UNIVERSITY OF NAIROBI

ON STOCK MARKET DYNAMICS AND OPTIONS PRICING BASED ON GARCH AND REGIME SWITCHING MODELS

BY

SEBASTIAN KAWETO KALOVWE

REG.NO: I80/52788/2018

A Research Thesis Submitted in Fulfillment of the Requirements for Award of the Degree of Doctor of Philosophy in Financial Mathematics of the University of Nairobi

2022

DECLARATION

I, the undersigned, declare that this thesis is my original work and has not been submitted elsewhere for examination, award of degree or publication. Where other people's work has been used, this has properly been acknowledged and referenced in accordance with the University of Nairobi's requirements.

Signature: ______________________ Date: 22-11-2022

Sebastian Kaweto Kalovwe

Reg. No I80/52788/2018

Department of Mathematics

Faculty of Science and Technology

University of Nairobi

This thesis is submitted for examination with our approval as research Supervisors.

	Signature	Date
Prof. Joseph Ivivi Mwaniki	______________________	24-11-2022

Department of Mathematics

University of Nairobi

P.O Box 30197-00100

Nairobi-Kenya

jimwaniki@uonbi.ac.ke

Prof. Richard Onyino Simwa	______________________	23-11-2022

Department of Accounting, Finance and Economics

KCA University

P.O Box 56808-00200

Nairobi-Kenya

rsimwa@kca.ac.ke

DEDICATION

I would like to dedicate this work to my beloved wife, Lucy, and children, Bosco, Mary and Grace, who have continued to bear with my busy schedule throughout the entire period of research. Their Support both financially and otherwise has been immeasurable.

ACKNOWLEDGEMENTS

I would like to thank Almighty God for giving me the mental and physical strength and above all his protection and care throughout the research. I sincerely appreciate my supervisors, Prof Joseph Ivivi Mwaniki and Prof Richard Simwa. Their constant supervision, diligent and objective criticisms in a great way made this research a success. Without their guidance, this fete would not have been achieved. Most importantly, my gratitude extends to my beloved wife, Lucy, and children, Bosco, Mary and Grace for their support and understanding.

ABSTRACT

The understanding of the linkage between stock returns, volatility and trading volume is paramount since it provides insights into the financial markets' micro-structure. The available literature reveals insufficient studies into modeling this correlation and most empirical studies have largely focused on developed markets than on emerging markets. GARCH model and its extensions have been utilized to model this relationship and to reproduce stylized features of financial time series. However, the model does not adequately describe the persistence of the financial markets' volatility. A model that can permit GARCH parameters to shift across regimes according to a Markov chain process is considered the solution to this problem, thus an attempt of this study is to put forward a regime-switching framework for modeling asset returns dynamics. The aim is to probe the dynamic correlation between stock returns, volatility and trade volume of both emerging and developed markets. In addition, the consequence of adding trade volume to the conditional variance equation of GARCH on volatility persistence is investigated. GARCH and regime-switching (RS) models are utilized to explore the link between stock returns, volatility and trade volume. The RS model is able to capture the structural changes in the variance process across regimes and its use extended to pricing European call options. The model is adapted to include GARCH effects and further implemented to pricing European call options. The estimated call options are compared with the corresponding Black-Scholes(B-S)' model estimates to establish the model with the best fit. The results reveal well-known features such as volatility clustering, heavy tails, leverage effects and a leptokurtic distribution. The developed markets are described with high volatility clustering and persistence compared with the emerging market and the volatility persistence is observed to decrease as the data changes frequency from daily to weekly. Furthermore, the volatility persistence is observed to dwindle after trade volume is included into the conditional variance equation of GARCH model. However, as the data frequency shifts from daily to weekly, mixed results emerge. The stock returns and volatility from the developed markets have a negative correlation, but the correlation in emerging market is positive. In addition, all the stock returns indices are characterized by regime shifts with heterogeneous conditional volatility, volatility clustering and varying responses to past negative returns. Furthermore, the volatility process stays longer in the high volatility regime of the developed markets before switching to regime 2 compared to the duration of stay in the same regime of the emerging markets. Finally, RS-GARCH model presents the best results when fitted to long-dated options data as compared to RS and B-S models whereas B-S model presents the best fit for short-dated options.

TABLE OF CONTENTS

List of Tables

List of Figures

LIST OF ABBREVIATIONS AND ACRONYMS

ADF	Argumented Dickey-Fuller
AR	Autoregressive
ARMA	Autoregressive Moving average
ARCH	Autoregressive Conditional Heteroscedasticity
BS	Black Scholes
EGARCH	Exponential Generalized Autoregressive Conditional Heteroscedasticity
FTSE	Financial Times Stock Exchange
GARCH	Generalized Autoregressive Conditional Heteroscedasticity
GBM	Geometric brownian motion
GED	Generalized Error Distribution
LM	Lagrange Multiplier
MDS	Martingale Difference Sequence
MLE	Maximum Likelihood Estimate
MS-GARCH	Markov switching Generalized Autoregressive Conditional Heteroscedasticity
NIG	Normal Inverse Gaussian
NSE	Nairobi Securities Exchange
SDE	Stochastic differential equation
S&P500	Standard and Poor Index
RS	Regime Switching
RS-GARCH	Regime Switching Generalized Autoregressive Conditional Heteroscedasticity

CHAPTER 1: INTRODUCTION

1.1 Background

The existing studies on financial modeling unveil that the connection between returns from stocks, volatility and the total traded shares have been broadly investigated in both established and developing stock markets. The understanding of this connection is paramount because it gives investors the insight of the financial market micro-structure. Abbondante et al. (2010) defines volatility as a statistical measure of asset returns dispersion, and refers to the total shares transacted within a specific period of time as trading volume(or volume). According to him, volatility may be calculated by determining the variance or standard deviation of an asset return and that a higher value is an indication of investment returns that are highly dispersed and this can be associated with a higher risk. Arguably, the success of stock market heavily relies on volatility in the sense that when the volatility reduces, the stock price may go up and vice versa. In other words, this implies that when volatility increases, market risk rises as well, and returns may fall. Moreover, Abbondante et al. (2010) contends that trading volume can be employed in technical analysis as an indicator of the direction of the stock price. Trading volume gives investors an estimate of the stock market value and to some extent it can confirm trend or trend reversal and it is likely that as trade volume increases, the stock prices generally would move in a similar direction. It is thus worth noting that a thorough understanding of how stock returns, volatility and volume are related is essential. According to Wiley and Daigler (1999), information flow into the market play a crucial role in determining the stock price and volume relationship, whereas, Karpoff (1987) on the other hand, argue that price-volume empirical relationship is paramount because of the fact that it gives insights into the understanding of the many theories that compete to widely spread ideas about the role played by the information that flows into the market. The investors' motivation to trade is mostly determined by their trading objective; it could be to speculate on market information or portfolios diversification to spread the risk, or the desire for liquidity. These varied trading motives stem from the interpretation of various data sources. As a consequence, trading volume may come from any of the investors with various information sets. Many empirical studies disclose that information flow into the market is linked to volatility and trading volume, as shown by the studies of Gallant et al. (1992), Lamoureux and Lastrapes (1990), and He and Wang (1995). All these studies, reports that the arrival of new information into the market causes changes in stock prices and as a consequence the returns on stock, volatility and trade volume are positively associated.

Moreover, investors are motivated by larger returns to investment and this results in capital flow, however, it is not easy to forecast returns in a volatile market environment, see Attari et al. (2012). According to Attari et al. (2012) and Glascock and Hsieh (2014), emerging stock markets are associated with extremely volatile stock returns emanating from the stock market having low volume. A market at the development stage of becoming a mature and developed system where growth is steady and political risks (the risk that an investment is likely to be adversely affected by political changes and instability in a country) are low, is known as an emerging market. These markets are in most cases located in undeveloped countries that are striving to achieve a steady business infrastructure. An investment in an emerging market has the tendency to be volatile and uncertain, and investors demand higher potential returns in exchange for the higher risk. In general, emerging markets are characterized by low income, rapid growth, high volatility currency swings, and high potential returns. Note that, an emerging market becomes a developed market if it has all the traits of a developed market. On the other hand, a developed market is a country that is most developed in terms of its economy and capital markets. That is, it is an economy (country) with a high level of economic activity characterized by high per capita income or per capita gross domestic product (GDP), high level of industrialization, developed infrastructure, technological advancement, and a relatively high rank in human development, health and education. In a developed market there are high levels of liquidity in debt and equity, political and financial stability (hence less risk), high economic development, and the market is open to foreign investment and this provides accessibility to global investors and hence encourages a higher volume of investment and transactions. In general, a developed market economy is characterized by high income, high human development rank, service sector domination, technological, and high level of infrastructure development.

A study by Girard and Biswas (2007) compared stock returns' volatility and trade volume in both established and developing markets and reported a negative relationship between the two markets. According to Al Samman and Al-Jafari (2015) trading volume is a crucial indicator that can be utilized to gauge the market strength because it includes information about the stock market's performance. Empirical studies have looked at the dynamic and contemporaneous correlation between returns from stocks and trading volume. In this regard, Lee and Rui (2000) report that in developed markets volume is granger-caused by returns and similarly, Mahajan and Singh (2009) unveils that volatility and volume are positively correlated in addition to disclosing that returns granger-cause trading volume (one-way causality). Furthermore, studies by Christie (1982) disclose that volatility and trading volume are negatively correlated while Rogalski (1978) on the other hand established a positive contemporaneous link between volume and absolute returns by utilizing month-on-month data. Recently, Jiranyakul (2016) investigated the dynamic association between returns on stocks, volatility and trade volume from the Thai Stock exchange and established that trade volume is paramount in dynamic relationships. Pertaining to the subprime mortgage crisis in the United States, the trade volume creates both returns and volatility. Moreover, the available literature reports a contemporaneous association between volatility and trade volume from the Thai Stock market. Despite this remarkable empirical and theoretical investigations on stock returns, volatility and trading volume correlations, blended results have been established in general. In addition, the focus of majority of these studies have been on the established stock markets rather than emerging stock markets which leave inadequate similar literature for emerging markets.

Lamoureux and Lastrapes (1990) in their study claim that the random arrival of information into the market causes price movements and consequently this leads to trade volume fluctuations. Despite information flow being unobserved, trade volume can be utilized as an exogenous variable for return series heteroscedasticity of variability of return series in question. When trade volume is factored into GARCH model, volatility persistence, as reflected by the ARCH and GARCH effects, may be minimized or eliminated. Moreover, Lamoureux and Lastrapes (1990) argue that when trade volume is added into the conditional variance equation, the conditional volatility persistence vanishes. Results similar to that of Lamoureux and Lastrapes (1990) have been established by studies of Miyakoshi (2002) and Omran and McKenzie (2000) who investigated the Australian equities and found that the volatility persistence decreased considerably after trade volume is utilized as a proxy for the information arrival rate. The studies of Huang and Yang (2001), Chen et al. (2001), Yüksel (2002), Salman (2002) and Ahmed et al. (2005) reports contrasting results that volatility persistence does not decrease when trade volume is incorporated into the conditional volatility equation.

Financial markets have been observed quite often to exhibit an abrupt change of their behavior, for instance, previous studies reveals that volatility on stock returns changes with time and the change tend to persist. Moreover, some stylized features of volatility for instance volatility clustering, leverage effects, higher volatility emanating from non-trading periods, etc, have been reported by many empirical studies. Volatility has been found to be a vital factor that influences option pricing despite the fact that it is a difficult factor to estimate, however, once estimated volatility may be utilized to determine future stock prices or the option prices. Black-Scholes model which is reported from the existing literature to have been broadly utilized in option pricing assumes that volatility is steady until the lapse time, which is not the case since volatility is known to change over time. Also, the model is deemed a single volatility regime model. Furthermore, the conditional variance of returns is time varying and researchers have paid attention to this stylized fact by building models that capture the changing variance, either in a continuous, see Heston (1993), or discrete, see Engle (1982), time setting. Unfortunately, both discrete and continuous time states of nature have given rise to incompleteness, for example, the multiplicity of equivalent martingale measures involves a continuum of equilibrium prices. As a result, the question of choosing the best model arises. For instance, an approach for pricing options was developed by Duan (1995) that considered a GARCH model with Gaussian innovations whereas Heston and Nandi (2000), basing their argument on the methodology by Duan (1995), considered a new conditionally Gaussian model to capture skewness in prices of options. They developed a nearly closed form formulation for pricing call options and used empirical data to verify its pricing performance. However, this model is conditionally Gaussian which means that it does not reflect the behavior of short-term equity options smiles. The model was later extended by Christoffersen et al. (2006) who utilized Inverse Gaussian distribution in order to increase the skewness effect. On the other hand, ARCH - type models have been widely used to model the varying volatility, however, in order to capture the sudden changes of financial market behavior such as the volatility swings, Regime switching models of Hamilton (1989) are utilized. According to an unobservable process that creates switching among a finite range of regimes, these model parameters assume various values in diverse time periods. In support of regime-switching models with autonomous mean and variance changes, Bollen et al. (2000) and Hardy (2001) suggested a regime-switching log-normal (RSLN) model in which log-returns follow a normal distribution with mean and variance that are dependent on the regime variable. To investigate the volatility of the valued-weighted Taiwan Stock Index returns, Li and Lin (2003) used the Hamilton and Susmel (1994) Markov-Switching ARCH model. Additionally, Markov regime-switching models have been used to show that volatility expectations in the German, Japanese, and US stock markets are regime-specific. Pricing options under regime switching model has been widely suggested, however, pricing of regime risk has been a problem. Regime risk is linked to the changing economic conditions and hence ought to be priced. Furthermore, valuing regime risk in a Markov regime switching model is a subject that is yet to be completely investigated in the literature. In fact, the impact of switching regimes in the underlying asset price dynamics on the behavior of option prices is addressed through direct pricing of regime risk. It also provides some insight into how macroeconomic factors affect option prices, which is particularly significant when pricing a long-maturity option because macroeconomic conditions might vary over time.

In conclusion, despite the fact that extensive research on stock returns, volatility, and trading volume exists, the majority of studies have focused mostly on developed stock markets, leaving analogous research on emerging markets lacking. However, this notwithstanding the documented evidence reveals results that are not in total consensus about the association between returns on stocks, volatility and trade volume. In addition, the issue of including trading volume in GARCH model's conditional variance equation and as to whether volatility persistence decreases or even vanishes altogether, has been found to disclose conflicting results. Another issue that emerges from the available literature is that the GARCH model and its extensions have been utilized to capture the stylized characteristics of financial time series for instance, heavy tails, clustering of volatility, leverage effects, long-memory, leptokurtic distribution, among others. The empirical researches report that stock returns volatility poses several issues that GARCH-type models fail to reflect well, see Bauwens et al. (2006). In particular, these models frequently exhibit a high level of conditional volatility persistence. Ardia et al. (2019) demonstrates that the GARCH-type model does not reflect the actual volatility change whenever there is regime changes in volatility dynamics. The volatility change in a market with regime switches can best be captured by considering a regime switching model because it permits the GARCH parameters to change over time. As a consequence, this study proposes to utilize GARCH-type and a regime-switching model to model the underlying asset's dynamics as well as to probe the contemporaneous and dynamic correlation between returns on stock, volatility and trade volume from both emerging and developed markets. This is further utilized to investigate how integrating trade volume into the GARCH-type models' conditional variance equation affects volatility persistence.

Moreover, modeling the underlying asset's dynamics using a regime switching model is extended to pricing European options and the Black-scholes model is used as a benchmark model for the option pricing. In fact, a regime-switching model for pricing European options is proposed in this paper when the underlying asset dynamics are dependent on market regimes. The formulation of this model is based on a geometric Brownian motion that is guided by a two-state continuous-time Markov chain and by an application of a change of measure, an option price is developed using risk-neutral valuation. The regime switching (RS) model is modified to include GARCH effects and dynamics resulting to regime-switching GARCH model, hereinafter referred to as RS-GARCH. The implementation of these two models is done by computing the European call option prices for some chosen stock market indices and the results compared with those from the famous Black-Scholes model. This comparison is necessary for establishing the best model for pricing the European options.

In general, in this study, the stock markets dynamics are modeled with the aim of probing the relationship that exists between stock returns, volatility and trading. Moreover, the dynamics of the stock returns' volatility in a regime switching market is investigated and applied in pricing the European options. The principal contributions of this study are that;

- The relationship between stock returns, volatility and trading volume is investigated using both GARCH-type and Regime switching models. A comparison of this relationship for both developed and emerging financial markets is carried out, such a comparative study that uses both GARCH-type and Regime switching models is missing from literature. Moreover, the relationship between asymmetric volatility and trade volume using GARCH-type models is investigated.

- The dynamics of stock returns' volatility is probed using the GARCH-type and Regime switching models. The GARCH-type models conditional to normal, student-t and GED distributions are utilized for both developed and emerging markets. The regime switching model is applied to the two market data sets to capture the stock market dynamics in a regime switching financial market setting. Moreover, a comparison of the stock market dynamics for the two sets of models is implemented and such a comprehensive comparison has not been carried out by earlier studies.

- The effect of trade volume on volatility asymmetry and volatility persistence when trade volume is added to the conditional variance equation of GARCH-type model is investigated under the three conditional distributions; normal, student-t and GED. The existing literature has not captured the relationship between trade volume and volatility asymmetry, and a comparison of the same for both emerging and developed markets is lacking. This study, therefore, has filled up this gap. Furthermore, this study adds to the existing literature by investigating the effect of trade volume on volatility persistence since earlier studies have reported conflicting results. This notwithstanding, the empirical studies existing in literature have not compared the case for developed and emerging markets in addition to the fact that many of these studies have focused more on modeling developed markets than emerging market. These two issues have been addressed by this study.

- The research by Hardy (2001) assumed that the stock returns are log-normally distributed, and utilized a regime switching model to price European options. In this study, this assumption is dropped and instead the European call options are priced using the regime switching model under the assumption that the returns on stocks follow a normal distribution. Moreover, a large data set with many observations as compared to the two market indices data considered by Hardy (2001) is utilized so as to increase the chance of getting better descriptions of the model in the options pricing. Finally, the regime switching model is adapted to include GARCH effects which results to a Regime switching GARCH (RS-GARCH) model and this model is utilized in pricing the European options. The results of the European call options estimated from both the RS and RS-GARCH models are compared with those of the Black-Scholes model and the model with better fit is determined.

1.2 Statement of the problem

Modeling stock returns is an important task in financial markets and the past couple of years have witnessed an increase in research that are geared towards modeling the financial time series of both emerging and developed stock markets. The majority of empirical studies on the dynamic and contemporaneous association between returns on stocks, volatility, and trade volume have focused on developed economies such as, the US, UK, Japan, and Hong Kong stock markets, as compared to emerging markets. These prior studies reports a link between returns on stocks, volatility and trade volume, however, some few studies have reported conflicting results pertaining this relationship. Most of these previous empirical studies generally support the mixture of distribution hypothesis (MDH) model in explaining the association between trade volume and stock returns in the setting of information arrival into the market. Many empirical studies that utilized the MDH model have explained the volatility persistence by adding trading volume in GARCH model as a proxy for information arrival. The existing literature reveals that there are very few studies that have utilized asymmetric GARCH models to investigate the connection between returns and trade volume. On the other hand, financial markets have been observed quite often to exhibit an abrupt change of behavior, for instance, previous studies have shown that stock returns volatility changes with time and the change tend to persist. Regime switching models proposed by Hamilton (1989) are utilized in modeling dynamic switches of the salient characteristics of financial time series, for instance, volatility asymmetry, among others. In the context of modeling the dynamics of the correlation between stock returns and trade volume, it is revealed that Regime switching models have hardly been used by prior studies to model this relationship. This study focuses on modeling the underlying asset returns using GARCH models and its extensions and Regime Switching models as well as applying these models in pricing the European options. Furthermore, the existing literature reveals limited information on whether regime-switching and RS-GARCH models outperform the famous Black-Scholes model in pricing European options. Therefore, a regime switching model for pricing European options is constructed and modified to include GARCH effects and dynamics resulting to regime-switching GARCH model.

1.3 Objectives of the study

1.3.1 Main objective

The overall objective of the study is to model the correlation between stock returns, volatility and trading volume of the emerging and developed financial markets and to price few financial derivatives such as European options.

1.3.2 Specific objectives

(i) To analyze stock returns volatility dynamics in both developed and emerging markets based on GARCH and Regime switching models.

(ii) To investigate the dynamic correlation of stocks returns, volatility and trade volume in emerging and developed financial markets.

(iii) To investigate the asymmetric relationship between trade volume and stock returns using GARCH models.

(iv) To price European options using regime switching and RS-GARCH models.

1.4 Scope of the study

This study probes the dynamic and contemporaneous correlation between stock returns, volatility and total number of shares traded in Nairobi securities exchange (NSE20), S&P500 and FTSE100 market indices. The dynamic structure of the underlying stock returns is modelled using GARCH models and their extensions, Regime switching, and MS-GARCH models conditioned with Gaussian, student-t, and generalized error distributions (GED) specifications. Application of these models is extended to the pricing of European options.

1.5 Significance of the study

Modeling financial market index as well as option pricing is significant in the following manner; first, it provides the understanding of the financial markets micro-structure. Secondly, it demonstrates the rate of information flow to the market, how the information is widely spread and how it influences stock returns by applying different models, for instance the GARCH models specify a symmetric volatility response to news. Finally, the use of exponential GARCH models gives new insight into the asymmetric effects of volatility, trading volume, and their impact on stock returns. The understanding of the dynamic structure of the underlying assets is the basis for application of the knowledge into pricing options.

CHAPTER 2: LITERATURE REVIEW

The relationship between stock returns, volatility and trade volume has previously been carried out extensively by many empirical studies. Karpoff (1987) argues that most of these early studies disclose a significant positive correlation between returns on stocks, volatility and trade volume. Lee and Rui (2000) utilized China stock market data to investigate the contemporaneous and causal correlation between stock returns, trade volume and volatility and unveiled that trade volume is not granger-caused by stock returns. Later, Lee and Rui (2002) carried a similar study utilizing data from New York, London and Tokyo stock markets and documented that returns on stock are not granger-caused by trading volume in each of the three markets. However, they reported that returns' volatility is positively related with trade volume across the three markets. These findings are not consistent with those of Chen et al. (2001) who examined the dynamic connection between stock returns, volatility and trade volume by utilizing data from nine different stock exchanges. This study reported that for some stock exchanges stock returns granger-caused trade volume, whereas for some other stock exchanges, trade volume granger-cause stock returns.

Mubarik and Javid (2009) conducted a similar research utilizing the ARCH and GARCH-M models on Pakistan Stock Market data and revealed that stock returns' volatility and trade volume are significantly correlated. The results suggest that the previous day trade volume has significant effect on the present-day return, that is, the previous day trade volume and returns explains the prevalent market returns. Moreover, studies of Khan and Ahmed (2009), and Al-Jafari and Tliti (2013) utilized Karachi Stock Exchange index and Amman Stock Exchange data and disclose that returns' volatility and trade volume have a significant positive correlation. In his work, Crouch (1970) investigated the connection between daily trading volume and the daily absolute changes of the stock market index and individual stocks and reported a positive relationship between them. Stock data recorded on monthly basis was analyzed by Rogalski (1978) who reported a positive relationship between the total number of traded shares and the absolute returns. In the emerging market situation, Brailsford (1996) reported a strong contemporaneous relationship between returns and volatility. Harris (1987) did employ several transactions to measure the total number of traded shares and reported that changes in total number of traded shares and changes in the square of returns were related.

The asymmetric effect of trading volume on return volatility through price formation process is also well documented in literature. Several empirical studies find asymmetric volatility to be a crucial factor in the understanding of trading volume-return volatility relationship. The asymmetry of volatility effect is largely associated with a greater rise in the volatility following an unexpected price fall compared to a price increase of the same magnitude; see Bollerslev et al. (1992), Patterson (2000) and Brooks et al. (2008). Furthermore, this asymmetry of volatility effects is due to price fluctuations and these changes are in most cases negatively related with volatility changes. Kroner and Ng (1998) argue that, the cause for this asymmetric effect is due to leverage effect and a rise in the information flow following unfavorable news. Moreover, increase in information flow due to unfavorable news leads to relative rise of the rate of information flow across firms which in turn affects the co-variances across stock returns. In terms of the asymmetric issue, ”bad (or unfavorable) news” refers to negative returns while during financial crises it refers to information with adverse effects across the integrated stock markets. Studies of Schwert (1989) and Nelson (1991) reported the asymmetric volatility behavior of stock returns using US data; Reyes (2001) estimated an asymmetric impact on volatility in the Tokyo Stock Exchange; Henry (1998) captured asymmetry of volatility using Hong Kong Stock market data; Sentana (1995) used the UK and US data to identify the asymmetric impact in the Stock market returns; Zakoian (1994) empirically tested the asymmetry of volatility behavior of French Stock data.

Kearns and Pagan (1993) reports the effect of asymmetric volatility in the emerging market stock returns was lower compared to the developed stock market returns. Frijns et al. (2010), in their study used traditional regression analysis methods whose findings documented contradictory results for volatility asymmetry in the Australian market volatility index, which was constructed using daily data of S&P500/ASX200 index options. GARCH models were utilized by Lamoureux and Lastrapes (1990) to model stock returns volatility and trading volume relationship who reported that persistence in the conditional variance vanishes with addition of trade volume as an exogenous variable. Their findings formed the foundation for further research by academicians and policy makers to investigate the correlation between returns, volatility and trading volume in developed markets and later in emerging markets. These later studies aimed at investigating whether trading volume explains volatility persistence. Omran and McKenzie (2000) and Miyakoshi (2002) both found similar results to those reported by Lamoureux and Lastrapes (1990) for the Australian Stock Exchange and the Tokyo Stock Exchange, respectively. Najand and Yung (1991) analyzed daily prices and volume of treasury bond’s future markets in their research and found that GARCH effects persisted even after volume is incorporated in their model’s conditional variance. Similarly Brailsford (1996) utilized Australian equities and found that after accounting for trading volume as a proxy for the information arrival rate, volatility persistence was significantly reduced.

Emenike and Opara (2014) conducted a study in Nigeria to look into the connection between stock returns' volatility and trading volume by fitting the daily All-share Index and closing trading volume of Nigerian Stock Exchange into GARCH (1, 1) and GARCH-M (1, 1) models. Their findings disclosed a significant positive relationship between trading volume and stock returns volatility. Moreover, they revealed that persistence in volatility does not disappear with inclusion of trading volume in the conditional variance equation of both GARCH (1, 1) and GARCH-M (1, 1) models. Additionally, studies by Chen et al. (2001) in nine developed markets, Huang and Yang (2001) in Taiwan, and Ali et al. (2005) for the Kuala Lumpur Stock Exchange markets have all documented that stock returns' persistence of volatility remains after inclusion of trade volume in the conditional variance equation of the GARCH model.

The relationship between trade volume and stock returns' volatility was investigated by Gulia (2016) who utilized GARCH model to analyze data from the Indian Stock Market. His study looked at the relationship of the price-volume changes and the effect on volatility persistence after adding trade volume to the basic ARCH variance equation and reported that negative returns have a lower price-volume change slope than positive returns. They further claimed that persistence of variance reduces when trade volume is included in the GARCH variance equation as an exogenous variable.

Moyo et al. (2018) evaluated the effects of trading volume as a proxy for the arrival of information on stock volatility, and the impact of adding trading volume into the conditional volatility equation on volatility persistence, using the EGARCH and TGARCH models. Their findings show a positive association between trading volume and stock returns, and that trading volume is a poor source of volatility on stock returns when used as a proxy for information flow. However, there is no observed change on volatility persistence when trading volume is added in the conditional variance equation.

For the Johannesburg Stock Exchange (JSE) in South Africa, Naik et al. (2018) revisited the association between equities' trade volume and returns' volatility. He looked at the volume-volatility link using EGARGH and Granger causality models, as well as the volatility persistence before and after trade volume is included in the volatility model as an exogenous variable. The researchers discovered a positive and contemporaneous association between transaction volume and market volatility and that volatility persistence never died off after the explanatory variable was included in the volatility model.

Poudel and Shrestha (2019) used an Autoregressive Distributed Lag technique to analyze the connection between trading volume and stock returns using the Nepalese Stock Market index (NEPSE index). He looked at the short and long term impact caused by trading volume on stock returns. His studies revealed that there was a long-term and short-term positive connection between trading volume and stock returns. As a result, he came to the conclusion that stock returns had a considerable impact on trade volume in the Nepalese Stock Market.

There is a wide discussion on option pricing that exists in literature. Black and Scholes (1973) developed the Black and Scholes model which was widely applied in scientific research in option pricing. The model was later extended to options on futures, see Black (1976). Even after academicians and practitioners could use this option pricing tool to get option prices, there was concern on how to make profits by being able to predict volatility direction or simply knowing how to accurately price options. Black (1976) argued that the simplest approach consisted fitting historical volatility into the Black and Scholes formula for futures prices. However, it was discovered that the model was inaccurate since it gave equal weights to all past rates of returns irrespective of the initial values. The other approach is the realized volatility estimator that was based on squared log returns summed over some time interval, see Black (1976). The approach also exhibited significant biases.

Further research was devoted to option pricing using GARCH models with Engle (1982) being the first one to propose the ARCH model. This model explains that future conditional variance is a function of the unconditional long-term variance, and of the returns of the past days. Bollerslev (1986) extended this model to GARCH model which has an additional parameter that takes into account low and high volatility. The question then was on how to implement these volatility estimators based on the theory of options pricing. Duan (1995) answered this question by proposing the GARCH option pricing model that utilized Monte-Carlo approach to price options. He reported that simulated asset prices followed a GARCH process hence heteroscedasticity of stock returns is captured leading to the reduction of pricing bias caused by the volatility smile. However, the model by Duan (1995) only corrects certain biases resulting from the Black-Scholes model and does not perform empirical tests of his approach.

Siu et al. (2004) proposed a model for pricing options based on GARCH assumption for underlying assets in the context of dynamic version of Gerber-Shiu's option-pricing model. They applied conditional Esscher Transforms to develop a martingale measure in the context of an incomplete market situation. The result was consistent with that of Duan (1995) under the assumption of conditional normality for the stock innovation. They, too, justified the pricing result based on the dynamic framework of utility maximization problems as per the Gerber-Shiu's option-pricing model.

Barone-Adesi et al. (2008) in their research proposed option pricing method that was based on GARCH models with filtered historical innovations. Their model outperformed other GARCH pricing models and Black Scholes models empirically for S&P500 index options. The model did explain implied volatility smiles by the negative asymmetry of the filtered historical innovations. Empirical evidence and the extent of deterioration of the delta hedging in the presence of large volatility shock were revealed by the study.

Under the regime switching Black-Scholes economy, Hao and Yang (2011) reported a Scenario-based risk estimate for a portfolio of European style derivative assets over a fixed period of time. In their study they derived a closed form expression for measuring risk for both barrier and vanilla European options and noted that this approach can be applied to a variety of exotic options as well. In his research, Duan (1995) developed a model for pricing options on an asset whose continuously compounded stock returns follow the generalized Autoregressive conditional heteroskedastic (GARCH) process. He utilized local risk neutralization approach to construct this model and unveiled that the GARCH option pricing model allows for the inference of implied GARCH parameters from the market data, and that the inferred parameter values can be employed in a similar way to the Black-Scholes model's implied volatility. In their research Dash et al. (2012) utilized the GARCH option pricing model for options traded on the National Stock Exchange, India. It was further reported that the GARCH (1,1) model was used to obtain volatility projections, and calculated option prices using these volatility projections in the Black-Scholes-Metron model. Their results indicated that the implied volatility (for both calls and puts) were overestimated, and that call and put option prices were predominantly overvalued. Further, they reported that put options were more overpriced than call options and that the overestimation of volatility and overvaluation of options prices increased with higher market capitalization and moderate/higher trading volume of the underlying stock.

Recent studies on option pricing have been carried, for instance, Biswas et al. (2018) used a regime switching stochastic model with a semi-Markov modulated square root mean reverting process to study European option pricing. In their study, they discovered the local risk-minimizing European-type vanilla options price, and it was proven that the price function adequately meet a non-local degenerate parabolic PDE, which is the general case of the Heston PDE model. A study by Deelstra et al. (2020) considers risk-neutral pricing of Vanilla, digital and down-and -out call options with the price of the underlying asset evolving according to the exponential of a Markov-modulated Brownian motion (MMBM) with two-sided phase-type jumps. Further, he documents that such options are closely related to the MMBM's first passage properties which he analyzed by randomizing the time horizon using Erlang distribution. Moreover, Nasri et al. (2020) used a regime switching Copula model to option prices where serial independence of error terms for time series is presented which was subjected to several time series analysis. Moreover, Kirkby and Nguyen (2020) utilized duality and FFT-based density projection implementation to establish a novel and efficient transform approach for pricing Asian options for general asset dynamics, such as regime switching Levy processes and stochastic volatility models with jumps.

In his research, Lin and He (2020) addressed the problem of pricing the European options using a regime switching finite moment log-stable model. Their model has the ability to represent the key features of asset returns, as well as the impact of regime transition, and it is compatible with market findings. Furthermore, they state that option prices are driven by a coupled FPDE (Fractional Partial Differential Equation) system in their model. They note that their model can mimic option prices emanating from other existing models with particular parameter settings while also bringing into significance pricing differences relative to current models by changing parameter values, indicating that it has the potential to be used in practice.

Godin et al. (2019) in his study reveals that Esscher transform approach for obtaining risk neutrality is famous in pricing options under regime switching models. However, this approach creates path-dependence in the dynamics of option price since under the physical measure, the underlying asset is integrated in a Markov process. To address the path-dependence they develop a novel and intuitive risk-neutral measure that integrates risk-aversion in regimes and consequently removing the path dependence side effects. Later on, the study of Godin et al. (2019) was complimented by Godin and Trottier (2021) who developed and utilized extended Girsanov principle which also removes the path-dependence effect. This model is easy to interpret in terms of consistence with hedging agents locally minimizing their risk-adjusted discounted squared hedging errors.

In general, it is noted from the existing literature that a comparative study of the relationship between stock returns, volatility and trade volume for both emerging and developed market is missing. The use of regime switching model to investigate this relationship as well as investigating the effect of adding trading volume to the GARCH-type conditional equation on volatility asymmetry is missing as well. Moreover, the literature does not document the use of regime switching model to model the relationship between trade volume and asymmetric volatility. The study of Hardy (2001) proposes use of regime switching model on pricing European options under the assumption that the returns are log-normally distributed, this study uses the regime switching model to price European options with the assumption that the stock returns are normally distributed. In addition, the effect of GARCH is included in the regime switching model and consequently a regime-switching GARCH model is derived and applied in European options pricing.

CHAPTER 3: METHODOLOGY

3.1 Introduction

This chapter presents the research methods and procedures followed in this study and it begins by giving a brief review of stylized facts of return series which is then followed by an outline of the underlying asset. In addition, models applied in analyzing the financial time series data sets and their parameter estimation methodologies are presented.

3.2 A review of stylized facts of returns

Financial time series is defined as a sequence of observations on financial data over a fixed period of time. Among the aims of analyzing a financial time series is to find physical models that can explain the empirically observed features of real life data. The observations of the time series is often assumed to exhibit a normal (Gaussian) distribution, however, this has always been disapproved by empirical studies of practically any financial time series whose results have revealed that most financial time series are non-stationary which means the mean, variance and auto-covariances of the time series vary with time. For instance, the presence of changing variance in stock returns were first captured by Engle (1982) using the ARCH model and later Bollerslev (1986) extended it to generalized ARCH (GARCH) model. Cont (2001) compiled and documented statistical properties that were observed to be common to a wide set of financial time series and referred to them as 'stylized facts'. We list some of the stylized facts as follows; *Heavy tails*-the returns for most data set seem to have a distribution(unconditional) that displays a power-law or Pareto-like tail; *Volatility clustering*-large volatility changes tend to be followed by large volatility changes of either sign or low volatility changes by low volatility changes; *Leverage effect*-changes in stock prices tend to be negatively correlated with changes in volatility, that is, volatility is high after negative shocks than after positive shocks of same magnitude; *Aggregational Gaussianity*-as the time scale for which the returns are calculated increases, the distribution of returns tends to be normal. That is, the shape of the distribution is different at different time scales; *Conditional heavy tails*-the time series's residual exhibits heavy tails even after correcting returns for volatility clustering, for example, through GARCH-type models. However, the tails are less heavy as compared to the unconditional distribution of returns, e.t.c.

Definition 3.2.1 *Let $\mathcal{F}$ be a set of subsets ω of Ω. Then $\mathcal{F}$ is a σ-algebra if the following conditions are satisfied* *(i)* $\phi \in \mathcal{F}$,

(ii) *if* $\omega \in \mathcal{F},$ *then* $\omega^c \in \mathcal{F}$,

(iii) *if* $\omega_1, \omega_2, ..., \omega_n, ... \in \mathcal{F},$ *then* $\bigcup_{i=1}^{\infty} \omega_i \in \mathcal{F}$ *i.e* ω_i *are countable unions*

Definition 3.2.2 *A probability measure* $\mathbb{P}$ *is a function* $\mathbb{P} \to [0, 1]$ *such that*

(i) $\mathbb{P}(\Omega) = 1$

(ii) $\mathbb{P}(\omega) + \mathbb{P}(\omega^c) = 1$

(iii) *if* $\omega_1, \omega_2, ..., \omega_n, ... \in \mathcal{F}$ *are disjoint, then* $P(\bigcup_{i=1}^{\infty} \omega_i) = \sum_{i=1}^{\infty} P(\omega_i)$

Definition 3.2.3 *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space where* $\Omega = \{X_1, ..., X_n\}$ *is the set of all possible outcomes,* $\mathcal{F}$ *is a* $\sigma - algebra$ *and* $\mathbb{P}$ *is a probability measure. A filtration on* $(\Omega, \mathcal{F}, \mathbb{P})$ *is an increasing family* $(\mathcal{F})_{t\geq 0}$ *of sub-*σ *-algebra of* $\mathcal{F}$*. That is, for each* t*,* $\mathcal{F}_t$ *is a* σ *-algebra included in* $\mathcal{F}$ *and* $\mathcal{F}_s \subset \mathcal{F}_t$ *if* $s \leq t$

Definition 3.2.4 *Let* $\{X_t : t = 0, 1, 2, 3, ...\}$ *be a sequence of random variables defined on some probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*. Then* $\{X_t, t\in T\}$ *is a stochastic process where* T *is an index set representing time and can be continuous or discrete. It thus follows that*

(i) $\mathbb{E}[X_t] = \mu_t$ *for* $t = 0, 1, 2, 3, ...$ *That is,* μ_t*, is the expected mean of the process at time* t*.*

(ii) $Cov(X_t, X_s) = \gamma_{t,s}$ *for* $t = 0, 1, 2, 3, ...$ *where* $Cov(X_t, X_s) = \mathbb{E}[(X_t - \mu_t)(X_s - \mu_s)] = \mathbb{E}[X_t X_s] - \mu_t \mu_s$ *is the auto covariance function.*

(iii) *The autocorrelation function,* $\rho_{t,s}$ *is defined by* $\rho_{t,s} = Corr(X_t, X_s)$ *for* $t = 0, 1, 2, 3, ...,$ *where* $Corr(X_t, X_s) = \frac{Cov(X_t, X_s}{\sqrt{Var(X_t)Var(X_s)}} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}}$

Note that (ii) *and* (iii) *are the linear dependence measures between random variables such that if the value of* $\rho_{t,s}$ *is close to* ± 1 *the linear dependence is strong otherwise the linear dependence is weak.*

Definition 3.2.5 *Let* $(X_t)_{t\geq 0}$ *be a stochastic process on* $(\Omega, \mathcal{F}, \mathbb{P})$ *. Then* $(X_t)_{t\geq 0}$ *is adapted to the filtration* $(\mathcal{F}_t)$ *if for each* $t > 0, X_t$ *is* $\mathcal{F}_t$ *-measurable, i.e.* X_t *is measurable with respect to* $\mathcal{F}_t$ *.*

Definition 3.2.6 *A process* X_t *is strictly stationary if the joint distribution of* $\{X_{t1}, X_{t2}, ..., X_{tn}\}$ *is the same as the joint distribution of* $\{X_{t_1 - k}, X_{t_2 - k}, ..., X_{t_n - k}\}$ *for time points,* $t_1, ..., t_n$ *and time lag k.*

Remark 3.2.1 *When $n = 1$, the distribution of X_t is the same as that of X_{t-k} for all t and k; that is, $X's$ are identically distributed. As a result, $\mathbb{E}[X_t] = \mathbb{E}[X_{t-k}]$ for all t and k, that is, the mean function is constant for all time. Moreover, $Var[X_t] = Var[X_{t-k}]$ for all t and k implying that the variance is constant over time.*

Remark 3.2.2 *If a process (X_t) is strictly stationary and its variance is finite, then the covariance function must only depend on the time lag.*

Definition 3.2.7 *A process (X_t) is weakly (or second-order) stationary if*

(i) $\mathbb{E}[X_t] = \mathbb{E}[X_{t-k}] = \mu_t$ *for all t, that is, the mean function is constant over time.*

(ii) $\gamma_{t,t-k} = \gamma_{0,k}$ *for all time t and lag k.*

3.3 Modeling the underlying asset

Suppose that $(X_t)_{t\geq 0}$ is a stochastic process on a probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$ that describes the stock market's uncertainty where $\mathbb{P}$ is a probability measure and $\mathcal{F}_t$ is a filtration that represents the available information upto time $t-1$ and is driven by a Brownian motion of the stochastic process. Further, denote the stock price at time t by S_t, which is adapted to the filtration $(\mathcal{F})_{t\geq 0}$ and define log returns as $X_t = \log_e S_t - \log_e S_{t-1}$, then the following definition results;

Definition 3.3.1 *Suppose X_t is a random variable with mean and variance conditional on the information set $\mathcal{F}_{t-1}$ (the σ-field generated by $X_{t-1}, j \geq 1$) comprising all the information upto time $t-1$, then under probability measure $\mathbb{P}$, the asset returns' model is defined as*

$$\begin{aligned} & X_t = \log_e\left(\tfrac{S_t}{S_{t-1}}|\mathcal{F}_{t-1}\right), \quad \implies \quad X_t = \mu_t + r_t, \;\; r_t = \sigma_t\varepsilon_t \\ & \sigma_t^2 = Var(X_t|\mathcal{F}_{t-1}) = \mathbb{E}\big[(X_t - \mu_t)^2|\mathcal{F}_{t-1}\big] \\ & \varepsilon_t \;\sim\; i.i.d(0,1) \quad \implies \quad \varepsilon_t|\mathcal{F}_{t-1} \;\sim\; \mathbb{D}(0,\sigma_t) \end{aligned} \tag{3.3.1}$$

where $\mathbb{D}$ represents the distribution that can be either normal or leptokurtic (students-t and GED), μ_t, σ_t and r_t are the conditional mean, conditional variance and the mean-corrected asset returns respectively.

3.4 The ARCH model

The ARCH model is a stochastic process with autoregressive conditional heteroscedasticity which is a postulate by Engle (1982), and it uses previous variances to explain future variances. The autoregressive property describes a method for giving feedback that incorporates previous observations into the present. Also, conditionality implies a dependence on the observations that have existed in the immediate past while heteroscedasticity means time-varying variance (volatility). In contrast to the ARMA models which focuses on modeling the first moment, ARCH models depends on the conditional second moments in modeling consideration. Let

$$X_t = \mathbb{E}_{t-1}[X_t] + r_t \; ; \quad r_t = \sigma_t \varepsilon_t \quad \text{, and } \; \varepsilon_t \sim i.i.d(0,1)$$

The ARCH(q) model of the process $(r_t), t \in \mathbb{Z}^+,$ is defined by

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i r_{t-i}^2 \quad \text{where} \quad \omega \geq 0, \;\; \alpha_i \geq 0, \;\; \text{for} \;\; i = 1, ..., p \tag{3.4.1}$$

In Equation (3.4.1) above, the mean and variance of the random variable r_t, are conditional on the information set $\mathcal{F}_{t-1}$ up to time $t-1$ and are defined as $\mathbb{E}[r_t|\mathcal{F}_{t-1}] = 0$ and $Var[r_t|\mathcal{F}_{t-1}] = \sigma_t^2$, respectively.

3.5 The GARCH Model

The GARCH model is a basic conceptual structure of Bollerslev (1986) and it is a generalization of Equation (3.4.1). Many researchers have largely used this model in the modeling and analysis of economic and financial data. The model posses some notable characteristics such as their capability to model volatility clustering as well as the ability to give account for the changing variance in time-series data. Suppose $(\varepsilon_t)_{t\in\mathbb{Z}}$ is an independent and identically distributed (i.i.d) sequence of random variables with mean zero i.e $\mathbb{E}[\varepsilon_t] = 0$ and unit variance, i.e, $Var(\varepsilon_t) = \mathbb{E}[\varepsilon_t^2] - (\mathbb{E}[\varepsilon_t])^2 = 1$ and let $P \in \mathbb{N} = \{1, 2, ...\}$ and $q \in \mathbb{N}_0 \cup \{0\}$. Further, let $\omega > 0$, $\alpha_i \geq 0$, for $i = 1, ..., p$, $\beta_j \geq 0$, for $j = 1, ..., q$ and $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ be positive parameters. A GARCH(p,q) process $(r_t)_{t\in\mathbb{Z}}$ with volatility process $(\sigma_t)_{t\in\mathbb{Z}}$ is then a solution to the equations

$$r_t = \sigma_t \varepsilon_t, \;\; \text{where} \;\; , \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i r_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \;\; , t \in \mathbb{Z} \tag{3.5.1}$$

where the process $(\sigma_t)_{t\in\mathbb{Z}}$ is non-negative. It is required that σ_t only depends on past innovations $(\varepsilon_{t-h})_{h\in\mathbb{N}}$, that is, it is demanded that σ_t is measurable with respect to the σ-algebra generated by $(\varepsilon_{t-h})_{h\in\mathbb{N}}$. Additionally, σ_t and r_t are not dependent on $(\varepsilon_{t+h})_{h\in\mathbb{N}_0}$ and $\sigma(\varepsilon_{t+h} : h \in \mathbb{N})$ respectively, for fixed t, and that p and q are the ARCH and GARCH process degrees respectively. $\alpha_i + \beta_j$ implies that the unconditional variance of r_t is finite, whereas its conditional variance σ^2 evolves over time. This study utilizes GARCH(1,1) model since its implementation is relatively simple and also from literature the model has given good results in modeling the changing variance. The model is expressed as follows

$$\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{3.5.2}$$

In Equation (3.5.2), ω is constant variance corresponding to the long run average, α_1 is the first order ARCH term that broadcasts volatility information from a previous time, and β_1, is the first-order GARCH term, which represents fresh information not available at the time of the prior forecast. The magnitude of α_1 and β_1 determine the extent of volatility persistence, that is, the closer the sum of α_1 and β_1 to 1, the more the shocks to volatility does not die off.

3.5.1 Conditional mean specification

The conditional mean $\mathbb{E}_{t-1}[X_t]$ can be specified as a constant or a low order autoregressive-moving average(ARMA) process to capture the autocorrelation generated by market microstructure impacts on non-trading effects. This, however, is contingent on the data frequency and the type of asset. Dummy variables linked with extreme or uncommon market events are frequently added to the conditional mean specification to remove these impacts if they occurred during the sample period. As a result, a typical conditional mean specification is of the form

$$\mathbb{E}_{t-1}[X_t] = \omega + \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{j=1}^{q} \theta_j r_{t-j} + \sum_{m=1}^{M} \beta_m' Y_{t-m} + r_t \tag{3.5.3}$$

where Y_t is a kx1 vector of explanatory variables.

3.5.2 Explanatory variables in the conditional variance

Exogenous variables can be introduced into the conditional GARCH(p,q) formula in the same way that they are introduced into the conditional mean equation, as shown below.

$$r_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i r_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \sum_{m=1}^{M} \delta_m' Z_{t-m} \tag{3.5.4}$$

where Z_t and δ are vectors of random variables and positive coefficients, respectively. These coefficients, such as trade volume and microeconomic news announcements help in the prediction of volatility, see Lamoureux and Lastrapes (1990). In financial investment, high returns are often associated with high risk. Although modern capital asset pricing theory does not infer such a simple link, it does imply that expected returns and risk, as measured by volatility, have certain interactions. Engle et al. (1987) suggested that the basic GARCH model be extended so that conditional volatility can create a risk premium that is included in the projected returns and referred to this extended GARCH model as GARCH-in-the-mean or GARCH-M.

3.6 The GARCH-M Model

In finance, a stock's return may be influenced by its volatility and the GARCH-in-mean model, abbreviated as GARCH-M, is the best way to model this phenomenon. The following is a simple GARCH-M(1,1) model:

$$X_t = \mu_t + \lambda_1 \sigma_t^2 + r_t, \text{ where } , r_t = \sigma_t \varepsilon_t, \text{ and } \sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{3.6.1}$$

where X_t, r_t and λ_1 are the log return series, the mean-corrected log return series and the risk premium parameter, respectively. A positive value of λ_1 implies that the stock return is positively correlated with its past volatility. The GARCH-M model formulated in Equation (3.6.1) implies presence of serial correlations in the return series $\{X_t\}$ which are caused by those in volatility process, $\{\sigma_t^2\}$. Therefore, another reason why stock returns have serial correlations is implied by the occurrence of the risk premium.

3.7 The exponential GARCH (EGARCH) Model

This model has the ability to capture asymmetric responses of time-varying variance to shocks and leverage effects, that is, a negative relationship between stock returns and volatility shocks. The model ensures that the variance is always positive and utilizes $\frac{r_{t-i}}{\sigma_{t-i}}$ as the standardized value. The EGARCH(p,q) model is expressed as follows;

$$\ln(\sigma_t^2) = \omega + \sum_{i=1}^{p} \{\alpha_i (|r_{t-i}| - \mathbb{E}|r_{t-i}|) + \gamma_i r_{t-i}\} + \sum_{j=1}^{q} \beta_j \ln(\sigma_{t-j}^2) \tag{3.7.1}$$

where γ_i is the asymmetric or leverage parameter that gives response to asymmetry. In most empirical cases, the value of γ_i is expected to be greater than 1, indicating that a negative shock can increase future volatility or uncertainty, whereas a positive shock decreases the effect on future uncertainty. A negative shock in financial market analysis usually means bad news, which leads to a more unpredictable future, whereas a positive shock means good news. As a result, investors, for example, would expect larger stock returns to compensate for the increased risk in their investment. This study employs EGARCH(1,1) model defined as follows;

$$\ln(\sigma_t^2) = \omega + \{\alpha_1(|r_{t-1}| - \mathbb{E}|r_{t-1}|) + \gamma_1 r_{t-1}\} + \beta_1 \ln(\sigma_{t-1}^2) \tag{3.7.2}$$

It should be noted that $\gamma_1 = 0$ implies symmetry(volatility is not asymmetric), $\gamma_1 < 0$ means bad news(negative shocks) increases volatility more than good news(positive shocks) and $\gamma_1 > 0$ implies that good news(positive shocks) increases volatility more than bad news(negative shocks). Moreover, the volatility persistence $\tilde{P}$ is in general estimated as $\tilde{P} = \sum_{j=1}^{q} \beta_j$.

3.8 Markov-Switching GARCH (MS-GARCH) Model

Suppose that the mean of r_t is zero and there is no serial correlation, that is, assume that $\mathbb{E}[r_t] = 0$ and $\mathbb{E}[r_t r_{t-i}] = 0$ for $i \neq 0$ and all $t > 0$. It is realistic to make this assumption if one is dealing with high-frequency returns whose conditional mean is usually assumed to be zero. In this study conditional variance process is allowed to switch regimes and in addition, the observed information set until time $t-1$ is denoted by $\mathcal{F}_{t-1}$. Thus, the general case of a Markov-Switching GARCH model is written as: $r_t|(\xi(t), \mathcal{F}_{t-1}) \sim D(0, \sigma_{k,t}^2, \epsilon_t)$. In this expression, $D(0, \sigma_{kt}^2, \epsilon_t)$ represents a continuous distribution whose mean is zero, has a changing variance with time, $\sigma_{k,t}^2$ and an additional shape parameters ϵ_k collected in the vector. The random variable $\xi(t)$ defined in the discrete space $1, ..., k$, represents the Markov-switching GARCH model. The standardized innovation is defined as $\eta_{k,t} = r_t/(\sigma_{k,t}) \sim D(0, 1, \epsilon_k)$. To model the dynamics of the state of random variables, it is assumed that the state $\xi(t)$ evolves according to an unobserved first order homogeneous Markov chain with a probability transition matrix of order kxk. In this study, two states are considered, that is, $k = 2$, so the transition probability matrix P is given by

$$P_{ij} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$$

where $P_{ij} = P[\xi(t) = j|\xi(t-1) = i]$ is the transition probability from state $\xi(t-1) = i$ to $\xi(t) = j$, $0 \leq P_{ij} \leq 1$, $\forall\ i, j \in \{1, 2\}$ and that $\sum_{j=1}^{2} P_{ij} = 1$, $\forall\ i \in \{1, 2\}$. Considering the parameterization of $\mathcal{D}(.)$, we've the variance of r_t^2 conditional on the realization of $\xi(t) = k$, that is, $\sigma_{k,t}^2 = \mathbb{E}[r_t^2|\xi(t) = k, \mathcal{F}_{t-1}]$

On the other hand, conditional variance dynamics is modeled under the assumption that the conditional variance of r_t follow a GARCH-type model, see Haas et al. (2004). Conditionally on regime $\xi(t) = k$, $\sigma^2_{k,t}$ is a function of the past observations, r_{t-1}, past variance, $\sigma^2_{k,t}$ and vector of parameters, θ_k which is regime-dependent, that is,

$$\sigma^2_{k,t} = h(r_{t-1}, \sigma^2_{k,t-1}, \theta_k). \tag{3.8.1}$$

Here $h(.)$ is a $\mathcal{F}_{t-1}$-function that defines the conditional variance filter and ensures that it is positive. The initial variance recursions, that is, $\sigma^2_{k,t}(k = 1, 2)$, are set equal to the unconditional variance in regime k. Depending on the form of $h(.)$, different scedastic specifications are obtained, and in our case GARCH and EGARCH scedastic specifications are considered as shown below. According to Bollerslev (1986), the GARCH model is given by

$$\sigma^2_{k,t} = \alpha_{0,k} + \alpha_{1,k} r^2_{t-1} + \beta_k \sigma^2_{k,t-1} \tag{3.8.2}$$

for $k = 1, ..., k$. We have $\theta_k = (\alpha_{0,k},\ \alpha_{1,k},\ \beta_k)'$ in this case and in order to ensure positivity, it demands that $\alpha_{0,k} > 0,\ \alpha_{1,k}\ \geq 0$ and $\beta_k \geq 0$. Moreover, covariance-stationarity is ensured in each regime by requiring that $(\alpha_{1,k} + \beta_k) < 1$.

Moreover, an EGARCH scedastic specification with regimes is expressed as

$$\ln\left(\sigma^2_{k,t}\right) = \alpha_{0,k} + \alpha_{1,k}\Big(|\eta_{k,t-1}| - \mathbb{E}[\eta_{k,t-1}|]\Big) + \alpha_{2,k} r_{t-1} + \beta_k \ln\left(\sigma^2_{k,t-1}\right), \text{ for } k = 1, 2. \tag{3.8.3}$$

The expectation $\mathbb{E}[|\eta_{k,t-1}|]$ is taken with respect to the distribution conditional on regime k. In this case, we have $\theta_k = (\alpha_{0,k}, \alpha_{1,k}, \alpha_{2,k}, \beta_k)'$. Covariance-stationarity in each regime is obtained by requiring that $\beta_k \geq 1$.

3.9 Conditional distributions

In this section, we provide the distributions that were utilized to model the financial log returns, each of which is standardized to have a zero mean and unit variance

3.9.1 Normal distribution

The probability density function of a random variable r_t with mean μ and variance σ is defined as

$$f(r_t) = \frac{1}{\sqrt{2\pi}\sigma} \exp -\frac{1}{2}\left(\frac{r_t - \mu}{\sigma}\right)^2 \quad \text{for} \quad -\infty < r_t < \infty \tag{3.9.1}$$

The mean, variance, skewness and kurtosis are defined as $\mathbb{E}(r_t) = 0$, $Var(r_t) = \sigma^2$, Skew$(r_t) = 0$, and Excess Kurt$(r_t) = 3$, respectively. Let $f(r_t|\mathcal{F}_{t-1})$ be a normal distribution with mean μ_t and variance σ_t^2, then the likelihood function of $f(r_t|\mathcal{F}_{t-1})$ is given by

$$f(r_1, ..., r_T; \theta) = f(r_1; \theta) \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left[-\frac{(r_t - \mu_t)^2}{2\sigma_t^2}\right] \tag{3.9.2}$$

Here, r_1 is the initial observation whose marginal density function is expressed as, $f(r_1; \theta)$. The maximum likelihood estimate of θ is obtained by maximizing Equation (3.9.2) as below;

$$L = \ln f(r_1, ..., r_T; \theta) = \ln f(r_1; \theta) - \frac{1}{2}\sum_{t=2}^{T}\left[\ln 2\pi + \ln \sigma_t^2 + \frac{(r_t - \mu_t)^2}{\sigma_t^2}\right] \tag{3.9.3}$$

3.9.2 Student-t distribution

Let x_v be a student-t distribution with v degrees of freedom, then $Var(x_v) = v/(v-2)$ for $v > 2$. If we define $\varepsilon_t = x_v/\sqrt{v/(v-2)}$, then the probability density function of ε_t is

$$f(\varepsilon_t|v) = \frac{\Gamma(v+1)/2}{\Gamma(v/2)\Gamma(v-2)\pi}\left[1 + \frac{\varepsilon_t^2}{v-2}\right]^{-(1+v)/2}, \quad v > 2 \tag{3.9.4}$$

where $\Gamma(x) = \int_0^\infty y^{x-1}e^{-x}dy$ is a Gamma function. The mean, variance, skewness and Excess kurtosis of the distribution are $\mathbb{E}(\varepsilon_t) = 0$, for $v \geq 2$, $Var(\varepsilon_t) = \frac{v}{v-2}$ for $v \geq 3$, Skew$(\varepsilon_t) = \frac{\mu_4}{\sigma^2} - 3$ for $v \geq 3$, and Excess Kurt$(\varepsilon_t) = \frac{6}{v-4}$ for $v \geq 5$, respectively. Moreover, the conditional likelihood function of r_t is given by

$$f(r_{m+1}, ..., r_T|\alpha, A_m) = \prod_{t=m+1}^{T} \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{(v-2)\pi}} \frac{1}{\sigma_t}\left[1 + \frac{r_t^2}{(v-2)\sigma_t^2}\right]^{(v+1)/2} \tag{3.9.5}$$

where $r_t = \sigma_t\varepsilon_t$, $v > 2$ and $A_m = (r_1, ..., r_m)$. The conditional log-likelihood function of the student-t distribution is provided as below for specified degrees of freedom v.

$$L = \ln(r_{m+1}, ..., r_T|\alpha, A_m) = -\sum_{t=m+1}^{T}\left[\frac{v+1}{2}\ln\left(1 + \frac{r_t^2}{(v-2)\sigma_t^2}\right) + \frac{1}{2}\ln \sigma_t^2\right] \tag{3.9.6}$$

3.9.3 The generalized error distribution(GED)

The GED is a symmetric distribution that can be both leptokurtic and platykurtic depending on the degree of freedom v and its likelihood function is given by

$$f(r_t, \mu_t, \sigma_t, v) = \frac{v e^{-\frac{1}{2}|\frac{r_t - \mu_t}{\sigma_t \lambda}|^v}}{\lambda 2^{(1+1/v)}\Gamma(1/v)}, \; 1 < r_t < \infty, \text{ where } \lambda = \Big[\frac{\Gamma(1/v)}{4^{(1/v)}\Gamma(3/v)}\Big]^{1/2} \tag{3.9.7}$$

and $v > 0$ is the degree of freedom or tail thickness parameter. The GED becomes a normal distribution if $v = 2$, and for $v < 1$, the density function has thicker tails compared with the normal density function, whereas it has thinner tails for $v > 2$. The mean, variance, skewness and Excess kurtosis of the distribution are $\mathbb{E}[r_t] = \mu$, $Var(r_t) = \alpha^2 \frac{\Gamma(3/v)}{\Gamma(1/v)}$, skew$(r_t) = 0$, and Excess kurt$(r_t) = \frac{\Gamma(5/v)\Gamma(1/v)}{\Gamma(3/v)^2} - 3$, respectively. The likelihood function of the above Equation (3.9.7) is maximized by

$$L = \ln(\frac{v}{\lambda}) - \frac{1}{2}|\frac{r_t - \mu_t}{\sigma_t \lambda}|^v - (1 + 1/v)\ln 2 + \ln[\Gamma(1/v)] \tag{3.9.8}$$

3.10 Estimation of model parameters

Parameter estimation is essential in financial time series modeling because it enables the researcher to measure and quantify different influences between the variables in question. In this study, the estimation of the model parameters has been done using the maximum likelihood estimation method.

3.10.1 The ARCH model parameter estimation

To estimate the ARCH model parameters by the maximum likelihood estimate approach under the assumption of normality, the likelihood function of an ARCH(p) model is expressed as

$$\begin{aligned} f(r_1, ..., r_n|\alpha) &= f(r_n|F_{n-1})f(r_{n-1}|F_{n-2})...f(r_1|F_0)f(r_1, ..., r_p|\alpha) \\ &= \textstyle\prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\Big[-\frac{r_t^2}{2\sigma_t^2}\Big] f(r_1, ..., r_p|\alpha) \end{aligned} \tag{3.10.1}$$

where $\alpha = (\omega, \alpha_1, ..., \alpha_p)'$ and $f(r_1, ..., r_p|\alpha)$ is the joint probability density function of $r_1, ..., r_p$. It is noted that the exact form of $f(r_1, ..., r_p|\alpha)$ is complicated, and hence it is dropped from the prior likelihood function since the sample is sufficiently large. As a result, the conditional likelihood function is considered as follows:

$$f(r_1, ..., r_n|\omega, \alpha_1, ..., \alpha_p) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\Big[-\frac{r_t^2}{2\sigma_t^2}\Big], \tag{3.10.2}$$

where σ_t^2 can be evaluated recursively. The estimates generated by maximizing the previous likelihood function are referred to as the conditional maximum likelihood estimates (MLEs) under normality. It is further noted that, to maximize the conditional likelihood is equivalent to maximizing its logarithm, which is easier to handle. It thus follows that the conditional log likelihood function is

$$\ln(r_1, ..., r_p|\omega, \alpha_1, ..., \alpha_p) = \sum_{t=1}^{n} \Big[-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma_t^2 - \frac{r_t^2}{2\sigma_t^2}\Big] \tag{3.10.3}$$

Here the first term $\ln(2\pi)$ does not involve any parameters, and thus the log likelihood function becomes

$$\ln(r_1, ..., r_p|\omega, \alpha_1, ..., \alpha_p) = -\frac{1}{2}\sum_{t=1}^{n} \Big[\ln\sigma_t^2 + \frac{r_t^2}{\sigma_t^2}\Big], \tag{3.10.4}$$

where $\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + ... + \alpha_p r_{t-p}^2$. Differentiating Equation (3.10.4) with respect to ω and equating to zero we get

$$\frac{\partial \ln}{\partial \omega} = -\frac{1}{2}\sum_{t=1}^{n} \Big[\frac{1}{\sigma_t^2} - \frac{r_t^2}{\sigma_t^4}\Big] = 0 \tag{3.10.5}$$

In a similar manner, the derivative of Equation (3.10.4) with respect to θ for $\theta = \alpha_1, \alpha_2, ..., \alpha_p$ and equating to zero is given as

$$\frac{\partial \ln}{\partial \theta} = -\frac{1}{2}\sum_{t=1}^{n} \Big[\sum_{i=1}^{p}(1 + [p-1]\alpha_i r_{t-i}^2)\Big(\frac{1}{\sigma_t^2} - \frac{r_t^2}{\sigma_t^4}\Big)\Big] = 0 \tag{3.10.6}$$

The estimates of ω and $\theta = \alpha_1, \alpha_2, ..., \alpha_p$ can be determined by solving Equations (3.10.5) and (3.10.6) recursively, respectively.

3.10.2 The GARCH model parameter estimation

To estimate the GARCH model as expressed in Equation (3.5.1), since ε_t is normally distributed with mean zero and conditional variance σ_t , that is, $\varepsilon_t \sim N(0, \sigma_t^2)$, then

$$p(\varepsilon_t|\varepsilon_{t-1}, ..., \varepsilon_0) = \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-\frac{\varepsilon_t^2}{2\sigma_t^2}} \tag{3.10.7}$$

The log-likelihood function of parameter vector $\theta = (\alpha_0, \alpha_1, ..., \alpha_p, \beta_0, \beta_1, ..., \beta_q)'$ is

$$(\theta) = \sum_{t=1}^{n} l_t(\theta) = \sum_{t=1}^{n}(-\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\sigma_t^2 - \frac{\varepsilon_t^2}{2\sigma_t^2}) \tag{3.10.8}$$

Taking first and second partial derivatives of Equation (3.10.8) w.r.t parameter vector θ and rearranging results to

$$\frac{\partial l_t(\theta)}{\partial\theta} = \sum_{t=1}^{n}(\frac{\varepsilon_t^2}{2\sigma_t^2} - \frac{1}{2\sigma_t})\frac{\partial\sigma_t}{\partial\theta} \tag{3.10.9}$$

$$\frac{\partial^2 l_t\theta}{\partial\theta\partial\theta'} = (\frac{\varepsilon_t^2}{2\sigma_t^2} - \frac{1}{2\sigma_t})\frac{\partial^2\sigma_t}{\partial\theta\partial\theta'} + (\frac{1}{2\sigma_t^2} - \frac{\varepsilon_t^2}{\sigma_t^3})\frac{\partial\sigma_t}{\partial\theta}\frac{\partial\sigma_t}{\partial\theta'}, \tag{3.10.10}$$

where $\frac{\partial\sigma_t}{\partial\theta} = (1, \varepsilon_{t-1}^2, ..., \varepsilon_{t-p}^2,\ \sigma_{t-1}, ..., \sigma_{t-q})' + \sum_{i=1}^{q}\beta_i\frac{\partial\sigma_{t-i}}{\partial\theta}$. Setting Equation (3.10.9) equal to zero becomes complex to solve analytically and calls for a numerical approach. For instance, the Newton-Raphson procedure can be used to solve this case. The iteration scheme for this method has the form

$$\vec{x}_{k+1} = \vec{x}_k - H_f^{-1}(\vec{x}_k)\nabla f(\vec{x}_k), \tag{3.10.11}$$

where $H_f^{-1}(\vec{x})$ is the Hessian matrix of the second partial derivative of f.

$$H_f(\vec{x})_{ij} = \frac{\partial^2 f(\vec{x})}{\partial x_i\partial x_j} \tag{3.10.12}$$

Thus it is observed that for the maximum likelihood problem, $\hat{\theta} = \underset{\theta\in\Theta}{argmax}L(\theta)$, the approximated value of θ, denoted as θ_k, can be obtained after the k^{th} iteration,

$$\theta_{k+1} = \theta_k - J^{-1}(\theta_k)\nabla L(\theta_k) \tag{3.10.13}$$

with $\nabla L = \frac{\partial L}{\partial\theta}$ and Fisher information matrix $J = E\big(\frac{\partial^2}{\partial\theta\partial\theta'}\big)$. Applying this method to estimate the GARCH(p,q) parameters we have

$$\nabla L(\theta) = \frac{1}{2}\sum_{t=1}^{n}(\frac{\varepsilon_t^2}{\sigma_t^2} - \frac{1}{\sigma_t})\frac{\partial\sigma_t}{\partial\theta} \text{ and} \tag{3.10.14}$$

$$J = \sum_{t=1}^{n}E\Big[(\frac{\varepsilon_t^2}{2\sigma_t^2} - \frac{1}{2\sigma_t})\frac{\partial^2\sigma_t}{\partial\theta\partial\theta'} + (\frac{1}{2\sigma_t^2} - \frac{\varepsilon_t^2}{\sigma_t^3})\frac{\partial\sigma_t}{\partial\theta}\frac{\partial\sigma_t}{\partial\theta'}\Big] = -\frac{1}{2}\sum_{t=1}^{n}E\Big(\frac{1}{\sigma_t^2}\frac{\partial\sigma_t}{\partial\theta}\frac{\partial\sigma_t}{\partial\theta'}\Big) \tag{3.10.15}$$

In this study, GARCH(1,1) model is employed and whose parameters $\theta = (\alpha_0, \alpha_1, \beta_1)'$ are estimated as follows:

$$\nabla L(\theta = \frac{1}{2}\sum_{t=1}^{n}(\frac{\varepsilon_t^2}{\sigma_t^2} - \frac{1}{\sigma_t})\frac{\partial\sigma_t}{\partial\theta}, \text{ and} \tag{3.10.16}$$

$$J = -\frac{1}{2}\sum_{t=1}^{n}E\Big(\frac{1}{\sigma_t^2}\frac{\partial\sigma_t}{\partial\theta}\frac{\partial\sigma_t}{\partial\theta'}\Big), \text{ where } \frac{\partial\sigma_t}{\partial\theta} = (1, \varepsilon_{t-1}^2\sigma_{t-1})' + \beta_1\frac{\partial\sigma_{t-1}}{\partial\theta} \tag{3.10.17}$$

3.10.3 MS-GARCH model parameter estimation

The maximum likelihood estimation technique is utilized to estimate the MS-GARCH as reported by Ardia et al. (2019). Let vector $\Psi = (\theta_1, \epsilon_1, ..., \theta_k, \epsilon_k, \mathbb{P})$ represent the model parameters, then the likelihood function is given by

$$L(\Psi|\mathcal{F}_T) = \prod_{t=1}^{T} f(r_t|\Psi, \mathcal{F}_{t-1}) \tag{3.10.18}$$

where $f(r_t|\Psi, \mathcal{F}_{t-1})$ is the probability density function of r_t given the past observations, $\mathcal{F}_{t-1}$ and the parameters of the model Ψ . Therefore, the conditional density of r_t , for MS-GARCH model is given by

$$f(r_t|\Psi, \mathcal{F}_{t-1}) = \sum_{i=1}^{K}\sum_{j=1}^{K} P_{ij} z_{i,t-1} f_D(r_t|\xi(t) = j, \Psi, \mathcal{F}_{t-1}) \tag{3.10.19}$$

where $z_{i,t-1} = P[\xi(t-1) = j|\Psi, \mathcal{F}_{t-1}]$ is the filtered probability of state i at time $t-1$ obtained via Hamilton filter, see Hamilton (1989). In Equation (3.10.19), the density of r_t in state $\xi(t) = k$ given Ψ and $\mathcal{F}_{t-1}$ is denoted by $f_D(r_t|\xi(t) = k, \Psi, \mathcal{F}_{t-1})$. In this study the standardized innovations, $\eta k, t$, of the model in each regime is assumed to be conditional to three distributions, that is, the normal, generalized error distribution(GED) and student-t distributions. The standardization is such that each distribution has zero mean and a unit variance. For notation purposes, the time and regime indices are dropped but the shape parameters are conditional on the regimes. The maximum likelihood estimator $\tilde{\Psi}$ is thus obtained by maximizing Equation (3.10.18).

In the next section, the knowledge of the stock returns volatility dynamics is applied in pricing options.

3.11 Option pricing

In this section the Black-Scholes model is derived using the change of measure technique. A regime-switching model is also derived for the European option pricing and which is further extended to incorporate GARCH effects resulting to a regime-switching GARCH model. These models are implemented using a real data as shown in Section 4.4 and the model with the best performance established by running the Root Mean Square Error (RMSE) test.

Hamilton (1989) developed the Markov Switching (Regime Switching) model which is among the most popular nonlinear time series models in literature that have been used to model dynamics of stock market data. The model involves multiple structures (equations) that can characterize the time series behaviors in different regimes and is able to capture more complex dynamic patterns if switching between these structures is permitted as well as it allows coefficients of the conditional mean and variance to vary according to some finite-valued stochastic process with states or regimes.

Definition 3.11.1 *Let $\{\xi(t), t \in [0,T]\}$ be a discrete-time random variable that can only assume an integer value $\{1,2,...,k\}$. The process is said to be in state i at time t if $\xi(t) = i$ and whenever in state i, the probability that it will next be in state j is P_{ij}. That is, the process is said to be a Markov chain if*

$$\begin{aligned} P\{\xi(t+1) = j|\xi(t) = i, \xi(t-1) = k, ...,\} \\ &= P(\xi(t+1) = j|\xi(t) = i) \\ &= P_{ij}, \ \textit{for all states} \ k, i, j, ..., \ \textit{and} \ t \geq 0 \end{aligned} \tag{3.11.1}$$

Remark 3.11.1 *Equation 3.11.1 describes an k-state Markov chain with $P_{i,j}$ representing the transition probabilities. Also, $0 \leq \ P_{i,j} \ \leq 1, \ i,j \geq 0, \ \sum_{j=1}^{k} P_{i,j} = 1, \ i,j = 1,2,...,k$*

3.11.1 Regime switching market dynamics

Let $(X_t)_{t\geq 0}$ be a stochastic process on a discrete time set $\mathcal{T} = \{0,...,T\}$ and $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, where $\mathbb{P}$ is a physical probability measure. Further, let $\xi(t)$ be a Markov chain with $\mathcal{M} = \{1,...,k\}$ states or regimes and transition probability P_{ij} . We consider a risk-free asset, $dS_t^* = rS_t^* dt$ that is assumed to be continuously compounded in value at a constant risk-free rate, r, (but not necessarily non-negative) over the trading time interval [0,T]. Taking $S_0^* = 1$, the price of the asset at time t, is given by $S_t^* = e^{rt}$ for $t \in [0,T]$. Consider a second market which is risky and whose price is defined as

$$S_t = S_0 \exp\Big(\sum_{j=1}^{t} X_t\Big), \ \ t \in \mathcal{T} \tag{3.11.2}$$

where S_0 is the initial price of the asset. The log-returns for the asset are given by

$$X_j = \mu_j + \sigma_j r_t, \ \ \text{where} \ \ r_t \ \sim \ i.i.d\, N(0,1). \tag{3.11.3}$$

The constants μ_j and σ_j represents the mean and standard deviation of log-returns under each regime and r_t is a sequence of i.i.d random variables with zero mean and unit variance under measure $\mathbb{P}$ and independent from the Markov chain $\xi(t)$. The filtration $\mathcal{F}_t$ on the probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$ contains all the information generated by all the underlying asset prices and realized regimes up to time T. However, the regimes are latent variables hence this filtration characterizes only a partial information to investors. Therefore, we assume that for $j \in \{1, ..., k\}$,

$$\mathbb{P}[\xi(t+1) = j|\mathcal{F}_t] = P_{ij} \tag{3.11.4}$$

where $P_{i,j}$ is the transition probability from state $i \to j$ of the Markov chain $\xi(t)$.

3.11.2 Risk-neutral measure

Suppose that $(\Omega, \mathcal{F}, \{\mathcal{F}_t : 0 \le t \le T\}, \mathbb{P})$ is a filtered probability space on which a standard Brownian motion $W = \{W_t : 0 \le t \le T\}$ is constructed. Let $S = \{S_t : 0 \le t \le T\}$, be the risky asset price process that follows a stochastic differential equation(SDE) defined by

$$dS_t = \mu S_t d_t + \sigma S_t dW_t \tag{3.11.5}$$

where μ and σ are constants representing the drift and volatility terms of the process, and W_t is a Wiener process that starts at zero, that is, $W_0 = 0$. Equation (3.11.5) has the solution

$$S_t = S_0 \exp\big(\int_0^t (\mu - \frac{\sigma^2}{2}) ds + \int_0^t \sigma dW_s \big) \tag{3.11.6}$$

Equivalently, this equation can also be expressed as

$$\begin{aligned} X_t = \log_e(\frac{S_t}{S_0}) \implies X_t &= (\mu - \frac{\sigma^2}{2})t + \sigma W_t \\ &= (\mu - \frac{\sigma^2}{2})t + \sigma\sqrt{t}\varepsilon_t), \ \ \varepsilon_t \sim i.i.d\, N(0,1) \end{aligned} \tag{3.11.7}$$

The solution of Equation (3.11.5) is derived using the Ito's formula as follows; let

$$\begin{aligned} f(t, S_t) = \log_e S_t, &\implies \frac{\partial}{\partial_t} f(t, S_t) = 0, \frac{\partial}{\partial S} f(t, S_t) = \frac{1}{S_t} \text{ and } \frac{\partial^2}{\partial^2 S} f(t, S_t) = -\frac{1}{S^2} \\ df &= \frac{1}{S_t}(\mu S_t d_t + \sigma S_t dW_t) - \frac{-1}{2S^2}(\mu S_t d_t + \sigma S_t dW_t)^2 \\ &= \mu dt + \sigma dW_t - \frac{\sigma^2}{2} dt = \big(\mu - \frac{\sigma^2}{2}\big) ds + \sigma dW_s \\ \int_0^t d(\log_e S_t) &= \int_0^t (\mu - \frac{\sigma^2}{2}) dt + \int_0^t \sigma dW_t \\ \log_e(\frac{S_t}{S_0}) &= (\mu - \frac{\sigma^2}{2})t + \sigma dW_t \implies S_t = S_0 \exp\big((\mu - \frac{\sigma^2}{2})t + \sigma W_t\big) \end{aligned} \tag{3.11.8}$$

Let $X_t(h)$ be the log return over the time interval $[t, t+h]$ where $h \geq 0$ is the length of time between the stock price observations, then

$$\begin{aligned}
X_t &= \log_e\big(\frac{S_{t+h}}{S_t}\big) \\
&= (\mu - \frac{\sigma^2}{2})[(t+h) - t] + \sigma W[(t+h) - t] \\
&= (\mu - \frac{\sigma^2}{2})h + \sigma W h \implies \quad X_t(h) \sim\ N((\mu - \frac{\sigma^2}{2})h, \sigma^2 h) \\
\text{when } h = 1, X_t &= (\mu - \frac{\sigma^2}{2}) + \sigma\varepsilon_t,\ \ \forall\ i \geq 0
\end{aligned} \tag{3.11.9}$$

Further, we consider a discounted stock price, $\tilde{S}_t$ defined by $\tilde{S}_t = S_0 e^{-rt}$. Utilizing Equation (3.11.6) we get

$$\begin{aligned}
\tilde{S}_t &= S_0 \exp(-\int_0^t r ds) \exp\Big(\int_0^t (\mu - \frac{\sigma^2}{2}) ds + \int_0^t \sigma dW_s\Big) \\
&= S_0 \exp\Big\{(\mu - r)t - \frac{\sigma^2}{2} t + \sigma W_t\Big\} \\
&= S_0 \exp\Big\{-\frac{1}{2}\sigma^2 t + \sigma\big(\frac{\mu - r}{\sigma} + W_t\big)\Big\}
\end{aligned} \tag{3.11.10}$$

Let

$$\tilde{W}_t = W_t + \frac{\mu - r}{\sigma} t \tag{3.11.11}$$

where $\tilde{W}_t$ is a Brownian motion under measure $\mathbb{Q}$ and $\frac{\mu - r}{\sigma}$ is the market price of a risk. Equation (3.11.10) thus becomes

$$\tilde{S}_t = S_0 \exp\Big\{-\frac{1}{2}\sigma^2 t + \sigma \tilde{W}_t\Big\} \tag{3.11.12}$$

By Ito's lemma we get the SDE

$$d\tilde{S}_t = \sigma \tilde{S}_t d\tilde{W}_t \tag{3.11.13}$$

Also, utilizing rules within Ito's calculus we differentiate $\tilde{W}_t = W_t + \frac{\mu - r}{\sigma} t$ with respect to t to get

$$d\tilde{W}_t = dW_t + \frac{\mu - r}{\sigma} dt \tag{3.11.14}$$

Using Equation (3.11.14), we re-arrange Equation (3.11.5) as follows;

$$\begin{aligned}
dS_t &= \mu S_t dt + \sigma S_t[d\tilde{W}_t - \frac{\mu - r}{\sigma} dt] \\
&= \mu S_t dt + \sigma S_t d\tilde{W}_t - \mu S_t dt + r S_t dt \\
&= r S_t dt + \sigma S_t d\tilde{W}_t
\end{aligned} \tag{3.11.15}$$

Notice that the drift term $\mu = r$ which implies risk neutrality and hence the discounted price process, $\tilde{S}_t$, is $\mathcal{F}$-martingale under measure $\mathbb{Q}$. This leads us to the following definition

Definition 3.11.2 *A risk neutral measure is a measure $\mathbb{P}$ that is equivalent to $\mathbb{Q}$ under which the discounted stock price process $\tilde{S}_t$ is a martingale.*

Remark 3.11.2 *If W_τ is a Brownian motion(BM) under measure $\mathbb{P}$ and then we shift the process by $Y(t) = \sigma t$, then the shifted process is BM under measure $\mathbb{Q}$, that is, $\tilde{W}_t = W_t - \sigma_t$ is BM under $\mathbb{Q}$ defined via $\frac{d\mathbb{Q}}{d\mathbb{P}} = e^{\sigma W_t - \frac{1}{2}\sigma^2 t}$. This is the Radon-Nikodym derivative.*

3.11.3 Derivation of the Black-Scholes formula by change of measure

The stock price, under the risk-neutral measure has the following dynamics

$$dS_t = rS_t dt + \sigma S_t d\tilde{W}_t \tag{3.11.16}$$

whose solution is given by

$$S_T = S_0 \exp\Big(\int_0^t (r - \frac{\sigma^2}{2})ds + \int_0^t \sigma d\tilde{W}_s\Big) \tag{3.11.17}$$

Suppose that the payoff of a call option on a stock at a future time T is $\varphi(S_T)$, where $\varphi(S_T)$ is a random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ describing the market. The payoff is defined as;

$$\begin{aligned} \varphi(S_T) &= Max(S_T - K, 0) = (S_T - K)_+, \quad \text{for } K \geq 0 \\ &= S_t \exp\Big(\int_t^T (r - \frac{\sigma^2}{2})ds + \int_t^T \sigma d\tilde{W}_s\Big) - K)_+ \end{aligned} \tag{3.11.18}$$

Moreover, suppose that the discount factor from now (time, t) until time T is $D_t = \exp(-\int_t^T rds)$, then today's price of the call option, C_t, under the risk-neutral measure is just the discounted expected value of its payoff. That is,

$$\begin{aligned} C_t &= \exp(-\int_t^T rds)\mathbb{E}^{\mathbb{Q}}[S_t \exp\Big(\int_t^T (r - \frac{\sigma^2}{2})ds + \int_t^T \sigma d\tilde{W}_s\Big) - K)_+|\mathcal{F}_t] \\ &= e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[\varphi(S_T)|\mathcal{F}_t] \\ &= e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[\varphi(S_T)I_{S_T > K}] \end{aligned} \tag{3.11.19}$$

where $\mathbb{Q}$ denotes the risk-neutral measure. This expression can be re-stated in terms of the physical measure $\mathbb{P}$ as follows;

$$C_t = e^{-r(T-t)}\mathbb{E}^{\mathbb{P}}\Big(\frac{d\mathbb{Q}}{d\mathbb{P}}\varphi(S_T)|\mathcal{F}_t\Big) \tag{3.11.20}$$

where $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is the Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$. Note that, here, $\tau = T - t$ and $I_{S_T \geq K}$ is an indicator function which takes value of 1 if the stock price at maturity is greater than the strike price. Equation (3.11.19) can thus be written as follows;

$$\begin{aligned} C_t &= e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[S_T I_{S_T>K} - K I_{S_\tau>K}] \\ &= e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[S_T I_{S_T>K}] - Ke^{-r\tau}\mathbb{E}^{\mathbb{Q}}[I_{S_T>K}] \end{aligned} \tag{3.11.21}$$

We first evaluate the second term in the right hand side (RHS) of Equation (3.11.21) as follows;

$$\begin{aligned} Ke^{-r\tau}\mathbb{E}^{\mathbb{Q}}[I_{S_T>K}] &= Ke^{-r\tau}\mathbb{Q}[S_T > K] \\ &= Ke^{-r\tau}\mathbb{Q}[S_t \exp((r - \frac{\sigma^2}{2})\tau + \sigma\tilde{W}_\tau) > K] \\ &= Ke^{-r\tau}\mathbb{Q}\Big[\ln S_t + (r - \frac{1}{2}\sigma^2)\tau + \sigma\tilde{W}_\tau > \ln K\Big] \\ &= Ke^{-r\tau}\mathbb{Q}\Big[\tilde{W}_\tau > \frac{\ln K - \ln S_t - (r - \frac{1}{2}\sigma^2)\tau}{\sigma}\Big] \end{aligned} \tag{3.11.22}$$

Since $\tilde{W}_\tau$ is normally distributed with mean zero and variance τ, that is, $\tilde{W}_\tau \sim N[0, \tau]$, it follows that $\frac{\tilde{W}_\tau}{\sqrt{\tau}} \sim N[0, 1]$. Therefore, Equation (3.11.22) becomes

$$Ke^{-r\tau}\mathbb{E}^{\mathbb{Q}}[I_{S_\tau>K}] = Ke^{-r\tau}N\Big[\frac{-\ln K + \ln S_t + (r - \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}\Big] \tag{3.11.23}$$

Next we evaluate the first term in the right hand side (RHS) of Equation (3.11.21) as below;

$$\begin{aligned} e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[S_T I_{S_T>K}] &= e^{-r\tau}\mathbb{E}^{\mathbb{Q}}\big[S_t e^{(r-\frac{1}{2}\sigma^2)\tau+\sigma W_\tau} I_{S_T>K}\big] \\ &= e^{-r\tau}\mathbb{E}^{\mathbb{Q}}\big[S_t e^{r\tau} e^{-\frac{1}{2}\sigma^2\tau+\sigma W_\tau} I_{S_T>K}\big] \\ &= S_t\mathbb{E}^{\mathbb{Q}}\big[e^{-\frac{1}{2}\sigma^2\tau+\sigma W_\tau} I_{S_T>K}\big] \end{aligned} \tag{3.11.24}$$

Using the above Remark (3.11.2), Equation (3.11.24) can be written as

$$e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[S_T I_{S_T>K}] = S_t\mathbb{E}^{\mathbb{P}}[\frac{d\mathbb{Q}}{d\mathbb{P}} I_{S_T>K}] = S_t\mathbb{E}[I_{S_T>K}] \tag{3.11.25}$$

Under the new measure the dynamics of the stock price S_T changes as follows;

$$\begin{aligned} S_T &= S_t e^{(r-\frac{1}{2})\tau+\sigma W_\tau} = S_t e^{(r-\frac{1}{2}\sigma^2)\tau+\sigma(\tilde{W_\tau}+\sigma\tau)} \\ &= S_t e^{(r+\frac{1}{2}\sigma^2)\tau+\sigma\tilde{W_\tau}} \end{aligned} \tag{3.11.26}$$

Now substituting Equations (3.11.26) to (3.11.24) results to

$$\begin{aligned} e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[S_T I_{S_T>K}] &= S_t\mathbb{E}^{\mathbb{Q}}[I_{S_T>K}] \\ &= S_t\mathbb{Q}\big[S_T > K\big] \\ &= S_t\mathbb{Q}\big[S_t e^{(r+\frac{1}{2}\sigma^2)\tau+\sigma\tilde{W_\tau}} > K\big] \\ &= S_t\mathbb{Q}\big[\ln S_t + (r+\frac{1}{2}\sigma^2)\tau + \sigma\tilde{W_\tau} > \ln K\big] \\ &= S_t\mathbb{Q}\Big[\tilde{W_\tau} > \frac{\ln K - \ln S_t - (r+\frac{1}{2}\sigma^2)\tau}{\sigma}\Big] \\ &= S_t\mathbb{Q}\Big[\frac{\tilde{W_\tau}}{\sqrt{\tau}} > \frac{\ln K - \ln S_t - (r+\frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}\Big] \\ &= S_t N\Big[\frac{-\ln K + \ln S_t + (r+\frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}\Big] \end{aligned} \tag{3.11.27}$$

Putting Equations (3.11.23) and (3.11.27) into Equation (3.11.21) the contingent claim price is given by

$$\begin{aligned} C_t &= S_t N\Big[\frac{-\ln K + \ln S_t + (r+\frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}\Big] + Ke^{-r\tau}N\Big[\frac{-\ln K + \ln S_t + (r-\frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}\Big] \\ &= S_t N(d_1) - Ke^{-r\tau}N(d_2). \end{aligned} \tag{3.11.28}$$

This is the Black-Scholes formula.

3.11.4 Regime-switching (RS) model

We consider a risk-free asset and a risky asset as discussed earlier in Section 3.11.1. These assets are tradable continuously over time in a finite time horizon $\tau = [t, T]$ where $T < \infty$. In order to describe uncertainty, we consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where $\mathbb{P}$ is a real world probability measure. According to Yao *et al.*(2006), in the regime switching world, one typically modulates the rate of return and the volatility by a finite-state Markov chain $\{\xi(t), t \geq 0\}$, which represents the market regime. Let $\{\xi(t), t \geq 0\}$ be continuous, finite-state, Markov process on $(\Omega, \mathcal{F}, \mathbb{P})$ with state space $\mathcal{M} = \{1, 2, ..., k\}$ and that S_t is the stock price at time t satisfying the stochastic differential equation defined by

$$\frac{dS_t}{S_t} = \mu_{\xi(t)}dt + \sigma_{\xi(t)}dW_t, \ \ \xi(t) = \{1, 2, ..., k\} \tag{3.11.29}$$

where $S_t > 0$ is the initial price, W_t is a standard Brownian motion independent of $\xi(t)$. The parameters $\mu_{\xi(t)}$ and $\sigma_{\xi(t)}$ denote the expected rate of return on the asset and volatility of the asset price respectively, and are assumed to be constant and distinct for each regime. In this way, $\xi(t)$ is regarded as a variable that chooses one of the states in $\mathcal{M}$ at time t of the market. Equation (3.11.29) is solved using Itô lemma to give

$$S_T = S_t e^{(\mu_{\xi(t)} - \frac{1}{2}\sigma^2_{\xi(t)})\tau + \sigma_{\xi(t)} W_\tau}, \ \text{for} \ \xi(t) = \{1, 2, ..., k\} \tag{3.11.30}$$

The stock price process in Equation (3.11.30) is assumed to exhibit regime switches. Let the price process undergo discrete shifts between regimes $\xi(t)$ and that it follows a first-order Markov chain, then the transition probability P_{ij} from state i at time t to state j at time $t-1$ is denoted by

$$P_{ij} = P\big\{\xi(t) = j | \xi(t-1) = i\big\} \ \text{for all} \ i, j = \{1, 2, ..., k\} \tag{3.11.31}$$

where $0 \leq P_{ij} \leq 1$ and $\sum_{j=1}^{k} P_{ij} = 1$. The transition matrix P_{ij} of the Markov chain is given by

$$P_{ij} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ . & . & \cdots & . \\ p_{k1} & . & \cdots & p_{kk} \end{bmatrix}$$

In particular, for a very small time interval $\delta > 0$,

$$P(T < \delta) = 1 - e^{-\lambda_i \delta} = 1 - (1 - \lambda_i \delta) = \lambda_i \delta \tag{3.11.32}$$

That is, in a short interval of time $\delta > 0$, the probability of leaving state i is approximately $\lambda_i \delta$. This means, λ_i is the transition rate out of state i and can formally be expressed as

$$\lambda_i = \lim_{\delta \to 0^+} \big[\frac{P(\xi(T) = j | \xi(t) = i}{\delta}\big]$$

Since P_{ij} is the probability of moving from state i to state j, the quantity $g_{ij} = \lambda_i P_{ij}$ is the transition rate from state i to state j.

Definition 3.11.3 (Generator matrix) *Define a generator matrix, $\mathbb{G}$ for a continuous-time Markov chain. The $(i,j)^{th}$ element of the transition matrix is given by*

$$g_{ij} = \begin{cases} \lambda_i P_{ij}, & i \neq j \\ -\lambda_i, & i = j \end{cases} \quad , \quad \textit{Also } g_{ii} = -\sum_{i \neq j} g_{ij} \tag{3.11.34}$$

The generator matrix of the Markov chain is thus given by

$$G = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \ldots & \lambda_{1k} \\ \lambda_{21} & \lambda_{22} & \ldots & \lambda_{2k} \\ . & . & \ldots & . \\ . & . & \ldots & . \\ . & . & \ldots & . \\ \lambda_{k1} & \lambda_{22} & \ldots & \lambda_{kk} \end{bmatrix}$$

and, for each state, the elements of the generator matrix satisfy the equation $\lambda_{jj} + \sum_{i=1, i\neq j}^{k} \lambda_{ij} = 0$. For small time interval $\delta > 0$ and i the probability to be in state j at time $t + \delta$ given that the random variable was in state i at time t is given by

$$P\big(X(t+\delta) = j | X(t) = i\big) = g_{ij}.\delta + O(\delta) \tag{3.11.35}$$

where $O(\delta)$ is the probability from state i to state j in more than one step but $\lim_{\delta \to 0} \left[\frac{O(\delta)}{\delta}\right] = 0$.

The model developed here postulates that the regimes are unobservable and thus state transitions until maturity is considered. Denote by $\xi(t)$, the variable representing the regime in which the process was at time t. When the change in log price X_t is in regime j, it is presumed to have been drawn from a normal distribution with mean, μ_j and variance, σ_j^2 . Hence, the probability density function of X_t conditional on $\xi(t)$ taking on the value j is given by

$$f(X_t | \xi(t) = j; \theta) = \frac{1}{\sigma_j \sqrt{2\pi}} \exp\left\{ -\frac{(X_t - \mu_j)^2}{2\sigma_j^2} \right\} \tag{3.11.36}$$

for $j = \{1, 2, ..., k\}$ and θ is a vector of parameters μ_j, σ_j for $j = 1, 2, ..., k$, that is, $\theta = \{\mu_j, \sigma_j\}$. Furthermore, it is presumed that the unobserved regime $\xi(t)$ have been generated by a probability distribution for which the conditional probability that $\xi(t) = j$ given the information upto time t is denoted by π_j , that is

$$P\{\xi(t) = j|\mathcal{F}_t; \theta\} = \pi_j \text{ for } j = \{1, 2, ..., k\} \tag{3.11.37}$$

Thus the vector θ also includes π_j for $j = 1, 2, ..., k$, that is, $\theta = \{\mu_j, \sigma_j, \pi_j\}'$. Next is to find the probability of the joint event that $\xi(t) = j$ and X_t falls within some time interval [t,T]. This is determined by integrating

$$p(X_t, \xi(t) = j; \theta) = f\big(X_t|\xi(t) = j; \theta\big)P\big(\xi(t) = j|\mathcal{F}_t; \theta\big) \tag{3.11.38}$$

over all values of X_t between t and T. Equation (3.11.38) is the joint probability density function of $\xi(t)$ and X_t and by utilizing Equations (3.11.36) and (3.11.37), it is expressed as

$$p(X_t, \xi(t) = j; \theta) = \frac{\pi_j}{\sigma_j\sqrt{2\pi}} \exp\Big\{ - \frac{\big(X_t - \mu_j\big)^2}{2\sigma_j^2}\Big\} \text{ for } j = \{1, 2, ..., k\} \tag{3.11.39}$$

The unconditional probability density function of X_t is determined by summing Equation (3.11.39) over all values for j:

$$\begin{aligned} f(X_t; \theta) &= \sum_{j=1}^{k} p(X_t, \xi(t) = j; \theta) \\ &= \frac{\pi_1}{\sqrt{2\pi}\sigma_1} \exp\Big\{ - \frac{\big(X_t - \mu_1\big)^2}{2\sigma_1^2}\Big\} + \ ... \ + \frac{\pi_j}{\sqrt{2\pi}\sigma_j} \exp\Big\{ - \frac{\big(X_t - \mu_j\big)^2}{2\sigma_j^2}\Big\} \end{aligned} \tag{3.11.40}$$

Again, based on the definition of conditional probability,

$$P\big[\xi(t) = j|X_t; \theta\big] = \frac{p\big(X_t, \xi(t) = j; \theta\big)}{f(X_t; \theta)} = \frac{\pi_j f(X_t|\xi(t) = j; \theta)}{f(X_t; \theta)} \tag{3.11.41}$$

Note that the magnitude in Equation (3.11.41) for each observation X_t in the sample can be calculated if the knowledge of the parameters θ is known by use of Equations (3.11.36) and (3.11.39). Equation (3.11.41) represent the probability, given the observed data, that the unobserved regime responsible for observation t was regime j. The probability of $\xi(t) = j$ given the information $\mathcal{F}_t$ upto time t, that is, $P[\xi(t) = j|\mathcal{F}_t]$ can be calculated by first calculating

$$P[\xi(t-1) = i|\mathcal{F}_{t-1}] \text{ from } P[\xi(t) = j|\mathcal{F}_{t-1}] \tag{3.11.42}$$

with the following equation

$$
\begin{aligned}
P[\xi(t) = j|\mathcal{F}_{t-1}] &= \sum_{j=1}^{2} P[\xi(t) = j, \xi(t-1) = i|\mathcal{F}_{t-1}] \\
&= \sum_{j=1}^{2} P[\xi(t) = j|\xi(t-1) = i] P[\xi(t-1) = i|\mathcal{F}_{t-1}]
\end{aligned}
\tag{3.11.43}
$$

where $P[\xi(t) = j|\xi(t-1) = i]$ is the transition probability as earlier defined in Equation (3.11.31). Next, adding the data X_t at the time t leads to the equation

$$
\begin{aligned}
P[\xi(t) = j|\mathcal{F}_{t-1}, X_t] &= \frac{f(\xi(t) = j, X_t|\mathcal{F}_{t-1})}{f(X_t|\mathcal{F}_{t-1})} \\
&= \frac{f(X_t|\xi(t) = j, \mathcal{F}_{t-1}) P[\xi(t) = j|\mathcal{F}_{t-1}]}{\sum_{j=1}^{2} f(X_t|\xi(t) = j, \mathcal{F}_{t-1}) P[\xi(t) = j|\mathcal{F}_{t-1}]}
\end{aligned}
\tag{3.11.44}
$$

Note that the probability, $P[\xi(t) = j|\mathcal{F}_{t-1}]$ for $t = 1, 2, ..., T$ can be obtained by repeating the calculations of Equations (3.11.43) and (3.11.44), and the results substituted into Equation (3.11.39).

In order to price options, it is key to derive a pricing model that produces no arbitrage (a situation where investors can make a guaranteed profit without incurring risk). To avoid arbitrage, the options are priced under risk neutral measure or martingale measure. Under the physical probability measure $\mathbb{P}$, we wish to have $e^{-\int_t^T r ds} S_t$ to be a martingale. Here, r is the risk free interest rate and it is assumed to be constant across the regimes. Since S_t denotes the price of a stock at time t which satisfies Equation (3.11.29), we assume that $\xi(t)$, and W_t are mutually independent; and $\sigma^2_{\xi(t)} > 0$, for all $\xi(t) \in \mathcal{M}$. Suppose that $\{\xi(u), W(u) : 0 \le u \le t\}$ denote the sigma field generated by $\mathcal{F}_t$. We note that all (local) martingales concerned are with respect to the filtration. It is clear that $\{W_t\}$ and $\{W_t^2 - t\}$, are both martingales. Define an equivalent measure $\mathbb{Q}$ under which the discounted stock process is a martingale. Let the risk-free rate be denoted by $r > 0$. Then for $0 \le t \le T$, let

$$
Z_t = \exp\Big(\int_t^T \beta_s dW_s - \frac{1}{2} \int_t^T \beta_s^2 ds \Big) \tag{3.11.45}
$$

where $\beta_s = \frac{\mu(\xi(u)) - r}{\sigma(\xi(u))}$. By Girsanov's theorem, the process

$$
\tilde{W}_t = W_t - \int_t^T \beta_s ds \implies d\tilde{W}_t = dW_t - \beta_t dt \tag{3.11.46}
$$

is a $\mathbb{Q}$-Brownian motion. Moreover, applying Ito's rule results to $\frac{dZ_t}{Z_t} = \beta_t dW_t$ where Z_t is a local martingale, with $\mathbb{E}[Z_t] = 1$, $0 \le t \le T$. An equivalent measure $\mathbb{Q}$ is defined via the Radon-Nikodym derivative, $\frac{d\mathbb{Q}}{d\mathbb{P}} = Z_T$. Combining Equations (3.11.29) and (3.11.46) gives

$$\frac{dS_t}{S_t} = rdt + \sigma_{\xi(t)} d\tilde{W}_t \text{ whose solution is } S_T = S_t \exp\left[(r - \frac{1}{2}\sigma^2_{\xi(t)})\tau + \sigma_{\xi(t)}\tilde{W}_\tau\right] \tag{3.11.47}$$

This model has two types of random sources, W_t and $\xi(t)$ and the inclusion of $\xi(t)$ makes the underlying market incomplete.

Let $\mathcal{R}$ denote the total time spent in regime $\xi(t) = j$ for $j = 1, 2, ..., k$ in the interval $[t, T]$ in n trials, given that at time t, the state is k. Denote the probability, $Pr(\mathcal{R} = \alpha_j)$, by p for $j = 1, 2, ..., k - 1$. For simplicity, we restrict ourselves to two regimes, i.e, $k = 2$, hence the transition matrix, discussed earlier, reduces to

$$P_{ij} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

The fraction of times spent by the Markov chain in each regime, as the numbers of transitions n become large, can be calculated using a time-average(invariant) distribution of the Markov chain, that is $\beta P = \beta$, where P is the transition matrix and β is the average fraction of the time spent in state $\xi(t) = j$ over n steps as n approaches infinity. Let $\beta = [\beta_1 \, \beta_2]$, then

$$[\beta P] = [\beta_1 \, \beta_2] \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = [\beta_1 \, \beta_2] \tag{3.11.48}$$

This results into two equations as $\beta_1 P_{11} + \beta_2 P_{21} = \beta 1$ and $\beta_1 P_{12} + \beta_2 P_{22} = \beta_2$ both of which simplify to $\beta_1 P_{12} = \beta_2 P_{21}($ since $P_{11} + P_{12} = 1$ and $P_{21} + P_{22} = 1$). Again, since β is a valid probability distribution, $\beta_1 + \beta_2 = 1$ and solving we get

$$\beta_1 = \frac{P_{21}}{P_{12} + P_{21}} \text{ and } \beta_2 = \frac{P_{12}}{P_{12} + P_{21}} \tag{3.11.49}$$

Since, $\tau = T - t$ is the trading period, the total time spent in regimes 1 and 2 can now be calculated as $\mathcal{R} = \beta_1 \tau$ and $\tau - \mathcal{R} = \tau - \beta_1 \tau$, respectively. In view of the research by Duan et al. (2002) and Hardy (2001) the distribution of log returns, $X_t = \log(\frac{S_T}{S_t})$, conditional on the total time spent in regime $\xi(t) = j,$ for $j = 1, 2$, k can be developed such that there exist a normal density function with mean μ^* and variance σ^{*2}, that is,

$$X_t | \mathcal{R} \sim N(\mu^*, \sigma^{*2}) \tag{3.11.50}$$

where $\mu^* = \frac{\mathcal{R}}{\tau}\mu_1 + \left(\frac{\tau-\mathcal{R}}{\tau}\right)\mu_2$ and $\sigma*^2 = \frac{\mathcal{R}}{\tau}\sigma_1^2 + \left(\frac{\tau-\mathcal{R}}{\tau}\right)\sigma_2^2$. Since p is the probability function for $\mathcal{R}$,

$$\begin{aligned} F_{X_t} = Pr\big[X_t \leq x\big] &= \sum_{j=1}^{k-1} Pr\big[X_t|\mathcal{R} = \alpha_j\big]p \\ &= \sum_{j=1}^{k-1} \phi_j\big(\frac{x-\mu*}{\sigma*}\big)p \end{aligned} \tag{3.11.51}$$

where $\phi()$ is the standard normal probability distribution function. This implies that the probability density function for X_t is

$$f_{X_t} = \sum_{j=1}^{k-1} \phi_j\big(\frac{x-\mu*}{\sigma*}\big)p \tag{3.11.52}$$

where $\phi()$ is the standard normal density function.

Now, define $C(K,T)$, the European call option (under regime switching world) with strike price K that matures after time T, and is valued at S_t at an initial time t. Since in a regime switching market, the parameter σ^2 switches regimes, we can define a parameter $\sigma*^2$ conditional on knowing $\mathcal{R}$, the total time spent in regime $\xi(t) = j$ for $j = 1, 2$. This implies that, the asset price $S_t|\mathcal{R}$ has a log-normal distribution with parameters that depend on $\mathcal{R}$, that is, the parameters are $\mu*$ and $\sigma*^2$ as defined earlier. Now, to derive a regime switching option pricing model, the Black-Scholes formulae is considered, where the parameter σ^2 is replaced with $\sigma*^2$ to give the desired model as below;

$$C(K,T) = \mathbb{E}^{\mathbb{Q}}\big[max(X_T - K)|\mathcal{R}\big] = S_t\phi(d_1) - e^{-r\tau}K\phi(d_2) \quad \text{where} \tag{3.11.53}$$

$$d_1 = \frac{\ln(\frac{S_t}{K}) + r\tau + \frac{1}{2}\big[\mathcal{R}\sigma_1^2 + (\tau-\mathcal{R})\sigma_2^2\big]}{\sqrt{\mathcal{R}\sigma_1^2 + (\tau-\mathcal{R})\sigma_2^2}}$$

$$d_2 = d_1 - \sqrt{\mathcal{R}\sigma_1^2 + (\tau-\mathcal{R})\sigma_2^2}$$

Proof 3.11.1 *Define $\varphi(S_T) = (S_T - K)$, with $K > 0$. Recall that for any $r, T, S_0 > 0$*

$$\begin{aligned} \mathbb{E}[e^{-r\tau}\varphi(S_T)] &= e^{-r\tau}S_0\Phi(d_1) - e^{-r\tau}K\Phi(d_2) \\ \textit{where } d_1 &= \log\frac{S_t}{K} + (r + \frac{1}{2}\sigma^2),\ d_2 = d_1 - \sigma\sqrt{\tau} \end{aligned} \tag{3.11.54}$$

Let $\mathcal{R}$ be the total time spent in regime $\xi(t) = j$ for $j = 1, 2$ in the interval $[t, T]$ in n trials. Then the distribution of log returns $X_t = \log(\frac{S_T}{S_t})$, conditional on the time spent in regime $\xi(t) = j$ for $j = 1, 2$ is a normal density function with mean μ^* and variance σ_t^{*2}, that is, $X_t|\mathcal{R}\ N(\mu^*, \sigma_t^{*2})$ with $\mu^* = \mathcal{R}\mu_1 + (\tau-\mathcal{R})\mu_2$ and $= \mathcal{R}\sigma_1^2 + (\tau-\mathcal{R})\sigma_2^2$. According to Hardy (2001), conditional on knowing $\mathcal{R}$, the asset price $S_n|\mathcal{R}$ has a log normal distribution with parameters dependent on $\mathcal{R}$. Substituting

μ^* *and* σ_t^{*2} *into Equation* (3.11.54) *completes the proof.*

3.11.5 The RS model parameter estimation

Equation (3.11.40) best describes the actually observed data X_t since the regime $\xi(t)$ is unobserved. If the regime variable $\xi(t)$ is i.i.d across different dates, t, then the log likelihood for the observed data can be calculated from Equation (3.11.40) as

$$L(\theta) = \sum_{t=1}^{T} \log f(X_t; \theta) \tag{3.11.55}$$

The maximum likelihood of θ is obtained by maximizing Equation (3.11.55) subject to the constrains that $\pi_1 + \pi_2 = 1$ and $\pi_j \geq 0$ for $j = 1, 2$. To obtain the maximum likelihood estimates(MLEs) of Equation (3.11.55) we form the Lagrangean

$$J(\theta) = L(\theta) + \lambda(1 - \pi_1 - \pi_2) \tag{3.11.56}$$

and set the derivative w.r.t θ equal to zero. From Equation (3.11.55), the derivative of the log likelihood is given by

$$\frac{\partial L(\theta)}{\partial \theta} = \sum_{t=1}^{T} \frac{1}{f(X_t; \theta)} \frac{\partial f(X_t; \theta)}{\partial \theta} \tag{3.11.57}$$

Note that from Equation (3.11.40)

$$\frac{\partial f(X_t; \theta)}{\partial \pi_j} = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\Big\{ - \frac{(X_t - \mu_j)^2}{2\sigma_j^2} \Big\} = f(X_t|\xi(t) = j; \theta) \tag{3.11.58}$$

Thus

$$\frac{\partial L(\theta)}{\partial \pi_j} = \sum_{t=1}^{T} \frac{1}{f(X_t; \theta)} f(X_t|\xi(t) = j; \theta) \tag{3.11.59}$$

Recalling Equation (3.11.41) the derivative in Equation (3.11.59) can be written as

$$\frac{\partial L(\theta)}{\partial \pi_j} = \pi_j^{-1} \sum_{t=1}^{T} P\big\{\xi(t) = j|X_t; \theta\big\} \tag{3.11.60}$$

From Equation (3.11.59) the derivative of Equation (3.11.56) with respect to (w.r.t) π_j is given by

$$\frac{\partial J(\theta)}{\partial \pi_j} = \pi_j^{-1} \sum_{t=1}^{T} P\big\{\xi(t) = j|X_t; \theta\big\} - \lambda = 0 \tag{3.11.61}$$

This implies that

$$\sum_{t=1}^{T} P\big\{\xi(t)=j|X_t;\theta\big\} = \lambda\pi_j \tag{3.11.62}$$

Summing Equation (3.11.62) over $j = 1, 2$ produces

$$\sum_{t=1}^{T} \Big[P\big\{\xi(t)=j|X_t;\theta\big\} + P\big\{\xi(t)=2|X_t;\theta\big\}\Big] = \lambda(\pi_1+\pi_2) \tag{3.11.63}$$

or $\sum_{t=1}^{T}[1] = \lambda(1) \implies T = \lambda$. Replacing λ with T in Equation (3.11.62) produces

$$\tilde{\pi_j} = T^{-1}\sum_{t=1}^{T} P\big\{\xi(t)=j|X_t;\tilde{\theta}\big\} \text{ for } j = 1, 2 \tag{3.11.64}$$

Next is to find the MLE of μ_j . From Equation (3.11.40) it follows that

$$\frac{\partial f(X_t;\theta)}{\partial \mu_j} = \frac{(X_t-\mu_j)}{\sigma_j^2} p(X_t, \xi(t)=j;\theta) \tag{3.11.65}$$

hence

$$\frac{\partial L(\theta)}{\partial \mu_j} = \sum_{t=1}^{T} \frac{1}{f(X_t;\theta)} \frac{(X_t-\mu_j)}{\sigma_j^2} p(X_t, \xi(t)=j;\theta) \tag{3.11.66}$$

Applying Equation (3.11.41) we have

$$\frac{\partial L(\theta)}{\partial \mu_j} = \sum_{t=1}^{T} \Big\{\frac{(X_t-\mu_j)}{\sigma_j^2}\Big\} P\{\xi(t)=j|X_t;\theta\} \tag{3.11.67}$$

Setting the derivative of Equation (3.11.56) w.r.t μ_j equal to zero implies that

$$\sum_{t=1}^{T} X_t P\{\xi(t)=j|X_t;\theta\} = \mu_j \sum_{t=1}^{T} P\{\xi(t)=j|X_t;\theta\} \tag{3.11.68}$$

Hence

$$\tilde{\mu_j} = \frac{\sum_{t=1}^{T} X_t P\{\xi(t)=j|X_t;\tilde{\theta}\}}{\sum_{t=1}^{T} P\{\xi(t)=j|X_t;\tilde{\theta}\}} \text{ for } j = 1, 2 \tag{3.11.69}$$

The estimate of σ_j^2 follows as below. From Equation (3.11.40) we have

$$\frac{\partial f(X_t;\theta}{\partial \sigma_j^2}) = \Big\{-\frac{1}{2}\sigma_j^{-2} + \frac{(X_t-\mu_j)^2}{2\sigma_j^4}\Big\} p(X_t, \xi(t)=j;\theta) \tag{3.11.70}$$

hence

$$\frac{\partial L(\theta)}{\partial \sigma_j^2} = \sum_{t=1}^{T} \frac{1}{f(X_t;\theta)} \Big\{-\frac{1}{2}\sigma_j^{-2} + \frac{(X_t-\mu_j)^2}{2\sigma_j^4}\Big\} P(\xi(t)=j|X_t;\theta) \tag{3.11.71}$$

Equating Equation (3.11.71) to zero leads us to finding the MLE of σ_j^2 ,that is,

$$\sum_{t=1}^{T}\left\{-\sigma_j^2+(X_t-\mu_j)^2\right\}P\{\xi(t)=j|X_t;\theta\}=0 \tag{3.11.72}$$

This implies that

$$\tilde{\sigma_j^2}=\frac{\sum_{t=1}^{T}(X_t-\mu_j)^2P\{\xi(t)=j|X_t;\tilde{\theta}\}}{\sum_{t=1}^{T}P\{\xi(t)=j|X_t;\tilde{\theta}\}}\ \text{for } j=1,2 \tag{3.11.73}$$

If we restrict the transition probability only by the conditions $P_{ij}>0$ and $(P_{i1}+P_{i2})=1$ for all i and j, then Hamilton (1990) showed that the MLEs for the transition probability is given by

$$\tilde{P_{ij}}=\frac{\sum_{t=2}^{T}P\{\xi(t)=j,\xi(t-1)=j|X_t;\tilde{\theta}\}}{\sum_{t=2}^{T}P\{\xi(t-1)=i|X_t;\tilde{\theta}\}}\ \text{for } j=1,2 \tag{3.11.74}$$

This implies that the estimated P_{ij} is essentially the number of times state i seems to have been followed by state j divided by the number of times the process was in state i.

3.11.6 Regime switching-GARCH (RS-GARCH) model

Consider a discrete-time economy whose stock price at time t is S_t, the one period asset returns under the physical measure $\mathbb{P}$ is defined as

$$\begin{aligned} X_t &= \ln S_t-\ln S_{t-1} \\ &= \mu_t+r_t,\ \ r_t=\sigma_t\varepsilon_t \end{aligned} \tag{3.11.75}$$

where $\varepsilon_t\sim\ N(0,1)$. The general GARCH model is defined as

$$r_t=\sigma_t\varepsilon_t,\ \ \sigma_t^2=\omega+\sum_{i=1}^{p}\alpha_i r_{t-i}^2+\sum_{j=1}^{q}\beta_j\sigma_{t-j}^2 \tag{3.11.76}$$

where p and q are the ARCH and GARCH process degrees, respectively, ε_t is an independent and identically distributed sequence of random variables whose mean and variance are zero and unit, respectively. It is required that $\omega>0,\ \alpha_i\geq 0,\ \beta_j\geq 0$ and $\sum_{i=1}^{p}\alpha_i+\sum_{j=1}^{q}\beta_j<1$. It is argued by Bauwens et al. (2006) that estimation of this model using daily or higher frequency data implies that the volatility persistence is very high and the model may not be covariance-stationary. This persistence may be as a result of regime changes in the GARCH parameters over time, see Mikosch and Stărică (2004), etc. To capture the regime shifts, a Regime-Switching GARCH model is considered since it allows the parameters to shift regime. Define an unobserved state variable at time t as $S_t\in\{1,2,..,k\}$

which selects the model parameters with probability $P_{ij} = P[\alpha(t) = j|\mathcal{F}_{t-1}]$ where $\mathcal{F}_{t-1}$ is the available information at time t. The RS-GARCH model can thus be defined as

$$\begin{aligned} X_t &= \mu_{\alpha(t)} + r_t, \text{ where } r_t = \sigma_{t,\alpha(t)}\varepsilon_t \text{ and } \varepsilon_t \sim N(0,1) \\ \sigma_{t,\alpha(t)} &= \omega_{\alpha(t)} + \sum_{i=1}^{p} \alpha_{i,\alpha(t)} r_{t-i}^2 + \sum_{j=1}^{q} \beta_k \sigma_{t-j}^2 \text{ for } \alpha(t) = \{1,2,...,k\} \end{aligned} \tag{3.11.77}$$

Let $\varphi(S_T)$ be the payoff of a European call option with exercise price K maturing at a future time T, where $\varphi(S_T)$ is a random variable on probability measure space $(\Omega, \mathcal{F}, \mathbb{P})$ describing the market. Then, under the RS-GARCH(p,q) specification, the today's price of the call option C_t under measure $\mathbb{Q}$ is given by

$$\begin{aligned} C_t &= e^{-(T-t)r}\mathbb{E}^{\mathbb{Q}}\Big[max(S_T - K, 0)|\mathcal{F}_t\Big] \\ &= e^{-(T-t)r}\mathbb{E}^{\mathbb{Q}}\Big[(S_T - K)^+|\mathcal{F}_t\Big] \end{aligned} \tag{3.11.78}$$

For simplicity we restrict the model to RS-GARCH(1,1).

3.11.7 RS-GARCH model parameter estimation

The parameter estimation of the RS-GARCH model is done via the maximum likelihood estimation technique reported by Ardia et al. (2019). Let $\theta = (\alpha_{0,k}, \alpha_{1,k}, \beta_k)'$ be the vector that represent the model parameters, then the likelihood function is given by

$$L(\theta|\mathcal{F}_T) = \prod_{t=1}^{T} f(r_t|\theta, \mathcal{F}_{t-1}) \tag{3.11.79}$$

where $f(r_t|\theta, \mathcal{F}_{t-1})$ is the probability density function of r_t given the past observations, $\mathcal{F}_{t-1}$ and the parameters of the model θ. Therefore, the conditional density of r_t, for RS-GARCH model is given by

$$f(r_t|\theta, \mathcal{F}_{t-1}) = \sum_{i=1}^{K}\sum_{j=1}^{K} P_{ij} z_{i,t-1} f_D(r_t|\alpha(t) = j, \theta, \mathcal{F}_{t-1}) \tag{3.11.80}$$

where $z_{i,t-1} = P[\alpha(t-1) = j|\theta, \mathcal{F}_{t-1}]$ is the filtered probability of state i at time $t-1$ obtained via Hamilton filter, see Hamilton and Susmel (1994). In Equation (3.11.80), the density of r_t in state $\alpha(t) = k$ given θ and $\mathcal{F}_{t-1}$ is denoted by $f_D(r_t|\alpha(t) = k, \theta, \mathcal{F}_{t-1})$. The maximum likelihood estimator $\tilde{\theta}$ is thus obtained by maximizing Equation (3.11.79).

3.12 Statistical tests

The statistical tests utilized in this studies are presented in this section.

3.12.1 Augmented Dickey-Fuller (ADF) test

The unit root test is critical in time series analysis for selecting techniques and procedures for further analysis. Presence of unit root implies that the time series is not stationary and it should be noted that a time series with unit root suffers spurious results in regression analysis. The general ADF unit root test is based on the following regression:

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \delta_1 \Delta Y_{t-1} + ... + \delta_p \Delta Y_{t-p} + \varepsilon_t \tag{3.12.1}$$

where ΔY_t is a time series with trend decomposition, t is the time trend, α is a constant, β is the coefficient on a time trend and p is the lag order of the autoregressive process. The number of augmenting lags(p) is determined by the Akaike Information Criterion (AIC). The test for the unit root is based on the null hypothesis $\gamma = 0$ versus the alternative hypothesis $\gamma < 0$. The computed test statistic

$$DF_r = \frac{\tilde{\gamma}}{SE(\tilde{\gamma})} \tag{3.12.2}$$

is then compared with the relevant critical value for the Dickey-Fuller test. Since this test is asymmetrical, the only concern is with negative values of the test statistic, DF_r. The null hypothesis is rejected if the test statistic is less(more negative) than the critical value and this implies no presence of unit root.

3.12.2 Jarque-Bera(JB) Test

The Jarque-Bera test determines whether the sample skewness and kurtosis are consistent with that of the normal distribution. Considering a sample of size n, the JB test is defined as

$$JB = \frac{n}{6}\left[S^2 + \frac{1}{4}(K-3)^2\right] \tag{3.12.3}$$

where S and K are the skewness and kurtosis of the sample, respectively and defined as follows;

$$S = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \tilde{x}\right)^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \tilde{x})^2\right)^{\frac{3}{2}}} \quad \text{and} \quad K = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \tilde{x}\right)^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \tilde{x})^2\right)^{2}}$$

The null hypothesis is that the excess kurtosis and skewness are both zero. A Sample drawn from a normal distribution is expected to have skewness and excess kurtosis of zero (the same as kurtosis of 3).

3.12.3 Ljung Box

This is a statistical test that determines whether the time series autocorrelations is different from zero. It is a portmanteau test since it examines the overall randomness based on a series of lags rather than assessing randomness at each individual lag. This test is defined as follows;

H_0 : The residuals are independently distributed (i.e the population correlations is zero)

H_a : The residuals are not independently distributed. The test statistic is:

$$Q(h) = n(n+2)\sum_{j=1}^{h}\frac{\tilde{\rho}_j^2}{(n-j)} \tag{3.12.4}$$

where n and $\tilde{\rho}_j$ are the sample size and sample autocorrelation at lag j, respectively, whereas m is the maximum number of lags being tested. The critical region is $Q > x^2_{1-\alpha,h}$ for α -level of significance, where $Q > x^2_{1-\alpha,h}$ is the $(1-\alpha)$-quantile of the chi-squared distribution with h degrees of freedom. If the p-value is less than the defined significance level, the null hypothesis is rejected . This shows that under the null hypothesis, the observed result is highly unlikely.

3.12.4 Lagrange Multiplier (LM) test

To derive the LM test, Equation (3.4.1) is rewritten as an AR(p) process for r_t^2 as below;

$$r_t^2 = \omega + \alpha_1 r_{t-1}^2 + ... + \alpha_{t-p}^2 + \mathcal{U}_t, \tag{3.12.5}$$

where $\mathcal{U}_t = r_t^2 - \sigma_t^2$ is a Martingale difference sequence (MDS) since $\mathbb{E}_{t-1}[\mathcal{U}_t] = 0$ and it is assumed that $\mathbb{E}[r_t^2] < \infty$. For the squared residuals r_t^2 , an ARCH model implies an AR model, therefore, according to Engle (1982), Equation(3.12.5) is used to construct a Lagrange Multiplier (LM) test for ARCH effects. The null hypothesis, H_0 of no ARCH effects (i.e $\alpha_1 = \alpha_2 = ... = \alpha_p = 0$) versus the alternative, H_a is tested based on the test statistic

$$LM = T.R^2 \tag{3.12.6}$$

The distribution of this test statistic is a Chi-Square with p degrees of freedom, where T is the sample size and R^2 is determined from Equation (3.12.5) using estimated residuals. To make a decision, reject H_0 if the test statistic LM is not significant.

3.12.5 Root Mean Square Error (RMSE)

A Root Mean Square Error (RMSE) is computed in order to compare the two models in terms of prediction of the option prices. It is computed by utilizing the model predicted option prices and the observed option prices. That is, the RMSE of a prediction model with respect to the observed option prices is defined as the square root of the mean squared error.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (P_i - O_i)^2}{n}} \tag{3.12.7}$$

where P_i is the predicted option prices, O_i the observed option prices and n is the number of observations. Note that, a small value of the RMSE implies a model with a better prediction.

CHAPTER 4: RESULTS AND DISCUSSION

4.1 Empirical data

The data set utilized in this study is in twofold. The first set of data is time series data consisting of S&P 500, FTSE100 and NSE20 indices that was randomly chosen from the New York stock exchange, London stock exchange and the Nairobi Securities Exchange respectively. The data for FTSE100 and S&P500 indices, which represents developed or established financial stock markets, was retrieved from Yahoo Finance website while that for NSE20 index, which represent a developing or emerging stock market, was obtained from the Nairobi Securities (NSE) Exchange at a fee. The financial time series consists of the daily and weekly stock prices and their respective trading volume for the three indices for the period 01/02/2001 to 31/12/2017. The second set of data is the daily closing stock prices as reported in the Russell 2000(RUT), Facebook(FB) and Google(GooG) indices for the period 01/02/2013 to 21/01/2022. The corresponding call options prices for the three markets is categorized as short-dated and long-dated, that is, call options prices expiring in 25 and 258 days respectively. The continuously compounded index returns are computed as in Equation (3.3.1) whereas the log trading volume (log volume), T_v , is computed in terms of logarithmic change as $T_v = \ln V_t - \ln V_{t-1}$ where V_t and V_{t-1} is trade volume at day t and $t-1$, respectively.

4.2 Descriptive statistics

The descriptive statistics of the daily FTSE100, S&P500 and NSE20 closing stock indices prices are presented in Table 4.1. The mean of the closing stock indices prices are positive and the S&P500 and NSE20 stock indices are right skewed while FTSE100 stock price index is left skewed implying that the stock indices have an asymmetric distribution. The skewness and kurtosis are significantly not the same as those of a normal distribution of 0 and 3, respectively and this is a suggestion that their distribution is leptokurtic, i.e, has fat tails. On the other hand the descriptive statistics of the second set of data , that is, RUT, FB and GooG are presented in Table 4.2. Their mean and skewness are positive implying a distribution that is asymmetric and skewed to the right

Table 4.1: Basic statistics for daily FTSE100, S&P500 and NSE20 stock prices

Index	Obs	Mean	Variance	Skewness	Kurtosis
FTSE100	4349	5679.672	904153.9	−0.210212	−0.719395
S&P500	4276	1436.096	195139.4	0.874500	−0.198456
NSE20	4272	3645.038	1470627	0.4760600	−0.593462

Table 4.2: Basic statistics for daily RUT, FB and GooG stock prices

Index	Obs	Mean	Variance	Skewness	Kurtosis
RUT	2281	1433.826	132556.4	1.003863	0.3636780
FB	2281	153.926339	7530.357353	0.641251	−0.250147
GooG	2281	1073.035	399962.8	1.467386	1.611128

The basic descriptive statistics for the FTSE100, NSE20 and S&P500 indices returns and log volume are presented in Table 4.3 which clearly indicate that the mean of the indices returns are all positive and close to zero. On the other hand, the mean of the log volume are positive except for the daily and weekly FTSE100 and the weekly S&P500 indices log volumes. The positive mean of index returns may be an indication that investors in these markets have realized a positive return rate on their investment. Further, the mean and variance of all the indices returns and log volume series are observed to increase as the data changes frequency from daily to weekly. The daily and weekly indices returns of the developed stock markets (FTSE100 and S&P500 indices) are all negatively skewed whereas the indices returns from the developing stock market (NSE20 index) has a positive skeweness. The log volume series are positively skewed except that for the daily FTSE100 and S&P500, and the weekly NSE20 indices.

Table 4.3: Descriptive statistics for index returns and log volume

Index	Variable	Level	Mean	Median	Variance	Std dev.	Skew	Ex.Kurt
FTSE100	X_t	Daily	0.00005	0.00026	0.00014	0.01185	-0.15894	3.75705
		Weekly	0.00024	0.00194	0.00059	0.02430	-1.11068	9.5557
	T_v	Daily	-0.00021	-0.00029	0.12324	0.35106	-0.01070	10.41753
		Weekly	-0.00135	-0.00249	0.18586	0.43111	0.00661	6.75700
S&P500	X_t	Daily	0.00017	0.00055	0.00014	0.01202	-0.21989	6.38927
		Weekly	0.00084	0.00203	0.00057	0.02379	-0.92811	5.09105
	T_v	Daily	0.00018	-0.00108	0.03575	0.18908	-0.09555	6.01709
		Weekly	-0.00108	-0.00176	0.06303	0.25109	0.09257	4.79695
NSE20	X_t	Daily	0.00016	0.00001	0.00007	0.00857	0.39295	8.32209
		Weekly	0.00076	0.00044	0.00066	0.02569	0.41755	2.87156
	T_v	Daily	0.00103	-0.00496	0.61150	0.78199	0.09298	-0.77629
		Weekly	0.00392	0.00584	0.77491	0.88029	-0.00651	-1.39927

The negative skewness implies that the distribution of the return series is left skewed whereas the positive skewness implies that the distribution is right skewed. Thus it can be concluded that indices returns from developed stock markets have a distribution with long tail to the left while the distribution of indices returns from developing stock market has a long tail to the right. Moreover, the all the time series exhibit a positive kurtosis except the daily and weekly NSE20 log volumes, which can be construed to imply that the distribution of the time series is leptokurtic. A less than or greater than three kurtosis implies flatness and peakedness, respectively, in the time series data. It is thus evident that the NSE20 weekly index returns, S&P500 daily log volume and NSE20 daily and weekly log volume have a flat distribution since they have a negative excess kurtosis whereas all the other indices returns and log volume have a peaked distribution.

The time series data is further described by plotting the empirical density versus the normal distribution as well as the quantile-quantile(qq) plots for the daily and weekly FTSE100, S&P500 and NSE20 indices returns and log volume as shown in Figure 1 to 4. The plots reveal that the empirical densities are different from that of normal distribution and hence the data does not follow normal distribution. This is confirmed by the values of skewness and excess kurtosis from Table 4.3 which are significantly not equal to those of normal distribution of zero and three. The Jarque-Bera(JB) statistic in Table 4.4 is in support of the non-normality of the indices returns and log volume.

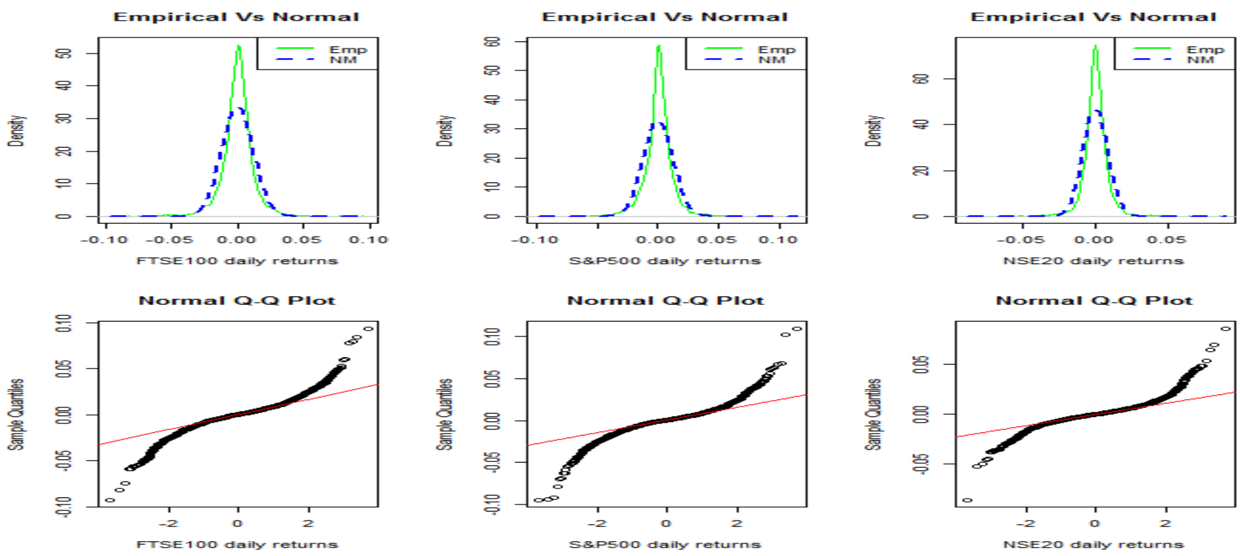

Figure 1: Empirical density versus normal distribution, and qq-plots for daily returns

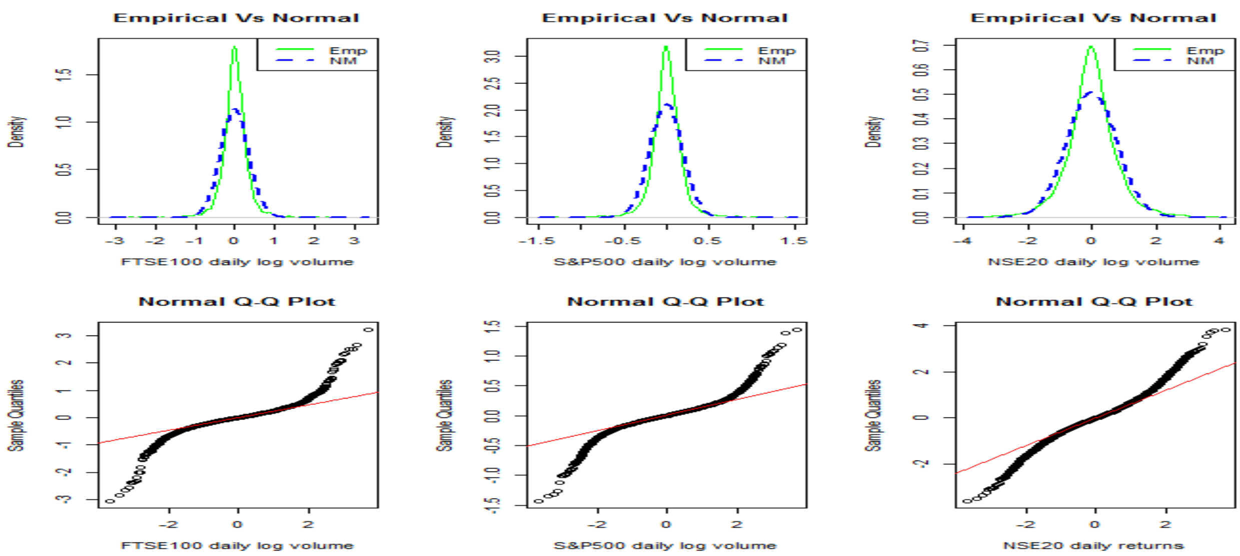

Figure 2: Empirical density versus normal distribution, and qq-plots for daily log volume

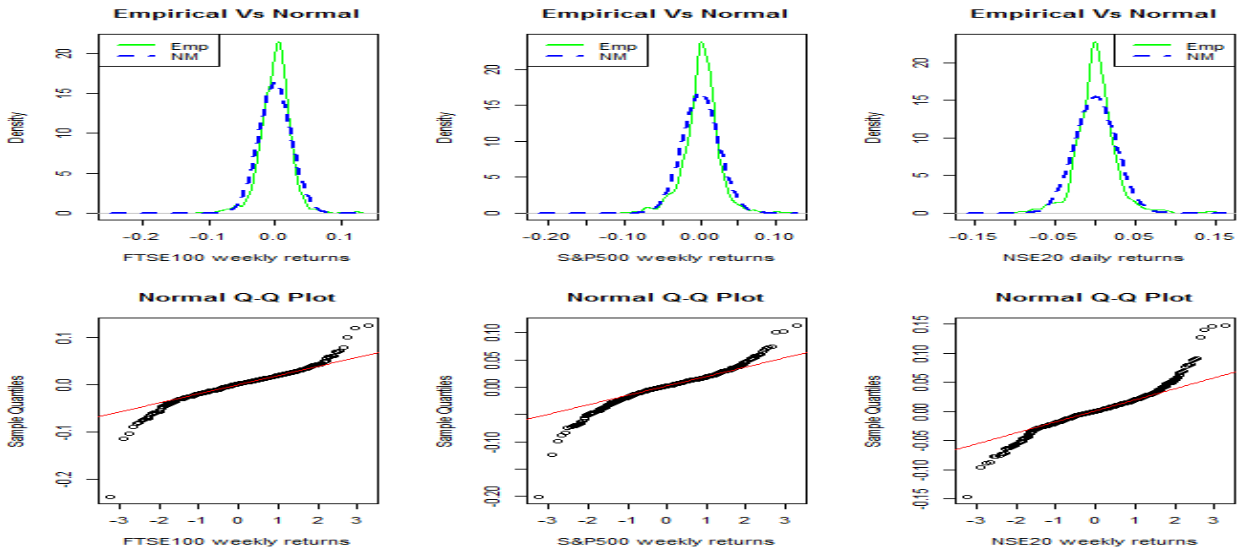

Figure 3: Empirical density versus normal distribution, and qq-plots for weekly returns

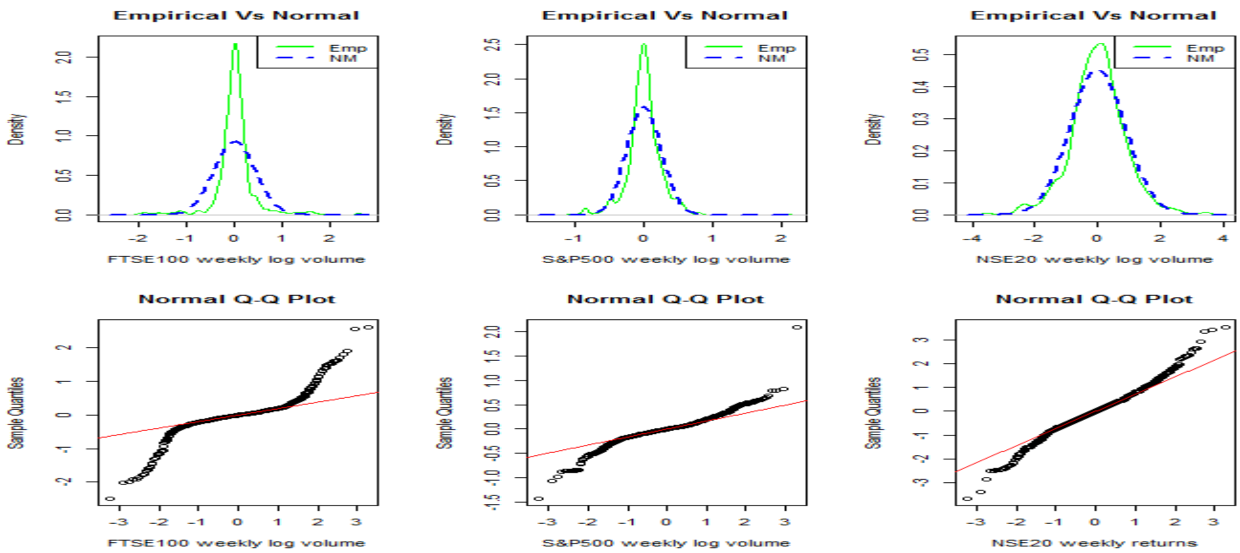

Figure 4: Empirical density versus normal distribution, and qq-plots for weekly log volume

4.3 Empirical findings and discussion

The daily and weekly time series plots for the stock indices, trading volume, stock indices returns and log volume are presented in Figure 5 to 8. It is evident from the stock price and trading volume plots that the series have a trend implying a non-constant mean and variance of the series and therefore it can be concluded that the data is not stationary. A clear discernible pattern of behavior can be inferred from the plots of indices returns in Figure 5 and 7. These plots reports evidence of the common properties of time series data, for instance, volatility clustering(that is, low volatility is preceded by low volatility and high volatility is preceded by high volatility), leverage effects, leptokurtic distribution, heavy tails and existence of outliers. In order to be able to analyze the data for desirable results, first-order difference is performed on the log of index returns to make the data series stationary, however, this is not enough for one to conclude that there is volatility clustering and therefore the data series are subjected to Ljung-Box and Lagrange Multiplier tests to inspect for autocorrelations and ARCH effects. The results of these tests and the ADF test for stationarity of indices returns and log volume are reported in Table 4.4. The null hypothesis that indices returns and log volume are not stationary is rejected at the 1% significance level, an indication that both indices returns and log volume are stationary. The Ljung-Box test results reports that all the data series have no autocorrelations and the Lagrange Multiplier test results indicate that all the series have ARCH effects since they are all statistically significant.

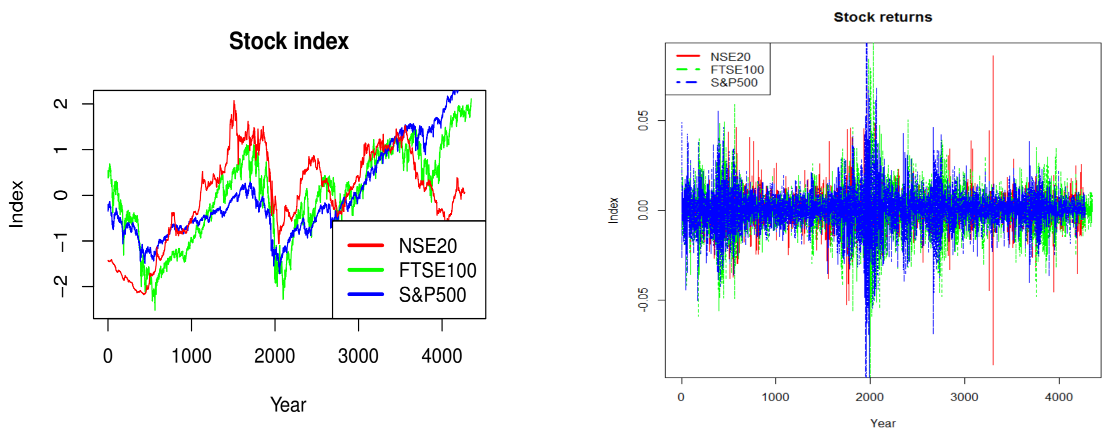

Figure 5: The daily stock indices and stock returns

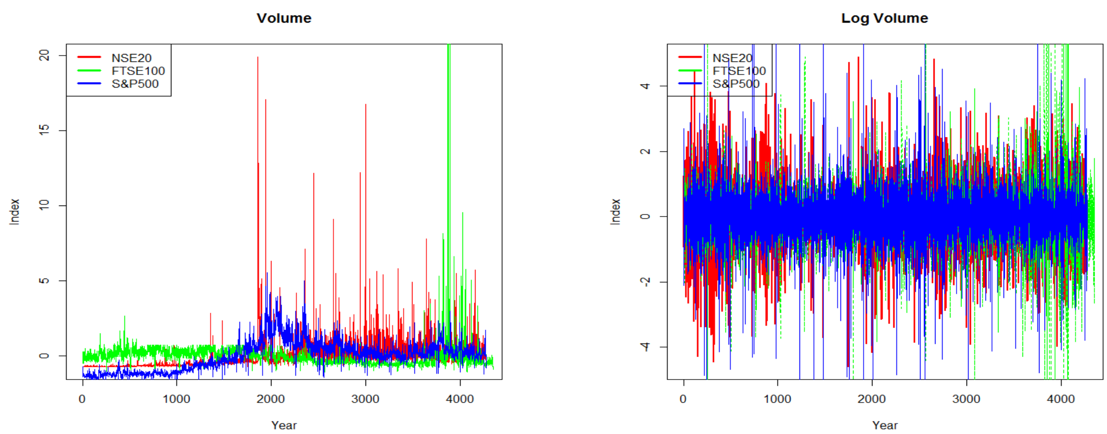

Figure 6: The daily trading volume and log volume

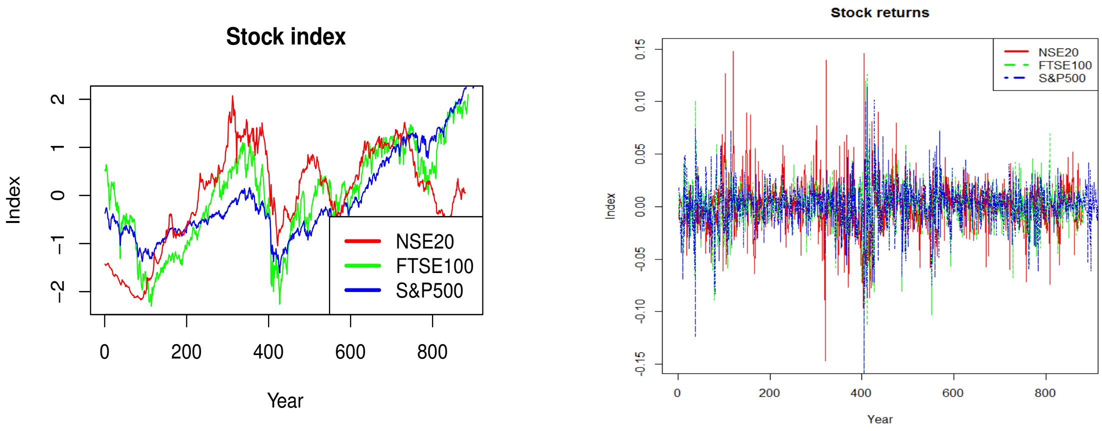

Figure 7: The weekly stock indices and stock returns

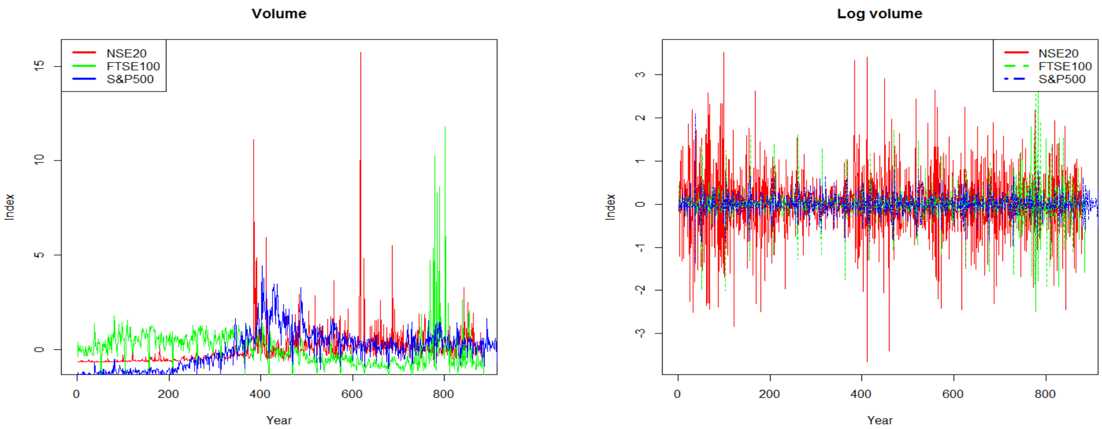

Figure 8: The weekly trading volume and log volume

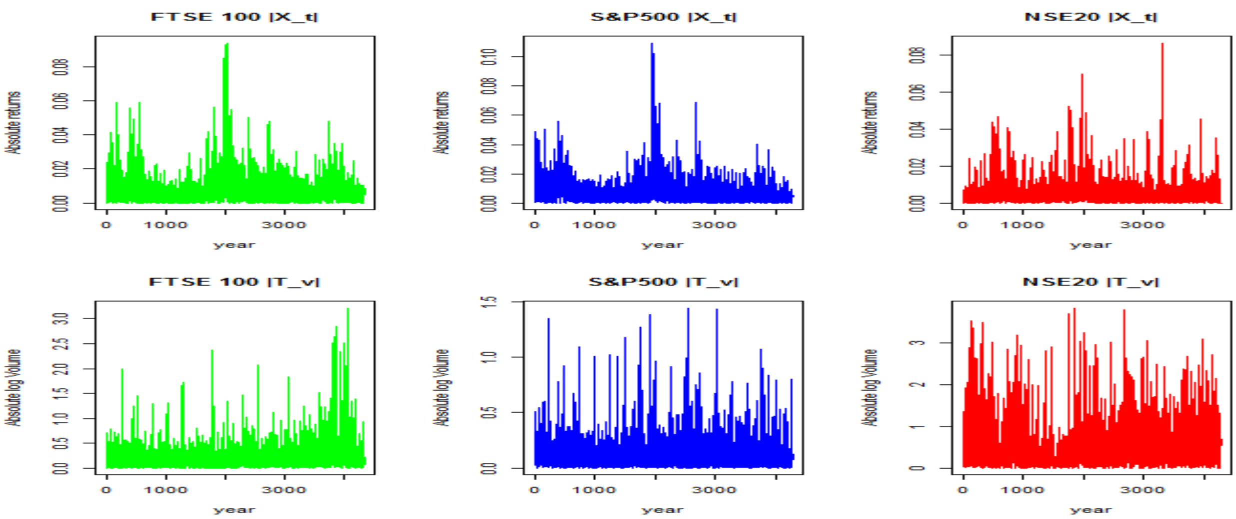

Figure 9: Daily absolute returns and log volume

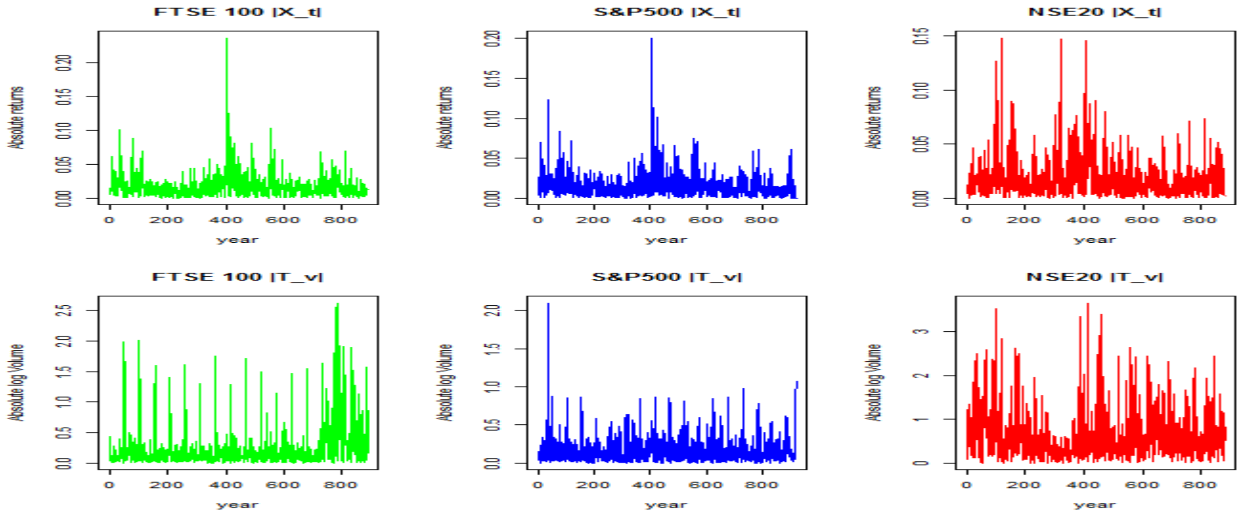

Figure 10: Weekly absolute returns and log volume

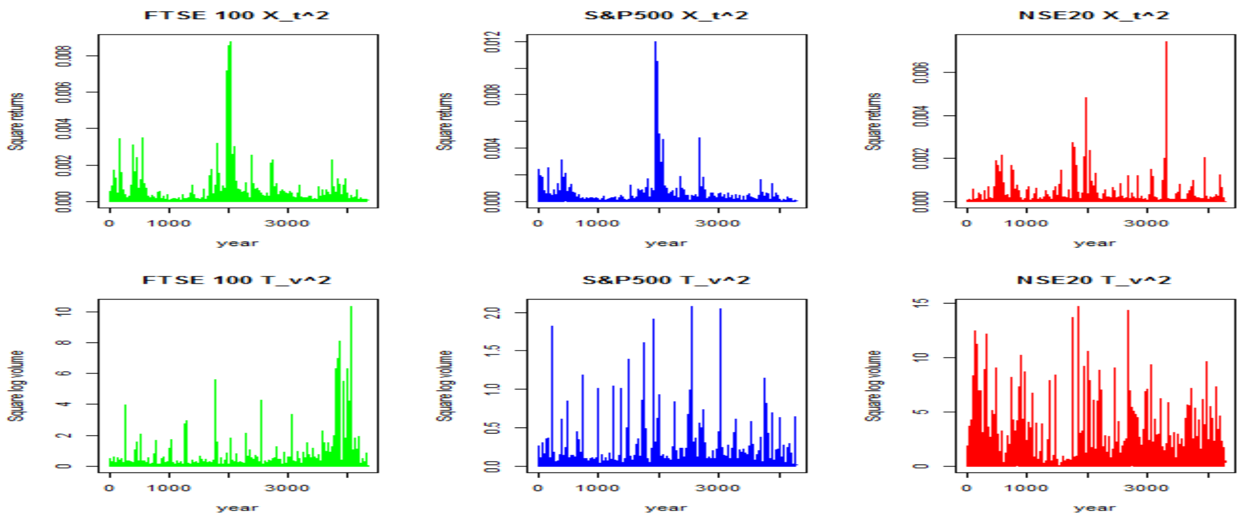

Figure 11: Daily square returns and log volume

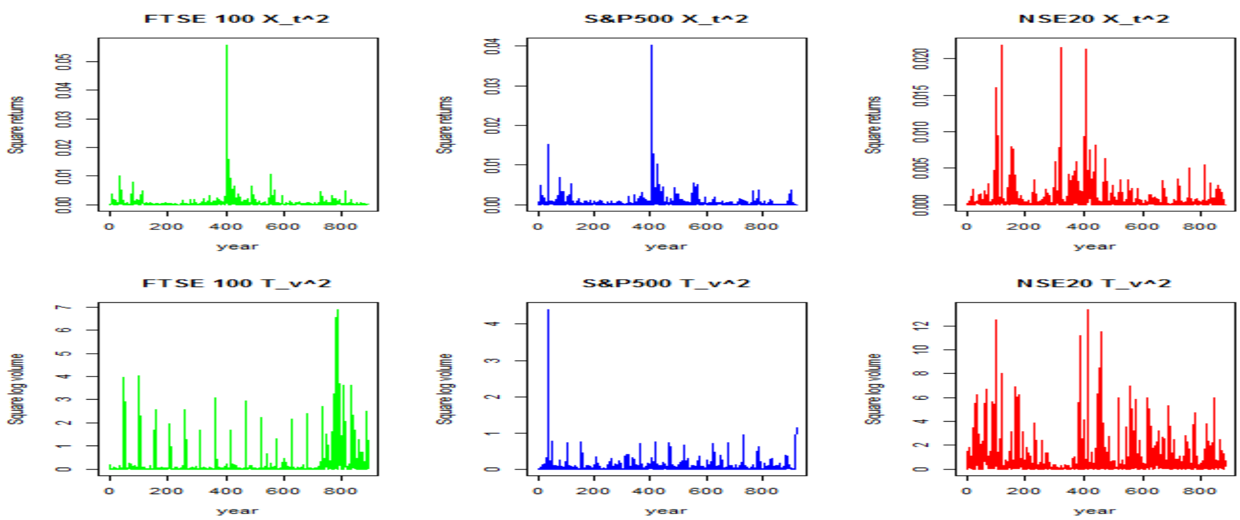

Figure 12: Weekly square returns and log volume

Table 4.4: Statistical tests for the indices returns and log volume

Index		Frequency	Ljung Test statistic	LM Test statistic	JB Test statistic	ADF Test statistic
FTSE100	X_t	Daily	73.83***	1,029***	83,001***	-29.99*
		Weekly	38.45***	113***	6,035***	-10.15*
	T_v	Daily	585.6***	846.5***	32652***	-29.30*
		Weekly	192.5***	178.8***	3535***	-14.31*
S&P500	X_t	Daily	61.44***	1,198***	15,757***	-20.44***
		Weekly	23.14*	147.2***	2,651***	-9.448*
	T_v	Daily	522.5***	456.7***	14508***	-28.06*
		Weekly	131.2***	95.04***	2341***	-13.54***
NSE20	X_t	Daily	814.1***	1,027***	22,949***	-17.84*
		Weekly	41.00***	86.75***	1,297***	-7.178*
	T_v	Daily	814.1***	476.0***	888.1***	-30.62*
		Weekly	41.00***	86.24***	95.09***	-13.62*

Note: The asterisks *, ** & *** implies that the statistics are significant at 10%,5% and 1% level of significance, respectively.

Further preliminary investigation of the data set is done by computing the squared returns for all the series and a test for heteroscedasticity and volatility clustering carried out. The unconditional variance can be approximated by the squared return of a particular time series if an assumption of zero mean is made. The sample graphics of the squared and absolute returns are presented in Figures 9 to 12 which further reveal volatility clustering. This particular stylized fact of financial asset returns is more present in the daily time series than it is in weekly time series. An investigation of heteroscedasticity is carried out for all the time series by calculating the autocorrelation (AC) whose graphics are presented in Figures 13 to 18.

Furthermore, the sample autocorrelation function(ACF) is utilized to describe linear dynamics of the data. It's worth noting that the sample ACF of a return series is crucial in linear time series analysis; in fact, the ACF defines a linear time series model. The graphics clearly show that the sample ACFs are relatively close to one another, implying that the serial correlation is very low.

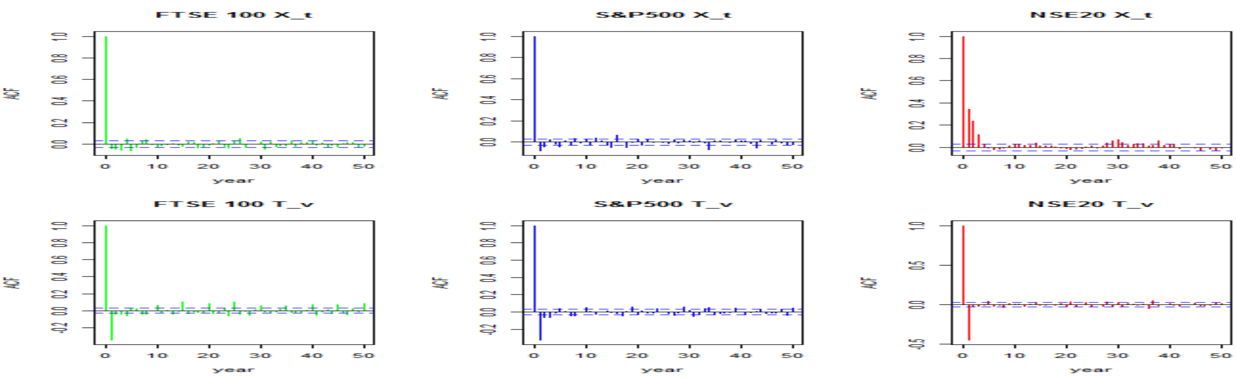

Figure 13: ACF for the daily returns and log volume

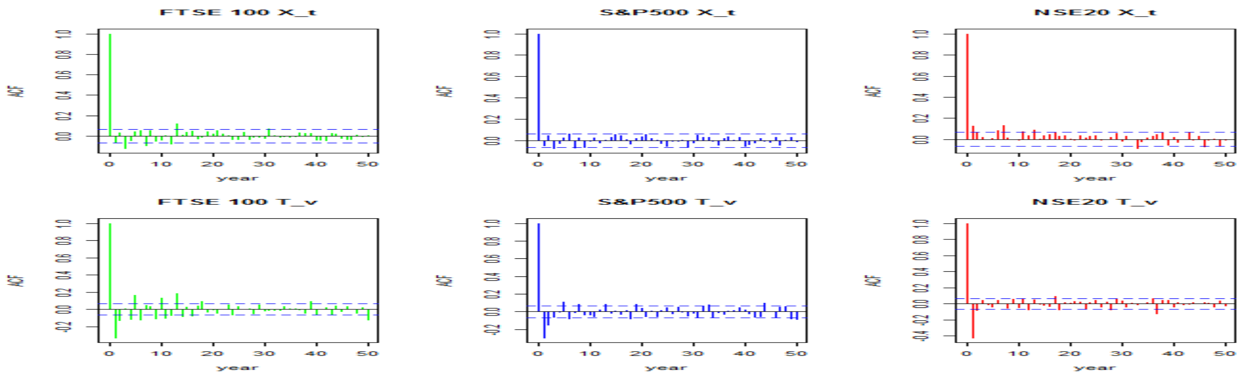

Figure 14: ACF for the weekly returns and log volume

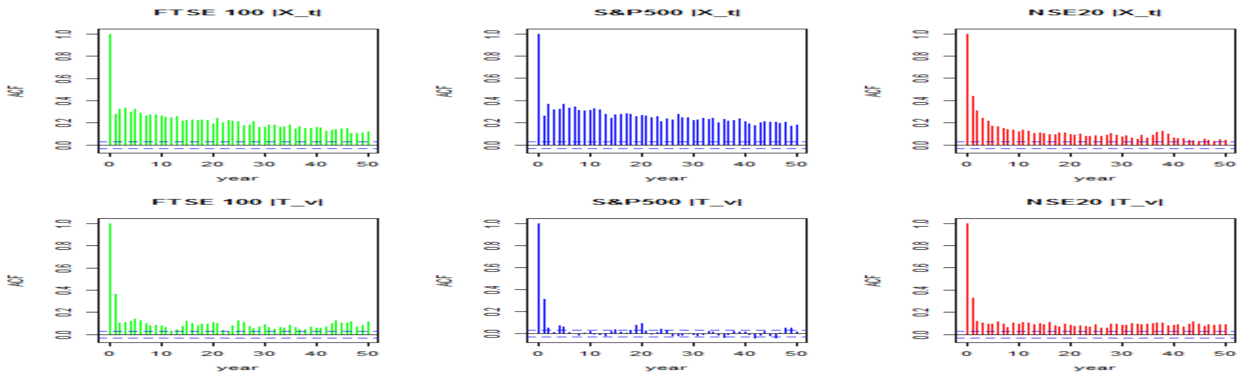

Figure 15: ACF for the daily absolute returns and log volume

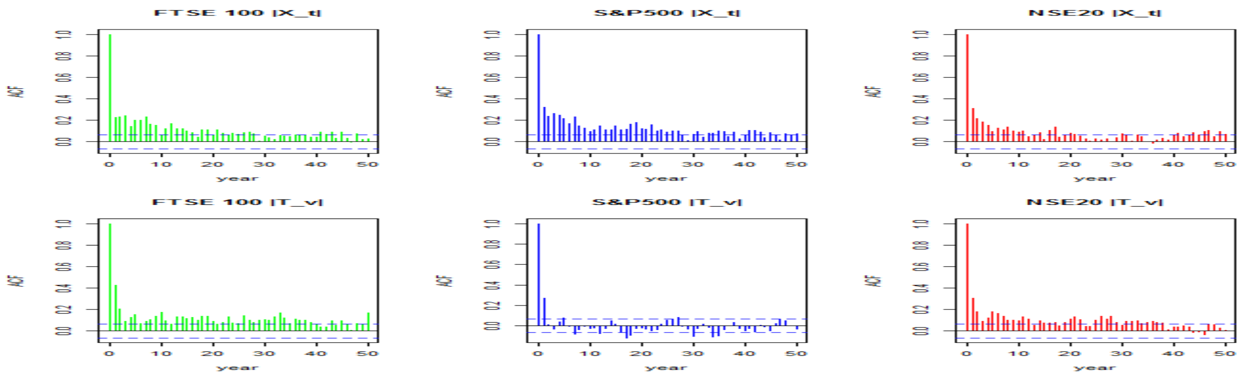

Figure 16: ACF for the weekly absolute returns and log volume

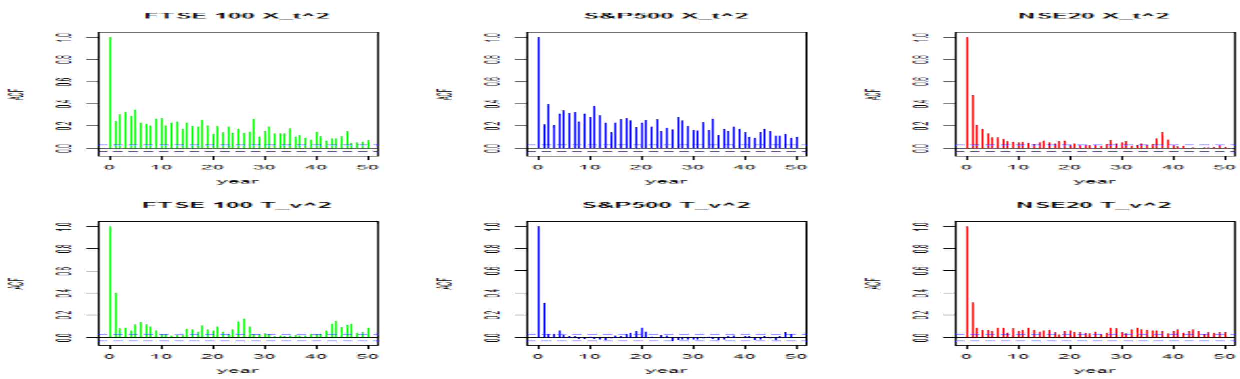

Figure 17: ACF for the daily square returns and log volume

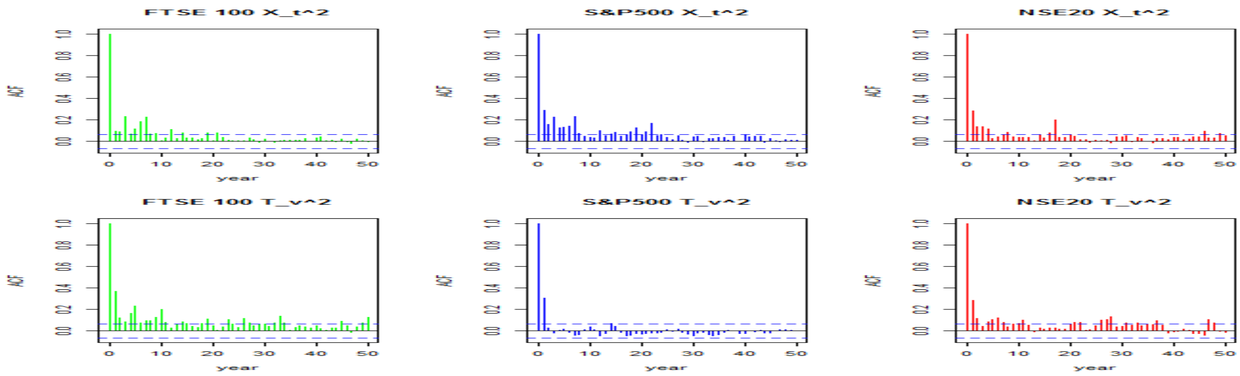

Figure 18: ACF for the weekly square returns and log volume

The GARCH(1,1) model parameters, as depicted by the equation

$$r_t = \sigma_t \varepsilon_t, \;\; \sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

, are estimated using the daily and weekly FTSE100, S&P500 and NSE20 indices returns and the results presented in Table 4.5 and 4.6. The mean μ_1 of the process is close to zero in all the daily and weekly indices returns. All the model coefficients, ω, α_1, and β_1, are significant at 1% confidence level for the all the daily and weekly indices returns. The non-negativity conditions of GARCH(1,1) model, that is, $\omega > 0, \; \alpha_1 \geq 0, \beta_1 \geq 0$ and that $\alpha_1 + \beta_1 < 1$ is respected by the parameter estimates. This condition is important for mean reverting (volatility persistence) process and the results reveals that the model is weakly stationary and the conditional volatility is mean reverting for all indices returns. It is noted that the weekly conditional volatility of indices returns tend to revert quickly towards the mean when compared to the daily conditional volatility. The parameter β_1 is a measure of volatility persistence and a high value implies that shocks to conditional variance take long to vanish or die off. Large values of β_1 are reported in both daily and weekly indices returns of the FTSE100 and S&P500 stock markets than in the NSE20 stock market daily and weekly indices returns. The implication is that the indices returns are characterized by volatility clustering, that is, small values of σ_{t-1}^2 are succeeded by small values of σ_t^2 and in a similar manner large values of σ_{t-1}^2 are succeeded by large values of σ_t^2 . It is therefore clear that the developed markets are characterized by high clustering of volatility compared to low clustering of volatility in the emerging market. This means that it takes a long time for the shocks to conditional variance to disappear in FTSE100 and S&P500 indices returns than in NSE20 index returns. This claim is further confirmed by the value of $\alpha_1 + \beta_1$ which is close to one and is high in developed markets than in emerging market for both daily and weekly indices returns. The value of $\alpha_1 + \beta_1$ measures persistence of volatility and if it nears unity, it implies high volatility persistence. It can be inferred that the daily and weekly NSE20 index returns reports relatively low values of $\alpha_1 + \beta_1$ compared to the daily and weekly FTSE100 and S&P500 indices returns. Moreover, the volatility persistence decreases as the data changes frequency from daily to weekly indices returns in all the stock indices. The coefficient α_1 is a measure of the extent to which the present time volatility shock feeds through into the volatility occurring in the next period. This value is large in the emerging market daily and weekly index returns than in the corresponding daily and weekly indices returns of the developed market stock indices. This means that, in comparison to developed stock market returns' volatility, the volatility of emerging stock market returns is influenced more by previous volatility than by comparable news from the previous period.

Table 4.5: The parameter estimates of GARCH(1,1) model for daily indices returns

	FTSE100			S&P500			NSE20		
Estimate	Norm	Std-t	Ged	Norm	Std-t	Ged	Norm	Std-t	Ged
μ_1	3.811×10^{-4}	4.453×10^{-4}	2.936×10^{-4}	5.254×10^{-4}	6.592×10^{-4}	4.281×10^{-4}	1.488×10^{-4}	4.780×10^{-5}	3.4×10^{-5}
	(0.0015)	(0.0001)	(0.0136)	(0.0001)	(0.0000)	(0.0002)	(0.1360)	(0.6090)	(0.7115)
ω	1.661×10^{-6}	1.475×10^{-6}	1.533×10^{-6}	1.627×10^{-6}	8.682×10^{-7}	1.118×10^{-6}	4.607×10^{-6}	9.358×10^{-6}	7.0×10^{-6}
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.00000)
α_1	0.1102	0.1101	0.1084	0.09749	0.09955	0.09766	0.2675	0.3642	0.3157
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.00000)
β_1	0.8778	0.8810	0.8800	0.8883	0.8989	0.8944	0.6920	0.5232	0.5986
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.00000)
$\alpha_1 + \beta_1$	0.988	0.9911	0.9884	0.98579	0.99845	0.9921	0.9595	0.8874	0.9143
Log L	14077	14120	14137	13928	14015	14039	15066	15284	15256.07
AIC	−6.4735	-6.4926	-6.5000	-6.5142	-6.5546	-6.5651	-7.0532	-7.1547	-7.1417
BIC	−6.4676	-6.4853	-6.4912	-6.5083	-6.5471	-6.5562	-7.0473	-7.1473	-7.1342

Note: The normal, students-t, and generalized error distributions are represented by Norm, Std-t, and Ged, respectively, and the p-values are placed in brackets. The log likelihood, Akaike Information Criterion, and Bayesian Information Criterion are abbreviated as Log L, AIC, and BIC respectively.

Table 4.6: The parameter estimates of GARCH(1,1) model for weekly indices returns

	FTSE100			S&P500			NSE20		
Estimate	Norm	Std-t	Ged	Norm	Std-t	Ged	Norm	Std-t	Ged
μ_1	1.47×10^{-3}	1.551×10^{-3}	9.894×10^{-4}	2.371×10^{-3}	2.55×10^{-3}	1.955×10^{-3}	4.046×10^{-4}	8.923×10^{-4}	5.0×10^{-4}
	(0.0163)	(0.0091)	(0.1081)	(0.0000)	(0.0000)	(0.0005)	(0.5950)	(0.1692)	(0.0831)
ω	2.911×10^{-5}	2.002×10^{-5}	2.120×10^{-5}	2.781×10^{-5}	1.871×10^{-5}	2.028×10^{-5}	1.3×10^{-4}	1.437×10^{-4}	1.0×10^{-4}
	(0.0050)	(0.0443)	(0.0404)	(0.0006)	(0.0254)	(0.0084)	(0.0000)	(0.0001)	(0.0002)
α_1	0.1907	0.1140	0.1342	0.224	0.1572	0.1671	0.3435	0.4196	0.3765
	(0.0000)	(0.0023)	(0.0006)	(0.0000)	(0.0003)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
β_1	0.7738	0.8482	0.8282	0.7379	0.8137	0.7935	0.4833	0.433	0.4512
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.00000)
$\alpha_1 + \beta_1$	0.9645	0.9622	0.9524	0.9619	0.9709	0.9604	0.8268	0.8526	0.8277
Log L	2152	2188	2189	2280	2301	2308	2072	2119	2116
AIC	-4.8499	-4.9289	-4.9274	-4.9594	-5.0041	-5.0148	-4.7072	-4.8107	-4.8032
BIC	-4.8283	-4.9019	-4.8950	-4.9383	-4.9778	-4.9832	-4.6855	-4.7835	-4.7760

Note: The normal, students-t, and generalized error distributions are represented by Norm, Std-t, and Ged, respectively, and the p-values are placed in brackets. The log likelihood, Akaike Information Criterion, and Bayesian Information Criterion are abbreviated as Log L, AIC, and BIC respectively.

Table 4.7 and 4.8 reports the GARCH-in mean model parameter estimates for the daily and weekly FTSE100, S&P500 and NSE20 indices return computed according to equation

$$X_t = \mu_t + \lambda_1\sigma_t^2 + r_t,\ \sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \beta_1\sigma_{t-1}^2.$$

The mean μ_1 of the process is very low and not significant for all daily and weekly indices returns whereas the risk parameter, λ_1 is positive and significant for the daily indices returns except for the S&P500 and NSE20 indices returns under the generalized error distribution. On the other hand the risk parameter, λ_1 is positive and not significant for the weekly indices returns. A positive risk parameter value indicates a positive association between returns and volatilities, as well as the fact that the mean of the return sequence is heavily influenced by previous innovations and conditional variance. In other words, utilizing the conditional variance as a proxy for risk of return has a positive relationship with the level of return. If the risk parameter is statistically significant, then it implies that the asset in question may not be risky to hold otherwise it is risky holding. The parameters α_1 and β_1 are significant at 1% significance level for all the daily and weekly indices returns and the sum of α_1 and β_1 is close to 1 and it decreases as the data changes frequency from daily to weekly indices return. This means the volatility persistence decreases as the frequency of the return series changes from daily to weekly in all stock indices. Since the value of β_1 is very low in the daily and weekly NSE20 index returns compared to the value for the daily and weekly FTSE100 and S&P500 indices returns, it is inferred that the persistence of volatility dies off quickly in emerging market than in developed market.

Table 4.7: The parameter estimates of GARCH-M(1,1) model for daily indices returns

	FTSE100			S&P500			NSE20		
Estimate	Norm	Std-t	Ged	Norm	Std-t	Ged	Norm	Std-t	Ged
μ_1	-0.0005	-0.0004	-0.0004	-0.0001	0.0002	0.0001	-0.0005	-0.0005	-0.0003
	(0.1749)	(0.2080)	(0.3395)	(0.7088)	(0.5756)	(0.8394)	(0.1045)	(0.1440)	(0.5092)
λ_1	0.1126	0.1032	0.1010	0.0872	0.0681	0.0714	0.1033	0.0772	0.0510
	(0.0099)	(0.0032)	(0.0210)	(0.0510)	(0.0693)	(0.3431)	(0.0286)	(0.0912)	(0.4201)
ω	2×10^{-6}	1×10^{-6}	2×10^{-6}	2×10^{-6}	1×10^{-6}	1×10^{-6}	5×10^{-6}	9×10^{-6}	7×10^{-6}
	(0.0852)	(0.5086)	(0.5341)	(0.0714)	(0.1950)	(0.3270)	(0.0000)	(0.0000)	(0.0000)
α_1	0.111	0.1107	0.1111	0.0975	0.1009	0.1004	0.2699	0.36283	0.3167
	(0.0000)	(0.0029)	(0.0059)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
β_1	0.8766	0.8801	0.8775	0.8884	0.8975	0.8921	0.6889	0.52552	0.5965
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\alpha_1+\beta_1$	0.9876	0.9908	0.9886	0.9859	0.9979	0.9913	0.9588	0.8883	0.9132
Log L	-14080	14123	14139	13930	14017	14040	15069	15285	15257
AIC	-6.4746	-6.4935	-6.5005	-6.5147	-6.5548	-6.5652	-7.0539	-7.1549	-7.1416
BIC	-6.4673	-6.4847	-6.4902	-6.5073	-6.5459	-6.5548	-7.0464	-7.1460	-7.1326

Note: The normal, students-t, and generalized error distributions are represented by Norm, Std-t, and Ged, respectively, and the p-values are placed in brackets. The log likelihood, Akaike Information Criterion, and Bayesian Information Criterion are abbreviated as Log L, AIC, and BIC respectively.

Table 4.8: The parameter estimates of GARCH-M(1,1) model for weekly index returns

	FTSE100			S&P500			NSE20		
Estimate	Norm	Std-t	Ged	Norm	Std-t	Ged	Norm	Std-t	Ged
μ_1	-0.0033	-0.0022	-0.0032	-0.0002	0.0003	-0.0003	-0.0044	0.0027	0.0016
	(0.1348)	(0.3774)	(0.1151)	(0.9072)	(0.8553)	(0.9000)	(0.1780)	(0.3118)	(0.1206)
λ_1	0.2595	0.2004	0.2752	0.1536	0.1296	0.1631	0.2120	-0.0778	-0.0491
	(0.0234)	(0.1160)	(0.0109)	(0.1147)	(0.2013)	(0.2814)	(0.1313)	(0.4881)	(0.3964)
ω	3×10^{-5}	3×10^{-5}	3.2×10^{-5}	3×10^{-5}	2×10^{-5}	2.4×10^{-5}	1.4×10^{-5}	1.5×10^{-5}	1.0×10^{-4}
	(0.0038)	(0.0319)	(0.0178)	(0.0005)	(0.0180)	(0.0086)	(0.0000)	(0.0002)	(0.0003)
α_1	0.2109	0.1326	0.1754	0.2313	0.1662	0.1966	0.3655	0.4234	0.3746
	(0.0072)	(0.0014)	(0.0003)	(0.0000)	(0.0002)	(0.0000)	(0.0000)	(0.0001)	(0.0000)
β_1	0.7477	.8210	0.7736	0.7316	0.8019	0.7658	0.4552	0.4342	0.4544
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\alpha_1+\beta_1$	0.9581	0.9536	0.9513	0.9629	0.9681	0.9624	0.8207	0.8576	0.8290
Log L	2155	2190	2179	2282	2303	2308	2074	2120	2116
AIC	-4.8536	-4.9294	-4.906	-4.9599	-5.0037	-5.0134	-4.7075	-4.8091	-4.8012
BIC	-4.8266	-4.8970	-4.8734	-4.9336	-4.9722	-4.9766	-4.6803	-4.7765	-4.7685

Note: The normal, students-t, and generalized error distributions are represented by Norm, Std-t, and Ged, respectively, and the p-values are placed in brackets. The log likelihood, Akaike Information Criterion, and Bayesian Information Criterion are abbreviated as Log L, AIC, and BIC respectively.

The parameter estimation of EGARCH(1,1) model expressed as below;

$$\ln(\sigma_t^2) = \omega + \{\alpha_1(|r_{t-1}| - \mathbb{E}|r_{t-1}|) + \gamma_1 r_{t-1}\} + \beta_1 \ln(\sigma_{t-1}^2)$$

is carried out using the daily and weekly FTSE100, S&P500 and NSE20 indices returns and the results presented in Table 4.9 and 4.10. It is evident that the conditional mean, μ, of the process is in general significantly close to zero. The parameters ω , α_1 , β_1 and γ_1 are all statistically significant at 1% confidence interval except α_1 which is not significant for both daily and weekly NSE20 index returns. The ARCH-term, α_1, which explains clustering of volatility is large in the emerging market index returns than for the developed markets indices returns. This implies that the developed market indices returns are characterized by high volatility clustering compared to low volatility clustering in the emerging market index returns. The GARCH term, β_1, explains volatility persistence and a high value is reported in the developed market indices returns than in the emerging market index returns. This implies that volatility persists for a long time in the daily and weekly FTSE100 and S&P500 indices returns than in emerging market, however, as the frequency of the returns changes from daily to weekly, volatility diminishes for all indices returns. These findings are in tandem with the results of both GARCH(1,1) and GARCH-M(1,1) models. The asymmetric or leverage parameter γ_1 is positive and high in both daily and weekly returns of the NSE20 index compared to the corresponding frequencies in both FTSE100 and S&P500 indices returns. In addition, the parameter γ_1 is positive and significant in all indices returns which means there is non-existence of leverage effects but asymmetric volatility is present among the indices returns and thus the impact of negative news does not outweigh positive news, that is, good news increases volatility more than bad news. It is also noted that the asymmetry parameter γ_1 is big in both daily and weekly NSE20 index returns than in FTSE100 and S&P500 indices returns which shows that volatility asymmetry is more in emerging market than in established markets. This means positive shocks affects volatility more than negative shocks in developing markets compared to developed markets. On the other hand, the ARCH effect coefficient, α_1, is negative except in the weekly NSE20 index returns which is an indication that the variance goes up more after negative returns than after positive returns. The results further imply a positive and significant relationship between the stock returns and conditional volatility since the value of β_1 is positive and significant.

Table 4.9: The parameter estimates of EGARCH(1,1) model for daily indices returns

	FTSE100			S&P500			NSE20		
Estimate	Norm	Std-t	Ged	Norm	Std-t	Ged	Norm	Std-t	Ged
μ_1	-7.5×10^{-5}	1.19×10^{-4}	-9.5×10^{-6}	1.81×10^{-4}	3.75×10^{-4}	1.2×10^{-4}	1.86×10^{-4}	6.9×10^{-5}	5.7×10^{-5}
	(0.3075)	(0.2739)	(0.3783)	(0.0420)	(0.0001)	(0.1916)	(0.0455)	(0.4498)	(0.5348)
ω	-0.1672	-0.1578	-0.1748	-0.17166	-0.1293	-0.1667	-0.8310	-1.3566	-1.1540
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
α_1	-0.1283	-0.1423	-0.1398	-0.1390	-0.1501	-0.1527	-0.0080	0.0015	-0.0021
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.4364)	(0.9198)	(0.8795)
β_1	0.9818	0.9834	0.9812	0.9815	0.9867	0.9823	0.9132	0.8631	0.8837
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
γ_1	0.1183	0.1183	0.1185	0.1024	0.1041	0.1052	0.4342	0.4985	0.4667
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Log L	14182	14212	14233	14033	14108	14134	15075	15282	15258
AIC	-6.5211	-6.5346	-6.5435	-6.5629	-6.5976	-6.6092	-7.0569	-7.1533	-7.1421
BIC	-6.5137	-6.5258	-6.5332	-6.5555	-6.5887	-6.5988	-7.0494	-7.1443	-7.1332

Note: The normal, students-t, and generalized error distributions are represented by Norm, Std-t, and Ged, respectively, and the p-values are placed in brackets. The log likelihood, Akaike Information Criterion, and Bayesian Information Criterion are abbreviated as Log L, AIC, and BIC respectively.

Table 4.10: The parameter estimates of EGARCH(1,1) model for weekly index returns

	FTSE100			S&P500			NSE20		
Estimate	Norm	Std-t	Ged	Norm	Std-t	Ged	Norm	Std-t	Ged
μ_1	1.0×10^{-5}	5.81×10^{-4}	-5.9×10^{-6}	8.61×10^{-4}	1.7×10^{-3}	7.46×10^{-4}	3.68×10^{-4}	7.82×10^{-4}	5.64×10^{-4}
	(0.9999)	(0.9036)	(0.0000)	(0.0778)	(0.0005)	(0.1602)	(0.6237)	(0.1047)	(0.4466)
ω	-0.5407	-0.4710	-0.5194	-0.7676	-0.5415	-0.6717	-1.0305	-1.3518	-1.2446
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0006)	(0.0021)	(0.0037)
α_1	-0.2538	-0.2191	-0.2408	-0.2242	-0.2149	-0.2301	0.0077	0.0091	0.0084
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.7616)	(0.8249)	(0.8203)
β_1	0.9305	0.9409	0.9335	0.9017	0.9326	0.9145	0.8595	0.8197	0.8355
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
γ_1	0.1319	0.1257	0.1353	0.2584	0.1901	0.1932	0.4315	0.5026	0.4653
	(0.0000)	(0.0002)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Log L	2197	2215	2223	2309	2321	2335	2073	2119	2116
AIC	-4.9484	-4.9864	-5.0017	-5.0192	-5.0429	-5.0720	-4.7054	-4.8086	-4.8007
BIC	-4.9214	-4.9540	-4.9638	-4.9930	-5.0114	-5.0352	-4.6782	-4.7760	-4.7681

Note: The normal, students-t, and generalized error distributions are represented by Norm, Std-t, and Ged, respectively, and the p-values are placed in brackets. The log likelihood, Akaike Information Criterion, and Bayesian Information Criterion are abbreviated as Log L, AIC, and BIC respectively.

The parameter estimates of MS-GARCH(1,1) model defined as

$$\sigma_{k,t}^2 = \alpha_{0,k} + \alpha_{1,k} r_{t-1}^2 + \beta_k \sigma_{k,t-1}^2 \text{ for } k = 1, 2$$

are carried out by use of the daily and weekly indices returns and reported in Table 4.11 and 4.12. The parameter estimates of MS-GARCH(1,1) model for the daily FTSE100, S&P500 and NSE20 indices returns are presented in Table 4.11. All the parameter estimates are significant at 5% α-level of significance across the regimes. The ARCH-term, $\alpha_{1,k}$, is generally high in regime 2 than in regime 1 for all indices returns implying high volatility clustering in regime 2 and low volatility clustering in regime 1. On the other hand, the GARCH-term,β_k is highest in regime 1 than in regime 2 in all indices returns across all the conditional distributions. It is thus evident that the evolution of the volatility process across the two regimes is heterogeneous and regime 1 is characterized by high volatility persistence. For instance, taking the case for the FTSE100 daily index returns fitted to the MS-GARCH(1,1) model conditioned under normal distribution, it is reported that $\alpha_{1,2} = 0.1183 > \alpha_{1,1} = 0.0087$, and that $\beta_1 = 0.9801 > \beta_2 = 0.8628$. The corresponding estimated values for the daily S&P500 and NSE20 indices returns are reported as $\alpha_{1,2} = 0.1258 > \alpha_{1,1} = 0.0362$, $\beta_1 = 0.9499 > \beta_2 = 0.8699$ and $\alpha_{1,2} = 0.2932 > \alpha_{1,1} = 0.0071$, $\beta_1 = 0.9463 > \beta_2 = 0.3565$, respectively. A similar case can be extracted for Generalized error and students-t distributions for the three indices returns. The sum of $\alpha_{1,k}$ and β_k, that is, $\alpha_{1,k} + \beta_k$, is less than one across the two regimes for all the conditional distributions, however, the sum values differ across the regimes for the indices returns. For instance, the S&P500 return series reports higher values in regime 2 than in regime 1 whereas FTSE100 and NSE20 indices returns reports a higher sum values in regime 1. The value $\alpha_{1,k} + \beta_k$ represents the volatility persistence in the regimes (for $k = 1, 2$) and it is thus clear that regime 1 has high volatility persistence for the FTSE100 and NSE20 indices returns whereas there is low volatility persistence in regime 1 for S&P500 index returns. In general, the two regimes are characterized by existence of heterogeneous unconditional volatility, volatility persistence as well as a varied reaction to the past negative returns.

Table 4.12 presents the parameter estimate results for MS-GARCH (1,1) model applied to the weekly indices returns. The parameters $\alpha_{1,k}$ and β_k are all significant at least at 5% α-level of significance for all indices returns. It is revealed that the GARCH term β_k is generally high in regime 1 than in regime 2 for the FTSE100 and S&P500 indices returns whereas the value is low in regime 1 than in regime 2 in the NSE20 index returns. This means that the volatility persists for long in regime 1 in the developed market indices returns than in regime 2 in the emerging market index returns.

Table 4.11: The parameter estimates of MS-GARCH(1,1) model for daily indices returns

	FTSE100			S&P500			NSE20		
Estimate	Norm	Std-t	Ged	Norm	Std-t	Ged	Norm	Std-t	Ged
$\alpha_{0,1}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0002)	(0.0008)	(0.0024)	(0.0000)	(0.0553)	(0.2633)	(0.0032)	(0.0078)	(0.0000)
$\alpha_{1,1}$	0.0087	0.0071	0.0072	0.0362	0.1107	0.0283	0.0071	0.0151	0.2737
	(0.0376)	(0.0594)	(0.0829)	(0.0000)	(0.1427)	(0.0000)	(0.1133)	(0.1632)	(0.0000)
β_1	0.9801	0.9828	0.9828	0.9499	0.7819	0.9705	0.9463	0.9301	0.6285
	(0.0000)	(0.0002)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\alpha_{0,2}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0001
	(0.0011)	(0.0072)	(0.0074)	(0.0000)	(0.0000)	(0.0000)	(0.0004)	(0.0001)	(0.0000)
$\alpha_{1,2}$	0.1183	0.1327	0.1432	0.1258	0.0927	0.0780	0.2932	0.5480	0.9999
	(0.0522)	(0.0296)	(0.0196)	(0.0000)	(0.00433)	(0.0000)	(0.0082)	(0.0000)	(0.0000)
β_2	0.8628	0.8477	0.8378	0.8699	0.8954	0.8900	0.3565	0.2862	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0040)	(0.0000)	(0.2321)
P_{11}	0.9805	0.9862	0.9884	0.9389	0.9959	0.9918	0.9624	0.9672	0.9695
P_{21}	0.0425	0.0210	0.0144	0.2074	0.0009	0.0220	0.1261	0.0392	0.1423
$\alpha_{1,1} + \beta_1$	0.9888	0.9899	0.9900	0.9861	0.8926	0.9988	0.9534	0.9452	0.9022
$\alpha_{1,2} + \beta_2$	0.9811	0.9804	0.9810	0.9957	0.9881	0.9680	0.6497	0.8342	0.9999
Log L	14119	14129	14133	13990	14022	13993	15270	15312	15294
AIC	-28222	-28239	-28247	-27963	-28023	-27967	-30523	-30605	-30564
BIC	-28171	-28175	-28183	-27912	-27959	-27903	-30473	-30541	-30488

Note: The normal, students-t, and generalized error distributions are represented by Norm, Std-t, and Ged, respectively, and the p-values are placed in brackets. The log likelihood, Akaike Information Criterion, and Bayesian Information Criterion are abbreviated as Log L, AIC, and BIC respectively.

Table 4.12: The parameter estimates of MS-GARCH(1,1) model for weekly index returns

	FTSE100			S&P500			NSE20		
Estimate	Norm	Std-t	Ged	Norm	Std-t	Ged	Norm	Std-t	Ged
$\alpha_{0,1}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0001
	(0.0067)	(0.0129)	(0.0102)	(0.0686)	(0.1327)	(0.1312)	(0.0287)	(0.0853)	(0.0025)
$\alpha_{1,1}$	0.0902	0.0740	0.0729	0.0096	0.0037	0.0033	0.0914	0.0474	0.4939
	(0.0184)	(0.0409)	(0.0249)	(0.2110)	(0.1466)	(0.2032)	(0.2231)	(0.2833)	(0.0029)
β_1	0.8506	0.8681	0.8744	0.9620	0.9874	0.9877	0.6071	0.6703	0.3298
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0008)	(0.0072)
$\alpha_{0,2}$	0.0011	0.0000	0.0002	0.0004	0.0001	0.0001	0.0002	0.0001	0.0000
	(0.1293)	(0.1321)	(0.1179)	(0.0740)	(0.0987)	(0.1344)	(0.2297)	(0.0716)	(0.1892)
$\alpha_{1,2}$	0.3372	0.0289	0.4251	0.2623	0.1555	0.1937	0.1914	0.1117	0.0308
	(0.3043)	(0.1884)	(0.3552)	(0.1114)	(0.1288)	(0.1327)	(0.3374)	(0.2259)	(0.2655)
β_2	0.6466	0.9690	0.5745	0.4152	0.7649	0.6719	0.7955	0.8731	0.9580
	(0.0000)	(0.0000)	(0.0000)	(0.0430)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
P_{11}	0.9594	0.9425	0.8189	0.9818	0.9858	0.9862	0.8761	0.8777	0.6993
P_{21}	0.9999	0.2902	0.6173	0.0386	0.0141	0.0158	0.3422	0.2008	0.4545
$\alpha_{1,1} + \beta_1$	0.9408	0.9421	0.9473	0.9716	0.9911	0.9910	0.6985	0.7177	0.8237
$\alpha_{1,2} + \beta_2$	0.9838	0.9979	0.9916	0.6775	0.6775	0.8656	0.9869	0.9848	0.9888
Log L	2186	2190	2202	2288	2293	2292	2118	2119	2124
AIC	-4356	-4359	-4381	-4561	-4566	-4588	-4565	-4217	-4224
BIC	-4318	-4312	-4323	-4522	-4518	-4517	-4182	-416	-4167

Note: The normal, students-t, and generalized error distributions are represented by Norm, Std-t, and Ged, respectively, and the p-values are placed in brackets. The log likelihood, Akaike Information Criterion, and Bayesian Information Criterion are abbreviated as Log L, AIC, and BIC respectively.

Further, regime 1 is characterized with low values of ARCH term, $\alpha_{1,k}$ as compared with regime 2 in all indices returns and this implies that regime 1 has low volatility clustering in comparison to regime 2. The probabilities of switching regimes are very low for all indices returns which imply that the process has the tendency to spend more time in regime 1 than in regime 2. That is, the volatility process takes long to switch from regime 1 to regime 2 than it takes to revert back once in regime 2.

The parameter estimates of MS-EGARCH (1,1) model presented in Table 4.13 and 4.14 are computed according to the equation

$$\ln\Big(\sigma_{k,t}\Big) = \alpha_{0,k} + \alpha_{1,k}\Big(|\eta_{k,t-1}| - \mathbb{E}[\eta_{k,t-1}|]\Big) + \alpha_{2,k} r_{t-1} + \beta_k \ln\Big(\sigma_{k,t-1}\Big), \text{ for } k = 1, 2.$$

Table 4.13 reports that all the parameter estimates are significant across the regimes except that $\alpha_{2,1}$ and $\alpha_{2,2}$ are not significant in both regimes for NSE20 index returns. The value of $\alpha_{1,k}$ in regime 1 is large than in regime 2 for FTSE100 index return whereas it is low in regime 1 compared to regime 2 for both the S&P500 and NSE20 indices returns. This implies that FTSE100 index returns have high volatility clustering in regime 1 compared to low volatility clustering in regime 1 for the S&P500 and NSE20 indices returns. Further, it is noted that $\alpha_{2,1}$ and $\alpha_{2,2}$ are in general negative in both regimes implying presence of leverage effects in the returns series across the two regimes. That is, when compared to a positive shock of the same magnitude, a negative shock has a higher influence on volatility, thus it can be argued that, this persistence of negative shocks, or volatility asymmetry, indicates that investors are more vulnerable to bad news than favorable news (the volatility spillover mechanism is asymmetric). The volatility persistence is high in regime 1 for both FTSE100 and NSE20 indices returns than in the same regime for S&P500 index returns as reported by high value of β_k.

Table 4.14 presents the parameter estimates of MS-EGARCH(1,1) for the weekly indices returns. A similar pattern of behavior as that for daily indices returns is noticed, however, they differ in the sense that NSE20 index returns does not show leverage effects in regime 2 and in addition the three markets indices exhibit high volatility persistence in regime 1 as reported by large values of β_1.

In general, in the daily and weekly FTSE100 indices returns, regime 1 is characterized by high conditional volatility, strong volatility reaction to past negative returns and low volatility persistence while the second regime exhibits low conditional volatility, weak volatility reaction to past negative returns and high volatility persistence. Leverage effects are reported in all the return series across all the regimes. Moreover, most of the transition probabilities reports that the volatility process has the tendency to spend more time in regime 1 than in regime 2, that is, the process has very high probability of staying in the same state compared to the probability of transiting to the other state.

Table 4.13: The parameter estimates of MS-EGARCH(1,1)model for daily indices returns

	FTSE100			S&P500			NSE20		
Estimate	Norm	Std-t	Ged	Norm	Std-t	Ged	Norm	Std-t	Ged
$\alpha_{0,1}$	-0.1547	-0.2088	-0.3185	-0.9472	-1.0937	-1.0277	-0.6693	-0.4198	-0.5139
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0066)	(0.0000)	(0.0017)
$\alpha_{1,1}$	0.0647	0.0689	0.0937	0.0621	0.0605	0.0638	0.1297	0.0400	0.0551
	(0.0000)	(0.0000)	(0.0000)	(0.0110)	(0.0412)	(0.0266)	(0.0117)	(0.0020)	(0.0091)
$\alpha_{2,1}$	-0.1486	-0.1609	-0.1759	-0.2635	-0.3547	0.3346	0.0023	-0.0095	-0.0050
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.4111)	(0.1562)	(0.2848)
β_1	0.9844	0.9786	0.9673	0.9052	0.8869	0.8933	0.9383	0.9611	0.9526
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\alpha_{0,2}$	-0.1516	-0.2016	-0.2019	-0.2024	-0.0842	-0.0861	-0.2757	-2.3101	-2.8062
	(0.0004)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0492)	(0.0000)	(0.0000)
$\alpha_{1,2}$	0.0461	0.0371	0.0813	0.0855	0.0731	0.0717	0.3747	0.6942	0.6037
	(0.0092)	(0.0417)	(0.0056)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\alpha_{2,2}$	-0.1600	-0.1686	-0.1395	-0.1389	-0.1142	-0.1113	-0.0114	0.0053	0.0048
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.3478)	(0.4335)	(0.4503)
β_2	0.9811	0.9755	0.9764	0.9767	0.9908	0.9907	0.9517	0.7455	0.6781
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
P_{11}	0.9846***	0.9954***	0.9977***	0.9936***	0.9941***	0.9954***	0.9437***	0.9678***	0.9658***
P_{21}	0.0764***	0.0237***	0.0062***	0.0062***	0.0032	0.0025	0.2761***	0.0452***	0.0751***
Log L	14218	14224	14222	14082	14137	14138	15285	15314	15297
AIC	-28418	-28425	-28420	-28144	-28251	-28253	-30550	-30603	-30570
BIC	-28353	-28348	-28344	-28081	-28175	-28176	-30486	-30527	-30494

Note: The normal, students-t, and generalized error distributions are represented by Norm, Std-t, and Ged, respectively, and the p-values are placed in brackets. The log likelihood, Akaike Information Criterion, and Bayesian Information Criterion are abbreviated as Log L, AIC, and BIC respectively. The asterisks *, ** & *** represents α-level significance at 10%, 5% and 1%, respectively

Table 4.14: The parameter estimates of MS-EGARCH(1,1) model for weekly index returns

	FTSE100			S&P500			NSE20		
Estimate	Norm	Std-t	Ged	Norm	Std-t	Ged	Norm	Std-t	Ged
$\alpha_{0,1}$	-0.8367	-1.8223	-1.2039	2.4664	-2.2089	-2.6942	-1.2462	-1.2628	-1.3276
	(0.0000)	(0.0000)	(0.0158)	(0.0000)	(0.0000)	(0.0000)	(0.0050)	(0.0058)	(0.0027)
$\alpha_{1,1}$	0.0506	0.0007	0.0990	0.0181	0.0278	0.0060	0.2481	0.2487	0.3376
	(0.0000)	(0.4963)	(0.0398)	(0.0000)	(0.3370)	(0.0000)	(0.0026)	(0.0037)	(0.0002)
$\alpha_{2,1}$	-0.2768	-0.4992	-0.3075	-0.3582	-0.3689	-0.3460	-0.0082	-0.0087	-0.0101
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.4113)	(0.4084)	(0.3812)
β_1	0.8961	0.7722	0.8544	0.7077	0.7379	0.6809	0.8564	0.8543	0.8423
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\alpha_{0,2}$	-0.7449	-0.1825	-0.4168	-1.2831	-0.7769	-1.4592	-0.2483	-0.2465	-0.0502
	(0.0000)	(0.0001)	(0.1012)	(0.0000)	(0.0027)	(0.0000)	(0.1474)	(0.1466)	(0.4021)
$\alpha_{1,2}$	-0.3299	0.0504	0.0561	0.1120	0.0611	0.1712	0.3379	0.3270	0.5794
	(0.0000)	(0.0558)	(0.1419)	(0.0000)	(0.2107)	(0.0000)	(0.0371)	(0.0414)	(0.0024)
$\alpha_{2,2}$	-0.4069	-0.1778	-0.2198	-0.2949	-0.2497	-0.3032	0.0098	0.0102	-0.0026
	(0.0000)	(0.0000)	(0.0006)	(0.0000)	(0.0000)	(0.0000)	(0.4251)	(0.4215)	(0.4823)
β_2	0.8914	0.9771	0.9428	0.8187	0.8933	0.7957	0.9428	0.9440	0.9412
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
P_{11}	0.9948***	0.9782***	0.9832***	0.9843***	0.9905***	0.9838***	0.8765***	0.8780***	0.9203***
P_{21}	0.0361***	0.0155	0.0277*	0.0144***	0.0102*	0.0301***	0.4402***	0.4271***	0.7501***
Log L	2218	2221	2216	2322	2326	2322	2122	2121	2123
AIC	-4416	-4419	-4408	-4624	-4628	-4620	-4223	-4219	-4221
BIC	-4368	-4361	-4351	-4576	-4570	-4562	-4176	-4162	-4164

Note: The normal, students-t, and generalized error distributions are represented by Norm, Std-t, and Ged, respectively, and the p-values are placed in brackets. The log likelihood, Akaike Information Criterion, and Bayesian Information Criterion are abbreviated as Log L, AIC, and BIC respectively. The asterisks *, ** & *** represents α-level significance at 10%, 5% and 1%, respectively

In order to explore the effect of trade volume on volatility persistence, Equation (3.5.2) is adapted to integrate log trade volume, T_v, as an exogenous variable in the conditional variance as shown below and the model parameter estimates are presented in Table 4.15 and 4.16.

$$r_t = \sigma_t \varepsilon_t,\ \ \sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \delta T_v\ .$$

The parameter estimates for the daily and weekly indices returns for the three markets are significant at 1% α-level of significance except that μ_1 is not significant for the NSE20 index returns. The ARCH term, α_1, is low for FTSE100 and S&P500 indices returns compared to the value for the NSE20 index returns which implies low volatility clustering in the developed markets and high volatility in the emerging market. Taking for instance the model estimates conditioned to normal distribution for the daily indices returns for the three markets, the value for α_1 are reported as 0.0900, 0.0963 and 0.2639 for the FTSE100, S&P500 and NSE20 indices returns respectively. Similarly, the α_1 parameter estimates are 0.1855, 0.1815 and 0.2432 for the weekly FTSE100, S&P500 and NSE20 indices returns respectively. On the other hand the β_1 value is high in both FTSE100 and S&P500 indices returns than in NSE20 index returns. The β_1 parameter estimate of the model for the daily indices returns under normal distribution are reported as 0.8984, 0.8886 and 0.6923 for the FTSE100, S&P500 and NSE20 indices returns respectively. Similarly, the values for the weekly indices returns under normal distribution are 0.7818, 0.7597 and 0.6168 for the FTSE100, S&P500 and NSE20 market indices respectively. In general, the established markets' indices returns are characterized by low volatility clustering and high volatility persistence as compared to high volatility clustering and low volatility persistence in the developing market index returns.

Table 4.15: GARCH(1,1) parameter estimates for daily indices returns with trading volume

	FTSE100			S&P500			NSE20		
Estimate	Norm	Std-t	Ged	Norm	Std-t	Ged	Norm	Std-t	Ged
μ_1	0.0004	0.0004	0.0004	0.0005	0.0006	0.0006	0.0001	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.8581)	(0.2013)
ω	1×10^{-6}	1×10^{-6}	1×10^{-6}	1×10^{-6}	1×10^{-6}	1×10^{-6}	5×10^{-6}	9×10^{-6}	7×10^{-6}
	(0.0000)	(0.0024)	(0.0015)	(0.0000)	(0.0000)	(0.0064)	(0.0000)	(0.0000)	(0.0000)
α_1	0.0900	0.09088	0.0912	0.0963	0.1102	0.0960	0.2639	0.3535	0.3104
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
β_1	0.8984	0.8980	0.8969	0.8886	0.8794	0.8940	0.6923	0.5355	0.6019
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
δ	1.9×10^{-5}	1.9×10^{-5}	1.9×10^{-5}	3.2×10^{-5}	6.9×10^{-5}	3.2×10^{-5}	4×10^{-6}	4×10^{-6}	4×10^{-6}
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0182)	(0.0000)
$\alpha_1+\beta_1$	0.9884	0.9889	0.9881	0.9849	0.9896	0.9900	0.9562	0.8890	0.9123
Log L	14108	14146	14150	13996	14099	14075	15072	15287	15260
AIC	-6.4872	-6.5041	-6.5061	-6.5455	-6.5933	-6.5821	-7.0555	-7.1558	-7.1429
BIC	-6.4798	-6.4953	-6.4973	-6.5381	-6.5844	-6.5732	-7.0480	-7.1468	-7.1340

Note: The normal, students-t, and generalized error distributions are represented by Norm, Std-t, and Ged, respectively, and the p-values are placed in brackets. The log likelihood, Akaike Information Criterion, and Bayesian Information Criterion are abbreviated as Log L, AIC, and BIC respectively.

In order to check the effect on volatility persistence in the three market indices after inclusion of lagged trading volume into the GARCH model the value of $\alpha_1 + \beta_1$ reported for GARCH(1,1) model is compared with that of GARCH(1,1) model with log volume included. A general finding is that for the daily indices returns, the volatility persistence decreased in both emerging and developed markets even though the normal distribution and student-t distributions reports contrary findings for the FTSE100 and NSE20 indices returns respectively after inclusion of log volume into the model. It is noted that the volatility persistence decreased for the weekly S&P500 index returns whereas it increased for the FTSE100 and NSE20 indices returns. It can be a general conclusion that inclusion of log trading volume on GARCH (1,1) model reports mixed results when data changes frequency from daily to weekly indices returns.

Table 4.16: GARCH(1,1) estimates for weekly index returns with volume

	FTSE100			S&P500			NSE20		
Estimate	Norm	Std-t	Ged	Norm	Std-t	Ged	Norm	Std-t	Ged
μ_1	0.0019	0.0022	0.0025	0.0021	0.0023	0.0023	0.0005	0.0009	0.0005
	(0.0000)	(0.1138)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.5298)	(0.1785)	(0.1797)
ω	2.8e-5	1.2e-5	1.2e-5	2.9e-5	1.9e-5	2.4e-5	9.3e-5	1.26e-4	1.10e-4
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0003)	(0.0003)	(0.0007)
α_1	0.1855	0.0789	0.0962	0.1815	0.1246	0.1533	0.2432	0.3545	0.2960
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0002)	(0.0002)
β_1	0.7818	0.8980	0.8815	0.7597	0.8344	0.7962	0.6168	0.5005	0.5443
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
δ	0.0001	0.0001	0.0001	0.0002	0.0002	0.0002	0.0001	0.0001	0.0001
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0125)	(0.0040)
$\alpha_1 + \beta_1$	0.9673	0.9769	0.9777	0.9412	0.9590	0.9495	0.8600	0.855	0.8403
Log L	2157	2192	2180	2295	2310	2309	2078	2122	2119
AIC	-4.8585	-4.9336	-4.9083	-4.9891	-5.0206	-5.0179	-4.7176	-4.8136	-4.8065
BIC	-4.8315	-4.9012	-4.8759	-4.9628	-4.9891	-4.9864	-4.6904	-4.7810	-4.7739

Note: The normal, students-t, and generalized error distributions are represented by Norm, Std-t, and Ged, respectively, and the p-values are placed in brackets. The log likelihood, Akaike Information Criterion, and Bayesian Information Criterion are abbreviated as Log L, AIC, and BIC respectively.

Table 4.17 presents GARCH-M(1,1) with volume included in the variance equation for daily indices returns computed according to equation

$$X_t = \mu_t + \lambda_1\sigma_t^2 + r_t, \text{ where } , r_t = \sigma_t\varepsilon_t, \text{ and } \sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \beta_1\sigma_{t-1}^2 + \delta T_v \, .$$

The parameters ω , α_1 , β_1 and δ are significant at 1% level of significance. The value of the risk parameter λ_1 is positive and decreases when trading volume is included into the conditional variance equation for developed markets but increases for the emerging market. This is an indication that indices returns and volatility have a positive association and that the risk of holding asset returns from developed markets is less compared with that of holding asset returns from emerging market.

Table 4.17: GARCH-M(1,1) estimates for daily indices returns with volume

	FTSE100			S&P500			NSE20		
Estimate	Norm	Std-t	Ged	Norm	Std-t	Ged	Norm	Std-t	Ged
μ_1	-0.000205	-0.00004	-0.000132	-0.000099	0.0004	0.000273	-0.000677	-0.000634	-0.000455
	(0.190415)	(0.8054)	(0.0000)	(0.31064)	(0.1830)	(0.1450)	(0.0222)	(0.5280)	(0.0000)
λ_1	0.0778	0.0670	0.0831	0.0783	0.0315	0.0535	0.1239	0.0995	0.0738
	(0.0039)	(0.0090)	(0.0000)	(0.0000)	(0.4421)	(0.0834)	(0.0037)	(0.5818)	(0.0000)
ω	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0905)	(0.0020)	(0.0000)	(0.0049)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
α_1	0.0891	0.0858	0.1160	0.096257	0.0943	0.1059	0.266110	0.348419	0.309233
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0132)	(0.0000)
β_1	0.8984	0.9031	0.8754	0.888953	0.9002	0.8803	0.687509	0.541778	0.602435
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
δ	0.000019	0.000019	0.000018	0.000032	0.000032	0.000027	0.000004	0.000004	0.000004
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.4319)	(0.0000)
$\alpha_1 + \beta_1$	0.9875	0.9889	0.9914	0.98521	0.9945	0.9862	0.953619	0.890197	0.911668
Log L	14111	14149	14151	13998	14067	14070	15075	15289	15261
AIC	-6.4882	-6.5048	-6.5061	-6.5459	-6.5779	-6.5793	-7.0566	-7.1564	-7.1431
BIC	-6.4794	-6.4946	-6.4958	-6.5370	-6.5675	-6.5689	-7.0476	-7.1460	-7.1327

Note: The normal, students-t, and generalized error distributions are represented by Norm, Std-t, and Ged, respectively, and the p-values are placed in brackets. The log likelihood, Akaike Information Criterion, and Bayesian Information Criterion are abbreviated as Log L, AIC, and BIC respectively.

Table 4.18: GARCH-M(1,1) estimates for weekly indices returns with volume

	FTSE100			S&P500			NSE20		
Estimate	Norm	Std-t	Ged	Norm	Std-t	Ged	Norm	Std-t	Ged
μ_1	-0.000988	0.003202	-0.000819	-0.0002	0.000333	-0.000341	-0.003108	0.002120	0.001136
	(0.6945)	(0.0000)	(0.7980)	(0.9088)	(0.8586)	(0.9058)	(0.3062)	(0.4194)	(0.8994)
λ_1	0.1475	-0.0740	0.1525	0.1536	0.1296	0.1631	0.1623	-0.0557	-0.0284
	(0.2559)	(0.0000)	(0.3418)	(0.1184)	(0.2109)	(0.3081)	(0.2238)	(0.6231)	(0.9416)
ω	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.000110
	(0.0077)	(0.0000)	(0.0406)	(0.0006)	(0.0185)	(0.0089)	(0.0002)	(0.0004)	(0.0007)
α_1	0.197387	0.080316	0.156225	0.231321	0.166210	0.196623	0.2513	0.358991	0.297316
	(0.0000)	(0.0000)	(0.0020)	(0.0000)	(0.0002)	(0.0000)	(0.0000)	(0.0002)	(0.0003)
β_1	0.767095	0.896484	0.801833	0.731552	0.8019	0.765727	0.6007	0.496704	0.543362
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
δ	0.000103	0.000147	0.000091	0.0000	0.0000	0.0000	0.0001	0.000063	0.000067
	(0.0105)	(0.0000)	(0.1105)	(0.9999)	(1.0000)	(1.0000)	(0.0000)	(0.0185)	(0.0081)
$\alpha_1 + \beta_1$	0.964482	0.9768	0.958058	0.962873	0.96811	0.96235	0.8520	0.855695	0.840678
Log L	2158	2194	2180	2282	2303	2302	2079	2122	2118
AIC	-4.8573	-4.9367	-4.9060	-4.9577	-5.0015	-4.9992	-4.7171	-4.8116	-4.8043
BIC	-4.8249	-4.8989	-4.8682	-4.9262	-4.9647	-4.9624	-4.6844	-4.7736	-4.7663

Note: The normal, students-t, and generalized error distributions are represented by Norm, Std-t, and Ged, respectively, and the p-values are placed in brackets. The log likelihood, Akaike Information Criterion, and Bayesian Information Criterion are abbreviated as Log L, AIC, and BIC respectively.

The positive correlation between indices returns and volume is further confirmed by the positive value of δ . Adding trading volume in the GARCH-M(1,1) equation slightly reduces the volatility persistence in both developing and developed markets although the Generalized error distribution and the student-t distributions reports an increase of the volatility persistence for the developing and developed markets, respectively. Table 4.18 shows the parameter estimates of the GARCH-M(1,1) with volume included in the variance equation for weekly indices returns. The parameters ω , α_1 , β_1 and δ are significant at 1% confidence interval except δ for the S&P500 index returns. The volatility persistence increases across the markets on adding trading volume into the equation of variance and the risky parameter decreases for the developed market whereas it increases for emerging market as is the case in the daily indices returns. The general finding is that trading volume explains volatility in both developed and emerging markets and that trading volume is positively related with indices returns.

The results of EGARCH(1,1) with trading volume included into the conditional variance equation for daily and weekly indices returns are computed by adapting Equation 3.7.2 to include trading volume, T_v, as an exogenous variable, as presented in the equation below, and presented in Table 4.19 and 4.20.

$$\ln(\sigma_t^2) = \omega + \{\alpha_1(|r_{t-1}| - \mathbb{E}|r_{t-1}|) + \gamma_1 r_{t-1}\} + \beta_1 \ln(\sigma_{t-1}^2) + \delta T_v$$

The parameters ω , α_1 , β_1 , γ_1 and δ are significant at 5% level of significance in both developed and emerging daily and weekly indices returns except that α_1 and δ are not significant for NSE20 and weekly S&P500 indices returns respectively. The asymmetry parameter γ_1 diminishes with inclusion of trading volume into the conditional variance equation for both daily and weekly indices returns respectively whereas the GARCH parameter β_1 , which is the measure for volatility persistence, increases with volume addition into the equation of conditional variance.

Table 4.19: EGARCH(1,1) estimates for daily indices returns with volume

	FTSE100			S&P500			NSE20		
Estimate	Norm	Std-t	Ged	Norm	Std-t	Ged	Norm	Std-t	Ged
μ_1	0.0001	0.0002	0.0002	0.0002	0.0004	0.0004	0.0002	0.0000	0.0000
	(0.5670)	(0.0725)	(0.1020)	(0.0256)	(0.0000)	(0.0000)	(0.0366)	(0.5909)	(0.5674)
ω	-0.1320	-0.1269	-0.1302	-0.1553	-0.1003	-0.1228	-0.8034	-1.2830	-1.1011
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
α_1	-0.1203	-0.1310	-0.1250	-0.1190	-0.1201	-0.1205	-0.0121	-0.0015	-0.0056
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.2281)	(0.9154)	(0.6816)
β_1	0.9856	0.9867	0.9864	0.9833	0.9898	0.9875	0.9162	0.8706	0.8891
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
γ_1	0.1087	0.1108	0.1096	0.1154	0.0970	0.1079	0.4250	0.4859	0.4562
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
δ	0.6804	0.6857	0.6816	1.2264	1.7073	1.4266	0.1716	0.1494	0.1607
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0004)	(0.0002)
Log L	14232	14253	14251	14139	14212	14196	15088	15288	15265
AIC	-6.5435	-6.5528	-6.5522	-6.6121	-6.6457	-6.6280	-7.0624	-7.1557	-7.1449
BIC	-6.5347	-6.5426	-6.5419	-6.6032	-6.6353	-6.6280	-7.0535	-7.1453	-7.1345

Note: The normal, students-t, and generalized error distributions are represented by Norm, Std-t, and Ged, respectively, and the p-values are placed in brackets. The log likelihood, Akaike Information Criterion, and Bayesian Information Criterion are abbreviated as Log L, AIC, and BIC respectively.

The trading volume parameter δ is greater than zero and this implies that trading volume significantly explains volatility and that trading volume and volatility have a positive relationship. The decrease of the asymmetry parameter with volume inclusion shows that volume leads to less asymmetric volatility on the underlying market and the increase in volatility persistence is a show that volume does not explain volatility. It can be construed to mean that trading volume inclusion into the conditional variance does not reduce volatility persistence hence it does not explain volatility and that trading volume is positively related with volatility. Moreover, it implies that bad news has bigger effect on conditional volatility than good news and this means existence of asymmetry in the market. As a result, trading volume influences the flow of information into the market, and it is also crucial in predicting volatility patterns in both developed and emerging markets.

Table 4.20: EGARCH(1,1) estimates for weekly indices returns with volume

	FTSE100			S&P500			NSE20		
Estimate	Norm	Std-t	Ged	Norm	Std-t	Ged	Norm	Std-t	Ged
μ_1	0.0004	0.0008	0.0001	0.0001	0.0017	0.0015	0.0003	0.0008	0.0005
	(0.0931)	(0.0000)	(0.0610)	(0.0794)	(0.0006)	(0.0004)	(0.6889)	(0.1251)	(0.4977)
ω	-0.4597	-0.4316	-0.4355	-0.7622	-0.5378	-0.6736	-0.8027	-1.1800	-1.0327
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0002)	(0.0003)	(0.0007)	(0.0013)
α_1	-0.2355	-0.2124	-0.2247	-0.2224	-0.2136	-0.2160	-0.0188	0.0026	-0.0075
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.4317)	(0.9454)	(0.8286)
β_1	0.9412	0.9459	0.9455	0.9024	0.9330	0.9157	0.8912	0.8437	0.8643
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
γ_1	0.1264	0.1252	0.1246	0.2582	0.1902	0.2351	0.3813	0.4761	0.4289
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
δ	0.4542	0.3551	0.4265	-0.1157	-0.1104	-0.1134	0.3677	0.3142	0.3386
	(0.0000)	(0.0173)	(0.0026)	(0.6721)	(0.7266)	(0.7288)	(0.0000)	(0.0002)	(0.0001)
Log L	2203	2218	2210	2309	2321	2319	2089	2127	2124
AIC	-4.9599	-4.9902	-4.9738	-5.0172	-5.0409	-5.0378	-4.7390	-4.8232	-4.8172
BIC	-4.9275	-4.9523	-4.9360	-4.9857	-5.0041	-5.0010	-4.7064	-4.7852	-4.7791

Note: The normal, students-t, and generalized error distributions are represented by Norm, Std-t, and Ged, respectively, and the p-values are placed in brackets. The log likelihood, Akaike Information Criterion, and Bayesian Information Criterion are abbreviated as Log L, AIC, and BIC respectively.

Table 4.21 reports the results of a simple Regime switching model for the returns-volume relationship, which are computed according to a simple Regime switching model defined as follows;

$$X_t = \mu_{\xi(t)} + \alpha_{\xi(t)} T_{v,\xi(t)} + \sigma_{\xi(t)} r_t \text{ where } r_t \sim N(0,1) \text{ for } \xi(t) = 1, 2 .$$

The probability of the volatility process to transit from regime 1 to 2 (P_{12}) for the FTSE100, S&P500 and NSE20 daily indices is 0.0099, 0.0092 and 0.0307 while that of reverting back(P_{21}) is 0.0232, 0.0223 and 0.1300, respectively. This means once the volatility process is in regime 1 it takes long to switch to regime 2 than it takes to fall back to regime 1 once in regime 2. In fact, the average duration of stay in regime i in a single run for the daily indices returns is computed using the formula $\frac{1}{(1-P_{ii})}$, where $i = 1, 2$ and it is reported as 101, 109 and 33 days compared to 43, 49 and 8 days in regime 2, respectively. Also, the transition probability of entering regime 2 from regime 1 is 0.0266, 0.0236 and 0.0634 for the FTSE100, S&P500 and NSE20 weekly indices returns respectively and it implies that the volatility process lingers in regime 1 than in regime 2 before transiting. The NSE20 index returns seems to have high probability of the volatility process to transit from regime 1 to regime 2 than that of reverting back to regime 1 compared to the case for the FTSE100 and S&P500 indices returns. Regime 1 of the developed market daily and weekly indices returns is characterized by high volatility and negative returns while regime 1 of the emerging market is characterized by high volatility and positive returns. On the other hand, regime 2 is low volatility regime and with positive returns for the daily and weekly FTSE100, S&P500 and NSE20 indices returns. Moreover, the daily and weekly indices returns for both established and developing markets and log trading volume are negatively related in both regimes except for the NSE20 weekly indices in regime 2.

Table 4.21: Regime switching model estimates for the returns-volume relationship

	FTSE100		S&P500		NSE20	
	Daily	Weekly	Daily	Weekly	Daily	Weekly
μ_1	-0.0009	-0.0060	-0.0001	-0.0045	0.0005	0.0013
	(0.0719)	(0.0865)	(0.0956)	(0.0608)	(0.4750)	(0.7103)
μ_2	0.0005***	0.0018*	0.0007***	0.0031***	0.0001	0.0006
	(0.0000)	(0.0101)	(0.0000)	(0.0000)	(0.3173)	(0.3914)
α_1	-0.0053***	-0.0064	-0.0047	-0.0303***	0.0004	0.0081*
	(0.0009)	(0.3248)	(0.0817)	(0.0005)	(0.6171)	(0.0404)
α_2	-0.0004	-0.0028	-0.0010	-0.0005	0.0000	-0.0002
	(0.3173)	(0.1198)	(0.1531)	(0.8676)	(1.0000)	(0.8242)
σ_1	0.0188	0.0428	0.0195	0.0364	0.0166	0.0451
σ_2	0.0071	0.0161	0.0070	0.0144	0.0050	0.0154
P_{12}	0.0099	0.0266	0.0092	0.0236	0.0307	0.0634
P_{21}	0.0232	0.1049	0.0223	0.0572	0.1300	0.2150

Note: The values $\{\mu_1, \sigma_1\}$and$\{\mu_2, \sigma_2\}$ stand for the mean and variance in regimes 1 and 2 respectively and P_{12} andP_{21} are the transition probabilities from regime 1 to 2 and from regime 2 to 1 respectively. The asterisks *, ** & *** indicate 10%, 5% and 1% significant levels, respectively.

Table 4.22 reports for parameter estimates of Regime Switching model for the FTSE100, S&P500 and NSE20 indices returns computed according to Equation

$$X_t = \mu_{\xi(t)} + \sigma_{\xi(t)} r_t \text{ where } r_t \sim N(0,1) \text{ for } \xi(t) = 1, 2 .$$

All the indices returns exhibit high volatility in regime 1 than in regime 2. A comparison of volatility across regimes for the daily FTSE100, S&P500 and NSE20 indices returns is reported as $\sigma_1 = 0.0189 > \sigma_2 = 0.0071$, $\sigma_1 = 0.0196 > \sigma_2 = 0.0070$ and $\sigma_1 = 0.0166 > \sigma_2 = 0.0050$, respectively. The same comparison for the weekly indices returns for the three markets infers similar trend of volatility as that for the daily indices returns. Clearly, it can be concluded that regime 1 is high volatility regime with negative returns for the developed markets and high volatility and with positive regimes for the NSE20 indices returns. The transition probabilities for the volatility process to switch from regime 1 to regime 2 are reported as 0.0098, 0.0092 and 0.0309 respectively for the daily FTSE100, S&P500 and NSE20 indices returns. These probabilities are lower than those of switching from regime 2 to regime 1. In fact, the probability of staying in one regime is very high, which means, once in one regime the volatility process takes long before switching to the next regime. In general, the probability of volatility process changing regimes from 1 to 2 is less than that of reverting back to regime 1 implying that while in regime 1 the volatility process takes long before switching to regime 2. The duration of stay in regime 1 is reported as 102, 109 and 33 days compared to 43, 45, and 8 days in regime 2 for the daily FTSE100, S&P500 and NSE20 indices, respectively. Furthermore, it is noted that the volatility increases as the indices returns changes frequency from daily to weekly in the respective regimes.

Table 4.22: Regime switching estimates for indices returns

	FTSE100		S&P500		NSE20	
	Daily	Weekly	Daily	Weekly	Daily	Weekly
μ_1	-0.0009	-0.0062	-0.0010	-0.0045	0.0005	0.0012
	(0.0719)	(0.0850)	(0.0956)	(0.0608)	(0.4047)	(0.7162)
μ_2	0.0004***	0.0018*	0.0007***	0.0031***	0.0001	0.0006
	(0.0000)	(0.0101)	(0.0000)	(0.0000)	(0.3173)	(0.3914)
σ_1	0.0189	0.0432	0.0196	0.0373	0.0166	0.0453
σ_2	0.0071	0.0161	0.0070	0.0144	0.0050	0.0152
P_{12}	0.0098	0.0274	0.0092	0.0220	0.0307	0.0651
P_{21}	0.0232	0.1101	0.0224	0.0532	0.1299	0.2122

Note: The values $\{\mu_1, \sigma_1\}$and$\{\mu_2, \sigma_2\}$ stand for the mean and variance in regimes 1 and 2 respectively and P_{12} andP_{21} are the transition probabilities from regime 1 to 2 and from regime 2 to 1 respectively. The asterisks *, ** & *** indicate 10%, 5% and 1% significant levels, respectively.

Furthermore, the Pearson correlation coefficient which measures the strength of linear dependence between any two variables is computed to investigate the relationship between volume and stock returns. Table 4.23 shows the correlation coefficients between index returns and their corresponding volume. The results disclose a negative relationship between index returns and volume for the FTSE100 and S&P500 indices compared to the NSE20 index which reports a positive relationship. These results are in agreement with the findings of the regime switching model discussed in the earlier section.

Table 4.23: Correlation coefficients for returns and trading volume

	FTSE100	S&P500	NSE20
Daily	-0.0504	-0.0364	0.0084
Weekly	-0.0702	-0.1208	0.0682

4.4 Option pricing

In this section the results and discussion of the European call options pricing are presented for the Regime-Switching, Black-Scholes and RS-GARCH models. The Black-scholes and RS-GARCH models are as presented earlier in Equations (3.11.28) and (3.11.77), respectively, whereas the Regime-switching model is stated as follows;

$$C(K,T) = \mathbb{E}^{\mathbb{Q}}\big[max(X_T - K)|\mathcal{R}\big] = S_t\phi(d_1) - e^{-r\tau}K\phi(d_2) \quad \text{where}$$

$$d_1 = \frac{\ln(\frac{S_t}{K}) + r\tau + \frac{1}{2}\big[\mathcal{R}\sigma_1^2 + (\tau - \mathcal{R})\sigma_2^2\big]}{\sqrt{\mathcal{R}\sigma_1^2 + (\tau - \mathcal{R})\sigma_2^2}}$$

$$d_2 = d_1 - \sqrt{\mathcal{R}\sigma_1^2 + (\tau - \mathcal{R})\sigma_2^2}$$

4.4.1 Empirical data

The data utilized here for analysis is the daily closing price as reported in the Russell 2000, Facebook and Google indices for the period January 2, 2013 to January 21, 2022. The indices returns, X_t , are computed as in Equation (3.11.75). The options data utilized from the three markets is in two sets, that is, call options prices expiring in 25 and 258 days.

4.4.2 Descriptive statistics

The basic statistics of the three indices returns are presented in Table 4.24. The daily mean returns are positive and are reported as 0.0361%, 0.1045% and 0.0867% for the RUT, FB and GooG indices returns respectively. This is an indication that the investment realized positive returns. Also, the variances are 0.000184, 0.000450 and 0.000255 for the RUT, FB and GooG indices returns respectively. Further, RUT index returns has a negative skewness of -1.2381 while FB and GooG indices returns have a positive skewness of 0.4031 and 0.4190 respectively.

Table 4.24: Basic statistics for RUT, FB and GooG indices returns

Index	Obs	Mean	Var	Std dev.	Skew	Ex.Kurt	JB
RUT	2280	0.000361	0.000184	0.013571	-1.238128	13.110355	25291.3***
FB	2280	0.001045	0.000450	0.021220	0.403136	14.683610	29830.3***
GooG	2280	0.000867	0.000255	0.015979	0.419047	6.844511	9294.8***

Note: The asterisks *** indicate 1% significant level.

All the indices returns have an excess kurtosis greater than 3, the value for normal distribution. This implies that the indices returns distribution has a thicker tail compared to that of normal distribution and this tail thickness may be due to the temporal volatility fluctuations of the indices returns. It is argued by Jarque and Bera (1980) that a sample drawn from a normal distribution has a skewness and excess kurtosis of zero. It is thus clear that the data is not normally distributed as per the inferences from the descriptive statistics. This is further confirmed by the JB test statistic which is significant at 1% confidence interval. The correlation coefficients between the three indices returns are presented in Table 4.25, and clearly there exists a significant positive relationship between them.

Table 4.25: Correlation coefficients for RUT,FB and GooG

Index	RUT	FB	GooG
RUT	1.0000	0.4687	0.5763
FB	0.4687	1.0000	0.5809
GooG	0.5763	0.5809	1.0000

4.4.3 Empirical findings and discussion

The plots of the stock prices and returns of RUT, FB and GooG indices returns are presented in Figure 19. The common properties of financial data are evident from the three indices returns, however FB and GooG indices returns are described by a pronounced alternating short and long spikes compared to the RUT indices returns. The alternating short and long spikes imply existence of volatility clustering in the underlying asset returns. The maximum likelihood estimates of the RS model are computed according to Equation (3.11.53) and presented in Table 4.26. The parameters are different across the two regimes, for instance, the mean in low volatility regime is 0.0008, 0.0015 and 0.0013 for the RUT, FB and GooG indices returns respectively, which is higher than the corresponding mean in the high volatility regime. Clearly, the low volatility regime has positive returns whereas high volatility regime has negative returns as indicated by the values of μ_1 and μ_2 . This shows that the underlying asset earns higher returns in regime 1(low volatility regime) which is associated with low risk as compared to regime 2(high volatility regime). The probability of switching from regime 1 to regime 2 is estimated at 0.0773, 0.2277 and 0.1931 for RUT, FB and GooG indices respectively. These transition probabilities are very low and the implication is that, once in regime 1 the volatility process lingers before transiting to regime 2. The average duration of stay in regime i in a single run is computed using the expression $\frac{1}{(1-P_{ii})}$ for $i = 1, 2$ and it is reported that once the process is in regime 1, it stays there for approximately 12, 4 and 5 days for RUT, FB and GooG indices respectively before switching regime. However, the volatility process stays longer in high volatility regime than in low volatility regime approximated at 79, 19 and 24 days for RUT, FB and GooG respectively. Figure 20 to 22 shows the smoothed probabilities for the RUT, FB and GooG indices returns and clearly the volatility process spends more time in regime 2 than in regime 1, which confirms the earlier findings of this study.

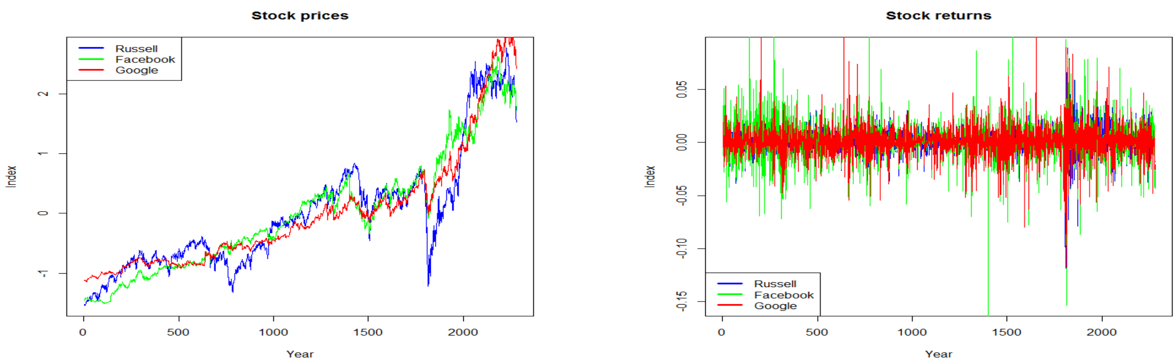

Figure 19: Stock prices and stock returns

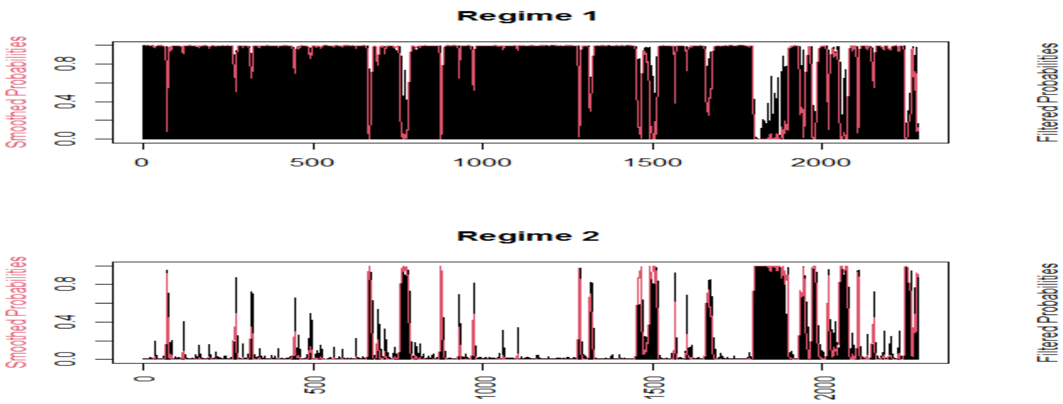

Figure 20: Smoothed probabilities for Russell 2000 index

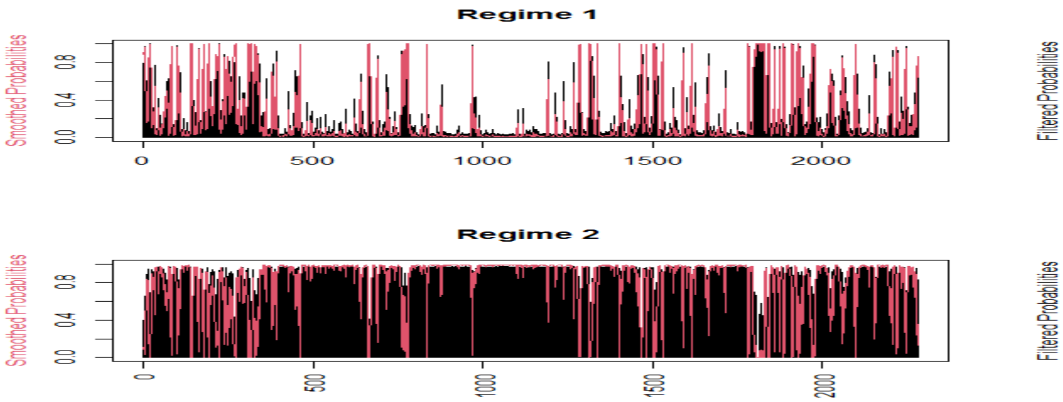

Figure 21: Smoothed probabilities for Facebook index

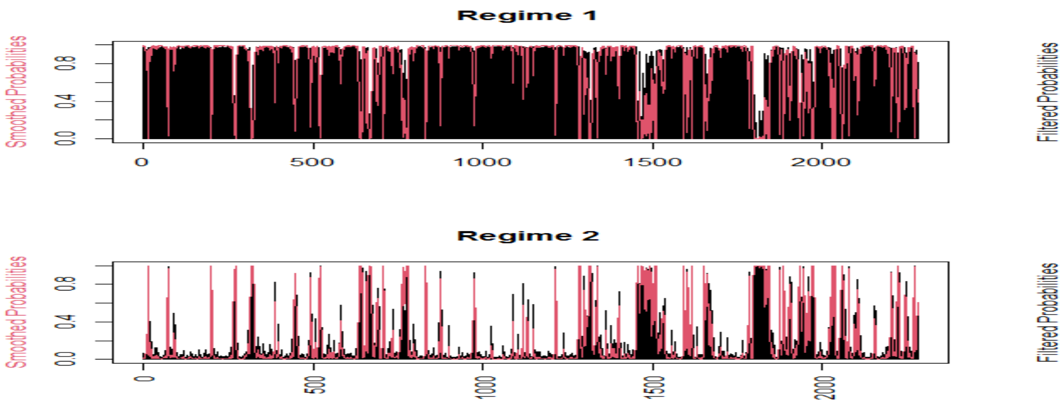

Figure 22: Smoothed probabilities for Google index

Table 4.26: Regime-switching(RS) model parameter estimates

Index	$\tilde{\mu_1}$	$\tilde{\mu_2}$	$\tilde{\sigma_1}$	$\tilde{\sigma_2}$	$\tilde{P_{12}}$	$\tilde{P_{21}}$
RUT	0.0008	−0.0002	0.0094	0.0273	0.0773	0.0126
FB	0.0015	−0.0009	0.0141	0.0394	0.2277	0.0515
GooG	0.0013	−0.0010	0.0105	0.0306	0.1931	0.0408

Moreover, the total time spent by the Markov chain in each regime are calculated from the transition probability transition estimates presented in Table 4.26. The following Markov chains for RUT, FB and GooG can be observed respectively follows;

$$R_{ij} = \begin{bmatrix} 0.9227 & 0.0773 \\ 0.0126 & 0.9874 \end{bmatrix}, F_{ij} = \begin{bmatrix} 0.7723 & 0.2277 \\ 0.0515 & 0.9485 \end{bmatrix}, \text{ and } G_{ij} = \begin{bmatrix} 0.8069 & 0.1931 \\ 0.0408 & 0.9485 \end{bmatrix}.$$

It can be observed that approximately 319.56, 483.81 and 397.71 days for RUT, FB and GooG are likely to be spent in regime one respectively, and also 1960.44, 1796.19 and 1882.29 days respectively, the chain is likely to spent in the second regime. Furthermore, the transition probabilities are plotted as smoothed probabilities and presented in Figures 20 to 22, and it can be inferred that the Markov chain spends more time in volatility regime(regime 2) than low volatility regime(regime 1). That is, the plots of smoothed probabilities are in agreement with the computations of the total time spent in each regime.

The RS-GARCH parameters are estimated by utilizing Equation (3.11.77) and reported in Table 4.27. The parameter estimates are significant at 5% significance level and indicate a heterogeneous volatility process across the regimes. In each regime, the unconditional variances are different and this is a confirmation of different volatility regimes for the underlying asset. The estimates of the conditional mean in regime 1, ω_1 , is lower than the corresponding estimates in regime 2, ω_2 , that is, $\omega_1 < \omega_2$ across all the indices returns. Moreover, the volatility dynamics are determined by the ARCH and GARCH parameters, i.e, α_{1i} and β_i for $i = 1, 2$ respectively. A large value of β_i indicates that the shock effects to future volatility take long to die off, that is, volatility is highly persistent whereas large values of α_{1i} display volatility reaction to the recent changes in price. Regime 1 has a low ARCH term and a high GARCH term and this means that the GARCH process is more reactive and less persistent in the low volatility regime than in high volatility regime. The persistence of volatility in each regime is calculated as $\alpha_{1i} + \beta_i$ for $i = 1, 2$ and for the process to be covariance stationary, it is required that $\alpha_{1i} + \beta_i < 1$. The calculated volatility persistence values for RUT, FB and GooG indices returns are reported as; $\alpha_{11} + \beta_1 = 0.9633$ versus $\alpha_{12} + \beta_2 = 0.9764$, $\alpha_{11} + \beta_1 = 0.9855$ versus $\alpha_{12} + \beta_2 = 0.2641$ and $\alpha_{11} + \beta_1 = 1.0986$ versus $\alpha_{12} + \beta_2 = 0.1789$ respectively. Inferences can be made that regime 1 for the FB and GooG has high volatility persistence than regime 2 whereas the process is explosive in the low volatility regime in the Google index. The volatility persistence is slightly higher in the high volatility regime than in low volatility regime for Russell 2000 index.

Table 4.27: RS-GARCH model parameter estimates

Index	ω_1	ω_2	α_{11}	α_{12}	β_1	β_2	P_{12}	P_{21}
RUT	7.4e-7	3.6e-5	0.0683	0.3086	0.8950	0.6678	0.3527	0.8414
FB	1.0e-7	0.0045	0.0370	0.2639	0.9485	0.0002	0.0319	0.6405
GooG	2.2e-6	0.0010	0.1401	0.1785	0.9585	0.0004	0.0398	0.3072

The estimated call option prices for some given strike prices for the RUT, FB and GooG indices are computed based on Equations (3.11.53) and (3.11.77) and presented in Table 4.28. The strike prices considered are for 25 and 258 days for the Russell 2000, Facebook and Google stock index markets and their initial stock prices S_0 are reported as 1987.92, 303.17 and 2601.84 respectively.

It is assumed that the markets are risk-free and that a risk-free interest rate of 6% per annum is utilized. It is noted that, for short-dated data the estimated option price values from the three models are dispersed and the dispersion is large in the RUT index as compared with FB and GooG indices. Again, the values are more or less similar for FB and GooG indices with the RS-GARCH model reporting the least values than BS and RS models except for the RUT index. Moreover, the estimated call option prices are significantly different from the actual market prices, however, in general the BS call option price estimates are closest to the actual market prices than the option prices estimates of the RS and RS-GARCH models. Therefore, it can be construed to mean that for short-dated data, the Black-Scholes model presents better results than the Regime switching and RS-GARCH models. On the other hand, it is revealed that for long-dated data, the estimated call option prices are more or less similar in the three models across the market indices. However, the call options price estimates from the RS-GARCH model are closer to the actual market prices than for BS and RS except for the RUT index. This means that RS-GARCH model presents better results than BS and RS models for long-dated data. Furthermore, Figure 23 and 24 presents the plots for strike prices versus the estimated call options (from the three models) and the observed option prices in the market indices.

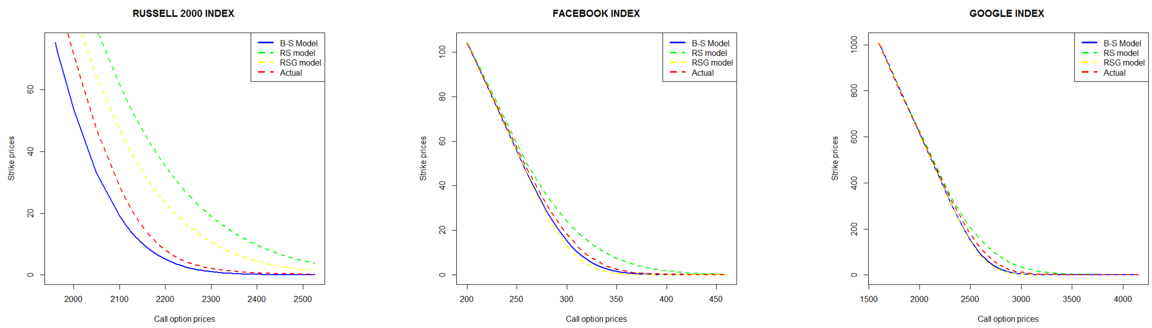

Figure 23: 25 days call option prices

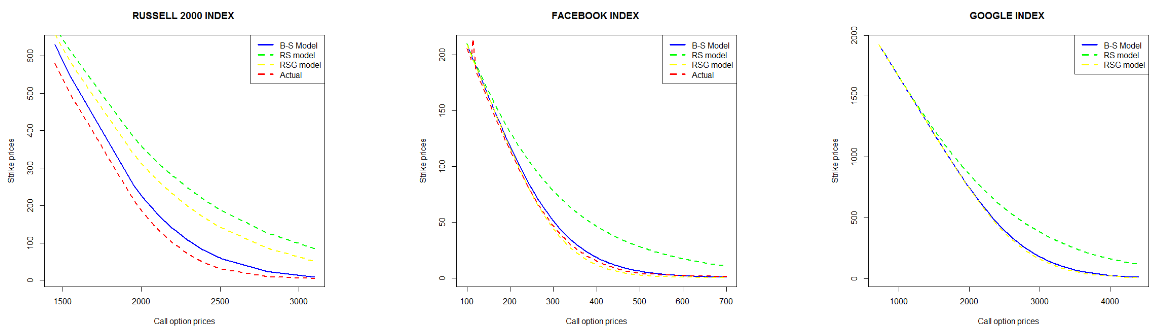

Figure 24: 258 days call option prices

Table 4.28: The Call Option prices

RUSSELL					FACEBOOK					GOOGLE				
		B-S	RS	RSG			B-S	RS	RSG			B-S	RS	RSG
Strike	Mkt	Call	Call	Call	Strike	Mkt	Call	Call	Call	Strike	Mkt	Call	Call	Call
25 days call option prices														
1960	94.2	75.36	121.47	106.22	200	103.83	104.36	104.50	104.36	1600	1005.35	1011.34	1011.35	1011.34
1965	91.2	72.40	118.82	103.50	260	47.55	45.57	49.95	30.79	1825	781.55	787.67	788.17	787.67
2000	71.4	53.70	101.37	85.74	280	31.55	28.50	35.65	21.39	2300	336.9	319.83	348.68	318.38
2050	47.35	33.12	79.71	64.23	285	27.85	24.78	32.52	31.34	2340	303.55	282.90	317.49	280.83
2100	28.7	19.06	61.69	46.99	290	4.58	18.19	26.81	16.00	2450	216.65	189.41	239.63	185.17
2110	25.7	16.93	58.50	44.02	295	21.13	18.19	26.81	16.00	2500	178.4	152.39	208.40	147.20
2115	24.25	15.93	56.95	42.59	300	18.23	15.35	24.23	13.03	2600	118.05	91.55	154.14	85.25
2120	22.9	14.99	55.44	41.20	305	16.18	12.82	21.84	10.45	2650	92.7	68.21	131.09	61.91
258 days call option prices														
1450	579.5	630.44	680.66	656.89	100	205.53	209.13	209.90	209.13	720	1892.15	1924.74	1924.98	1924.74
1500	537.5	586.32	645.00	618.56	105	200.85	204.43	205.42	204.42	740	1872.55	1905.93	1906.23	1905.93
1550	496	543.18	610.60	581.54	110	195.98	199.74	200.97	199.72	760	2045.55	1887.12	1887.49	1887.12
1950	214.5	254.11	382.18	337.00	115	186.40	195.04	196.57	195.02	780	2200.00	1868.31	1868.77	1868.31
1990	192.5	232.11	363.78	317.68	120	181.65	190.35	192.22	190.32	800	1793.95	1849.50	1850.07	1849.50
2000	187	226.82	359.30	312.99	125	176.9	185.67	187.92	185.62	820	1847.75	1830.70	1831.38	1830.70
2010	182	221.62	354.87	308.36	130	172.25	180.99	183.66	180.92	840	1842.75	1811.89	1812.71	1811.89
2090	143.5	183.04	321.07	273.26	135	167.45	176.33	179.47	176.23	860	2090.8	1793.08	1794.06	1793.08

The call options price estimate values reported by the three models seem to be dispersed especially with the RUT index for short-dated data. The plots of the RS-GARCH model appear to have similar estimates with that of the Black-Scholes model since they are closer to the actual price plot than that of the Regime switching(RS) model. However, the Black-Scholes(BS) model presents slightly better estimates than the RS and RS-GARCH model since the plot is closest to the actual market plot across the three market indices. As for the long-dated data plots for the estimated call option prices, the RS-GARCH model plot is slightly closer to the actual price plots, however, the plots for the RUT index are a bit dispersed.

A comparison of the models is done by employing the Root Mean Square Error(RMSE) test and the results are presented in Table 4.29. The results portray RS-GARCH model as a better model than the Black-Scholes and RS model in estimating the 258 days option prices for Facebook and Google indices. On the other hand, for the 25 days option prices estimation the Black-Scholes model gives better estimates followed by RS-GARCH and RS models in that order.

Table 4.29: RMSE for the Black-Scholes, Regime-switching and RS-GARCH models

	25 days options			258 days options		
Index	BS	RS	RSG	BS	RS	RSG
RUT	5.159	21.48	11.65	29.78	160.0	115.4
FB	1.685	4.175	2.939	4.081	24.55	3.433
GooG	14.07	29.96	17.59	286.0	305.6	285.3

Therefore, the general findings is that for short-dated data, the BS model presents better options price estimates than the RS and RS-GARCH models. On the other hand for the long-dated data, the RS-GARCH model is slightly better than the BS and RS models. Furthermore, it can be argued that when the Regime switching model is adjusted to include GARCH effects it gives a better pricing model since in both short and long dated data, the RS-GARCH model performs better than the RS model.

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

It is reported in the previous literature on financial time series modeling that stock returns, volatility and trading volume have a paramount relationship. The knowledge of this relationship gives information about the micro-structure of financial markets. It also aids in determining the rate at which information enters the market and its impact on stock returns. The empirical studies on the stock returns, volatility and trading volume relationships have reported blended results, for instance, whereas some studies found a positive association between stock returns and volatility, others found a negative correlation. In addition, most of these empirical studies have dealt much on developed stock markets as compared to the emerging stock markets and this leaves inadequate literature for the emerging stock markets. It is further noted that the random arrival of information into the market causes variations in stock prices which in turn leads to changes in trade volume. When trade volume is included in the conditional variance of the GARCH equation, it can be utilized as an exogenous variable to explain the effect on volatility persistence.

The goal of this study is to model the dynamic and contemporaneous linkage between volatility, stock returns, and trading volume of the developing and developed stock markets. Additionally, volatility dynamics of stock returns is explored since it is considered an essential factor in many economic and financial fields like price assets, risk management and portfolio allocation. Many financial markets experience drastic structural changes in stock prices and this leads to a changing volatility of the financial market and the change is also persistent. This pattern of behavior of volatility makes it not easy to be estimated, however, once estimated it may be utilized to determine future stock prices and in pricing options. To model this dynamics of the underlying financial assets and the structural changes in volatility, GARCH model and its extensions and regime switching model are utilized. Furthermore, the study investigates the impact of incorporating trading volume into the GARCH model's conditional variance equation on volatility persistence. Another issue addressed by this study is the European option asset pricing which has been widely dealt with by utilizing Black-Scholes model. A regime-switching (RS) model for the European option pricing is developed for an underlying asset whose dynamics depend on market regimes. The formulation of the model is founded on a geometric Brownian motion governed by a continuous-time Markov chain with two states. The model is further modified to include GARCH effects which results to a RS-GARCH model. These two models are implemented by utilizing stock market indices and call option data to estimate the European call prices. The estimated European call option prices of these two models are compared with the Black-Scholes' European call option price estimates in a bid to establish the validity of the European regime-switching option model.

The results reveals that the behavior of the three indices returns and volume have the common properties of many financial time series. A considerable level of excess kurtosis is exhibited and this can be related to the time-dependence in conditional variance and the asymmetric distribution of returns. The asymmetric behavior in the conditional variance can be attributed to leverage effects. These two properties are an indication of non-normality of the indices returns and volume series and existence of conditional heteroscedasticity. The non-normality of the data series is confirmed by the JB- statistic and the empirical density plots which are different from that of normal distribution as well as the quantile-quantile plots. In addition, there is strong evidence of the common stylized facts of financial time series, for instance, clustering of volatility, leptokurtic distribution, heavy tails and leverage effects. The distribution of indices returns from developed markets is skewed to the left whereas that of emerging market is skewed to the right.

The stock market dynamics and the contemporaneous relationship between stock returns and trading volume is further investigated by utilizing GARCH model and its extensions. The estimates of GARCH (1,1) model reveals that the weekly conditional returns volatility tends to revert quickly towards the mean compared to the daily conditional volatility, that is, the volatility persistence decreases as the data changes frequency from daily to weekly returns. In addition, the developed market is characterized by high volatility clustering and volatility persistence when compared to the emerging market indices. That is, large changes in volatility tend to be succeeded by large volatility changes of either sign or low volatility changes by low changes, and shocks to conditional variance takes long time to die off in developed markets than in emerging markets. The ARCH term, α_1 is large in the emerging market than in developed market which indicates that stock returns volatility in the emerging market is much affected by past volatility than by the related news from the previous period compared to the developed market.

The estimates of GARCH-M (1,1) model show that the indices returns and their corresponding volatility are positively related as indicated by the positive risk parameter. The positive risk parameter also implies that the mean of return depends largely on past innovations and conditional variance. Moreover, the developed markets has high volatility clustering and a high persistence of volatility than the emerging market. The volatility persistence decreases as the data changes frequency from daily to weekly.

The results of estimates of EGARCH (1,1) parameters reports no leverage effects in the return series but the returns are asymmetric which implies that the impact of negative news does not outweigh positive news. The ARCH coefficient value is negative and this is an indication that the volatility rises more after negative returns than after positive returns. It is observed that the volatility asymmetry is large in emerging market than in developed markets and this shows that volatility is increased by positive shocks than by negative shocks in emerging market than in developed markets. In addition the results reveal that the stock returns and the conditional volatility have a positive relation.

In order to model the market regime dynamics, the MS-GARCH (1,1) and MS-EGARCH(1,1) models are fitted to the data. The MSGARCH (1,1) results reports existence of regimes that are heterogeneous and this is implied by different parameter values across the regimes. The results indicate that all the daily and weekly indices returns across the three markets are characterized by low volatility clustering in the first regime as opposed to high volatility clustering in the second regime. In general, the two regimes are described by a heterogeneous conditional volatility, volatility persistence as well as varied reaction to the previous negative returns. Furthermore, it is noted that the volatility process stays in one regime for a long time before switching to the next regime as reported by high state probabilities. This also means that the volatility process reverts back from regime 2 to regime 1 at a higher rate than it enters the regime. The results of MS-EGARCH (1, 1) parameter estimate show that regime 1 for the daily and weekly returns is characterized by high conditional volatility, strong volatility reaction to past negative returns and low volatility persistence. Regime 2 on the other hand exhibits low conditional volatility, weak volatility reaction to past negative returns and high volatility persistence. The results reports leverage effect (this is not reported by EGARCH (1, 1) model results) in all indices returns and across the regimes. The leverage effect is an indicator of risk associated with trading in the stock markets in question. It is also revealed that the indices returns and conditional volatility are positively related in the two regimes. In addition, it is reported that the volatility process stays in one regime for a long time before switching to the next regime.

The change on volatility persistence of the three market indices after inclusion of lagged trade volume into the GARCH model is investigated. The volatility persistence reported by the GARCH (1,1) model before and after inclusion of log volume are compared. The general finding is that for the daily indices returns the volatility persistence decreased in both developed and developing markets. However, the normal and student-t distributions reports contrary findings for the FTSE100 and NSE20 indices returns respectively. It is noted that the volatility persistence decreased for the weekly

S&P500 index returns whereas it increased for the FTSE100 and NSE20 indices returns. Therefore it can be argued in general that inclusion of log trading volume on GARCH(1,1) model reports mixed results when data changes frequency from daily to weekly indices returns. Moreover, the volatility persistence slightly decreases on addition of volume into the GARCH-M model's conditional variance equation in daily indices returns for both developed and developing markets although the results reported by GED distribution are different. As for the weekly returns, the persistence slightly increases and arguably trade volume explains volatility in both developed and developing markets and that trade volume and indices returns are positively related. Also, inclusion of trading volume in the conditional variance equation of EGARCH model does not reduce the volatility persistence but instead it increases, however, the asymmetry of volatility is decreased for both daily and weekly indices returns across the markets. Also, the volume parameter δ is positive which implies that trade volume has no significant explanation of volatility and that trade volume and volatility are positively related. Further, it shows that bad news has more impact on conditional volatility than good news and this indeed means the market is characterized by volatility asymmetry. Thus information flow into the market is affected by trading volume in addition to trade volume being important in prediction of volatility dynamics across the two categories of markets.

The stock returns, volatility and trade volume dynamic relationship is investigated by implementing the Regime switching (RS) model on the developed and emerging stock market indices data. The results reveal that regime 1 of the developed market indices is described by high volatility and negative returns whereas that of emerging market by high volatility and positive returns. A negative association between stock returns and log volume is revealed in regime 1 of the developed market and stock returns and log volume are observed to be positively correlated in regime 1 in the emerging market. Regime 2 of both developed and developing markets reports low volatility and positive returns for both daily and weekly indices returns. Moreover, stock returns and log trade volume are negatively related for all market indices except for the daily NSE20 indices returns. It is also observed the volatility process lingers in regime 1 before crossing to regime 2 than it takes to fall back once it has entered regime 2.

The relationship between the indices returns and the corresponding trade volume is also carried out using the Pearson correlation coefficient test. The results reveal that stock returns and trading volume from the developed markets are negatively related whereas for emerging market the relationship is positive. This finding supports the results of the Regime Switching(RS) model presented in our earlier discussion which reported that stock returns and trading volume have a negative relationship in the developed market indices in both high and low volatility regimes and a positive relationship in the emerging market index.

This study also derived a regime-switching model for pricing European call options which allows the volatility process to follow a regime-switching process that is governed by a Markov chain process. The model is further modified to incorporate GARCH effects in the regimes so as to price the European call options. The implementation of the models is carried out in pricing the European options derived from the Russell 2000, Facebook and Google market indices for both short and long dated options. A model comparison is done by calculating the Root Mean Square Error (RMSE) for each model and the model with the least RMSE is the best model for pricing the European call options. The results indicate that the financial time series for the three markets exhibit the common stylized facts of financial data such as volatility clustering, heavy tails among others. The parameter estimates of the models indicate that the market indices have distinct regimes. In this case two regimes are used, high and low volatility. It is clear that the low volatility regime has high volatility persistence for Facebook and Google markets, whereas, the volatility process is explosive in low volatility regime for Google market index. The results show that RS-GARCH is the best model compared with Black-Scholes and RS models when applied to long-dated options contract. However, when short-dated options contract are used, Black-Scholes model out performs the RS and RS-GARCH models.

5.2 Recommendations

The following recommendations are made based on our discussions and findings.

- The dynamics of the underlying asset was carried out using regime switching models with two regimes. Moreover, the dynamic relationship between stock returns, volatility and trading volume was carried out using a regime switching model with two regimes. We recommend modeling the stock market dynamics and this relationship using a regime switching model with more than two regimes.

- A regime switching model for pricing options with two regimes was derived and implemented, hence we recommend a case for more than two regimes.

- A RS-GARCH model was found to outperform both regime switching and the Black-Scholes models in pricing long-dated data. We recommend more research using different data sets in order to support the findings in this study.

REFERENCES

Abbondante, P. et al. (2010). Trading volume and stock indices: A test of technical analysis. *American Journal of Economics and Business Administration*, 2(3):287–292.

Ahmed, H. J. A., Hassan, A., and Nasir, A. M. (2005). The relationship between trading volume, volatility and stock market returns: A test of mixed distribution hypothesis for a pre and post crisis on kuala lumpur stock exchange. *Investment Management and Financial Innovations*, 3(3):146–158.

Al-Jafari, M. K. and Tliti, A. (2013). An empirical investigation of the relationship between stock return and trading volume: Evidence from the jordanian banking sector. *Journal of Applied Finance and Banking*, 3(3):45.

Al Samman, H. and Al-Jafari, M. K. (2015). Trading volume and stock returns volatility: Evidence from industrial firms of oman. *Asian Social Science*, 11(24):139.

Ali, A. H., Hassan, A., and Nasir, A. M. (2005). The relationship between trading volume, volatility and stock market returns: a test of mixed distribution hypothesis for a pre-and post crisis on kuala lumpur stock exchange. *Investment Management and Financial Innovations*, 2(3):146–158.

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*, 91(4).

Attari, M. I. J., Rafiq, S., and Awan, H. M. (2012). The dynamic relationship between stock volatility and trading volume. *Asian Economic and Financial Review*, 2(8):1085–1097.

Barone-Adesi, G., Engle, R. F., and Mancini, L. (2008). A garch option pricing model with filtered historical simulation. *The review of financial studies*, 21(3):1223–1258.

Bauwens, L., Preminger, A., and Rombouts, J. V. (2006). Regime switching garch models. *Available at SSRN 914144*.

Biswas, A., Goswami, A., and Overbeck, L. (2018). Option pricing in a regime switching stochastic volatility model. *Statistics & Probability Letters*, 138:116–126.

Black, F. (1976). Studies of stock market volatility changes,‘proceedings of the american statistical association. business and economic statistics. pages 177–181.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. volume 81, pages 631–654.

Bollen, N. P., Gray, S. F., and Whaley, R. E. (2000). Regime switching in foreign exchange rates:: Evidence from currency option prices. *Journal of Econometrics*, 94(1-2):239–276.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.

Bollerslev, T., Chou, R. Y., and Kroner, K. F. (1992). Arch modeling in finance: A review of the theory and empirical evidence. *Journal of econometrics*, 52(1-2):5–59.

Brailsford, T. J. (1996). The empirical relationship between trading volume, returns and volatility. *Accounting & Finance*, 36(1):89–111.

Brooks, C. et al. (2008). Rats handbook to accompany introductory econometrics for finance. *Cambridge Books*.

Chen, G.-m., Firth, M., and Rui, O. M. (2001). The dynamic relation between stock returns, trading volume, and volatility. *Financial Review*, 36(3):153–174.

Christie, A. A. (1982). The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of financial Economics*, 10(4):407–432.

Christoffersen, P., Heston, S., and Jacobs, K. (2006). Option valuation with conditional skewness. *Journal of Econometrics*, 131(1-2):253–284.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative finance*, 1(2):223.

Crouch, R. L. (1970). The volume of transactions and price changes on the new york stock exchange. *Financial Analysts Journal*, 26(4):104–109.

Dash, M., Dagha, J. H., Sharma, P., and Singhal, R. (2012). An application of garch models in detecting systematic bias in options pricing and determining arbitrage in options. *Journal of CENTRUM Cathedra: The Business and Economics Research Journal*, 5(1):91–101.

Deelstra, G., Latouche, G., and Simon, M. (2020). On barrier option pricing by erlangization in a regime-switching model with jumps. *Journal of Computational and Applied Mathematics*, 371:112606.

Duan, J.-C. (1995). The garch option pricing model. *Mathematical finance*, 5(1):13–32.

Duan, J.-C., Popova, I., and Ritchken, P. (2002). Option pricing under regime switching. *Quantitative Finance*, 2(2):116.

Emenike, K. O. and Opara, C. C. (2014). the relationship between stock returns volatility and trading volume in nigeria. *Verslo Sistemos ir Ekonomika*, 4(2).

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.

Engle, R. F., Lilien, D. M., and Robins, R. P. (1987). Estimating time varying risk premia in the term structure: The arch-m model. *Econometrica: journal of the Econometric Society*, pages 391–407.

Frijns, B., Tallau, C., and Tourani-Rad, A. (2010). The information content of implied volatility: evidence from australia. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 30(2):134–155.

Gallant, A. R., Rossi, P. E., and Tauchen, G. (1992). Stock prices and volume. *The Review of Financial Studies*, 5(2):199–242.

Girard, E. and Biswas, R. (2007). Trading volume and market volatility: Developed versus emerging stock markets. *Financial Review*, 42(3):429–459.

Glascock, J. L. and Hsieh, H.-C. S. (2014). The causal relationships between stock returns, trading volume, and volatility. *International Journal of Managerial Finance*.

Godin, F. and Trottier, D.-A. (2021). Option pricing in regime-switching frameworks with the extended girsanov principle. *Insurance: Mathematics and Economics*.

Godin, F., Trottier, D.-A., et al. (2019). Option pricing under regime-switching models: Novel approaches removing path-dependence. *Insurance: Mathematics and Economics*, 87:130–142.

Gulia, S. (2016). Testing of relationship between trading volume, return and volatility. *Amity Global Business Review, i*, pages 96–103.

Haas, M., Mittnik, S., and Paolella, M. S. (2004). 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*, pages 357–384.

Hamilton, J. D. (1990). Analysis of time series subject to changes in regime. *Journal of econometrics*, 45(1-2):39–70.

Hamilton, J. D. and Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of econometrics*, 64(1-2):307–333.

Hao, F. and Yang, H. (2011). Coherent risk measure for derivatives under black-scholes economy with regime switching. *Managerial Finance*.

Hardy, M. R. (2001). A regime-switching model of long-term stock returns. *North American Actuarial Journal*, 5(2):41–53.

Harris, L. (1987). Transaction data tests of the mixture of distributions hypothesis. *Journal of Financial and Quantitative Analysis*, 22(2):127–141.

He, H. and Wang, J. (1995). Differential information and dynamic behavior of stock trading volume. *The Review of Financial Studies*, 8(4):919–972.

Henry, O. (1998). Modelling the asymmetry of stock market volatility. *Applied Financial Economics*, 8(2):145–153.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, 6(2):327–343.

Heston, S. L. and Nandi, S. (2000). A closed-form garch option valuation model. *The review of financial studies*, 13(3):585–625.

Huang, B.-N. and Yang, C.-W. (2001). An empirical investigation of trading volume and return volatility of the taiwan stock market. *Global Finance Journal*, 12(1):55–77.

Jarque, C. M. and Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics letters*, 6(3):255–259.

Jiranyakul, K. (2016). Dynamic relationship between stock return, trading volume, and volatility in the stock exchange of thailand: does the us subprime crisis matter? *Trading Volume, and Volatility in the Stock Exchange of Thailand: Does the US Subprime Crisis Matter*.

Karpoff, J. M. (1987). The relation between price changes and trading volume: A survey. *Journal of Financial and quantitative Analysis*, pages 109–126.

Kearns, P. and Pagan, A. R. (1993). Australian stock market volatility: 1875–1987. *Economic Record*, 69(2):163–178.

Khan, A. and Ahmed, M. S. (2009). Trading volume and stock return: The impact of events in pakistan on kse 100 indexes. *International Review of Business Research Papers*, 5(5):373–383.

Kirkby, J. L. and Nguyen, D. (2020). Efficient asian option pricing under regime switching jump diffusions and stochastic volatility models. *Annals of Finance*, 16(3):307–351.

Kroner, K. F. and Ng, V. K. (1998). Modeling asymmetric comovements of asset returns. *The review of financial studies*, 11(4):817–844.

Lamoureux, C. G. and Lastrapes, W. D. (1990). Persistence in variance, structural change, and the garch model. *Journal of Business & Economic Statistics*, 8(2):225–234.

Lee, B.-S. and Rui, O. M. (2002). The dynamic relationship between stock returns and trading volume: Domestic and cross-country evidence. *Journal of Banking & Finance*, 26(1):51–78.

Lee, C. F. and Rui, O. M. (2000). Does trading volume contain information to predict stock returns? evidence from china's stock markets. *Review of quantitative finance and accounting*, 14(4):341–360.

Li, M.-Y. L. and Lin, H.-W. W. (2003). Examining the volatility of taiwan stock index returns via a three-volatility-regime markov-switching arch model. *Review of Quantitative Finance and Accounting*, 21(2):123–139.

Lin, S. and He, X.-J. (2020). A regime switching fractional black–scholes model and european option pricing. *Communications in Nonlinear Science and Numerical Simulation*, 85:105222.

Mahajan, S. and Singh, B. (2009). The empirical investigation of relationship between return, volume and volatility dynamics in indian stock market. *Eurasian Journal of Business and Economics*, 2(4):113–137.

Mikosch, T. and Stărică, C. (2004). Nonstationarities in financial time series, the long-range dependence, and the igarch effects. *Review of Economics and Statistics*, 86(1):378–390.

Miyakoshi, T. (2002). Arch versus information-based variances: evidence from the tokyo stock market. *Japan and the World Economy*, 14(2):215–231.

Moyo, E., Waititu, A. G., and Ngunyi, A. (2018). Modelling the effects of trading volume on stock return volatility using conditional heteroskedastic models. *Journal of Finance and Economics*, 6(5):193–200.

Mubarik, F. and Javid, A. Y. (2009). Relationship between stock return, trading volume and volatility: Evidence from pakistani stock market. *Asia Pacific Journal of Finance and Banking Research*, 3(3).

Naik, P. K., Gupta, R., and Padhi, P. (2018). The relationship between stock market volatility and trading volume: Evidence from south africa. *The Journal of Developing Areas*, 52(1):99–114.

Najand, M. and Yung, K. (1991). A garch examination of the relationship between volume and price variability in futures markets. *The Journal of Futures Markets*, 11(5):613.

Nasri, B. R., Remillard, B. N., and Thioub, M. Y. (2020). Goodness-of-fit for regime-switching copula models with application to option pricing. *Canadian Journal of Statistics*, 48(1):79–96.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pages 347–370.

Omran, M. and McKenzie, E. (2000). Heteroscedasticity in stock returns data revisited: volume versus garch effects. *Applied Financial Economics*, 10(5):553–560.

Patterson, K. D. (2000). *An introduction to applied econometrics: a time series approach*. Macmillan.

Poudel, R. B. and Shrestha, S. R. (2019). Stock return and trading volume relation in nepalese stock market: An ardl approach. *SEBON JOURNAL*.

Reyes, M. G. (2001). Asymmetric volatility spillover in the tokyo stock exchange. *Journal of Economics and Finance*, 25(2):206–213.

Rogalski, R. J. (1978). The dependence of prices and volume. *The Review of Economics and Statistics*, pages 268–274.

Salman, F. (2002). Risk-return-volume relationship in an emerging stock market. *Applied Economics Letters*, 9(8):549–552.

Schwert, G. W. (1989). Why does stock market volatility change over time? *The journal of finance*, 44(5):1115–1153.

Sentana, E. (1995). Quadratic arch models. *The Review of Economic Studies*, 62(4):639–661.

Siu, T. K., Tong, H., Yang, H., et al. (2004). On pricing derivatives under garch models: a dynamic gerber-shiu approach. *North American Actuarial Journal*, 8:17–31.

Wiley, M. K. and Daigler, R. T. (1999). A bivariate garch approach to the futures volume-volatility issue. In *Article Presented at the Eastern Finance Association Meetings*.

Yüksel, A. (2002). The performance of the istanbul stock exchange during the russian crisis. *Emerging Markets Finance & Trade*, pages 78–99.

Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and control*, 18(5):931–955.

Appendix A: List of published/Unpublished papers as part of this study

In this study, parts of the research work and results reported in this thesis have been published as peer-reviewed journals and conference proceedings as listed below.

List of published papers

- Kalovwe, S.K., Mwaniki, J.I., Simwa, R.O. (2022). **On stock market asymmetric volatility and trading volume.** *Far East Journal of Theoretical Statistics*, 66 , 89-104
- Kalovwe, S. K., Mwaniki, J. I., Simwa, R. O. (2022). **European option pricing under the regime-switching GARCH model.** *Financial Mathematics and Applications*, 7(1), 1-12
- Kalovwe, S. K., Mwaniki, J. I., Simwa, R. O. (2021). **Modeling returns and trading volume in Regime switching world.** *Financial Mathematics and Applications*, 6(1),1-14
- Kalovwe, S. K., Mwaniki, J. I., Simwa, R. O. (2021). **On Stock Returns Volatility and Trading Volume of the Nairobi Securities Exchange Index**. RMS: *Research in Mathematics Statistics*, 8(1), 1889765.

Paper presented at conference and published online

- 11th -12th October 2022: KCA University: Presented a research paper titled, ?European option pricing based on regime switching models? during the first interdisciplinary conference on innovations and sustainability.
- 24th -26th April 2019: Machakos University: Presented a research paper titled, ?modeling stock returns volatility using Regime switching models? during the 2nd Annual International Conference.

List of papers submitted for publication

- On Regime-Switching European Option Pricing − paper submitted to the Journal of Probability and statistics on 16^{Th} February, 2022 and is under review stage.

Appendix B: Some R codes used in data analysis

R codes for RUSSELL 2000 index 25 days options

BS-MODEL

```
BS_RUT<-function(r,so,k,days){
 sigma<- 0.01357104

  #mu<-mean(da)
  k<-opt_rut25[,3]
  R<-r*0.01/252
  d1<-(1/(sigma*sqrt(days)))*(log(so/k)+(R
  d2<-d1-sigma*sqrt(days)
  Ecal<-so*pnorm(d1)-exp(-R*days)*k*pn
  Eput<--so*pnorm(-d1)+exp(-R*days)*k*
  Edelta<-pnorm(d1)
  price<-cbind(days,r,so,k,Ecal,Eput,Edelta
  return(price)
 }
```

RS model

```
RS_RUT<-function(r1,so,k,days){
  sigma1<- 0.009365936
  sigma2<- 0.027349583
  #mu<-mean(da)
  k<-opt_rut25[,3]
  n<-25
  p21<- 0.01255443
  p12<- 0.07732007
  p1<-p21/(p12+p21)
  R<-n*p1
  r<-r1*0.01/252
  A1<-sqrt(R*(sigma1^2)+(n-R)*(sigma2^
  B1<-R*0.5*(sigma1^2)+(n-R)*0.5*(sigm
  d1<-(1/(A1))*(log(so/k)+(n*r+B1))
  d2<-d1-A1
  Ecal<-so*pnorm(d1)-exp(-r*days)*k*pnc
  Eput<--so*pnorm(-d1)+exp(-r*days)*k*p
  Edelta<-pnorm(d1)
  price<-cbind(days,r,so,k,Ecal,Eput,Edelta
  return(price)
}
```

RSG model

```
 RSG_RUT<-function(r1,so,k,days){
sigma1<- 0.004499527
  sigma2<- 0.03924072
  #mu<-mean(da)
  k<- opt_rut25[,3]
  n<-25
  p21<- 0.8414
  p12<- 0.3527
  p1<-p21/(p12+p21)
  R<-n*p1
  r<-r1*0.01/252
  A1<-sqrt(R*(sigma1^2)+(n-R)*(sig
ma2^2))
  B1<-R*0.5*(sigma1^2)+(n-R)*0.5*(
sigma2^2)
  d1<-(1/(A1))*(log(so/k)+(n*r+B1))
  d2<-d1-A1
  Ecal<-so*pnorm(d1)-exp(-r*days)*k
*pnorm(d2)
  Eput<--so*pnorm(-d1)+exp(-r*days)
*k*pnorm(-d2)
  Edelta<-pnorm(d1)
  price<-cbind(days,r,so,k,Ecal,Eput,E
delta)
  return(price)
 }
```

R codes for Facebook index 25 days options

BS-MODEL

```
BSFB<-function(r,so,k,days){
sigma<- 0.02122042
  #mu<-mean(da)
  k<-opt_fb25[,3]
  R<-r*0.01/252
  d1<-(1/(sigma*sqrt(days)))*(log(s
  d2<-d1-sigma*sqrt(days)
  Ecal<-so*pnorm(d1)-exp(-R*days)
  Eput<--so*pnorm(-d1)+exp(-R*da
  Edelta<-pnorm(d1)
  price<-cbind(days,r,so,k,Ecal,Eput
  return(price)
 }
```

RS model

```
RSFB<-function(r1,so,k,days){
  sigma1<- 0.01409633
  sigma2<- 0.03948836
  #mu<-mean(da)
  k<- opt_fb25[,3]
  n<-25
  p21<- 0.0514713
  p12<- 0.227733
  p1<-p21/(p12+p21)
  R<-n*p1
  r<-r1*0.01/252
  A1<-sqrt(R*(sigma1^2)+(n-R)*(sig
  B1<-R*0.5*(sigma1^2)+(n-R)*0.5
  d1<-(1/(A1))*(log(so/k)+(n*r+B1)
  d2<-d1-A1
  Ecal<-so*pnorm(d1)-exp(-r*days)
  Eput<--so*pnorm(-d1)+exp(-r*da
  Edelta<-pnorm(d1)
  price<-cbind(days,r,so,k,Ecal,Eput
  return(price)
}
```

RSG Model

```
RSGFB<-function(r1,so,k,da
ys){ sigma1<-0.002623124
  sigma2<- 0.07852753
  #mu<-mean(da)
  k<-opt_fb25[,3], n<-25, p
21<- 0.6405, p12<- 0.0
319, p1<-p21/(p12+p21)
  R<-n*p1, r<-r1*0.01/252
A1<-sqrt(R*(sigma1^2)+(n-R)
*(sigma2^2)) B1<-R*0.5*(sig
ma1^2)+(n-R)*0.5*(sigma2^2)
  d1<-(1/(A1))*(log(so/k)+(n*r
+B1)) d2<-d1-A1
  Ecal<-so*pnorm(d1)-exp(-r*d
ays)*k*pnorm(d2)
  Eput<--so*pnorm(-d1)+exp(-r
*days)*k*pnorm(-d2)
  Edelta<-pnorm(d1)
  price<-cbind(days,r,so,k,Ecal,
Eput,Edelta)
  return(price)
 }
```

R codes for Google index 25 days options

BS-MODEL

```
BSGG<-function(r,so,k,days){
sigma<- 0.01597882

  #mu<-mean(da)
  k<-opt_goog25[,3]
  R<-r*0.01/252
  d1<-(1/(sigma*sqrt(days)))*(log(s
  d2<-d1-sigma*sqrt(days)
  Ecal<-so*pnorm(d1)-exp(-R*days)
  Eput<--so*pnorm(-d1)+exp(-R*da
  Edelta<-pnorm(d1)
  price<-cbind(days,r,so,k,Ecal,Eput
  return(price)
 }
```

RS model

```
RSGG<-function(r1,so,k,days){
  sigma1<- 0.01046077
  sigma2<- 0.03057581
  #mu<-mean(da)
  k<- opt_goog25[,3]
  n<-25
  p21<- 0.04083742
  p12<- 0.193102
  p1<-p21/(p12+p21)
  R<-n*p1
  r<-r1*0.01/252
  A1<-sqrt(R*(sigma1^2)+(n-R)*(sig
  B1<-R*0.5*(sigma1^2)+(n-R)*0.5
  d1<-(1/(A1))*(log(so/k)+(n*r+B1)
  d2<-d1-A1
  Ecal<-so*pnorm(d1)-exp(-r*days)
  Eput<--so*pnorm(-d1)+exp(-r*da
  Edelta<-pnorm(d1)
  price<-cbind(days,r,so,k,Ecal,Eput
  return(price)
 }
```

RSG Model

```
RSGGG<-function(r1,so,k,days)
{ sigma1<- 0.008995463
  sigma2<- 0.03568906
  #mu<-mean(da), k<- opt_g
oog25[,3], n<-25, p21<- 0
.3072, p12<- 0.0398
  p1<-p21/(p12+p21)
  R<-n*p1, r<-r1*0.01/252
  A1<-sqrt(R*(sigma1^2)+(n-R)
*(sigma2^2)), B1<-R*0.5*(sig
ma1^2)+(n-R)*0.5*(sigma2^2)
  d1<-(1/(A1))*(log(so/k)+(n*r
+B1))
  d2<-d1-A1
  Ecal<-so*pnorm(d1)-exp(-r*d
ays)*k*pnorm(d2)
  Eput<--so*pnorm(-d1)+exp(-r
*days)*k*pnorm(-d2)
  Edelta<-pnorm(d1)
  price<-cbind(days,r,so,k,Ecal,
Eput,Edelta)
  return(price)
 }
```

R codes for Density/Q-Q plots

DENSITY PLOTS

```
>  curve(dnorm(x,mean=mean(fd),sd=sd(fd)),add=T,lwd=3,lty=c(4),col=c(2))
> par(mfrow=c(2, 3))
> plot(density(fd),main="Emp Density vs Normal",xlab="FTSE100 daily returns")
> curve(dnorm(x,mean=mean(fd),sd=sd(fd)),add=T,lwd=3,lty=c(4),col=c(2))
> par(mfrow=c(2, 3))
> plot(density(fd),main="Emp Density vs Normal",xlab="FTSE100 daily returns")
> curve(dnorm(x,mean=mean(fd),sd=sd(fd)),add=T,lwd=3,type="l",col=c(2))
> legend("topright",legend=c('Emp','NM'),col=c(1,3),lwd=c(3,3),tit)
> legend("topright",legend=c('Emp','NM'),col=c(1,2),lwd=c(3,3),tit)

> plot(density(sd),main="Emp Density vs Normal",xlab="S&P500 daily returns",col="1")
> curve(dnorm(x,mean=mean(sd),sd=sd(sd)),add=T,lwd=3,type="l",col=c(2))
> legend("topright",legend=c('Emp','NM'),col=c(1,2),lwd=c(3,3),tit)

> plot(density(nd),main="Emp Density vs Normal",xlab="NSE20 daily returns",col="1")
> curve(dnorm(x,mean=mean(nd),sd=sd(nd)),add=T,lwd=3,type="l",col=c(2))
> legend("topright",legend=c('Emp','NM'),col=c(1,2),lwd=c(3,3),tit)

> qqnorm(fd,xlab="FTSE100 daily returns")
> qqline(fd,col="red")

> qqnorm(sd,xlab="S&P500 daily returns")
> qqline(sd,col="red")
> qqnorm(nd,xlab="NSE20 daily returns")
> qqline(nd,col="red")
```

DENSITY PLOTS

```
> par(mfrow=c(2, 3))
> plot(density(fd),main="Emp Density vs Normal",xlab="FTSE100 daily returns")
> curve(dnorm(x,mean=mean(fd),sd=sd(fd)),add=T,lwd=3,type="l",col=c(2))
> legend("topright",legend=c('Emp','NM'),col=c(1,3),lwd=c(3,3),tit)
> legend("topright",legend=c('Emp','NM'),col=c(1,2),lwd=c(3,3),tit)
> plot(density(sd),main="Emp Density vs Normal",xlab="S&P500 daily returns",col="1")
> curve(dnorm(x,mean=mean(sd),sd=sd(sd)),add=T,lwd=3,type="l",col=c(2))
> legend("topright",legend=c('Emp','NM'),col=c(1,2),lwd=c(3,3),tit)
> plot(density(nd),main="Emp Density vs Normal",xlab="NSE20 daily returns",col="1")
> curve(dnorm(x,mean=mean(nd),sd=sd(nd)),add=T,lwd=3,type="l",col=c(2))
> legend("topright",legend=c('Emp','NM'),col=c(1,2),lwd=c(3,3),tit)
> qqnorm(fd,xlab="FTSE100 daily returns")
> qqline(fd,col="red")
> qqnorm(sd,xlab="S&P500 daily returns")
> qqline(sd,col="red")
> qqnorm(nd,xlab="NSE20 daily returns")
> qqline(nd,col="red")
> par(mfrow=c(2, 3))
> plot(density(fw),main="Emp Density vs Normal",xlab="FTSE100 daily returns")
> curve(dnorm(x,mean=mean(fw),sd=sd(fw)),add=T,lwd=3,type="l",col=c(2))
> par(mfrow=c(2, 3))
> plot(density(fw),main="Emp Density vs Normal",xlab="FTSE100 weekly returns")
> curve(dnorm(x,mean=mean(fw),sd=sd(fw)),add=T,lwd=3,type="l",col=c(2))
> legend("topright",legend=c('Emp','NM'),col=c(1,3),lwd=c(3,3),tit)
> legend("topleft",legend=c('Emp','NM'),col=c(1,3),lwd=c(3,3),tit)
```

DENSITY PLOTS

```
> par(mfrow=c(2, 3))
> plot(density(fw),main="Emp Density vs Normal",xlab="FTSE100 weekly returns")
> curve(dnorm(x,mean=mean(fw),sd=sd(fw)),add=T,lwd=3,type="l",col=c(2))
> legend("topleft",legend=c('Emp','NM'),col=c(1,3),lwd=c(3,3),tit)
> plot(density(sw),main="Emp Density vs Normal",xlab="S&P500 weekly returns",col="1")
> curve(dnorm(x,mean=mean(sw),sd=sd(sw)),add=T,lwd=3,type="l",col=c(2))
> legend("topleft",legend=c('Emp','NM'),col=c(1,3),lwd=c(3,3),tit)
> plot(density(nw),main="Emp Density vs Normal",xlab="NSE20 weekly returns",col="1")
> curve(dnorm(x,mean=mean(nw),sd=sd(nw)),add=T,lwd=3,type="l",col=c(2))
> legend("topleft",legend=c('Emp','NM'),col=c(1,3),lwd=c(3,3),tit)
> par(mfrow=c(2, 3))
> plot(density(fw),main="Emp Density vs Normal",xlab="FTSE100 weekly returns")
> curve(dnorm(x,mean=mean(fw),sd=sd(fw)),add=T,lwd=3,type="l",col=c(2))
> legend("topleft",legend=c('Emp','NM'),col=c(1,2),lwd=c(3,3),tit)
> plot(density(sw),main="Emp Density vs Normal",xlab="S&P500 weekly returns",col="1")
> curve(dnorm(x,mean=mean(sw),sd=sd(sw)),add=T,lwd=3,type="l",col=c(2))
> legend("topleft",legend=c('Emp','NM'),col=c(1,2),lwd=c(3,3),tit)
> plot(density(nw),main="Emp Density vs Normal",xlab="NSE20 weekly returns",col="1")
> curve(dnorm(x,mean=mean(nw),sd=sd(nw)),add=T,lwd=3,type="l",col=c(2))
> legend("topleft",legend=c('Emp','NM'),col=c(1,2),lwd=c(3,3),tit)
> qqnorm(fw,xlab="FTSE100 weekly returns")
> qqline(fw,col="red")
> qqnorm(sw,xlab="S&P500 weekly returns")
> qqline(sw,col="red")
> qqnorm(nw,xlab="NSE20 weekly returns")
> qqline(nw,col="red")
```