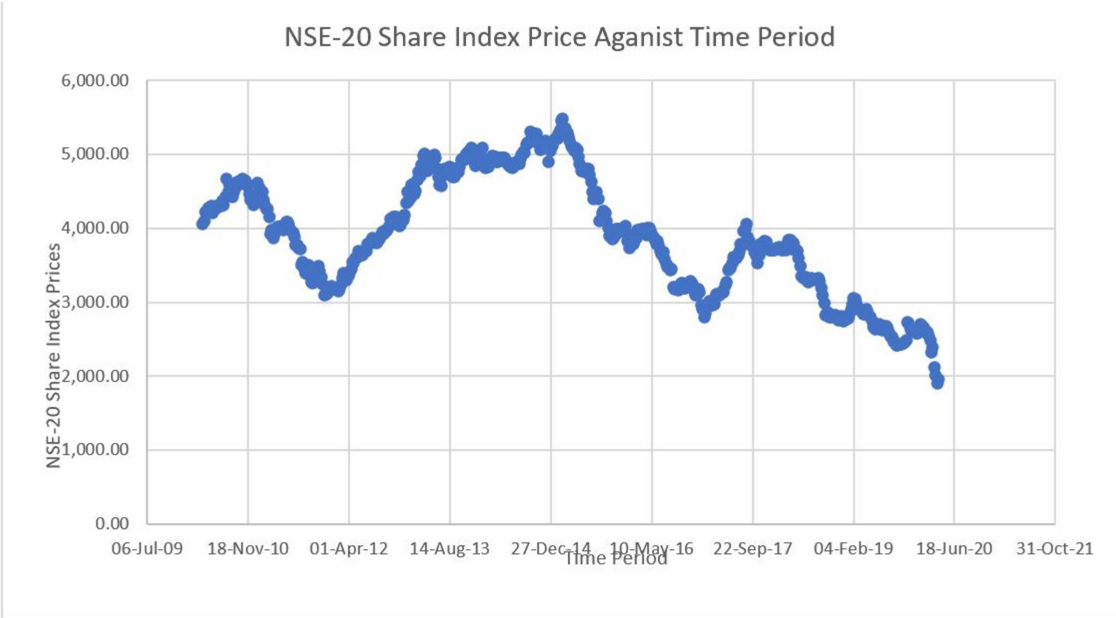

Figure 4.3.1: Fitting NSE-20 Share Index Trading Against Time

It is important to summarizes the relevant information on the empirical log-normal distributions of securities index returns under different consideration. The vital statistics that needs to reported includes mean, variance, standard deviation, minimum as well as maximum level of return during the trading period, kurtosis and coefficients of skewness.

Mean	Variance	Std	Max Return	Min Return	Skewness	Kurtosis
5988.565	655354.3	809.5395	5491.37	1917.67	-0.0381	2.03802

Table 4.9: Sample Moments Distributions of the NSE 20 Share Index

From statistics analysis from Gil-Alana et al. [2015], the third central moment is referred as an asymmetry measure or skewness within the specified distribution. From results illustrated in table 4.9, it is important to note the data distribution is not symmetrical as it is negatively skewed with the skewness coefficient of -0.0381. In addition, the kurtosis that measures of general peakedness or just flatness near the center of the distribution degree. The ratio is obtained by getting the ratio of the 4th central moment before dividing it with the variance square. For a normal distribution, the ratio should be equal to 3. Any ratio that is greater than 3 does indicate that more values within the neighborhood of its mean (which is more peaked compared to the normal or Gaussian distribution). From the analysis of our NSE-20 share index data, it is termed as a heavy tailed, which is an excess kurtosis. This is shown by the coefficient of kurtosis of 2.038016. We can conclude that our data is mesokurtic.

3.2 Non-normality Assumption Test

For purposes of testing for normality, we will assume that NSE- 20 share index Prices are following a Gaussian distribution with mean of θ and variance of σ^2. The pdf will be given by;

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} exp\{-\frac{1}{2}(\frac{lny-\theta}{\sigma})^2\} \tag{4.3.1}$$

The distribution of NSE-20 share Index derived from Kenyan Nairobi Securities Exchange market, are assumed not to follow a normal distribution to take into consideration the presence of extreme events such the existing presence of Covid-19 that massively affects the behavioral of investors. Using the procedure of [Rosenkrantz, 2003], which as assumption that all data available at NSE is not normal. We will obtain the MLEs of the normal distribution as well as their corresponding standard errors. The results are shown below;

Parameters	Estimates	Standard Error
θ	3888.886	0.0189152
σ	809.5395	0.0134072

Table 4.10: Estimates Parameters of the Normal Distribution

From the above table 2, the estimation of the parameters are as follows; $\theta = 3888.886$ and $\sigma = 809.5395$ with the standard error of 0.0189152 and 0.0134072 respectively. While these are parameters for the normal distribution, we can use the same parameters to obtain the parameters of the log-normal distribution as in equation (0.5) and (0.6) respectively.

$$e^{(\theta+\frac{\sigma^2}{2})} = 3888.886 \tag{4.3.2}$$

$$[E(y)]^2(e^{\sigma^2}-1) = 655354.3 \tag{4.3.3}$$

Solving equation 0.0.9 and 0.0.10 simultaneously;

$$(e^{\sigma^2}-1) = \frac{655354.3}{3888.886^2} = 0.0433342$$

$$e^{\sigma^2} = 1.0433342 \Longrightarrow \sigma^2 = ln1.0433342$$

$$\sigma = \sqrt{0.04242} \Longrightarrow \sigma = 0.206 \quad (4.3.4)$$

Replacing back the value of σ to equation 0.0.9 to obtain the value of θ,

$$\theta = ln(3888.886) - \frac{0.206}{2}$$

$$\theta = 8.256 \quad (4.3.5)$$

From equation (4.3.4) and (4.3.5), we can derive that the value of NSE-20 share Index is following a log-normal distribution with parameters, $\theta = 8.256$ and $\sigma = 0.206$) i.e. $Y \sim LG(\theta = 8.256, \sigma = 0.206)$.

On testing of the correlation between the different NSE-20 share Index prices during day tradings, the results are in the Table below:

Residuals				
Min	1Q	Median	3Q	Max
-49.852	-0.197	0.030	0.277	36.77
Coefficients	Estimate	Std.Error	t-value	$P(>\|t\|)$
Intercept	-0.3292837	0.5445662	-0.58340	0.59978
Open	1.0000781	0.00001419	7045.95	$< 2.23e^{-17}$
High				
Low				
Change %				
Residual	Error:	2.6 on 620	degrees	of Freedom
Multiple Error:	1	Adjusted R-squared:	1	
F-statistic:	$2.524e^{-7}$	on 2 and 620 df	p-value:	$2.22e^{-16}$

Table 4.11: Relationship between Daily Lows and Highs of the NSE-20 Share Index

3.3 Fitting a Log-Normal Distribution

With the data available at the NSE website, we are capable of fitting the data with the log-normal distribution after estimation of the parameters to obtain new fitted data. The data is then used to model the new prices of NSE-20 share index, which is following a long normal distribution with the estimated levels of volatility of the market compared to findings of [Vitali and Mollah, 2010] paper.

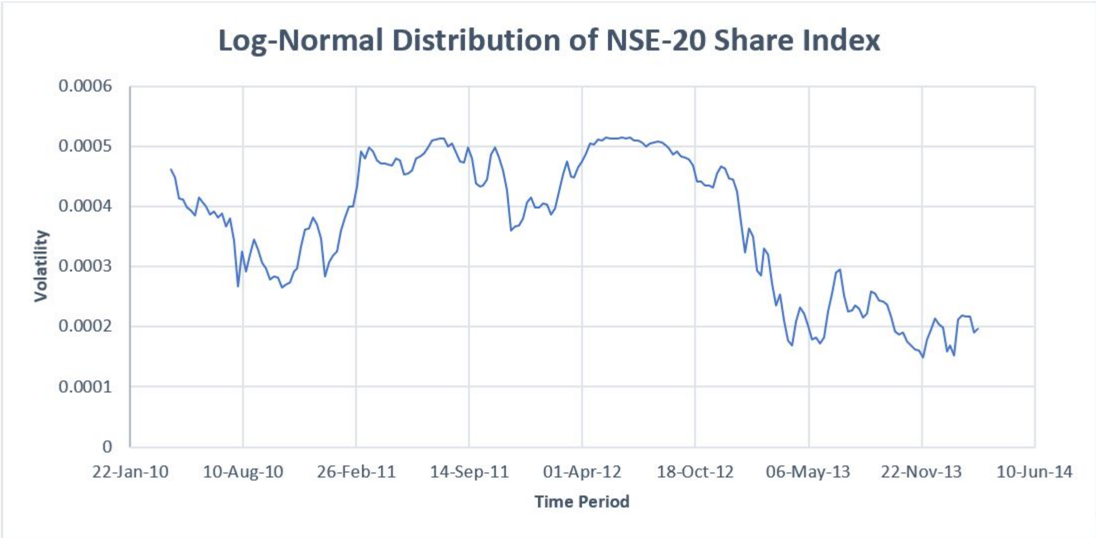

Figure 4.3.2: Estimated NSE-20 share Index volatility with Time

From Table 4.11, we can see that the daily volatility of the NSE-20 share Index is stochastic during the duration of trading within the market. The values changes daily as the value of the securities index changes during the trading periods with the mean volatility obtained in equation (3.6.11).

Chapter 5

Conclusions and Recommendations

5.1 Conclusions

From the test carried out, the results strongly indicate the fact of NSE is not efficient in the weak form because high volatility levels, delay in electronic trading when trading orders are made and poor information flow within the securities exchange market that can lead to insider tradings. The price changes of most of the securities are not independent thus making technical analysis to be more viable during the process. It offers a chance for investors to make preparations in advance for the NSE activities. The essential part of knowledge of unfolding in securities market is that during the period of all available information provided on several securities that are absorbed to reflect in the prices, thus investors will have purchased or sold securities by taking advantage of all existing favorable prices in the process.

The research project used a log-normal distribution, we were finding returns in terms of logarithms price index as opposed to the simple price index changes. The argument is that the natural logarithms of price is the return, that has a continuous compounding, which you earn when hold the specific security for a particular day.

The results from the research has shown from Figure 4.3.2 that variability of security price index changes that makes it an increasing price index function. However, by taking the natural logarithms can neutralize the effects on price index levels. From the descriptive analysis, the historical data of NSE-20 share price index data showing that it heavy tailed and non-symmetrical as assumed by most papers. With skewness coefficient of and low kurtosis, the data is mesokurtic which disputes the previous assumptions.

5.2 Recommendations

On the recommendations, a test on the semi-strong form of efficient market hypothesis can be done to ascertain whether the information that is available in the market shows the correct pricing of the prices of stock traded in the exchange market. This would surely help the investors make an informed decision from the NSE 20-share index whenever they want to make an investment from the market for possibility of getting higher returns. However, with strong form EMH, it is important that the traders take care of the insider trading that is a massive problem when dealing with the issues of trading. With no cases of insider trading, it would be possible to have an efficient market where the information available in the market reflects the demand and supply of traders thus the prices. With the assumption that volatility of the NSE-20 Share Index is constant over the time of trading is flawed since it changes over time when making investments.

From our results, we can affirm that log-normal distribution fits the NSE-20 share Index data when modeling the market trends especially when taking into considerations the cases of extreme events during the period of study. Since the trends of NSE-20 share Index are heavy tailed, it offers a clear evidence that the changes in price over a unit of time interval is not a constant variable. We can conclude that the future predictions of the NSE 20 Share Index returns should be based on a log normal distribution as opposed to a normal distribution. In addition, we recommend an investigation of the inverse gamma normal mixture distribution if it can best fit the returns of NSE 20 Share Index return as well as other normal mixtures.

5.3 Room for Further Research

One of the simple assumptions underlying is on the Random Walk theory and thus Efficient Market Hypothesis is stated that when securities prices are random then its distribution should be Gaussian distributed. Any Gaussian distribution is a huge relief as we need only two measures of mean as well as variance, when describing the entire distribution. The normal distribution is among the basic assumptions underlying during capital asset pricing model. All histogram prices of the prices are computed with the curve for normal distributions have been plotted in order to determining whether the price values distribution fits well the Gaussian distribution.

A non symmetric distribution has in many more cases have a tail moving towards one end of the distribution as opposed to other end known as called skewed. If the tail is towards larger values, the distribution is skewed to the right. If the tail is towards smaller values, the distribution is always skewed to the left. In addition, kurtosis always indicates an extent for a given standard deviation, recordings that cluster around a given central point. Whenever the cases cluster less compared to the case of Gaussian distribution, then it is simple to term the Gaussian distribution as platokurtic.

While relaxing some of these assumptions, it offers an opportunity for researchers to explore on how to model Exchange traded funds using statistical distributions and modeling them as heavy tail before testing them based on the information provided. The values for skewness as well as kurtosis are always zero whenever all observed Gaussian distribution is exactly normal which can be extrended to another distribution with a heavy tail property to take into consideration extreme events of worldwide pandemics such as Covid-19 pandemic. Modeling of NSE-20 share Index can also be modeled using other distributions that takes into account extreme events that can be experienced in stock market such as pandemics that has huge impact on ETFs.

References

Irma Adelman et al. Theories of economic growth and development. *Theories of economic growth and development.*, (2nd ed.), 1965.

Mayowa Gabriel Ajao and Richard Osayuwu. Testing the weak form of efficient market hypothesis in nigerian capital market. *Accounting and Finance Research*, 1(1): 169–179, 2012.

Osamah M Al-Khazali, David K Ding, and Chong Soo Pyun. A new variance ratio test of random walk in emerging markets: a revisit. *Financial review*, 42(2):303–317, 2007.

Antonios Antoniou, Nuray Ergul, and Phil Holmes. Market efficiency, thin trading and non-linear behaviour: evidence from an emerging market. *European Financial Management*, 3(2):175–190, 1997.

I Antoniou, Vi V Ivanov, Va V Ivanov, and PV Zrelov. On the log-normal distribution of stock market data. *Physica A: Statistical Mechanics and its Applications*, 331 (3-4):617–638, 2004.

Mohamed El Hedi Arouri, Fredj Jawadi, and Duc Khuong Nguyen. *The Dynamics of Emerging Stock Markets*. Springer, 2010.

Kyoung-hun Bae and Daejin Kim. Liquidity risk and exchange-traded fund returns, variances, and tracking errors. *Journal of Financial Economics*, 2020.

Charles Cao, Bing Liang, Andrew W Lo, and Lubomir Petrasek. Hedge fund holdings and stock market efficiency. *The Review of Asset Pricing Studies*, 8(1):77–116, 2018.

EJ Chesire. Testing the weak form of efficient market hypothesis at nairobi stock exchange. *Nairobi: University of Nairobi*, 2014.

Thomas E Copeland, John Fred Weston, Kuldeep Shastri, et al. *Financial theory and corporate policy*, volume 4. Pearson Addison Wesley Boston, 2005.

Rui Dias, Paula Heliodoro, Nuno Teixeira, and Teresa Godinho. Testing the weak form of efficient market hypothesis: Empirical evidence from equity markets. *International Journal of Accounting, Finance and Risk Management*, 5(1):40, 2020.

John P Dickinson and Kinandu Muragu. Market efficiency in developing countries: A case study of the nairobi stock exchange. *Journal of Business Finance & Accounting*, 21(1):133–150, 1994.

S Drożdż, M Forczek, J Kwapień, P Oświe, R Rak, et al. Stock market return distributions: From past to present. *Physica A: Statistical Mechanics and its Applications*, 383(1):59–64, 2007.

Eugene F Fama. The behavior of stock-market prices. *The journal of Business*, 38(1): 34–105, 1965.

Gary L Gastineau. Exchange-traded funds. *Handbook of finance*, 1, 2008.

Luis A Gil-Alana and Peter M Robinson. Testing of unit root and other nonstationary hypotheses in macroeconomic time series. *Journal of Econometrics*, 80(2):241–268, 1997.

Luis A Gil-Alana, Shinhye Chang, Mehmet Balcilar, Goodness C Aye, and Rangan Gupta. Persistence of precious metal prices: A fractional integration approach with structural breaks. *Resources Policy*, 44:57–64, 2015.

Christopher J Green, Victor Murinde, and Rose W Ngugi. Does the revitalisation process really enhance stock market microstructure? evidence from the nairobi stock exchange. *African Finance Journal*, 4(1):32–61, 2002.

Keith Jefferis and Graham Smith. Capitalisation and weak-form efficiency in the jse securities exchange. *South African Journal of Economics*, 72(4):684–707, 2004.

Peter Kimani Kairu. *Some legal aspects of the Nairobi Stock Exchange*. PhD thesis, 1976.

Evans Kirui, Nelson HW Wawire, and Perez O Onono. Macroeconomic variables, volatility and stock market returns: A case of nairobi securities exchange, kenya. *International Journal of Economics and Finance*, 6(8):214–228, 2014.

Sifunjo E Kisaka, Ngugi W Rose Sifunjo, Pokhariyal Ganesh, and Wainaina Gituro. An analysis of the efficiency of the foreign exchange market in kenya. *Economics Bulletin*, 14(2):1–13, 2008.

Chipo Mlambo and Nicholas Biekpe. The efficient market hypothesis: Evidence from ten african stock markets. *Investment Analysts Journal*, 36(66):5–17, 2007.

Chipo Mlambo, EV d M Smit, and N Biekpe. Testing the random walk hypothesis on thinly-traded markets: the case of four african stock markets. *African finance journal*, 5(1):16–35, 2003.

Muinde Patrick Mumo et al. Effects of macroeconomic volatility on stock prices in kenya: A cointegration evidence from the nairobi securities exchange (nse). *International Journal of Economics and Finance*, 9(2):1–14, 2017.

Joab Odhiambo, Patrick Weke, and Philip Ngare. Modeling kenyan economic impact of corona virus in kenya using discrete-time markov chains. *Journal of Finance and Economics*, 8(2):80–85, 2020a.

Joab Odhiambo, Patrick Weke, and Jusper Wendo. Modeling of returns of nairobi securities exchange 20 share index using log-normal distribution. 2020b.

Joab O Odhiambo, Philip Ngare, Patrick Weke, and Romanus Odhiambo Otieno. Modelling of covid-19 transmission in kenya using compound poisson regression model. *Journal of Advances in Mathematics and Computer Science*, pages 101–111, 2020c.

Cliff Osoro and A Jagongo. Investors perceptives on the nasi and the nse 20 share index as performance measurement indicators at the nairobi securities exchange in kenya. *International Journal of Humanities and Social Science*, 3(18):153–162, 2013.

Peter M Robinson. Efficient tests of nonstationary hypotheses. *Journal of the american statistical association*, 89(428):1420–1437, 1994.

Dominik M Rösch, Avanidhar Subrahmanyam, and Mathijs A Van Dijk. The dynamics of market efficiency. *The Review of Financial Studies*, 30(4):1151–1187, 2017.

Walter A Rosenkrantz. Why stock prices have a lognormal distribution. *Department of Mathematics and Statistics, University of Massachusetts at Amhers*, 2003.

Abu Towhid Muhammad Shaker. Testing the weak-form efficiency of the finnish and swedish stock markets. *European Journal of Business and Social Sciences*, 2(9): 176–185, 2013.

Daniel N Simons and Samuel Laryea. Testing the efficiency of african markets. *Available at SSRN 874808*, 2005.

Fallaw Sowell. Modeling long-run behavior with the fractional arima model. *Journal of monetary economics*, 29(2):277–302, 1992.

Federica Vitali and S Mollah. Stock market efficiency in africa: Evidence from random walk hypothesis. *South Western Finance*, pages 1–54, 2010.

Hua Wang and Liao Xu. Do exchange-traded fund flows increase the volatility of the underlying index? evidence from the emerging market in china. *Accounting & Finance*, 58(5):1525–1548, 2019.

Appendices

R Codes

```
library("readxl")
setwd("C:\\Users\\Home user\\Documents\\cavin project")
NSE <- read.csv(file = "nse.csv")
NSE
View(NSE)
mean(NSE$NSE.20,na.rm = TRUE)
var(NSE$NSE.20,na.rm = TRUE)
nsedata=na.omit(NSE$NSE.20)
nsedata
View(nsedata)
mean(nsedata)
var(nsedata)
plot(nsedata)
fit=arfima(nsedata)
fit
library(arfima)
p <- arfima(nsedata, order = c(1, 0, 1), numeach = c(3, 3))
p
u=arfima(nsedata, order = c(0, 0, 0), lmodel = c("FD", "FGN", "PLA", "NONE"))
acff=acf(nsedata, maxlag = 8)
pacfRes <- pacf(nsedata)
args(stationaryTest)
library(hwwntest)
hwwn.test(nsedata)
```

```
hwwn.test(nsedata, lowlev = 0, plot.it = FALSE, stopeveryscale = FALSE, n.cdf.grid = 1000, mc.method = p.adjust.methods, mac.spread=10)
library(beyondWhittle)
gibbs_ar(data=nsedata)
#fractinal#
library(fracdiff)
library(LongMemoryTS)
T<-4294 d<-c(0.4, 0.2, 0.3)
data<-FI.sim(T, q=3, rho=0, d=d)
X=nsedata
Y<-data[,3] cper<-cross.Peri(X, Y)
pmax<-max(Re(cper),Im(cper))
pmin<-min(Re(cper),Im(cper))
plot(Re(cper[1,,]), type="h", ylim=c(pmin,pmax))
lines(Im(cper[1,,]), col=2)
plot(Re(cper[2,,]), type="h", ylim=c(pmin,pmax))
lines(Im(cper[2,,]), col=2)
T<-4294
d=c(-0.8:10)
d
series<-fracdiff.sim(n=T, d=d)$series
local.W(series,m=floor(1+T^0.65))
gibbs_ar(data, ar.order, Ntotal, burnin, thin = 1, print_interval = 500, numerical_thresh = 1e-07, adaption.N = burnin, adaption.batchSize = 50, adaption.tar = 0.44, full_lik = F, rho.alpha = rep(1, ar.order), rho.beta = rep(1, ar.order), sigma2.alpha = 0.001, sigma2.beta = 0.001) library(tseries)
adf.test(nsedata)
adf.test(nsedata, nlag = 6, output = TRUE)
lm(formula = testmat)
library(testmat)
ur.za(nsedata, model = c("intercept", "trend", "both"), lag=NULL)
new 2
3/4 "q()" 3/5
median(3,5,6)
```

```
x=c(20,3,10,5,6)
x
mean(x)
var(x)
median(x)
plot(x,type="l", col="red",main="marks",ylab="stream",xlab="new index")
hist(x)
barplot(x, main="marks", ylab="cat marks", xlab = "stream") lab=c("j","t","k","m","o")
lab
pie(x, labels =lab,main= "marks")
set.seed(6533)
sim <- arfima.sim(1000, model = list(phi = .2, dfrac = .3, dint = 2))
fit <- arfima(x, order = c(1, 2, 0)) fit
library("readxl")
setwd("C:\\\Users\\Home user\\Documents\\cavin project")
NSE <- read.csv(file = "nse.csv")
NSE View(NSE)
mean(NSE$NSE.20,na.rm = TRUE)
var(NSE$NSE.20,na.rm = TRUE) nsedata=na.omit(NSE$NSE.20)
nsedata View(nsedata)
mean(nsedata)
var(nsedata)
plot(nsedata)
fit=arfima(nsedata)
fit library(arfima)
p <- arfima(nsedata, order = c(1, 0, 1), numeach = c(3, 3)) p u=arfima(nsedata, order =
c(0, 0, 0), lmodel = c("FD", "FGN", "PLA", "NONE"))
acff=acf(nsedata, maxlag = 8)
pacfRes <- pacf(nsedata)
args(stationaryTest)
library(hwwntest)
hwwn.test(nsedata)
hwwn.test(nsedata, lowlev = 0, plot.it = FALSE, stopeveryscale = FALSE, n.cdf.grid
= 1000, mc.method = p.adjust.methods, mac.spread=10) library("SciViews")
```

```
data=ln(nsedata)
data
library(hwwntest)
d<-c(0.4, 0.2, 0.3)
data=data
X<-data[,1:2]
Y<-data[,3]
cper<-cross.Peri(X, Y)
pmax<-max(Re(cper),Im(cper))
pmin<-min(Re(cper),Im(cper))
plot(Re(cper[1,,]), type="h", ylim=c(pmin,pmax))
lines(Im(cper[1,,]), col=2) plot(Re(cper[2,,]), type="h", ylim=c(pmin,pmax)) l
ines(Im(cper[2,,]), col=2)
library(beyondWhittle)
gibbs_ar(data=nsedata)
#fractinal#
ibrary(fracdiff)
library(LongMemoryTS)
T<-4294 d<-c(0.4, 0.2, 0.3)
data<-FI.sim(T, q=3, rho=0, d=d)
X=nsedata
Y<-data[,3]
cper<-cross.Peri(X, Y)
pmax<-max(Re(cper),Im(cper))
pmin<-min(Re(cper),Im(cper))
plot(Re(cper[1,,]), type="h", ylim=c(pmin,pmax))
lines(Im(cper[1,,]), col=2)
plot(Re(cper[2,,]), type="h", ylim=c(pmin,pmax))
lines(Im(cper[2,,]), col=2)
T<-4294
d=c(-0.8:10)
d
series<-fracdiff.sim(n=T, d=d)$series
```

```
local.W(series,m=floor(1+T^0.65))
gibbs_ar(data, ar.order, Ntotal, burnin, thin = 1, print_interval = 500, numerical_thresh = 1e-07, adaption.N = burnin, adaption.batchSize = 50, adaption.tar = 0.44, full_lik = F, rho.alpha = rep(1, ar.order), rho.beta = rep(1, ar.order), sigma2.alpha = 0.001, sigma2.beta = 0.001) library(tseries)
adf.test(nsedata)
adf.test(nsedata, nlag = 6, output = TRUE)
lm(formula = testmat)
library(testmat)
ur.za(nsedata, model = c("intercept", "trend", "both"), lag=NULL)
session 5
3/4
"q()" 3/5
median(3,5,6)
x=c(20,3,10,5,6)
x
mean(x)
var(x)
median(x)
plot(x,type="l", col="red",main="marks",ylab="stream",xlab="new index")
hist(x)
barplot(x, main="marks", ylab="cat marks", xlab = "stream")
lab=c("j","t","k","m","o")
lab
pie(x, labels =lab,main= "marks")
set.seed(6533)
sim <- arfima.sim(1000, model = list(phi = .2, dfrac = .3, dint = 2))
fit <- arfima(x, order = c(1, 2, 0))
fit
library("readxl")
setwd("C:\\Users\\Home user\\Documents\\cavin project")
NSE <- read.csv(file = "nse.csv")
NSE
View(NSE)
```

```
mean(NSE$NSE.20,na.rm = TRUE)
var(NSE$NSE.20,na.rm = TRUE)
nsedata=na.omit(NSE$NSE.20)
nsedata View(nsedata)
mean(nsedata)
var(nsedata)
plot(nsedata)
fit=arfima(nsedata)
fit
library(arfima)
arfima0(nsedata, order = c(0, 0, 0), lmodel = c("FD", "FGN", "PLA", "NONE"))
fit2 <- arfima(nsedata, order = c(2, 0, 1), back=TRUE)
fit2
distance(fit2)
fit1 <- arfima(nsedata, order = c(1, 0, 1), numeach = c(3, 3), dmean = FALSE)
fit1
plot(tacvf(fit1), maxlag = 30, tacf = TRUE)
p <- arfima(nsedata, order = c(1, 0, 1), numeach = c(3, 3))
p
u=arfima(nsedata, order = c(0, 0, 0), lmodel = c("FD", "FGN", "PLA", "NONE"))
acff=acf(nsedata, maxlag = 8)
pacfRes <- pacf(nsedata)
args(stationaryTest)
library(hwwntest)
hwwn.test(nsedata)
hwwn.test(nsedata, lowlev = 0, plot.it = FALSE, stopeveryscale = FALSE, n.cdf.grid
= 1000, mc.method = p.adjust.methods, mac.spread=10) library("SciViews")
data=ln(nsedata)
data
library(hwwntest)
d<-c(0.4, 0.2, 0.3) data=data
X<-data[,1:2]
Y<-data[,3]
```

```
cper<-cross.Peri(X, Y)
pmax<-max(Re(cper),Im(cper))
pmin<-min(Re(cper),Im(cper))
plot(Re(cper[1,,]), type="h", ylim=c(pmin,pmax))
lines(Im(cper[1,,]), col=2)
plot(Re(cper[2,,]), type="h", ylim=c(pmin,pmax))
lines(Im(cper[2,,]), col=2)
library(beyondWhittle)
gibbs_ar(data=nsedata)
#fractinal#
library(fracdiff) l
ibrary(LongMemoryTS)
T<-4294 d<-c(0.4, 0.2, 0.3)
data<-FI.sim(T, q=3, rho=0, d=d)
X=nsedata
Y<-data[,3] cper<-cross.Peri(X, Y)
pmax<-max(Re(cper),Im(cper))
pmin<-min(Re(cper),Im(cper))
plot(Re(cper[1,,]), type="h", ylim=c(pmin,pmax))
lines(Im(cper[1,,]), col=2)
plot(Re(cper[2,,]), type="h", ylim=c(pmin,pmax))
lines(Im(cper[2,,]), col=2)
T<-4294 d=c(-0.8:10)
d series<-fracdiff.sim(n=T, d=d)$series
local.W(series,m=floor(1+T^0.65))
gibbs_ar(data, ar.order, Ntotal, burnin, thin = 1, print_interval = 500, numerical_thresh = 1e-07, adaption.N = burnin, adaption.batchSize = 50, adaption.tar = 0.44, full_lik = F, rho.alpha = rep(1, ar.order), rho.beta = rep(1, ar.order), sigma2.alpha = 0.001, sigma2.beta = 0.001) library(tseries)
adf.test(nsedata)
adf.test(nsedata, nlag = 6, output = TRUE)
View(nsedata)
lm(formula = testmat)
library(testmat)
```

```
ur.za(nsedata, model = c("intercept", "trend", "both"), lag=NULL)
library("fractal")
DFA(nsedata, detrend="poly1", sum.order=0, overlap=0, scale.max=trunc(length(nsedata)/2),
scale.min=NULL, scale.ratio=2, verbose=FALSE)
```