

**Discrete-Time Markov Chains for a Multivariate Stochastic Autoregressive
Volatility**

NGASSIMA Fanny Laure

(MF300-0004/16)

**A thesis submitted to Pan African University Institute for Basic Sciences,
Technology and Innovation in partial fulfilment of the Degree of Master of
Science in Mathematics (Financial Option)**

2018

Declaration

I, Ngassima Fanny Laure (Reg. No: MF300-0004/16) hereby declare that this thesis entitled “Discrete-Time Markov Chains for a Multivariate Stochastic Autoregressive Volatility” is my original work and has not been presented for a degree in any other University. I also confirm that:

This work was done wholly or mainly while in candidature for a research degree at the noble University.

Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

I have acknowledged all main sources of help.

Signature:.....

Date:.....

Ngassima Fanny Laure

This thesis has been submitted for examination with my approval as University Supervisor.

1. Signature.....

Date.....

Dr. Joseph K. Mung’atu

Jkuat University of Nairobi, Kenya

2. Signature.....

Date.....

Dr. Mbele Bidima Martin Le Doux

University of Yaoundé I, Cameroon

Dedication

To my mum Yuego Micheline, Dad Pierre Flambeau Ngayap, Loving Brother and Sisters Ngayap William, Nyieng Edine, Djiloh Honorine, Djom Angele, my soulmate Tsague Kitio Doniale and darling Cedric Raffick MBA EYIMI.

Acknowledgement

In the journey of an academician, there is need for commitment, hard work, endurance, focus, patience and team work in order to reach the destination SUCCESS. These however, need to be coupled with the works of our giants on whom we base to attain the ultimate goal. At this moment in time, I would like to humbly acknowledge the great works of my giants that have made me reach this point in my academic journey.

I give thanks to God the provider for giving me the gift of life, the will, the patience and making this work a success.

I am highly indebted to the African Union for offering me a scholarship to study M.Sc. Financial Mathematics at the Pan African University, Institute for Basic Sciences, Technology and Innovation without which my dream would not have been realized.

I acknowledge the important role of my supervisors Dr. Joseph Mung'atu and Dr. Mbele Bidima for their guidance and motivation in the struggle of researching in this field. I also extend my sincere gratitude to the administrative and teaching staff of PAUISTI and JKUAT for their commitment and support offered to us during the whole study period. I specially give my respectful thanks to Prof. Gabriel Magoma, PAUISTI Director, and our Coordinator Prof. George Orwa for their encouragement.

I wish to thank my classmates and friends; Ayo Eunice, Adesokan Ibrahim, Ibrahim Sawadogo, Stephane Kentim, Dennis Odhiambo, teferi dereje Wirtu, Milton Tonui, Atem Bol, Lamin Kabareh for the determination, devotion and support to each other towards this achievement. Special thanks go to my relatives and close friends for their continued encouragement and support. May the Almighty reward you abundantly.

Finally, I shall be failing in my obligation if I do not express my sincere appreciation to my family who missed me while I was away on study leave. Your patience, support, love and affection showed to me always kept me stronger to reach this ultimate goal. My success is our success! Je vous remercie infiniment!

Abstract

Modelling and Forecasting volatility is one of the fundamental areas of research in Financial Mathematics, and thus has been the focus of many researchers; also, financial markets are known to be far from deterministic but stochastic and hence random models tend to perfectly model the markets. This study used appropriate Discrete-time Markov models to predict the multivariate stochastic Autoregressive volatility of an equity portfolio on a stock market. Therefore, the idea of modelling volatility as a stochastic process for an accurate forecast using the Markov chain on the financial data sets are based on the risks that often affect investment opportunities and the risk factors for prices changing that investors are most concerned about making decisions. The results provided more accuracy on forecasting price volatility on stock markets. We used a 3-state Discrete-Time Markov Chain (DTMC) for a portfolio of two stocks for the same sector and we compared the used model (fitted on a portfolio) to the multivariate GARCH models using real data from a stock market. The modified and generalized model provided more suitable volatility smiles compared to the Multivariate Generalized Autoregressive Conditional Heteroscedasticity (MGARCH) models and showed that working in a multivariate frame is most relevant especially when the number of state is bigger.

Résumé

La modélisation et prévision de la volatilité est l'un des domaines fondamentaux de la recherche en mathématiques financières et a donc été au centre de nombreuses recherches; aussi, les marchés financiers sont connus pour être non seulement déterministes mais stochastiques et donc les modèles aléatoires tendent à modeler parfaitement les marchés. Cette étude a utilisé les modèles de Markov à temps discret approprié pour prédire la volatilité autorégressive stochastique multivariée d'un portefeuille d'actions sur un marché boursier. L'idée de modéliser la volatilité en tant que processus stochastique pour une prévision précise utilisant la chaîne de Markov sur les ensembles de données financières repose sur les risques qui affectent souvent les opportunités d'investissement et les facteurs de risque de variation des prix. Les résultats fournissent plus de précision sur la prévision de la volatilité des prix sur les marchés boursiers. Nous avons utilisé une chaîne de Markov à temps discret (DTMC) à 3 états pour un portefeuille de deux actions du même secteur et nous avons comparé le modèle utilisé (appliqué à un portefeuille) aux modèles GARCH multivariés en utilisant les données réelles d'un marché boursier. Le modèle modifié fournit de meilleurs sourires de volatilité par rapport aux modèles multicritères d'autorétérosion conditionnelle autorégressive généralisée (MGARCH) et a montré qu'il est plus favorable de travailler dans un cadre multivarié surtout lorsque le nombre d'état est plus grand.

Contents

Declaration	i
Dedication	ii
Acknowledgement	iii
Abstract	iv
Résumé	v
Contents	vi
List of Figures	viii
List of Tables	ix
List of Abbreviations, Symbols and Definitions	x
List of Abbreviations	x
List of Symbols	xi
List of Definitions	xi
Chapter 1	1
INTRODUCTION	1
1.1 Background of the study	1
1.2 Statement of the problem	2
1.3 Justification of the study	3
1.4 Objectives of the Study	4
1.4.1 General Objective	4
1.4.2 Specific Objectives	4
1.5 Significance of the study	4
1.6 Scope of the study	5
1.7 Organization of the study	5
Chapter 2	6
LITERATURE REVIEW	6
2.1 Generalized ARCH Models	6
2.1.1 ARCH (p) Model	6
2.1.2 GARCH (p, q) Model	7
2.1.3 Symmetric GARCH Models	9
2.1.4 Asymmetric GARCH Models	12
2.2 Multivariate GARCH Models	13
2.2.1 Basic Idea	13

2.2.2 Generalizations of the univariate to Multivariate.....	14
2.3 A review on Discrete-Time Markov Chain.....	20
Chapter 3	22
METHODOLOGY	22
3.1 Proposed Model.....	22
3.2 Parameterization Strategies	29
3.3 Models with Asymmetric Volatility.....	30
3.4 Model with time varying volatility persistence	32
3.5 Estimation Method.....	33
3.5.1 Simulated Maximum Likelihood (SML) Method	33
3.6 Models Selection Criteria	35
3.6.1 The Akaike Information Criterion.....	35
3.6.2 The Bayesian Information Criterion.....	36
Chapter 4	37
DATA ANALYSIS AND RESULTS	37
4.1 Introduction.....	37
4.2 Descriptive Statistics	37
4.2.1 Assets returns	37
4.2.2 Summary statistics of NSE returns series data	38
4.2.3 ACF and PACF	40
4.2.4 Testing for Stationarity	43
4.2.5 Testing for Autocorrelation and Heteroscedasticity	43
4.3 Empirical Results	44
4.3.1 Selection of GARCH (p, q) Model.....	44
4.3.2 Selection of MGARCH (p, q) Model	46
4.3.3 Discrete Stochastic Autoregressive Volatility Models	50
4.4 Models Comparison	53
4.5 Forecasting Performance	54
Chapter 5	55
CONCLUSION AND RECOMMENDATIONS	55
5.1 Conclusion	55
5.2 Recommendations	56
References	57
Appendix.....	60

List of Figures

Figure 4.1 Equity and KCB daily prices and returns distributions (Jan 2010-Dec 2016)	39
Figure 4.2 ACF and PACF of Equity and KCB Assets Returns	41
Figure 4.3 ACF and PACF of Equity and KCB Squared Returns.....	42
Figure 4.4 Estimated Dynamic Conditional Correlation of Stocks Indices: Equity and KCB.....	48
Figure 4.5 Estimated Constant Conditional Correlation of Stocks Indices: Equity and KCB.....	49
Figure 4.6 Estimated Dynamic Conditional Covariance of Stocks Indices: Equity and KCB.....	49
Figure 4.7 Prediction of Conditional Covariance between Equity and KB and One-step prediction of volatility over the sample.....	54

List of Tables

Table 4.1 Summary statistics of NSE return series	38
Table 4.2 ADF test results	43
Table 4.3 Ljung-Box test for residuals of the NSE asset returns series	44
Table 4.4 Unconditional correlation coefficient of the returns series Error! Bookmark not defined.	
Table 4.5 GARCH (p, q) models for the residual returns.....	45
Table 4.6 Estimation results of AR (2)-GARCH (1, 1) model.....	45
Table 4.6 a Estimation results of AR (2)-GARCH (1, 1) model (Equity)	45
Table 4.6 b Estimation results of AR (2)-GARCH (1,1) model (KCB)	46
Table 4.7 Univariate ARCH results. Log-likelihood values.....	47
Table 4.8 Multivariate GARCH results: Bivariate Log-likelihood values	47
Table 4.9 Estimation of the coefficients of the DCC model for the stocks indices.....	48
Table 4.10 Estimation of the coefficients of the CCC model for the stocks indices.....	50
Table 4.11 Estimation results for linear discrete SARV (1, M) models	51
Table 4.11. a DSARV (1, 3) parameters estimates	51
Table 4.11. b DSARV (1, 10) parameters estimates.....	51
Table 4.11. c BIC and R^2 of volatility forecasts	51
Table 4.12 Model selection criteria for linear, log linear and log-quadratic discrete MSARV (1, M) models	52
Table 4.13 Pairwise Comparison-Diebold Mariano test for stock indices.....	54

List of Abbreviations, Symbols and Definitions

List of Abbreviations

ACF	Autocorrelation Function
ADF	Augmented Dickey Fuller
AIC	Akaike Information Criterion
ARCH	Autoregressive conditional Heteroscedasticity
ARMA	Autoregressive Moving Average
BIC	Bayesian Information Criterion
CCC	Constant Conditional Correlation
DCC	Dynamic Conditional Correlation
DMSARV	Discrete Multivariate Stochastic Autoregressive Volatility
DSARV	Discrete Stochastic Autoregressive Volatility
DVGARCH	Diagonal Vector Generalized Autoregressive Conditional Heteroscedasticity
EGARCH	Exponential Generalized Autoregressive Conditional Heteroscedastic
GARCH	Generalized Autoregressive Conditional Heteroscedasticity
i.e.	that is to say
i.i.d	independent and identically distributed
MLE	Maximum Likelihood Estimation
MSE	Mean Squared Error
NSE	Nairobi Securities Exchange
OLS	Ordinary Least Squared
PACF	Partial Autocorrelation Function
QMLE	Quasi Maximum Likelihood Estimation
SE	Standard Error
SMLE	Simulated Maximum Likelihood Estimation

List of Symbols

\mathbb{Z}	Set of integers
\mathbb{N}	Set of natural numbers
\forall	for all
\square	end of the proof

List of Definitions

Definition 1: Time Series (TS)

A TS model for a single risk factor is a stochastic process $\{\varepsilon_t\}_{t \in \mathbb{Z}}$, i.e. a family of random variables indexed by the integers and defined on some probability space (Ω, F, P) .

Definition 2: Stationary

A TS $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ is stationary if $(\varepsilon_{t_1}, \dots, \varepsilon_{t_n}) = (\varepsilon_{t_1+k}, \dots, \varepsilon_{t_n+k})$, $\forall t_1, \dots, t_n, k \in \mathbb{Z}$ and $\forall n \in \mathbb{N}$

Definition 3: Covariance Stationary (CS)

A TS $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ is CS if the first two moments exist and satisfy $\mu(t) = \mu$, $t \in \mathbb{Z}$, $\gamma(t, s) = \gamma(t+k, s+k)$, $t, s, k \in \mathbb{Z}$

Definition 4: ACF

ACF of a CS process $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ is defined as $\rho(h) = \rho(\varepsilon_h, \varepsilon_0) = \gamma(h)/\gamma(0)$, $\forall h \in \mathbb{Z}$

Definition 5: White Noise (WN)

A TS $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ is a WN process if it is covariance stationary with ACF $\rho(h) = 1$ if $h = 0$, and $\rho(h) = 0$ if $h \neq 0$

Definition 6: Martingale Difference

$\{\varepsilon_t\}_{t \in \mathbb{Z}}$ is known as a martingale difference sequence with respect to the filtration $\{F_t\}_{t \in \mathbb{Z}}$ if $E|\varepsilon_t| < \infty$, ε_t is F_t and $E\varepsilon_t | F_{t-1} = 0$, $\forall t \in \mathbb{Z}$

Definition 7: Stochastic Process (SP)

A SP is a family of random variables $X_{(t)} : t \in T$ defined on a given probability space, indexed by the time variable t , where t varies over an index set T .

Definition 8: Transition Probability Matrix (*TPM*)

The conditional probabilities $P_{ij}^{(m)} = \Pr(X_{t+m} = j | X_t = i)$ are called transition Probability and can be arranged in the form of a $n \times n$ matrix known as the *TPM*. It is the matrix consisting of the one-step transition probabilities P_{ij} and where the m -step transition probabilities is the probability of transitioning from state i to state j in m steps. $P_{ij}^{(m)} = \Pr(X_{t+m} = j | X_t = i)$

Definition 9: Markov Chain

A Markov chain is a special type of Stochastic Process; it is a stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event.

Chapter 1

INTRODUCTION

1.1 Background of the study

The stock market is an attractive area for wealth creation through investments; in fact, it is a market where companies can offload securities (stocks or bonds) in order to find financing. However, the profitability-risk couple is a primary criterion for each financial analyst in order to realize profits; to mitigate this risk, the players in the financial market must resort to modelling volatility.

The interest in volatility has always given rise to a number of researches, and is motivated by important reasons including: the increasing number of companies using risk management tools, the number of derivatives traded in global financial markets, the increasing need to invest, the instability of macroeconomic factors and so on. Thus, for a company to mitigate the level of exposure to a financial risk, it must be able to assess the volatility of each of the assets on the stock market. Considering the importance of volatility, or the price movement in finance, it is necessary to model volatility in order to make forecasts for the future. For this purpose, several models for forecasting volatility have been recorded in literature including the ARCH model proposed by Engle (1982) and its extension GARCH model by Bollerslev (1986). These have become very popular and useful in that they enable researchers and analysts to estimate the variance of a series at a particular point of time. Since then, there have been a great number of empirical applications of modelling and forecasting volatility of financial time series by employing these models and more other new models such as the Discrete Stochastic Autoregressive Volatility of Adriana and Kirby (2014) in which they make use of the famous Markov chain. Andrei Markov (1906) discovered Markov chain or Markov process in his first publication on Markov chains with finite state space. Random walks on the integers and the Gambler's ruin problem are Markov processes in discrete-time studied in the 1800 years.

In financial mathematics, a Markov's chain is simply a discrete-time or continuous-time stochastic process with a discrete or continuous state space. The peculiarity of this model is

that it has "no memory" that means, in determining its future state, it only takes into account the state of the last moment (that is, only the present moment can influence the future) and not from the past moment or the process history. However, this is not the case in many other processes such as AR, ARMA processes, where the determination or the approximation of the future depends on the previous states especially when the order is different from one.

Moreover, as shown in (Norris, 1997) the class of Markov chains is rich enough to serve in several applications; making them the first and most important examples of random processes. An in-depth analysis of the Markov chain in forecasting volatility in a stock market has the potential to provide very useful information to investors who wish to invest in these markets, also, to researchers who intend to use it for one reason or for another (Simeyo et Al. (2015)). In the world of finance, the theory of volatility has proved to be very useful in the analysis of many problems, especially the problems related to investment decisions. The analysis and study of Markov process may provide useful insight into the qualitative and quantitative analysis of some models.

Financial times series usually exhibit stylized characteristics. Knowledge on volatility of stocks returns is a crucial area of concern that needs special attention to compete favourably with developed stock markets.

1.2 Statement of the problem

The Discrete-Time Markov chain introduced in (Norris, 1997) proved to be somewhat useful in modelling volatility. Therefore, due to the importance of the study on volatility, there have been numerous research and various approaches on it. Among them, the two main approaches are deterministic models and stochastic models; however, deterministic models assume that, volatility at a particular time follows a deterministic function of the past whereas stochastic models assume that the volatility follows certain random process, and use ranges of values for variables in the form of probability distributions. Several authors have thoroughly modelled volatility using different deterministic models, but less attention has been given to the Markov chains, especially discrete-time Markov chain due to their property that the next value of the process depends on the current value but it is conditionally independent of its behaviour in the past.

Generally, stock markets are known to be far from deterministic and hence, random models tend to perfectly model the market. The most specific type of a mathematical object known as a stochastic or random process is the Markov process. Notwithstanding, Markov chain methods have been used to develop a class of Discrete Stochastic Autoregressive models in forecasting volatility for individual stock and stock index returns and it has been shown the performance of that models comparing to the GARCH(1,1) model (Adrianas & Kirby, 2014). However, the study has not carried out in forecasting volatility for portfolio returns that is a multivariate scenario for the simple reason that individuals or investors for a purpose of wealth creation need to diversify and maximise their investments. Today, globalization has resulted in higher international economic integration, investors and financial institutions are interested in knowing financial markets integration and how financial volatilities together move over time across several markets. Empirical results show that working with separate univariate models is much less relevant than multivariate modelling framework (Mgr Milan, 2014).

1.3 Justification of the study

Forecasting stock returns variations is one of the fundamental areas of research in Finance, and thus, has been the focus of many researchers; this is due to its potential ability to assess the level of exposure to a financial risk, and to increase the trend of investment decisions.

Therefore, a successful prediction of the stock's future price could yield significant profit and, developing better frameworks for risk modelling and forecasting is of huge benefit in financial industry because of the minimization of the risk or the maximization of the investments returns. Only to consider the characteristics of the evolution on the history situation of the event itself, and to predict changes of the internal states by calculating the state transition probability; thus, Markov model has broad applicability in prediction of the stock market (Zhang & Zhang, 2009)

1.4 Objectives of the Study

1.4.1 General Objective

The general objective was to develop a discrete Multivariate Stochastic Autoregressive Volatility model using Markov Chains process with application to the NSE stock returns.

1.4.2 Specific Objectives

- i. To fit the Multivariate GARCH models to the NSE data and measure portfolio volatility.
- ii. To generalize a discrete time Multivariate stochastic autoregressive volatility model using Markov chain process.
- iii. To fit the Multivariate stochastic autoregressive volatility model to the stock indices returns in the NSE data.

1.5 Significance of the study

The study is highly significant considering the importance of forecasting in the world of finance especially in financial markets, which are somewhat considered as risk areas. Volatility plays an important role for stock prices prediction; it refers to the measure for price fluctuation of a specific financial instrument over time. It is therefore, a very important factor as mentioned above, that can deeply affect investment decisions and concerns every other participant or player in the financial markets.

The prices movement on the financial markets is the factor that investors care the most about, the more the prices are variant, the greater the risk of investing. In stock markets, the price movement affect investment decisions making it important to model volatility of returns using the stochastic models, which assume that volatility follows certain random process. Therefore, the study will provide useful insight and help us in understanding the suitability of the Discrete-time Markov chain in forecasting the stock prices movement.

1.6 Scope of the study

The study will be limited only to the numerical investigation on the discrete-time Markov process. We use daily closing prices for Equity and KCB stocks of the NSE market for a period from 2010 to 2016. In particular, underlying model, theorems, model assumptions and time will be considered.

However, we shall not only do the analysis of some models available in the literature for modelling and forecasting volatility, but we will also apply the Discrete Multivariate Stochastic Autoregressive Volatility model to Stock market data.

1.7 Organization of the study

The rest of this work is organized as follows: Chapter two presents the literature review relating to our research objectives which provides the related works that have been done. Chapter three discusses the methodology in which we develop the model. Chapter four gives data analysis and discussion and the conclusion and recommendations for further study are given in the last Chapter.

Chapter 2

LITERATURE REVIEW

This chapter presents the research on modelling and forecasting volatility of stock returns. It reviews the various models that have been used and studied to model and forecast volatility by different authors both theoretically and empirically.

2.1 Generalized ARCH Models

In recent years, a variety of models which apparently forecast changes in stock market prices have been introduced, and have played an important role to help people forecast the future.

The ARCH class of model introduced by Engle (1982) and its generalization, GARCH models by Bollerslev (1986) are the most and widely used methodologies in modelling and forecasting volatility of financial time series. The literature of ARCH-type models is developed and we will use it as benchmark models for comparison of the Discrete Multivariate Stochastic Autoregressive Volatility model. In this chapter we will study different univariate and multivariate GARCH models. We will also use the Quasi Maximum Likelihood Estimation which is the common estimate method for this type of models.

2.1.1 ARCH (p) Model

Engle (1982) introduced this model for forecasting volatility, with the following specifications. Let $\{e_t\}_{t \in \mathbb{Z}}$ be a sequence of independent and identically distributed random variable such that $e_t \sim N(0,1)$. $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ follows an ARCH (p) process if:

$$\begin{aligned}r_t &= \mu_t + \varepsilon_t \\ \varepsilon_t &= \sigma_t e_t \\ \sigma_t^2 &= \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2\end{aligned}\tag{2.1}$$

with r_t , the return at time t , μ_t denotes the average return at time t , p is the order of the process, $\omega > 0$, and $\alpha_i \geq 0$ for $i = 1, \dots, p$ in order to ensure non-negativity of the conditional variance (Engle, 1982).

Now, let $\{F_t\}_{t \in \mathbb{Z}}$ denote the sigma algebra representing the information set given at time t denoted by $F_t = \sigma(\{\varepsilon_t, \dots, \varepsilon_{t-1}\})$.

Equation (2.1) ensures that σ_t is measurable w.r.t F_{t-1} provided that $E(|\varepsilon_t|) < \infty$, we have

$$\begin{aligned} E(\varepsilon_t / F_{t-1}) &= E(\sigma_t e_t | F_{t-1}) \\ &= \sigma_t E(e_t | F_{t-1}) \\ &= \sigma_t E(e_t) = 0 \end{aligned} \tag{2.2}$$

A simple interpretation of this, is that the ARCH process has a martingale difference property w.r.t $\{F_t\}_{t \in \mathbb{Z}}$. We then assume that the process is a covariance stationary white noise, with $E(|\varepsilon_t|) < \infty$ and

$$\begin{aligned} \text{var}(\varepsilon_t | F_{t-1}) &= E(\sigma_t^2 e_t^2 | F_{t-1}) \\ &= \sigma_t^2 \text{var}(e_t | F_{t-1}) \\ &= \sigma_t^2 \text{var}(e_t) = \sigma_t^2 \end{aligned} \tag{2.3}$$

This means that the conditional standard deviation is a continually changing function of previous squared values of the process. However, despite the simplicity and flexibility of the ARCH model, one of its disadvantages is that, it often requires many parameters to be estimated to accurately describe the process of volatility of an asset return (Tsay, 2010). This can present difficulties when using the model to accurately describe the data set, that is why Bollerslev introduced the Generalized ARCH model in order to solve this problem.

2.1.2 GARCH (p, q) Model

Bollerslev (1986) introduced an extension of the ARCH model, with the following specifications. Let $\{e_t\}_{t \in \mathbb{Z}}$ a sequence of i.i.d random variables such that $e_t \sim N(0, 1)$. Here, $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ is said to be GARCH (p, q) process if:

$$\begin{aligned}\varepsilon_t &= \sigma_t e_t \\ \sigma_t^2 &= \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2\end{aligned}\quad (2.4)$$

where p and q are the orders of the process, ω , α_i , and β_j are the parameters to be estimated, with $\omega > 0$, $\alpha_i \geq 0$ (for $i=1, \dots, p$) and $\beta_j \geq 0$ (for $j=1, \dots, q$). These are the necessary conditions for the variance to be positive (Chong et Al, 1999).

If we let $q=0$, the process reduces to an ARCH (p) process and for $p=q=0$, ε_t is simply white noise. However, the short run dynamics of the resulting volatility process is determined by the size of the parameters α_i and β_j .

Large ARCH coefficients, α_i imply that volatility reacts significantly to markets movements, while large GARCH coefficients β_j indicate that shocks are persistent on the stocks market (Perrelli, 2001).

We can also write the variance σ_t^2 of equation (2.4) in terms of the lag-operator L where ($L\varepsilon_t = \varepsilon_{t-1}$) we get:

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2 \quad (2.5)$$

where

$$\alpha(L) = \alpha_1 L + \alpha_2 L^2 + \dots + \alpha_p L^p \text{ and } \beta(L) = \beta_1 L + \beta_2 L^2 + \dots + \beta_q L^q \quad (2.6)$$

Moreover, if $1 - (\alpha_1 x + \alpha_2 x^2 + \dots + \alpha_p x^p) = 0$

that means if the roots of the characteristic equation lie outside the unit circle and the process $\{\varepsilon_t\}$ is stationary, then we can write the variance equation of equation (2.4) as

$$\sigma_t^2 = \frac{\omega}{1 - \beta(1)} + \frac{\alpha(L)}{1 - \beta(L)} \varepsilon_t^2 \quad (2.7)$$

Now let $\tilde{\omega} = \frac{\omega}{1 - \beta(1)}$, and γ_i the coefficients of L^i in the expansion of $\frac{\alpha(L)}{1 - \beta(L)}$ then, we

obtain the following transformation of equation (2.7)

$$\sigma_t^2 = \tilde{\omega} + \sum_{i=1}^{\infty} \gamma_i \varepsilon_{t-i}^2 \quad (2.8)$$

We have demonstrated that the GARCH (p, q) can also be written as an ARCH(∞) process with a fractional structure of the coefficients. This clearly means that ε_t is also a martingale difference and the conditional variance of ε_t is given by

$$\sigma_t^2 = \frac{\omega}{1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j} \quad (2.9)$$

with $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$ to ensure the stationarity of the conditional variance.

GARCH models have been extended by many others authors in order to fill the gaps in the main GARCH model.

2.1.3 Symmetric GARCH Models

Our study is focused on one of the several symmetric GARCH models that have been used in the literature; this is the Standard GARCH (1, 1) model which will be discussed in the subsequent section. This model has the particularity that the conditional variance depends only on the magnitude.

GARCH (1, 1) model

In the GARCH (1, 1) model, the dynamics show up in the ACF of the squared returns and the ACF is like that of the ARMA (1, 1) process. If $\alpha + \beta$ is close to one then, the ACF will decay slowly indicating a relatively slowly changing conditional variance. In this model, the conditional variance is presented as a linear function of its own lags. It is a particular case of GARCH (p, q) where $p = q = 1$. The basic univariate GARCH (1, 1) is given by

$$\text{mean Equation} \quad r_t = \mu + \varepsilon_t \quad , \quad \varepsilon_t \sim N(0, \sigma_t^2) \quad (2.10)$$

$$\text{volatility Function} \quad \sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (2.11)$$

where $\omega > 0$, $\alpha_1 \geq 0$, $\beta_1 \geq 0$ and r_t is the return of the asset at time t , μ denotes the average return, σ_t^2 is the conditional variance and ε_t is the residual returns as defined in equation (2.4).

The size of parameters α and β determine the short-run dynamics of the volatility time series and if $\alpha_1 + \beta_1 = 1$, then any shock will lead to a permanent change in all future values. Hence, shock to the conditional variance is “persistence”.

We can write the variance equation as a stochastic recurrence equation (SRE) by substituting the first equation (2.4) in equation (2.11), we obtain

$$\begin{aligned}\sigma_t^2 &= \omega + \alpha_1 (\sigma_{t-1}^2 e_{t-1}^2) + \beta_1 \sigma_{t-1}^2 \\ &= \omega + \sigma_{t-1}^2 (\alpha_1 e_{t-1}^2 + \beta_1)\end{aligned}\quad (2.12)$$

This can be written in the following form

$$\mathbf{X}_t = \mathbf{A}_t \mathbf{X}_{t-1} + \mathbf{B}_t \quad (2.13)$$

where $\{\alpha_t\}$ and $\{e_t\}$, $\forall t \in \mathbb{Z}$ are sequences of i.i.d random variables and with $\mathbf{X}_t = \sigma_t^2$, $\mathbf{X}_{t-1} = \sigma_{t-1}^2$, $\mathbf{A}_t = \alpha_1 e_{t-1}^2 + \beta_1$, and $\mathbf{B}_t = \omega$

These following conditions are sufficient to get a solution

$$\mathbb{E}(\ln^+ |\mathbf{B}_t|) < \infty \text{ and } \mathbb{E}(\ln |\mathbf{A}_t|) < \infty \quad (2.14)$$

and the meaning of $\mathbb{E}(\ln^+ |\mathbf{B}_t|)$ is $\max\{0, (\ln |\mathbf{B}_t|)\}$.

By iteration n times we get from equation (2.13) the following expression

$$\begin{aligned}X_t &= A_t (A_{t-1} X_{t-2} + B_{t-1}) + B_t \\ &= B_t + \sum_{i=1}^n B_{t-i} \prod_{j=0}^{n-1} A_{t-j} + X_{t-k-1} \prod_{i=0}^n A_{t-i}\end{aligned}$$

Conditions given in (2.14) ensure that the middle term on the right hand side converges absolutely and the last term disappears as shown by the following expression

$$\frac{1}{n+1} \sum_{i=0}^n \ln |A_{t-i}| \rightarrow \mathbb{E}(\ln |A_t|) < 0$$

and by the strong law of large numbers, this yields

$$\prod_{i=0}^n |A_{t-i}| = \exp\left(\sum_{i=0}^k \ln |A_{t-i}|\right) \rightarrow 0$$

Hence, the unique solution of equation (2.13) is given by

$$X_t = B_t + \sum_{i=1}^{\infty} B_{t-i} \prod_{j=0}^{i-1} A_{t-j} \quad (2.15)$$

In this case, the sum $\sum_{i=1}^{\infty} B_{t-i}$ also converges absolutely almost surely. Then the general solution of equation (2.12) becomes

$$\sigma_t^2 = \omega \left(1 + \sum_{i=1}^{\infty} \prod_{j=1}^i \alpha_1 e_{t-1}^2 + \beta_1 \right) \quad (2.16)$$

Then, the solution of the GARCH (1, 1) defining equations is given by

$$\varepsilon_t = e_t \sqrt{\omega \left(1 + \sum_{i=1}^{\infty} \prod_{j=1}^i \alpha_1 e_{t-1}^2 + \beta_1 \right)} \quad (2.17)$$

Proposition 2.1

A GARCH (1, 1) process is covariance stationary white noise process if and only if $\alpha_1 + \beta_1 < 1$. And the variance of the covariance stationary process is constant and then given by

$$E(\varepsilon_t^2) = \frac{\omega}{1 - \alpha_1 - \beta_1}$$

Proof 2.1

From equation (2.11),

$$\begin{aligned} \varepsilon_t^2 &= \sigma_t^2 e_t^2 \\ &= (\omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2) e_t^2 \\ \varepsilon_t^2 &= \omega e_t^2 + \alpha_1 e_t^2 \varepsilon_{t-1}^2 + \beta_1 e_t^2 \sigma_{t-1}^2 \end{aligned}$$

Let's assume the covariance-stationarity, then it follows from the previous equation of ε_t^2 and

$$E(e_t^2) = 1 \text{ that}$$

$$E(\varepsilon_t^2) = \omega + \alpha_1 E(\varepsilon_{t-1}^2) + \beta_1 E(\sigma_{t-1}^2)$$

$$\text{Since } E(\varepsilon_t^2) = E(\varepsilon_{t-1}^2) = E(\sigma_{t-1}^2) = \sigma^2$$

then it is clearly shown that $\sigma^2 = E(\varepsilon_t^2) = \frac{\omega}{1 - \alpha_1 - \beta_1}$ with the condition that $\alpha_1 + \beta_1 < 1$ to ensure stationarity.

2.1.4 Asymmetric GARCH Models

The problem of leverage effects in stock returns is a very important aspect to be studied. Since the previous symmetric GARCH models have not been able to capture this relevant aspect of the stock returns, we propose some models which have been introduced in the literature in order to solve this problem, while being able to capture this asymmetry: they are called asymmetric models.

These models include, EGARCH model (Nelson, 1991), TGARCH (Glosten et Al, 1993) ... but our study will focus on the EGARCH model which is most common and where the stocks of the same magnitude, positive or negative have different effects on future volatility.

EGARCH Model

This model is based on the logarithmic expression of the conditional variability. We use this model to capture the asymmetric responses of the time varying volatility and returns at the same time, whenever the parameter values are negative, the model ensures that the conditional variance is always positive (Suliman and Winker, 2012), this means that there is no need for parameter restrictions to impose non negativity. The model was developed by Nelson (1991) and hence, the following equation

$$\ln \sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \frac{|\varepsilon_{t-i}| + \delta_i \varepsilon_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^q \beta_j \ln \sigma_{t-j}^2 \quad (2.18)$$

where δ is the asymmetric response parameter.

The EGARCH (p, q) conditional variance model includes q past log conditional variances that compose the GARCH component polynomial.

In most empirical cases, δ is expected to be negative so that a “negative shock” increases “future volatility”, while a positive shock eases the effect on future volatility (Harvey and Genaro, 2013).

Therefore, for an EGARCH (1, 1) model where $p = q = 1$ given by

$$\ln \sigma_t^2 = \omega + \beta_1 \ln \sigma_{t-1}^2 + \alpha_1 \left\{ \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| - \sqrt{\frac{\pi}{2}} \right\} - \delta \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \quad (2.19)$$

The left hand side is the log of the conditional variance. The coefficient δ is known as the asymmetry or leverage term. The presence of leverage effects can be tested by the hypothesis that $\delta < 0$, the impact is symmetric if $\delta \neq 0$ and $\pi = 22/7$.

2.2 Multivariate GARCH Models

2.2.1 Basic Idea

Today, globalization has resulted in higher international economics integration, investors and also financial institutions are interested in knowing financial markets integration and how financial volatilities together move over time across several markets or assets.

Empirical results show that working with separate univariate models is much less relevant than multivariate modelling framework. Let's assume that

Assumption₁. Assets pricing depends on the covariance of the assets in a portfolio. Hence, it is important to consider the co-movements in the portfolio.

Assumption₂. Financial volatilities move together more or less closely over time across assets and markets (correlation coefficient).

Assumption₃. Recognizing the previous feature through a Multivariate model should lead to more relevant empirical models than working with separate univariate models.

Assumption₄. In financial applications, extending from univariate to multivariate modelling opens the door to better decision tools in various areas such as asset pricing models or portfolio selection.

MGARCH models were initially developed in the late of 1980s and the first half of the 1990s. The most common application of these class of models is to estimate the volatility effects among different markets or assets. In MGARCH models, covariance matrix need by definition to be positive definite, therefore imposing positive definiteness is one of the features that needs to be taken into account in its specifications. One possibility is to derive conditions under which the conditional variance matrices implied by the model are positive definite, but this is often not

feasible in practice. In this case, an alternative is to formulate the model in a way that positive definiteness is implied by the structure (in addition to some simple constraints).

Before we start with the definitions, we will introduce some basic multivariate framework of GARCH models.

Let's consider a stochastic vector* process X_t with dimension k . Let F_t be the non-decreasing collection of σ -fields generated by past of the series X_t that means:

$$F_t = \sigma(X_t, X_{t-1}, \dots)$$

Assume that conditional covariance matrix H_t of X_t is measurable with respect to F_{t-1} . The MGARCH framework is then given by

$$X_t = \sqrt{H_t} z_t \quad (2.20)$$

where $H_t = [\sigma_{ij}^2]_t$ is $k \times k$ symmetric positive definite matrix for all t . $\sqrt{H_t}$ may be obtained by Cholesky decomposition of H_t . z_t is a k dimensional i.i.d vector process with zero mean and unit variance*.

Hence, z_t is independent of F_{t-1} , it follows that $\text{cov}(z_t | F_{t-1}) = \text{cov}(z_t) = I_n$. The process X_t is then a k dimensional vector martingale difference sequence*. That means

$$\begin{aligned} E(X_t | F_{t-1}) &= 0 \\ \text{cov}(X_t | F_{t-1}) &= \sqrt{H_t} \text{cov}(z_t | F_{t-1}) \sqrt{H_t} = H_t \end{aligned} \quad (2.21)$$

The demonstrations can be referred in the univariate section. The information set F_t contains both lagged values of the squares and cross-product of X_t and elements of the conditional covariance matrices up to time t .

2.2.2 Generalizations of the univariate to Multivariate

The extension from univariate GARCH to the Multivariate introduced above requires considering k -dimensional stochastic process with zero mean random variables X_t and covariance matrix H_t as shown previously.

a. VGARCH Model

Bollerslev et Al (1988) proposed a VGARCH model which is a straightforward generalization of the univariate GARCH model.

Every conditional variance and covariance is function of all lagged conditional variance and covariance, as well as lagged squared returns and cross products of returns. The VGARCH is defined as follow:

Definition 2.1

A VGARCH (p, q) process is a martingale difference sequence X_t , relative to a given filtration F_t , whose conditional covariance matrix $H_t = \text{cov}(X_t | F_{t-1})$ satisfy, $\forall t \in \mathbb{Z}$

$$\text{Vech}(H_t) = \omega + \sum_{i=1}^p A_i \text{vech}(X_{t-i} X_{t-i}') + \sum_{j=1}^q B_j \text{vech}(H_{t-j}) \quad (2.22)$$

where $\text{vech}(\cdot)$ is the operator that stocks the lower triangular portion of a symmetric square $k \times k$ matrix into a $(k(k+1)/2)$ -dimensional vector. ω is an $(k(k+1)/2)$ dimensional vector, A_i and B_i are square parameter matrices of order $(k(k+1)/2)$

For a purpose of explanation, let's consider a bivariate VGARCH (1, 1) model with $k = 2$ and we denote $\sigma_t^2 = \text{vech}(H_t)$, the equation (2.29) becomes

$$\sigma_t^2 = \begin{pmatrix} \sigma_{11,t}^2 \\ \sigma_{12,t}^2 \\ \sigma_{22,t}^2 \end{pmatrix} = \begin{pmatrix} \omega_{1,t} \\ \omega_{2,t} \\ \omega_{3,t} \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{bmatrix} X_{1t-1}^2 \\ X_{1t-1} X_{2t-1} \\ X_{2t-1}^2 \end{bmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{bmatrix} \sigma_{11t-1}^2 \\ \sigma_{12t-1}^2 \\ \sigma_{22t-1}^2 \end{bmatrix}$$

We can notice that, from this we immediately see equivalency between the VEC and VECH representation.

In VEC representation all the covariance equations appear twice, because there is an equation for $\sigma_{ij,t}^2$ as well as for $\sigma_{ji,t}^2$. This is because all the off-diagonal terms appear twice within each equation. (i.e. both of the terms $\sigma_{ij,t-1}^2$ and $\sigma_{ji,t-1}^2$ appear in each equation). We can then remove

this redundant terms without affecting the model doing so, dimensions of matrices A_i and B_i become $(k(k+1)/2)$ instead of k^2 .

This model is general and flexible, and the coefficients are also directly interpretable, but it has some drawbacks in applications, the higher number of parameters which equals $(p+q)(k(k+1)/2)^2 + k + (k+1)/2$, however the model will be practicable in practice in our study since we use the bivariate case. Another disadvantage is, there exists only sufficient conditions on the parameters to ensure that conditional variance matrices H_t are positive definite almost surely $\forall t$.

The restrictions of the model are introduced by Bollerslev, Engle and Wooldridge (1988) such that, each component of the covariance matrix H_t depends only on its own past and past values of $X_t X_t'$ as in equation (2.22), that means in the diagonal representation, it is assumed that the matrices A_i and B_i are diagonal, we call it a diagonal VEGARCH model.

b. Diagonal VGARCH Model

This so called DVGARCH model will reduce the number of parameters to $(p+q+1)(k(k+1)/2)$ and therefore it is still possible to obtain conditions for positive definiteness of $H_t \forall t$.

To illustrate the bivariate case, the DVGARCH model is simply:

Letting $h_t = \sigma_t^2$,

$$\sigma_t^2 = \begin{pmatrix} \sigma_{11,t}^2 \\ \sigma_{12,t}^2 \\ \sigma_{22,t}^2 \end{pmatrix} = \begin{pmatrix} \omega_{1,t} \\ \omega_{2,t} \\ \omega_{3,t} \end{pmatrix} + \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix} \begin{bmatrix} X_{1,t-1}^2 \\ X_{1,t-1} X_{2,t-1} \\ X_{2,t-1}^2 \end{bmatrix} + \begin{pmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{pmatrix} \begin{bmatrix} \sigma_{11,t-1}^2 \\ \sigma_{12,t-1}^2 \\ \sigma_{22,t-1}^2 \end{bmatrix}$$

We have

$$\sigma_{11,t}^2 = \omega_{1,t} + a_{11} X_{1,t-1}^2 + b_{11} \sigma_{11,t-1}^2$$

$$\sigma_{12,t}^2 = \omega_{2,t} + a_{22} X_{1,t-1} X_{2,t-1} + b_{22} \sigma_{12,t-1}^2$$

$$\sigma_{22,t}^2 = \omega_{3,t} + a_{33} X_{2,t-1}^2 + b_{33} \sigma_{22,t-1}^2$$

In the bivariate model illustrated here, there are three free parameters in each of the A_1 and B_1 matrices and nine parameters (including constants). In the general k – variate DVGARCH model there are $(k(k+1)/2)$ free parameters in each matrix.

However, the DVGARCH representation seems to be too restrictive since no interaction is allowed between the different conditional variances and covariance.

In order to derive a sufficient condition for the DVGARCH for H_t to be positive definite , we write the known DVGARCH model in a matrix representation yields

$$H_t = \tilde{W} + \tilde{A} \odot X_{t-1} X_{t-1}' + \tilde{B} \odot H_{t-1} \quad (2.23)$$

where \odot denotes the element-by-element product of the two matrices. \tilde{W} , \tilde{A} , and \tilde{B} are all $k \times k$ parameter matrices. Using Cholesky decomposition of the parameter matrices and from the properties of a Hadamard product yields

$$H_t = \tilde{W}\tilde{W}' + \tilde{A}\tilde{A}' \odot X_{t-1} X_{t-1}' + \tilde{B}\tilde{B}' \odot H_{t-1} \quad (2.24)$$

where $\tilde{W}\tilde{W}'$, $\tilde{A}\tilde{A}'$, and $\tilde{B}\tilde{B}'$ are all positive semi-definite and therefore, H_t is positive definite $\forall t$, since the initial covariance matrix H_0 is also positive definite.

By writing the parameters matrices in the form of Cholesky decomposition, the positive semi-definiteness is guaranteed in estimation without imposing any further restrictions.

By definition we have the operator L as mentioned previously in the univariate case, where $L^i X_t = X_{t-i}$ and by convention, $A(L) = A_1 L + A_2 L^2 + \dots + A_p L^p$ and $B(L) = B_1 L + B_2 L^2 + \dots + B_q L^q$. Now let z_t our k – dimensional i.i.d vector process with mean zero and unit variance, knowing that z_t is independent of F_{t-1} , it follows that $\text{cov}(z_t / F_{t-1}) = \text{cov}(z_t) = I_n$. There exists a VGARCH process X_t such that $X_t = \sqrt{H_t} z_t$ where $H_t = \text{cov}(X_t / F_{t-1})$ and $F_t = \sigma(X_t, X_{t-1}, \dots)$.

Assuming that X_t is doubly infinite sequence, yields to the following equation for conditional covariance matrix, by rewriting equation (2.29) as

$$vech(\mathbf{H}_t) = \sum_{i=1}^{\infty} \mathbf{B}(\mathbf{L})^{i-1} \left[\omega + \mathbf{A}(\mathbf{L}) vech(\mathbf{X}_t \mathbf{X}_t') \right] \quad (2.25)$$

Proposition 2.2

Equation (2.32) is indeed a VGARCH model.

Proof 2.2

$$\begin{aligned} vech(\mathbf{H}_t) &= \omega + \mathbf{A}(\mathbf{L}) vech(\mathbf{X}_t \mathbf{X}_t') + \sum_{i=2}^{\infty} \mathbf{B}(\mathbf{L})^{i-1} \left[\omega + \mathbf{A}(\mathbf{L}) vech(\mathbf{X}_t \mathbf{X}_t') \right] \\ &= \omega + \mathbf{A}(\mathbf{L}) vech(\mathbf{X}_t \mathbf{X}_t') + \mathbf{B}(\mathbf{L}) \sum_{i=1}^{\infty} \mathbf{B}(\mathbf{L})^{i-1} \left[\omega + \mathbf{A}(\mathbf{L}) vech(\mathbf{X}_t \mathbf{X}_t') \right] \\ &= \omega + \mathbf{A}(\mathbf{L}) vech(\mathbf{X}_t \mathbf{X}_t') + \mathbf{B}(\mathbf{L}) vech(\mathbf{H}_t) \quad \square \end{aligned}$$

In the VGARCH representation and even DVGARCH representation, the restrictions can be difficult to check, let alone impose during estimation. We then propose a new parameterization which easily imposes these restrictions and which eliminates very few if any interesting models allowed by the VGARCH representation. Necessary conditions for the conditional variance matrix is presented by Engle et Al (1995), to be positive definite, such conditions are often difficult to impose during the optimization of the log-likelihood function; Bollerslev (1990) suggested a CC-MGARCH model that can overcome these difficulties.

c. Constant Conditional Correlation Model

Introduced for the first time by Bollerslev (1990), the conditional correlation matrix in this class of models is time invariant. We then choose a GARCH-type model for each conditional variance and we model the conditional correlation matrix, based on the conditional variances.

Since the conditional correlation matrix is time invariant, the conditional covariance is therefore proportional to the product of the corresponding conditional standard deviations.

Hence,

Definition 2.3

The CCC (p, q) process is a martingale difference sequence \mathbf{X}_t , relative to a given filtration \mathbf{F}_t , whose conditional covariance matrix $\mathbf{H}_t = \text{cov}(\mathbf{X}_t | \mathbf{F}_{t-1})$ satisfy

$$H_t = D_t R D_t = \rho_{ij} (\sigma_{iit} \sigma_{jtt}) \quad (2.26)$$

where
$$D_t = \text{diag}(\sigma_{11t}, \dots, \sigma_{kkt}) \quad (2.27)$$

and
$$R = (\rho_{ij}) \quad (2.28)$$

is a symmetric positive definite matrix with $\rho_{ii} = 1$, $\forall i$ then off diagonal elements of the conditional covariance matrix are defined as $[H_t]_{ij} = \sigma_{it} \sigma_{jt} \rho_{ij}$ for $i \neq j$, $1 \leq i, j \leq k$. σ_{iit}^2 is defined as univariate GARCH (p, q) model

$$\sigma_t^2 = \omega + \sum_{i=1}^p A_i X_{t-i}^2 + \sum_{i=1}^q B_i \sigma_{t-i}^2 \quad (2.29)$$

where ω is $k \times 1$ vector, A_i and B_i are diagonal $k \times k$ matrices. See Francq and Zakoian (2010).

d. Dynamic Conditional Correlation Model

A generalization of the CCC model was proposed by Engle (2002), the so-called DCC is a new class of multivariate models where conditional correlation matrix is time-dependent. These models are flexible like the previous univariate GARCH and parsimonious parametric models for the correlations.

Definition 2.4

The DCC process is a martingale difference sequence X_t , relative to a given filtration F_t , whose conditional covariance matrix $H_t = \text{cov}(X_t | F_{t-1})$ satisfy

$$H_t = D_t R_t D_t \quad (2.30)$$

where

$$D_t = \text{diag}(\sigma_{1t}, \dots, \sigma_{kt}) \quad (2.31)$$

and R_t is $k \times k$ time varying correlation matrix of X_t , σ_{iit}^2 is defined as univariate GARCH (p, q) model.

$$\sigma_{it}^2 = \omega_i + \sum_{j=1}^{p_i} \theta_{ij} X_{t-j}^2 + \sum_{j=1}^{q_i} \phi_{ij} \sigma_{t-j}^2$$

where ω_i , θ_{ij} , and ϕ_{ij} are non-negative parameters for $i = 1, \dots, k$, with the usual GARCH restriction for non-negativity and stationary being imposed, such as non-negativity of variances and $\sum_{j=1}^{p_i} \theta_{ij} + \sum_{j=1}^{q_i} \phi_{ij} < 1$.

In bivariate case, the number of parameters to be estimated equals $(k + 1)(k + 4) / 2$. Note that H_t , being a covariance matrix has to be positive definite, D_t is positive definite since all the diagonal elements are positive, this ensure R_t to be positive definite. Also, all the elements in the correlation matrix R_t have to be equal or less than one by definition; See Engle (2002).

2.3 A review on Discrete-Time Markov Chain

The concept of Markov chains came into being at the beginning of the 20th century. The Russian mathematician Andrei Markov introduced them in a paper published in 1906. Since that time, many scientists working in fields as varied as statistical physics, biology, mathematics and mathematics finance, have been able to integrate the key properties of the Markov chains in their respective scope. In the literature, the first paper that we have been able to identify, involving the Markov chains in a financial context is Pye (1966), the problem is the evaluation of cash flows in an uncertain environment of interest rates. The author uses Markov chains to model the evolution of the short-term rate for a given period. This employment is based on the Markov hypothesis of the evolution of these rates.

In other words, the author assumes that the short-term rate that will be effective in a period depends only on the current short-term rate for a period. Under these assumptions, the evaluation of financial flows is therefore carried out only by simple matrix operations.

Simeyo et Al (2015) used a derived initial state vector and a transition matrix to predict the states of Safaricom share price accurately. They concluded that the Markov chain predict method is purely a probability forecasting method, as their predicted results were simply expressed probability of certain state of stock prices in the future rather than be in absolute state. Their study showed how Markov model fits the data and its ability to predict trend.

Kobbacy and Nicol (1994) also use Markov chains to model interest rates. The transition matrix of the Markov chain has been obtained with historical data. We can thus see that; the Markov chain can be used as much in a context of historical probability as in a more theoretical context. The Markov chains have also been privileged tool to analyse the concept of random walk, for example McQueen and Thorley (1991) in which the authors use a transition matrix to test the hypothesis that annual returns follow a random walk.

A discrete time stochastic process is a family of random variables $\{v_t, t \in N\}$ defined on a given probability space and indexed by the parameter $t \in N = \{0, 1, 2, \dots\}$.

If the state space of a Markov process is discrete, the Markov process is a discrete-time stochastic process and is called a Markov chain. We will consider Markov chains having finite number of states:

$$M = \{1, 2, 3, \dots, m\} \quad (2.32)$$

with $m = 3$. Therefore, a first-order discrete-time Markov chain (DTMC) having $m = 3$ discrete states satisfies the following relationship:

$$\text{Prob}(v_{t+1} = \sigma_k / v_1 = \sigma_h, \dots, v_{t-1} = \sigma_i, v_t = \sigma_j)$$

$$P_{ij} = \text{Prob}(v_{t+1} = \sigma_i / v_t = \sigma_j)$$

where $v_t \in \{\sigma_1, \sigma_2, \dots, \sigma_N\}$ for all t . The conditional probabilities $\text{Prob}(v_{n+1} = \sigma_{n+1} / v_n = \sigma_n)$ are called the single-step transition probability of the Markov chain. They give the conditional of making a transition from state i to state j when the time parameter increases from n to $n + 1$. These probabilities are independent of n and are written as:

$$P_{ij} = \text{Prob}(v_{t+1} = \sigma_i / v_t = \sigma_j) \quad \forall i, j \in M \quad (2.33)$$

The matrix P , formed by placing P_{ij} in row i and column j for all i and j , is called the transition probability matrix (t.p.m). Therefore, the elements of the matrix P should satisfy the two following properties:

$$(1) \quad 0 \leq P_{ij} \leq 1 \quad \forall i, j \in M$$

$$(2) \quad \sum_{i=1}^m P_{ij} = 1 \quad \forall j \in M$$

Chapter 3

METHODOLOGY

3.1 Proposed Model

In this work, we extended the model of Adriana and Kirby (2014) by developing a discrete multivariate stochastic autoregressive volatility model. Stochastic modelling is a form of financial modelling that includes one or more random variables. Among the stochastic models, one kind of process called Markov process is a specific type of a stochastic process, which is a mathematical object usually defined as a collection of random variables, this has been studied by several independent researchers. Two related modelling strategies are typically followed in specifying the dynamics of the volatilities; the volatilities can be assumed to be a non-linear function of past returns, as shown previously in the ARCH type models also, the volatility process is a function of an exogenous shock as well as past volatilities as shown below. The stochastic volatility model can easily be parsimoniously extended to include multiple assets. The standard stochastic volatility model is defined as

$$r_t = v_t \varepsilon_t \quad t = 1, \dots, T, \quad \varepsilon_t \sim N(0,1)$$

where r_t is the return for the interval $t-1$ to t , $v_t > 0$ is the return volatility for period t , and ε_t is a white-noise error that is independent of v_{t-j} for all $j \geq 0$.

A discrete time stochastic process is a family of random variables $\{v_t, t \in N\}$ defined on a given probability space and indexed by the parameter $t \in N = \{0, 1, 2, \dots\}$. We can assume that v_{t+1} follows a first-order Markov process, i.e.

$$\begin{aligned} & \Pr(v_{t+1} = \sigma_k / v_1 = \sigma_h, \dots, v_{t-1} = \sigma_i, v_t = \sigma_j) \\ &= \Pr(v_{t+1} = \sigma_k / v_t = \sigma_j) \quad \forall t \end{aligned}$$

where σ is the standard deviation, and also that the transition probabilities for this process are time invariant, that is:

$$\begin{aligned} & \Pr(v_{t+n+1} = \sigma_k / v_{t+n} = \sigma_i) \\ &= \Pr(v_{t+1} = \sigma_k / v_t = \sigma_i) \quad \forall n, t \end{aligned}$$

That means we model volatility as a time-homogeneous, first-order Markov Chain with $m = 3$ states, which are respectively (Decrease-Stable-Increase). It was first decided that stock prices could either fall (this is bearishness in a downtrend), remain unchanged (that is what we called Stable; this is an indecision in a sideways market) or rise (this is bullishness in an uptrend). When a stock market is on the rise, it is considered to be an up and coming economy, it is often considered as the primary indicator of a country's economy strength and development. A rise in share prices is usually associated with increased business investment and vice versa.

Volatility is simply conditional standard deviation and is computed by the given formula, then

$$\text{its unconditional: } \hat{\sigma} = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (r_t - \mu)^2}$$

As mentioned before, r_t is the return of an asset at time t and μ is an average return over T periods. We use conditional variance, σ^2 , as a measure of volatility but variance and standard deviation are already connected by a simple relationship.

From that, suppose $\mathbf{S} = (S_1, \dots, S_k)'$ denotes a vector of log-prices for k financial assets, and $\mathbf{r}_t = (r_{1t}, \dots, r_{kt})'$ denotes a vector of the observed log-returns for k financial assets at time t for $t = 1, \dots, T$. We assume that the conditional mean of \mathbf{r} is zero for expositional purposes. Note that the conventional first-order Markov Chain model for k financial data sets of m states has m^k states.

Let $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \dots, \varepsilon_{kt})$, $\mathbf{V}_t = (v_{1t}, \dots, v_{kt})$. A Multivariate model of returns is then defined as:

$$\mathbf{r}_t = \mathbf{V}_t \boldsymbol{\varepsilon}_t \quad (3.1)$$

where $\mathbf{V}_t' \mathbf{V}_t = \boldsymbol{\Omega}_t$ is the $k \times k$ volatility matrix of \mathbf{r}_t , and $\boldsymbol{\varepsilon}_t$ is a $k \times 1$ vector of White-noise error which are independent of \mathbf{V}_{t-j} for all $j \geq 0$, and k the number of assets. The model of returns has to focus both the distribution of the shocks $\boldsymbol{\varepsilon}_t$ and the functional form of the volatilities $\boldsymbol{\Omega}_t$. In order to illustrate the key elements of our strategy, we assume that the dynamics of volatility is governed by a first-order Markov chain properly parameterized.

Let's assume that $\boldsymbol{\Omega}_{t+1}$ follows a first-order Markov process, i.e.

$$\mathbf{P}(\mathbf{V}_{t+1} = \boldsymbol{\sigma}_k / \mathbf{V}_1 = \boldsymbol{\sigma}_h, \dots, \mathbf{V}_{t-1} = \boldsymbol{\sigma}_i, \mathbf{V}_t = \boldsymbol{\sigma}_j)$$

$$= P(\mathbf{V}_{t+1} = \boldsymbol{\sigma}_k / \mathbf{V}_t = \boldsymbol{\sigma}_j), \quad \forall t$$

In fact, we assume here that the volatilities that will be effective in a period depends only on the current volatilities for a period.

Commonly the return shocks $\boldsymbol{\varepsilon}_t$ are assumed to be normal and \mathbf{r}_t is conditionally normal, while the unconditional distribution of \mathbf{r} is non-normal, and can exhibit the expected stylized facts about returns, such as fat tails and volatility clusters. The variance of the shock returns is not constant over time or the volatility is clustering.

In the SV case, the volatilities are a dynamic latent variable and estimation is nontrivial since the volatilities have to be integrated out of the joint density for returns and volatilities. The stochastic volatility specification has several advantages over the GARCH class of models e.g. they are much more closely integrated with microeconomic theory (Anderson,1994).

The expression of the volatility is:

$$\mathbf{V}_{t+1} = \boldsymbol{\sigma}' \mathbf{x}_{t+1} \quad (3.2)$$

and

$$\mathbf{x}_{t+1} = \mathbf{P}' \mathbf{x}_t + \mathbf{e}_{t+1} \quad (3.3)$$

where \mathbf{x}_t is the state of today and $\boldsymbol{\sigma}'$ an $k \times 1$ vector of σ_m that specifies the volatilities mass points and where each $\sigma_m = (\sigma_1, \sigma_2, \dots, \sigma_M)'$.

We can represent an M-states Markov chain in terms of a $M \times 1$ vector \mathbf{x}_t whose each j^{th} element equals 1 if the process is in state $j \in \{1, 2, \dots, M\}$ at time t and 0 otherwise; see Hamilton (1994). We have the state-transitions described by a VAR (1) process as below:

$$\mathbf{x}_{t+1} = \mathbf{P}' \mathbf{x}_t + \mathbf{e}_{t+1}$$

where \mathbf{P} is a $M \times M$ transition matrix with $\mathbf{P}_{\sigma_k \sigma_j} = P(\mathbf{V}_{t+1} = \boldsymbol{\sigma}_k | \mathbf{V}_t = \boldsymbol{\sigma}_j) \forall k, j \in M$ and \mathbf{e}_{t+1} is a vector martingale difference sequence, i.e., $E(\mathbf{e}_{t+1} / \mathbf{x}_1, \dots, \mathbf{x}_{t-1}, \mathbf{x}_t) = 0$.

If we assume that for the 3-variate case, we work with 3-states (Decrease-Stable-Increase) then, the probability transition matrix is of the form:

$$\mathbf{P} = \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix}$$

Calculating the state Probabilities of the posterior closing days

Let us denote the state probabilities matrix in different periods and for the portfolio of stocks by $n_{k,i}$.

Where $k = 1, 2$ (the number of stocks), $i = 1, \dots$ (the different periods) and $n_{k,(i+1)} = n_{k,i} * P_{(ij)}^k$

Thus, $n_{k,(i+n)} = n_{k,i} * P_{k,(ij)}^{(n)}$

Depending on the state of the closing share prices on the last day “of trading” period under study (i.e the 1756th day) with no follow-up information, it will be regarded as the initial matrix. Suppose they are all at “**Stable**” state

$$n_{k,(0)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

By virtue of the initial state matrix, we can predict state probabilities and transition matrix of various closing date in future. For example, we can obtain the state probabilities closing prices on the 1757th day as:

$$n_{k,(1)} = n_{k,(0)} * P_{k,(ij)}$$

State probabilities matrix of closing prices on 1758th day will be $n_{k,(2)} = n_{k,(1)} * P_{k,(ij)}$ and so on.

Initial state vector-matrix of portfolio shares prices

Each closing day is taken as a discrete time unit and the closing share prices of each stock of the portfolio are divided into three states as Decrease (D), Stable (S), Increase (I).

If we let $X_{1k} = D$, $X_{2k} = S$, $X_{3k} = I$ or more precisely, let $X_{11} = D^1$, $X_{21} = S^1$, $X_{31} = I^1$ (for the first stock (Equity)). Then, $X_{12} = D^2$, $X_{22} = S^2$, $X_{32} = I^2$ (for the second one (KCB)), where $X_{i,k}$ are the number observations for the shares prices in the named states gathered over the

period of study, then the state space is: $E \begin{pmatrix} X_{11} & X_{21} & X_{31} \\ X_{12} & X_{22} & X_{32} \end{pmatrix}$

Note that a state probability is the possibility size of emergence of a variety of states. Therefore, the state matrix here is denoted by:

$$n_{(i,k)} = \begin{pmatrix} X_{11}/1756 & X_{21}/1756 & X_{31}/1756 \\ X_{12}/1756 & X_{22}/1756 & X_{32}/1756 \end{pmatrix}$$

Where $i = 1, 2, 3$ (is the number of states) and $k = 1, 2$ (is the number of stocks)

Then, the probabilities for share prices that decrease (D) are stable (S) and increase (I) in the portfolio is:

$$n_{(0)} = \begin{pmatrix} 0.3963 & 0.2340 & 0.3695 \\ 0.4322 & 0.1771 & 0.3906 \end{pmatrix}$$

Establishment of the three states Transition Matrix

Let the transition matrix involve three states; the states are the fact that from a given started point each stock Decrease, doesn't change (Stable) or Increase. Let's now compile the transitions from one state to another for the given portfolio stocks from our data panel.

The transitions from one state to another are:

Stock	Decrease	Stable	Increase	Sum
Decrease	n_{11}	n_{12}	n_{13}	n_{1j}
Stable	n_{21}	n_{22}	n_{23}	n_{2j}
Increase	n_{31}	n_{32}	n_{33}	n_{3j}

We have:

Equity	Decrease	Stable	Increase	Sum
Decrease	279	148	269	696
Stable	159	114	138	411
Increase	261	148	240	649

KCB	Decrease	Stable	Increase	Sum
Decrease	259	79	421	759
Stable	205	106	0	311
Increase	296	126	264	686

Each number in these tables refers to the number of times a transition has occurred from state i to state j for each stock. We computed the transition probability in order to get the desired transition matrix for the portfolio of stocks. We used the previous constructed transition probability matrix and in conjunction with the probabilities values of the present state of the

system in order to determine the probability of the next states. Here $P_{ij} = \frac{n_{ij}}{\sum n_{\cdot j}}$ and

$P_{ij}^k = \Pr(\mathbf{V}_{t+1} = \sigma_i | \mathbf{V}_t = \sigma_j)$, we have the following transition probability matrix:

$$P_{ij}^k = \begin{pmatrix} 0.40 & 0 & 0.21 & 0 & 0.39 & 0 \\ 0.39 & 0 & 0.28 & 0 & 0.34 & 0 \\ 0.40 & 0 & 0.23 & 0 & 0.37 & 0 \\ 0 & 0.34 & 0 & 0.10 & 0 & 0.55 \\ 0 & 0.66 & 0 & 0.34 & 0 & 0.00 \\ 0 & 0.43 & 0 & 0.18 & 0 & 0.38 \end{pmatrix}$$

The model allows for leverage effects and time varying correlation. Hence, it is more flexible than those other models, the estimated model should be the same as the Multivariate model of Danielson (1998) and Harvey et Al (1994).

Assume that in the equation of the multivariate model of returns, the return shocks $\boldsymbol{\varepsilon}_t$ are multivariate normal. From the definition of the $k \times k$ matrix of volatilities Ω_t , where the covariance matrix, ϕ_t is defined by:

$$\phi_t = \Omega_t' \Gamma \Omega_t$$

where Γ is the matrix of correlation coefficient defined by $\mathbf{E}[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t']$ and:

- $-1 < \Gamma_{ij} < 1 \quad \forall i \neq j$
- $\Gamma_{ij} = 1 \quad \forall i = j$

The covariance matrix will always be positive definite since Γ is positive definite. In the MSV model, variances and correlations are instantaneous stochastic variables.

In the discrete state-space framework, the autocorrelation function of volatility is determined by how we parameterize the transition matrix of the Markov chain. The most transparent way to obtain a multivariate first order autoregressive model is to specify P such that it is immediately apparent. We have:

$$\Omega_{t+1} = \zeta + \varphi\Omega_t + \eta_{t+1} \quad (3.4)$$

This is an AR (1) model of volatility, where φ is the first order autocorrelation coefficient of volatility and, $\eta_{t+1} = \sigma' e_{t+1}$ is white noise. We can set

$$P = \varphi I_M + (1 - \varphi) 1_M \pi' \quad (3.5)$$

where $\varphi \in [0,1)$, I_M denotes an $M \times M$ identity matrix, 1_M is the $M \times 1$ vector of 1's and π is the $M \times 1$ vector of ergodic probability with $\pi = E(x_t)$, with x_t denoting the state.

By substitution of equation (3.5) into equation (3.3) and simplification, we obtain the volatility process in equation (3.4) with $\zeta = (1 - \varphi)\sigma'\pi$. This is a straightforward approach for formulating multivariate first order discrete stochastic volatility model.

Therefore, we will refer equations (3.2), (3.3) and (3.5), as a Discrete Multivariate Stochastic Autoregressive Volatility (1, M) model, this designator conveys its two most important features. Volatility follows a discrete AR (1) process, and allows for M different realizations of volatility. We then assume that the distribution of volatility has discrete support for achieving the computational tractability of our approach.

The choice of N and M controls the degree of approximation error, if we wanted to approximate a continuous stochastic AR volatility process in which the marginal distribution of volatility is log normal, we could parameterize π and σ by specifying the mean and variance of the log normal distribution. Formulating a multivariate higher-order of the previous DMSARV model will require some modifications to the methods developed above and this is an interesting avenue for future research.

3.2 Parameterization Strategies

Since M is small (i.e. ≤ 3), the multivariate stochastic volatility specifications developed below are slightly parameterized because each π_k and σ_k have M elements for $k = 1, 2$. We will impose two additional restrictions on the parameter space for more specifications. Let parameterize π_k and σ_k as:

$$\sigma_{jk} = \delta_k + \gamma_k j, \quad k = 1, 2 \text{ and } j = 1, 2, \dots, M \quad (3.6)$$

where each $\gamma > 0$ and $\delta > -\gamma$

Volatility mass points are evenly spaced along a line. Now, π_k can be parameterized as:

$$\pi_{jk} = \frac{(M-1)!}{(j-1)!(M-j)!} \omega_k^{j-1} (1-\omega_k)^{M-j} \quad (3.7)$$

$j = 1, \dots, M$ with $M = 3$ and $k = 1, 2$, where $\omega \in (0, 1)$.

By imposing these two restrictions, we obtain a class of models that have only 8 parameters regardless the size of the state space.

By extending this approach to a log linear specification, this yields

$$\log \sigma_{jk} = \delta_k + \gamma_k j \quad j = 1, \dots, M \quad (3.8)$$

where $\gamma > 0$ and the value of δ is unrestricted, one in which the mass points of the log-volatility are evenly spaced along a line. Changing how to parameterize σ has no effect on the basic time series of volatility, Ω_t still follow a discrete AR (1) process.

Therefore, knowing that parameterization strategies of σ is an important factor for the approximation of the marginal distribution of volatility, however, the evidence from the realized-volatility literature suggested that the marginal distribution of volatility is much closer to log normal than to normal. Then, we use the parameterization which is more in line with log-normality and expect it to better fit the data.

Another parameterization of σ that offers greater flexibility in the positioning of the volatility mass points is done by replacing the linear functions in the previous equations with polynomials of any order less than M , That is

$$\log \sigma_{jk} = \delta_k + \gamma_k j + \beta_k j^2 \quad (3.9)$$

where $\gamma > 0$ and the values of δ and β are unrestricted, allowing the mass points of log-volatility to take on a quadratic configuration, provided that $M = 3$.

The normal distribution has the fourth moment equal to 3, although some papers have shown that the distribution of market returns have sample fourth moments larger than 3. Also, prices movements are negatively correlated with volatility, this means that the volatility of shock tends to increase when the stocks prices fall, decrease when the stock prices rise and null when the stock prices are stable, since the stock market prices are highly fluctuating.

The simplest way to estimate volatility is taking daily squared returns, Unfortunately, this method gives an inaccurate estimation of volatility (Taylor, 1986); we then calculate daily returns using the closing price of each asset in the end of a trading session.

3.3 Models with Asymmetric Volatility

There is an asymmetric relation between stock prices changes and the volatility of future stock returns as shown previously. Therefore, the source of this asymmetry has been explained in the literature of Adriana and Kirby (2014); the common explanations are known as the leverage hypothesis and the volatility feedback hypothesis.

It is known that the leverage hypothesis asserts that a fall in the stock market price leads to an increase in financial leverage, which makes the stock a riskier investment and that can create a decreasing need to invest, and causes its volatility to increase, while the volatility feedback hypothesis asserts that the risk premium demanded by investors increases whenever they expect volatility to increase, and this increase in the risk premium immediately causes a decrease of the stock prices.

In that case, in order to capture these effects, we may allow the transition probability for the volatility process to be variant over the time.

Let consider the following model

$$\begin{aligned} \Omega_{t+1} &= \sigma' x_{t+1} \\ x_{t+1} &= P_t' x_t + e_{t+1} \end{aligned} \quad (3.10)$$

with P_t a time varying transition matrix, denoted by

$$P_t = \phi I_M + (1 - \phi) 1_M \pi_t' \quad (3.11)$$

Here π_t is no more $E(x_t)$ but $f(r_{it}, r_{it-1}, \dots, r_{i1})$ with $i = 1, \dots, k$. The time-varying transition probability still requires some parameterization decisions and since it's well-known that the volatility follows a discrete AR (1) process, and that P_t depends on predetermined and exogenous variables (see Diebold et Al. (1994), for further discussion). The transition probabilities for Ω_{t+1} are a function of only the lagged returns which are those predetermined variables. Let,

$$\Omega_{t+1} = \omega_t + \phi \Omega_t + \eta_{t+1} \quad (3.12)$$

This model is described by a discrete AR (1) process with T-varying intercept, where $\omega_t = (1 - \phi) \sigma' \pi_t$. This process can capture asymmetric volatility effects because it allows the expected value of Ω_{t+1} knowing Ω_t to be correlated with r_t and $r_{t-1}, r_{t-2}, \dots, r_1$. The correlation between returns and volatility like that implied by the leverage and volatility-feedback hypothesis, is generated by having negative returns in periods t and earlier to be associated with changes in π_t that increase the value of ω_t . Another parameterization of π_t is

$$\pi_{jt}^k = \frac{(M-1)!}{(j-1)!(M-j)!} (\omega_t^k)^{j-1} (1 - \omega_t^k)^{M-j} \quad (3.13)$$

$$j = 1, \dots, M$$

Let's specify a binomial inspired parameterization for π_t

$$\pi_{jt} = \frac{(M-1)!}{(j-1)!(M-j)!} \omega_t^{j-1} (1 - \omega_t)^{M-j} \quad (3.14)$$

$$j = 1, \dots, M$$

where the time varying parameter ω_t is:

$$\omega_t = \frac{\exp(\eta + \psi r_t + \psi \rho r_{t-1} + \psi \rho^2 r_{t-2} + \dots + \psi \rho^{t-1} r_1)}{1 - \exp(\eta + \psi r_t + \psi \rho r_{t-1} + \psi \rho^2 r_{t-2} + \dots + \psi \rho^{t-1} r_1)} \quad (3.15)$$

with $\rho \in [0,1)$

In this case, the sign of ψ controls the strength and direction of the asymmetric volatility effect. If we set $\psi < 0$, that means, it gives us a model in which negative returns are associated with increases in expected future volatility. On the other side, ρ controls the rate at which this asymmetric volatility response diminishes with time. If at time t , a negative return tells us that the volatility is expected to increase in the future, then this expected increase could be entirely transitory if $\rho = 0$, if $\rho = 0.5$ then it could be moderately persistent and highly persistent of and only if it is close to one i.e. $\rho = 0.9$

3.4 Model with time varying volatility persistence

We would like to formulate a model that displays time-varying volatility persistence, allowing for time-varying transition probabilities, we therefore allow the transition probabilities matrix for the volatility process to vary over the time by assuming it is selected in a stochastic manner for each t .

If we suppose that $\{y_t\}_{t=1}^{\infty}$ is a stochastic process with discrete support such that $Y_t \in \{1, 2\}$ for all t and $y_t = (y_{1t}, y_{2t})'$ a $k \times 1$ vector. If we let Y_t be generated by a time-homogenous ergodic and irreducible 3-state Markov chain, then we can express the transition probabilities for Y_{t+1} as

$$\begin{aligned} \Pr(Y_{t+1} = j / Y_t) \\ = (1 - \varphi) \left((1 - \bar{\omega})(2 - j)(3 - j) + \bar{\omega}(j - 1)(j - 2) \right) + \varphi 1_{[b_{t=j}]} \end{aligned}$$

where $\varphi \in [0,1)$ and $\bar{\omega} \in (0,1)$

Let $x_t^{(y)}$ denote a 3×1 vector whose j^{th} element equals one if $Y_t = j$ and zero otherwise.

To obtain a general multivariate stochastic volatility model for volatility that displays time-varying volatility persistence, we assume that the joint transition probabilities of Ω_{t+1} and Y_{t+1} are given by

$$\begin{aligned} \Pr(\Omega_{t+1} = \sigma_k, Y_{t+1} = j / \Omega_t, Y_t) \\ = \left((1 - \phi_j) \pi_k + \phi_j 1_{[\Omega_t = \sigma_k]} \right) \Pr(Y_{t+1} = j / Y_t) \end{aligned}$$

with $\phi_j \in [0, 1]$ for $j \in \{1, 2, 3\}$, and $\phi_1 > \phi_2 > \phi_3$.

3.5 Estimation Method

As S.V models typically do not have a closed-form expression for the likelihood function, the estimation of the parameters for the wide range of univariate and multivariate S.V models has attracted significant attention in the literature. An important concern for the choice of a particular estimation method lies in its efficiency. We applied the Simulated Maximum Likelihood techniques to the estimation of multivariate Stochastic Autoregressive Volatility models.

3.5.1 Simulated Maximum Likelihood (SML) Method

The SML method introduced by Danielsson and Richard (1993) depends on Monte Carlo integration to evaluate the likelihood. The likelihood function of multivariate stochastic volatility models involves high-dimensional integration, which is difficult to calculate numerically. Nevertheless, estimation of the parameters can be based on evaluating high-dimensional integrals with simulation methods and then maximizing the likelihood function, resulting in the so-called SML estimators. There are several ways to perform SML estimation for multivariate stochastic volatility models, the most usual approach to SML is the importance sampling method. The basic idea of this method is to approximate first the integrand by a multivariate normal distribution using the so-called Laplace approximation and then draw samples from this multivariate normal distribution.

Let $f(\mathbf{R}_{t+1} | \Omega_{t+1}, I_t; \boldsymbol{\theta})$ be the joint probability density function of \mathbf{R}_{t+1} conditional on observing both Ω_{t+1} and $I_t = \{\mathbf{R}_t, \mathbf{R}_{t-1}, \dots, \mathbf{R}_1\}$, with $\boldsymbol{\theta}$ a vector of unknown parameters

which is estimated by maximum likelihood. In order to fit our model in equation [3.2], [3.3], and [3.5], let's assume that

$$\mathbf{R}_{t+1} | \Omega_{t+1}, I_t \sim N(0, \sigma_{t+1}^2) \text{ where } \Omega_{t+1} = \mathbf{V}'_{t+1} \mathbf{V}_{t+1}$$

$$\text{And } \mathbf{V}_{t+1} = \boldsymbol{\sigma}' \mathbf{x}_{t+1}$$

$$\mathbf{x}_{t+1} = (\phi \mathbf{I}_N + (1 - \phi) \mathbf{1}_N \boldsymbol{\pi}')' \mathbf{x}_t + \mathbf{e}_{t+1}$$

$\boldsymbol{\mathcal{G}} = (\delta, \gamma)'$ since the normal distribution is determined by its mean and variance, in the case where σ is parameterized as $\sigma_j = \delta + \gamma j$, $j = 1, \dots, M$.

Now, let $\mathbf{x}_{t+1|t} = E(\mathbf{x}_{t+1} / I_t)$ denotes the expectation of the $M \times 1$ vector \mathbf{x}_{t+1} given the period t information set. Hamilton (1989) show that $\mathbf{x}_{t+1|t}$ is given by

$$\mathbf{x}_{t+1|t} = \mathbf{P}' \left(\frac{\mathbf{x}_{t|t-1} \odot \boldsymbol{\eta}_t}{\mathbf{1}'_N (\mathbf{x}_{t|t-1} \odot \boldsymbol{\eta}_t)} \right)$$

where $\boldsymbol{\eta}_t = (\eta_{1t}, \dots, \eta_{Mt})'$ is a $M \times 1$ vector with j^{th} element,

$$\eta_{jt} = f(\mathbf{R}_t | \Omega_t = \sigma_j, I_{t-1}; \boldsymbol{\mathcal{G}})$$

Then we can write the log likelihood function as

$$L(\boldsymbol{\theta}) = \sum_{t=1}^T \log \mathbf{1}'_N (\mathbf{x}_{t|t-1} \odot \boldsymbol{\eta}_t)$$

where $\boldsymbol{\theta}$ contains both parameters that determine the transition probabilities and those contained in $\boldsymbol{\mathcal{G}}$, with \mathbf{P} parameterized as in equation (3.5) and $\boldsymbol{\pi}$ is as in equation (3.7). We then use a quasi-Newton method to find the value of $\boldsymbol{\theta}$ that maximizes $L(\boldsymbol{\theta})$ and we compute standard errors using the second-derivative estimate of the information matrix. In order to select the model that fit better the data, we measure the performance of the out-of-sample variance forecasts produced by various models, and we require a proxy for the unobserved variance of daily returns.

For example, if we want to evaluate one-step ahead forecasts, we might fit a regression of the following form

$$R\Omega_{t+1} = a + b\hat{\sigma}_{t+1|t}^2 + \varepsilon_{t+1}$$

where $R\Omega_{t+1}$ is the realized variance joint variance for period $t + 1$ and $\hat{\sigma}_{t+1|t}^2$ is constructed using maximum likelihood estimates of the model parameters, and models are rank using the regression R-squared.

To conduct formal comparisons of the various models under study, we will use either Akaike Information Criterion or Bayesian Information Criterion tests.

3.6 Models Selection Criteria

In financial modelling, one of the main challenges is to select an adequate model from the panoply of models in order to achieve the goal of accurate volatility forecasting in a given data set. The choice of a good model in the application of time series analysis is crucial, since there is not a perfect or a unique model. Models selection criteria provide useful tools in this regard and assess whether one model is better than another one in fitting data. Ideally, a criterion will identify candidate models that are either too simplistic to accommodate the data or unnecessarily complex, the most common modes selection criteria are the AIC and the BIC. We then use them to select the best model in each family of candidate models. The model with the least values of these information criteria is the best.

3.6.1 The Akaike Information Criterion

Akaike (1974) introduced the AIC as an extension to the ML principle and this was the first model selection criterion to gain widespread acceptance. The AIC is defined by

$$AIC = -2(\log \text{likelihood}) + 2n$$

where n is the number of parameters. One of the advantages of AIC is that it provides an asymptotically unbiased estimator of the expected Kullback discrepancy between the generating model and the fitted approximating model. However, it is not consistent and not enough to have an accurate decision on the choice of the best model.

3.6.2 The Bayesian Information Criterion

Schwarz (1978) later, also introduced the BIC which is useful for model comparison in its own right. It is defined by

$$\text{BIC} = -2\ln(\text{likelihood}) + n \ln T$$

where T is the number of observations or equivalently, the sample size and n denotes the number of parameters. BIC penalizes more complex models (especially those with many parameters) relative to simpler models. This definition permits multiple models to be compared at once, the model with the highest posterior probability is the one that minimize BIC. BIC provides a large-sample estimator of a transformation of the Bayesian posterior probability associated with the approximating model. It is consistent but not asymptotically efficient, reason why in our study we use both AIC and BIC for exact results.

Chapter 4

DATA ANALYSIS AND RESULTS

4.1 Introduction

Having explored the general theory of the family of models under study in chapters 2 and 3, this chapter is dedicated to fitting both family of models, the GARCH models and the MSV models to the NSE data. Therefore, in the first section we display a description of the data, where the general statistical features of the NSE data are investigated and the rest of the sections discuss the application of MGARCH models together with DMSARV models in real life data. R software was used to analyse the data.

4.2 Descriptive Statistics

The data sequences are generated by the same source. Daily closing prices of NSE Equity and KCB shares data over a period of 7 years extending from 01/01/2010 to 31/12/2016 with 1756 observations were used. The Equity and KCB shares are the most traded and most profitable companies trading in NSE market. They track the daily performance of the most capitalized companies in the sector of Banking among the eight (08) segments listed on the NSE. The choice of 7 years represents an attempt to balance the potential adverse impact of a phenomena (such as occasional structural breaks) against the desire for precise parameters estimates; we could obviously use a much longer sample period for the stocks. In order to make forecasts, the full sample was divided into two parts, in sample and out-of-sample observations. The sample period is January 1, 2010 to December 31, 2016 (1756 observations) while the out of sample covers one year: January 1, 2017 to December 31, 2017 (200 observations).

4.2.1 Assets returns

Most financial studies involve returns instead of prices of assets to forecast volatility. This is because the return of an asset is a complete and scale-free summary of the investment opportunity for average and aware investors, and returns series are easier to handle than price series because return series have more attractive statistical properties. (Giot and Laurent, 2001).

We used the daily percentage returns for the stock indices namely Equity and KCB stocks, in order to fit the discrete MSARV models.

Let P_t and P_{t-1} denote the closing asset prices of NSE assets at the current (t) and previous (t-1) day respectively. The NSE All share returns (log-returns or continuously compounded returns) at any time are given by:

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right)$$

4.2.2 Summary statistics of NSE returns series data

In order to describe the behaviour of NSE return series, we have drawn descriptive statistics table for the returns. The data are in log-difference form. The skewness, kurtosis, Kolmogorov test for normality, and correlation coefficients are used as the diagnostic tools under this study.

This is implemented by using the estimated mean, μ and the standard deviation, σ . The null hypothesis of normality is rejected if the p-valued of the Kolmogorov statistic is less than the significance level.

Table 4. 1 Summary statistics of NSE return series

Statistics	Equity Group	KCB Group Ltd
Observations	1756	1756
Max	0.0946	0.0878
Min	-0.1022	-0.1121
Mean	0.00043	0.00022
Variance	0.00037	0.00031
SD	0.0192	0.0176
Skewness	-0.093	-0.402
Kurtosis	7.3388	7.541
P-normal	< 5%	< 5%
Correlation coef		0.2409
Corr coef for squared returns		0.2803

The summary of the descriptive statistics for the NSE returns series are shown in Table 4.1. As it is expected for a time series of returns the mean is close to zero.

The return series are both negatively skewed, this indicates a distribution with an asymmetric tail extending toward more negative values. The kurtosis is greater than three for the normal distribution, this indicates that the underlying distribution of the returns are leptokurtic or heavy tailed. The series fail the Kolmogorov normality test statistic which rejects

normality at the 1% confidence level in both cases; that means they have positive excess kurtosis which confirms that the returns are effectively leptokurtic or heavy tailed. We can observe that the standard deviation of the daily returns shows little variation across the indices. We found out that the stock indices are a bit volatile and the least volatile stock index is the KCB Group, because it is the smallest from our sample in terms of market capitalization.

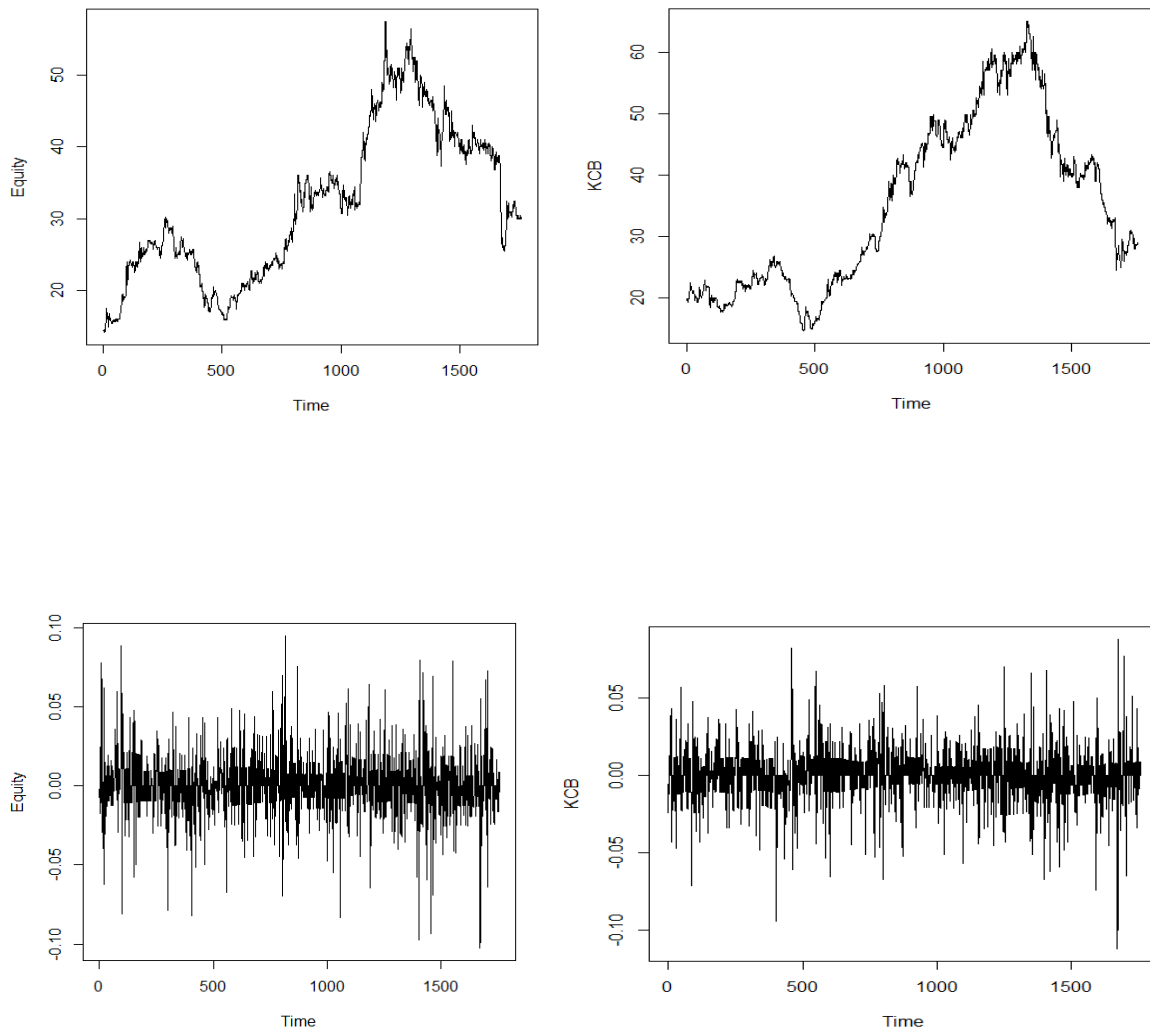


Figure 4.1 Equity and KCB daily prices and returns distributions (Jan 2010-Dec 2016)

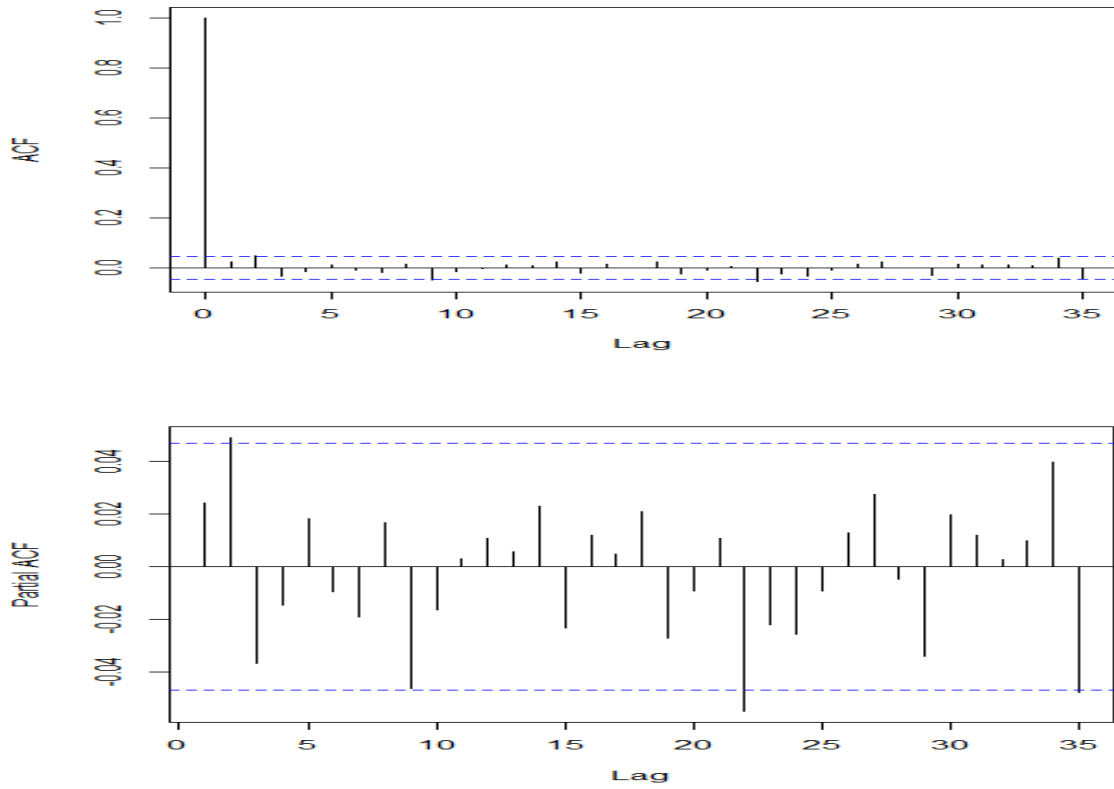
Figure 4.1 shows that, the stocks prices are non stationary while the return series are mean-stationary with a mean return of zero, there is also volatility clustering in the returns series. We can see that for the whole period of study the volatility is very high. We can interpret this as a result of macro economic factors like inflation and exchange rates.

4.2.3 ACF and PACF

Figure 4.2 shows that ACF of both return series are not significant, thus the returns are stationary. Hence, there is no need to test the mean effect (ARMA test). Moreover, the ACF of the return series in Figure 4.2.a and Figure 4.2.b show no evidence of serial correlation except at lag 0 and the only significant lag of the PACF of the return series in the same figures are lag 2. It is shown from the correlograms in figure 4.2 that p is obtained from the PACF while q is obtained from the ACF. Hence, the only competitive lag order (p, q) of the ARMA model is $(2, 0)$ which is an AR (2), that means it is the one that fits best to the NSE returns data. This means that the residuals of the ARMA $(2, 0)$ model can be used to test for ARCH effects before we apply the GARCH models.

Figure 4.3 shows that, the ACF of the squared returns are almost surely significant, this means that there is an ARCH effect. The PACF are both significant at lag 1 and lag 2. Using the ACF and PACF, the probable lag orders (p, q) of the GARCH (p, q) model are GARCH $(1, 1)$, GARCH $(1, 2)$, GARCH $(1, 3)$, GARCH $(2, 1)$, GARCH $(2, 2)$, GARCH $(2, 3)$.

4.2.a



4.2.b

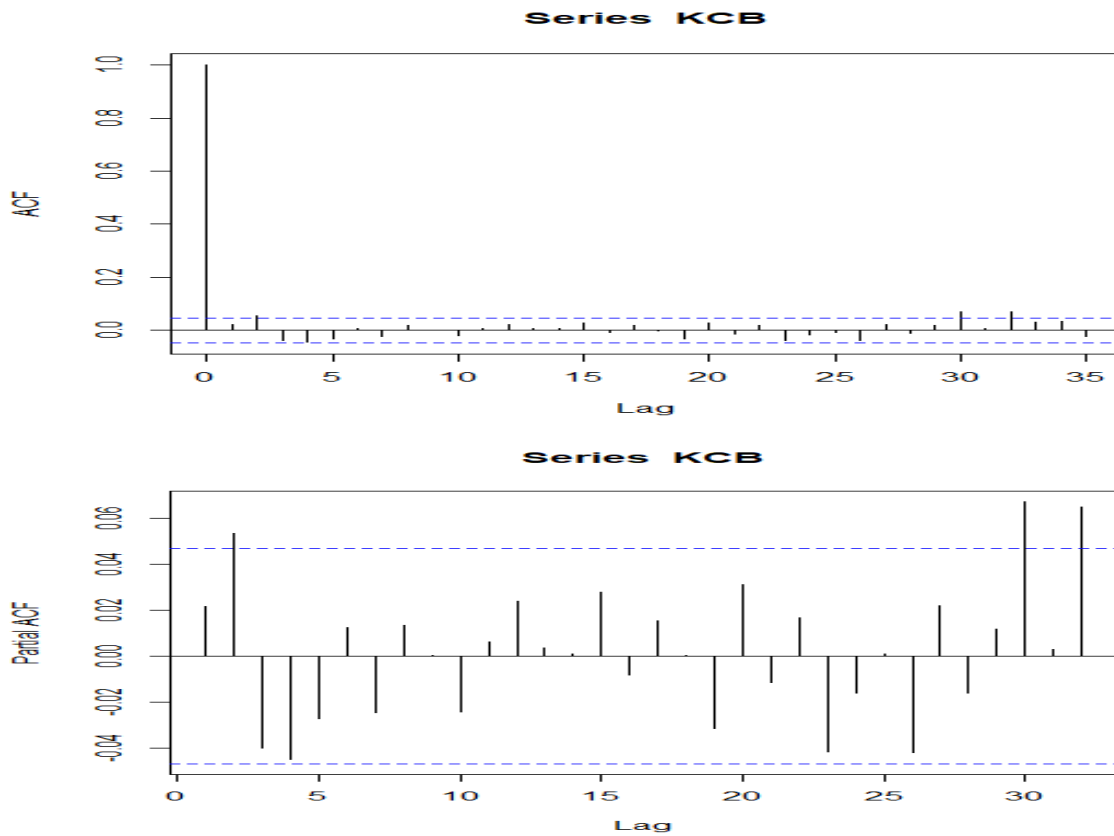
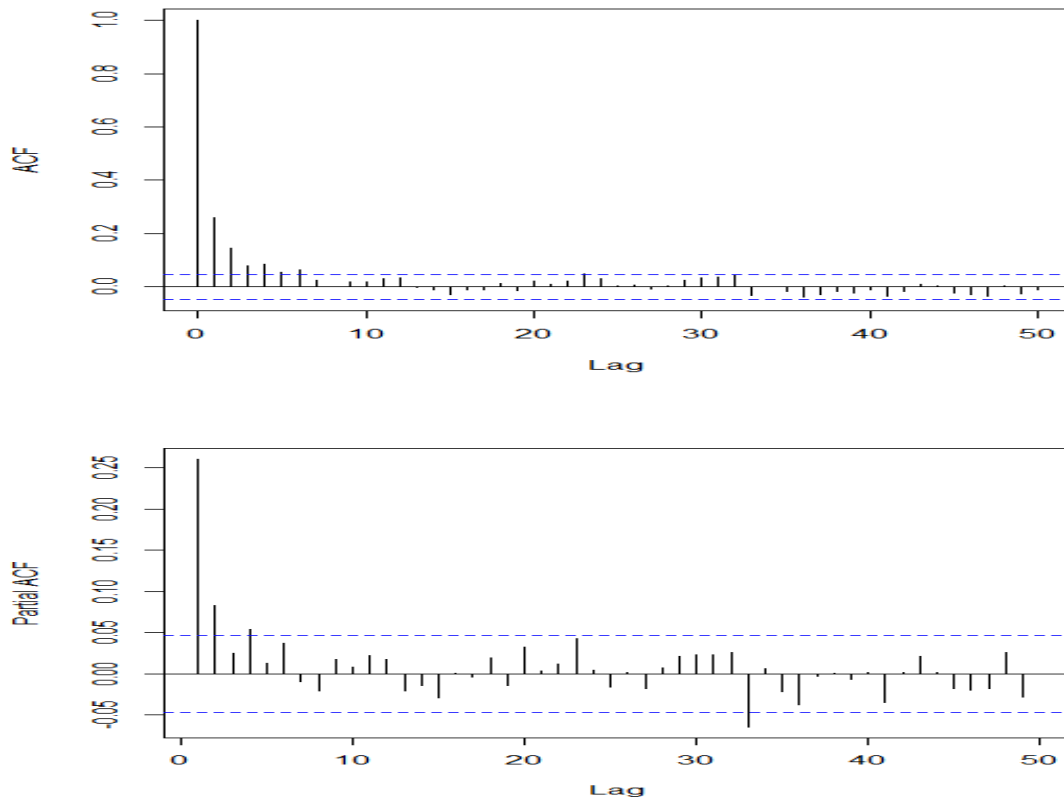


Figure 4.2 ACF and PACF of Equity and KCB Assets Returns

4.3.a



4.3.b

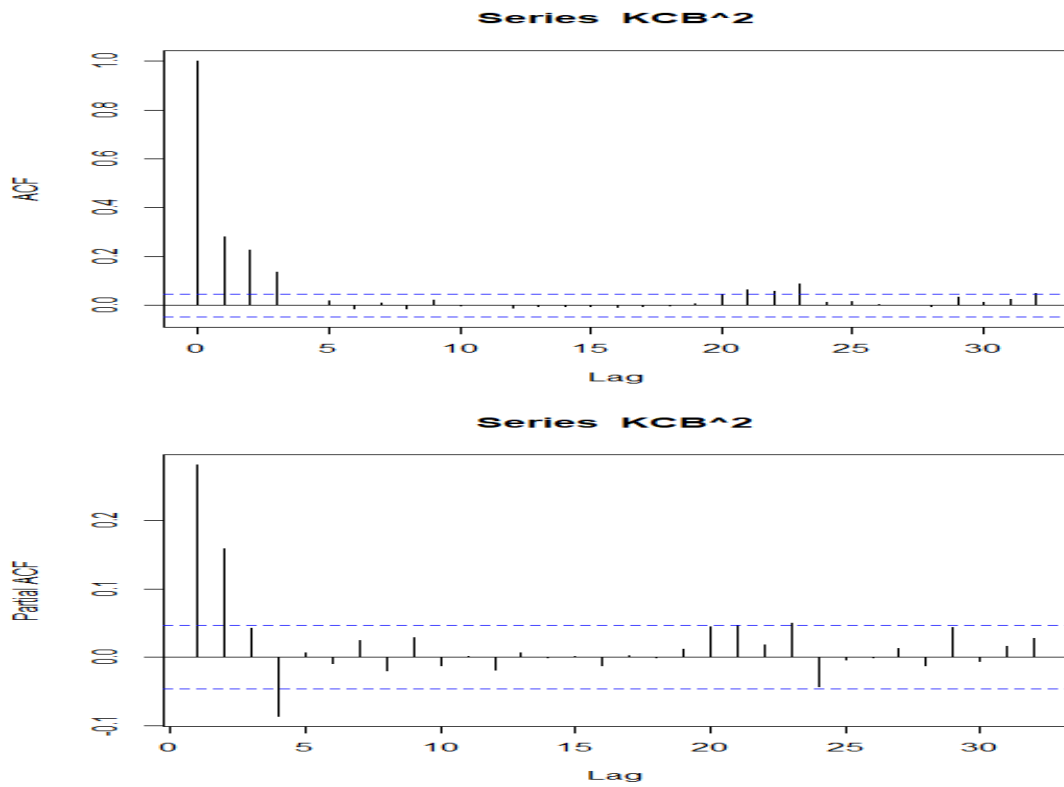


Figure 4.3 ACF and PACF of Equity and KCB Squared Returns

4.2.4 Testing for Stationarity

Before starting with the estimation of the parameters of our models, it is required to check whether the series are stationary. Under this study, ADF test by Dickey and Fuller (1981) was used to investigate the stationarity of our series. The test includes a constant term without trend.

Table 4.2 ADF test results

Series	ADF Statistic	P-Value (5%)
Equity Ret. Series	-13.517(10)	0.01
KCB Ret. Series	-13.295(10)	0.01

Table 4.2 shows the ADF test results for the two return series, where the Null hypothesis of a unit root is rejected at the level of significance, this means that they are both stationary with respect to the mean over the specified period.

4.2.5 Testing for Autocorrelation and Heteroscedasticity

Before GARCH models are applied to any data set, it is also required to test for conditional heteroscedasticity or ARCH effects, even though it has already been shown that the data has an ARCH effect based on the ACF of squared returns plots, we still need to confirm this by a test. To this effect, the Ljung-Box test is used in this study. This test was used to examine the existence of serial correlation, it checks whether the data are autocorrelated based on a number of lags m and this was done through the Ljung-Box Q-statistics given by

$$Q_m = T(T+2) \sum_{i=1}^m (T-i)^{-1} r_i^2$$

where r_i is the sample autocorrelation coefficient, T is the sample size and m is the max lag length. The Null hypothesis that all r_i are zero is rejected if the value of the test statistic Q is larger than the critical Q-statistic from the chi-square distribution at the given level of significance.

In order to test for ARCH effects, ARMA (2, 0) model for the conditional mean in the return series was employed as an initial regression. Then, the Null hypothesis that there are no ARCH effects in the residual series up to lag 36 were tested. The results are summarized in Table 4.3.

Table 4.3 Ljung-Box test for residuals of the NSE asset returns series

lag	Ljung Box of squared residuals				Ljung Box of residuals			
	Equity		KCB		Equity		KCB	
	Test Stat.	Prob.	Test Stat.	Prob.	Test Stat.	Prob.	Test Stat.	Prob.
12	198.53	$< 2.2e - 16$	264.09	$< 2.2e - 16$	13.628	0.3251	16.961	0.15
24	209.35	$< 2.2e - 16$	294.9	$< 2.2e - 16$	26.334	0.33	26.481	0.32
36	224.12	$< 2.2e - 16$	304.53	$< 2.2e - 16$	37.073	0.4193	53.594	0.029

As the ARCH effect in the residuals was already shown by the plot ACF of the squared returns, the results in Table 4.3 confirm the presence of ARCH effects in the residuals returns series of the mean equation, and an evidenced serial correlation. This means that the variance of the stocks returns series is non-constant and so GARCH models can be applied.

4.3 Empirical Results

The main findings of this thesis can be summarized as follows, we first provide a general view of our estimation results, then we take a look at each model separately, finally, we provide a comparison of all models and choose the suitable one.

4.3.1 Selection of GARCH (p, q) Model

We base the choice of the GARCH (p, q) model on AIC and BIC tests specially. The idea is to have a parsimonious model that best describes the data among the precedent candidate GARCH (p, q) models as indicated by ACF and PACF in figure 4.3.

Note that GARCH (1, 0) is similar to ARCH (1) model, in this case GARCH (1, 1) model is preferred to the ARCH (1) model because ARCH (1) model has few parameters and so can not describe adequately the volatility process of the assets data returns (Tsay, 2010).

After comparison of the different GARCH (p, q) models under study, GARCH (1,1) model revealed to be the best model to capture the ARCH effect in the residual returns of the AR(2) model, with least value of BIC and AIC. The results are shown in table 4.4.

Table 4.4 GARCH (p, q) models for the residual returns

Model	AIC		BIC	
	Equity	KCB	Equity	KCB
GARCH(1, 1)	-5.1738	-5.3427	-5.15517	-5.324
GARCH(1, 2)	-5.1725	-5.3415	-5.1507	-5.3196
GARCH(1, 3)	-5.1714	-5.3403	-5.1465	-5.3153
GARCH(2, 1)	-5.1729	-5.3414	-5.1511	-5.3196
GARCH(2, 2)	-5.1734	-5.3403	-5.1485	-5.3154
GARCH(2, 3)	-5.1721	-5.3391	-5.1441	-5.3111

A composite model was developed for the ARMA (2, 0) and GARCH (1, 1) models and the estimated parameters are shown in Table 4.6. Table 4.6 shows that all the estimated parameters of the ARMA (2, 0)-GARCH (1,1) are not statistically significant, thus AR (2)-GARCH (1, 1) model is refined to the simple GARCH (1, 1) model by dropping all the AR (2) parameters.

Table 4.6 Estimation results of AR (2)-GARCH (1, 1) model**Table 4.6 a Estimation results of AR (2)-GARCH (1, 1) model (Equity)**

Parameters	Estimate	Std.Error	t-Value	Pr(> t)
μ	1.77e-04	4.10e-04	0.432	0.666
ar1	-1.858e-02	2.92e-02	-0.635	0.5252
ar2	6.04e-02	2.66e-02	2.276	0.0229*
ω	1.05e-02	1.65e-05	6.338	2.33e-10***
α	2.34e-01	3.67e-02	6.378	1.17e-10***
β	4.92e-01	6.12e-02	8.038	8.88e-16***

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The variance equation is given by

$$\sigma_{1,t}^2 = 0.0105 + 0.234\varepsilon_{t-1}^2 + 0.492\sigma_{1,t-1}^2$$

This result supports that of Bollerslev et al (1992) where he states that simple GARCH models capture most of the variability in most stabilized series for volatilities even over long sample periods. The sum $\alpha + \beta = 0.7256$ and 0.7056 for Equity and KCB respectively, determine the rate at which the response function decays on daily basis. The half-life was also investigated;

the half-life volatility measures the time required for the volatility to move half way back towards its unconditional mean (Engle and Patton,2001). The half-life was estimated using the relation

$$\tau = \frac{\log((\alpha + \beta) / 2)}{\log(\alpha + \beta)}$$

Fitting the values, the estimated half-life is approximately 3 days in each case; this means that any shocks to this volatility takes approximately 3 days to return half-way back without any further shocks to this volatility.

Table 4.6 b Estimation results of AR (2)-GARCH (1,1) model (KCB)

Parameters	Estimate	Std.Error	t-Value	Pr(> t)
μ	1.62e-04	3.77e-04	0.43	0.6668
ar1	-2.886e-02	2.92e-02	-0.987	0.3234
ar2	7.08e-02	2.65e-02	2.671	0.0075**
ω	9.4e-05	1.4e-05	6.729	1.71e-01***
α	2.31e-01	3.70e-02	6.24	4.37e-10***
β	4.75e-01	5.96e-02	7.976	1.55e-15***

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The variance equation is given by

$$\sigma_{2,t}^2 = 0.000094 + 0.2306\varepsilon_{t-1}^2 + 0.4752\sigma_{2,t-1}^2$$

4.3.2 Selection of MGARCH (p, q) Model

Empirical results show that working with separate univariate models is much less relevant than multivariate modelling framework. We have investigated the empirical evidence of conditional volatilities and believe that such approach provides a comprehensive picture of world stock market co-movements. The selected models used in the empirical application were respectively the Dynamic Conditional Correlation of Engle, Constant Conditional Correlation of Bollerslev and the MDSARV models.

a. Estimation results

Several ARCH models were estimated with the data. In each model, conditional normality was assumed since the purpose of the estimation is a comparison with discrete multivariate stochastic autoregressive volatility models, the final univariate ARCH results are reported in Table 4.7.

Table 4.7 Univariate ARCH results. Log-likelihood values

Model	Number of Parameters	Equity	KCB
Observations		1756	1756
GARCH	3	4548.656	4696.95
EGARCH	4	4670.046	4821.672

And multivariate results are in Table 4.8. In all univariate cases, the EGARCH model outperforms the GARCH (1, 1) model in log-likelihood values with the largest difference in likelihood values (121.4) and (124.7) for Equity and KCB stock index respectively.

Table 4.8 Multivariate GARCH results: Bivariate Log-likelihood values

Model	Number of Parameters	Stock Indices
DVGARCH	9	9274.12
VGARCH	21	8877.301
CCC	9	9274.197
DCC	9	9279.317

For the multivariate models, the DCC model has the best performance. The CCC model, almost like DVGARCH performs significantly worse than the DCC model, and finally the VGARCH model is far the worst model for the stock indices. The difference in likelihood values is therefore indicative of which model fits better.

We continue with the DCC (1, 1) model. The DCC estimates of the conditional correlations between the volatilities and also estimates of the GARCH parameters are presented in Table 4.9. As the estimates of both dcc_a (the impact of past shocks on current conditional correlations) and dcc_b (the impact of previous dynamic conditional correlations) are statistically significant, this clearly indicates that the conditional correlations are not constant. The estimate of dcc_a is generally low and close to zero, while the estimate of dcc_b is high and close to one. The conditional correlations between the stocks indices are dynamic.

Table 4.9 Estimation of the coefficients of the DCC model for the stocks indices

DCC-Garch	ω_1	ω_2	θ_{11}	θ_{21}	θ_{22}
Est.	0.000193	0.000167	0.000102	0.229975	0.000096
Std. Error	0.00041	0.00041	0.000028	0.052596	0.000028
	ϕ_{11}	ϕ_{21}	ϕ_{22}	dcc_a	dcc_b
Est.	0.503067	0.225838	0.474611	0.018665	0.615445
Std. Error	0.1004	0.058419	0.110881	0.014236	0.214576

These findings are consistent with the plot of Dynamic correlations of the stocks indices in figure 4.4 which change over time. Figure 4.4 displays the estimated volatilities based on the DCC model. At first sight the graphs of these methods seem to imply very similar volatilities. The second model, which we considered was the CCC model, that allows contemporaneous dependence through conditional correlations.

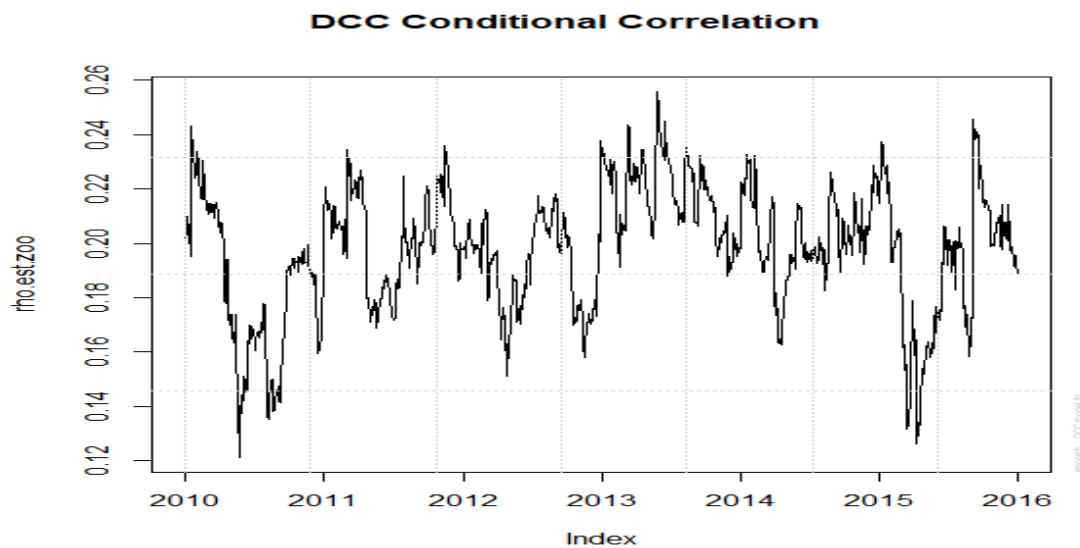


Figure 4.4 Estimated Dynamic Conditional Correlation of Stocks Indices: Equity and KCB

DCC correlation seems to be much more stable in estimates of stock market data, the smoother volatilities are provided by DCC model, and CCC estimates are more volatile than the other multivariate models. It is clear that correlations have changed greatly over the 7 years' period and exhibit time-dependence.

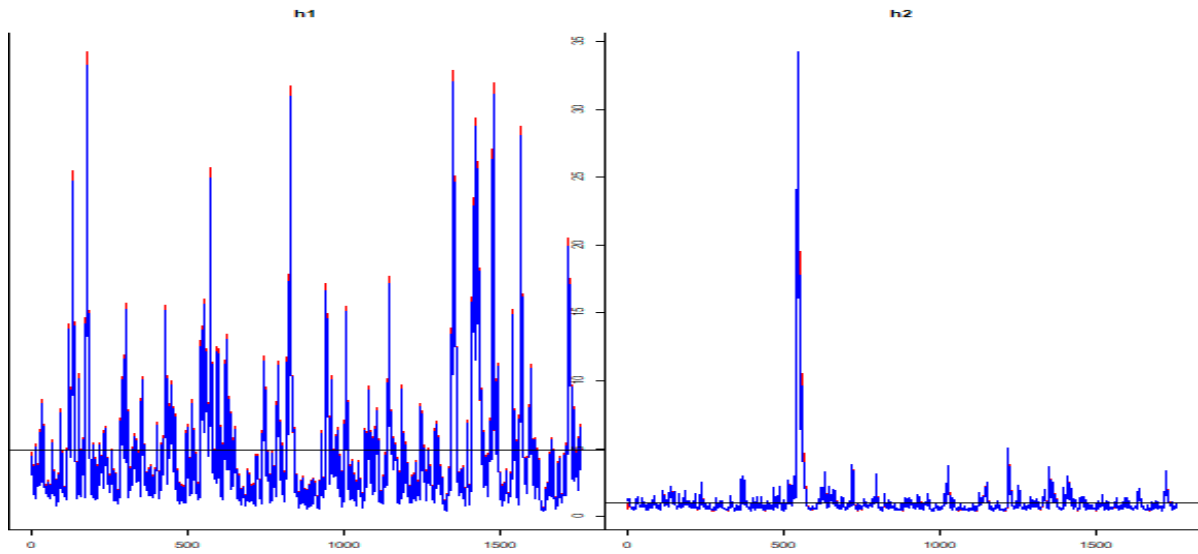


Figure 4.5 Estimated Constant Conditional Correlation of Stocks Indices: Equity and KCB

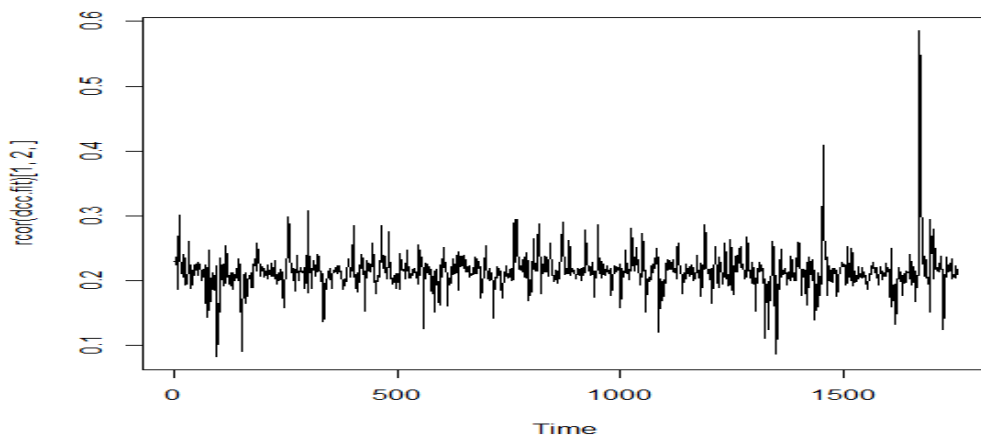


Figure 4.6 Estimated Dynamic Conditional Covariance of Stocks Indices

We can see from figure 4.4 and 4.5 quite similar level of dependence between Equity and KCB. The correlations between Equity and KCB are positively correlated and seem to be stable during a long period. From our sample, we finally notice that pair Equity and KCB represents for investors a good ability of portfolio diversification.

Table 4.10 Estimation of the coefficients of the CCC model for the stocks indices

CCC-Garch	ω_1	ω_2	A_{11}	A_{22}
Est.	0.09458	0.12559	0.19338	0.21159
Std. Error	0.01073	0.01738	0.005017	0.02295
	B_{11}	B_{22}	ρ	loglik.
Est.	0.70677	0.659882	0.39867	
Std. Error	0.009672	0.02856	0.02	9274.197

4.3.3 Discrete Stochastic Autoregressive Volatility Models

All the estimation in this thesis were performed by R programming software and we used the DCC and CCC models as benchmarks Multivariate models which have been revealed to fit better the data. To address the issue of model selection, we measured the performance of the in-sample and the out of sample variance forecasts produced by the various models. We required a proxy for the unobserved variance of daily returns. Fitting Mincer and Zarnowitz regressions is a common strategy for evaluating the forecasting performance of volatility models (Calvet and Fisher, 2004) and (Fleming and Kirby, 2013). If we are evaluating one-step-ahead forecasts for example, we might fit a regression of the form

$$\mathbf{RV}_{t+1} = \beta_0 + \beta_1 \hat{\sigma}_{t+1}^2 + e_{t+1} \quad (4.1)$$

where \mathbf{RV}_{t+1} is the realized variance for period $t+1$ variance based on the period t information set. Unbiased forecasts correspond to the hypothesis $\beta_0 = 0$ and $\beta_1 = 1$. And $\hat{\sigma}_{t+1}^2$ is constructed using maximum likelihood estimates of the model parameters and models are ranked using the regression R-squared. We used the Diebold and Mariano test of equal predictive accuracy to conduct formal comparisons. For example to compare model i and j

under a specified loss function $L(\sigma_{t+1}^2, \hat{\sigma}_{t+1}^2)$. Our null hypothesis is $E(e_{t+1}^{(ij)}) = 0$. Where,

$$e_{t+1}^{(ij)} = L(\sigma_{t+1}^2, \hat{\sigma}_{it+1}^2) - L(\sigma_{t+1}^2, \hat{\sigma}_{jt+1}^2) \quad (4.2)$$

denotes the loss differential for period $t+1$. To implement the test, we used the MSE loss function, we fitted to daily percentage returns, the regressions are estimated via OLS and the forecasts are for one-day horizon.

Let's start with the linear discrete SARV model. Results from estimation of DSARV (1, M) model of the data are presented in table 4.11. We fitted first-order discrete SARV models to daily percentage returns for the two stocks first individually with T=1756. All the specifications employ the linear parameterization of σ given by $\sigma_j = \delta + \gamma j$ with $j=1,2,\dots,M$. Table 4.11.a and 4.11.b report maximum likelihood parameter estimates for M=3 and M=10. And table 4.11.c reports the in-sample and out-of-sample model selection criteria for all values of M from 3 to 10.

Table 4.11 Estimation results for linear discrete SARV (1, M) models

Table 4.11. a DSARV (1, 3) parameters estimates

Parameter	Equity		KCB	
	Est.	Std.E	Est.	Std.E
ω	0.22	0.046	0.24	0.03
δ	-0.131	0.041	-0.392	0.066
γ	1.63	0.051	1.78	0.069
ϕ	0.91	0.014	0.922	0.015

Table 4.11. b DSARV (1, 10) parameters estimates

Parameter	Equity		KCB	
	Est.	Std.E	Est.	Std.E
ω	0.12	0.012	0.23	0.015
δ	0.51	0.023	0.121	0.028
γ	0.56	0.015	0.53	0.018
ϕ	0.93	0.009	0.938	0.009

Table 4.11. c BIC and R^2 of volatility forecasts

Model	Equity		KCB	
	BIC	R^2	BIC	R^2
DSARV(1,3)	4921.67	0.442	5090.51	0.4331
DSARV(1,4)	5002.58	0.5512	5025.16	0.5606
DSARV(1,5)	4928.4	0.6713	4977.65	0.658
DSARV(1,6)	5034.72	0.7374	5055.28	0.743
DSARV(1,7)	4987.67	0.802	4989.12	0.8149
DSARV(1,8)	4957.72	0.844	4956.42	0.8403
DSARV(1,9)	4859.71	0.8487	5021.37	0.87
DSARV(1,10)	4985.99	0.8695	5000.37	0.8907

The former is the BIC obtained by fitting the model and the latter is the R-squared for a regression of daily realized variances on the variance forecasts produced by the fitted model. The parameter estimates for the $M = 3$ display the expected characteristics, the estimates of δ and γ provide clear evidence of time-varying volatility for both stocks while the estimate of ω the ergodic probability of the high volatility state are all below 0.5. This implies that the process spends more time in the low-volatility state, and the estimates of ϕ is high for both stocks, this indicates a strong persistence in volatility. Increasing the volatility mass point to $M = 3$ changes all the estimates and the BIC decreases monotonically with M in each case; this suggests that it is suitable to work with $M \geq 3$.

The question is whether first-order discrete MSARV models capture the dynamics of volatility.

Table 4.12 Model selection criteria for linear, log linear and log-quadratic discrete MSARV (1, M) models

Model	Linear		Log-Linear		Log quadratic	
	BIC	R^2	BIC	R^2	BIC	R^2
DMSARV(1,3)	10073.74	0.4078	10054.82	0.897	10016.83	0.971
DMSARV(1,4)	9984.37	0.5818	9957.53	0.941	9915.89	0.985
DMSARV(1,5)	9977.49	0.6806	9947	0.961	9897.388	0.989
DMSARV(1,6)	9971.65	0.748	9935.94	0.973	9897.115	0.993
DMSARV(1,7)	9969.22	0.8	9924.11	0.98	9897.065	0.995
DMSARV(1,8)	9967.643	0.841	9907.74	0.9838	9850.042	0.997
DMSARV(1,9)	9925.37	0.8685	9889.73	0.987	9824	0.997
DMSARV(1,10)	9916.64	0.8933	9842.88	0.989	9816.065	0.997

From Table 4.12, we see how changing the parameterization of σ affects the performance of the model. We fitted the linear parameterization in equation (3.6), the log-linear parameterization in equation (3.8) and the log quadratic in (3.9), then we notice that from $M \geq 3$ the BIC values are decreasing and the parameterization for log volatility values lower than those for the volatility itself for every $M \geq 3$, this is because there is gain in moving to log parameterization. BIC values match with higher R-squared values, and the R-squared values are increasing with M . Moreover, we notice that almost all the values of R-squared are closed to one, this suggests that the model fits the data well; for example with $M \geq 3$, the R-squared improves from linear to log quadratic parameterization. It is worth noting that, since the BIC values diminish quickly as the value of M increases, this means there is benefit of increasing the number of states. We can conclude that the parameterization σ impacts on the

performance of the discrete multivariate SARV and shows the advantage of working with a portfolio of stocks instead of a single stock.

Finally, the model selection criteria suggest that there is more benefit fitting the log quadratic parameterization for the bivariate model, since for every number of state chosen from $M = 3$ the BIC decreases. We took a look at the dynamics of estimated conditional volatilities using all three models. We studied volatility dynamics of the returns by utilizing MGARCH and MDSARV models and then we reported statistically significant cross market effects as evidence of linkages and measured the extent of the linkages by the estimated time-varying correlations. The next part of this chapter is focused on comparison of the multivariate models.

4.4 Models Comparison

To conduct suitable pairwise comparison for the selected multivariate models, we used Diebold and Mariano to evaluate the forecast accuracy of the models by comparing the out-of-sample forecasting performance of selected first-order discrete multivariate model to that of the two benchmarks models above. We considered the discrete MSARV (1, 10) model, our benchmarks are a DCC and CCC models which are nonlinear combinations of univariate GARCH models.

Therefore, Table 4.13 reports the results of t-statistics for pairwise tests of equal predictive accuracy under MSE loss. We considered forecast horizons one day and one week; this is 5 trading days and the loss differentials for day $t + 1$ are computed as follow:

$$e_{t+1,H}^{(ij)} = \left(\mathbf{RV}_{t+1,H} - \hat{\sigma}_{it+1,H}^2 \right)^2 - \left(\mathbf{RV}_{t+1,H} - \hat{\sigma}_{jt+1,H}^2 \right)^2$$

With $H \in \{1, 5\}$, the forecast horizon, $i = (1, 2)$ and $j \in \{1, 2\}$ indicates the benchmark models. The null hypothesis for the test is $E\left(e_{t+1,H}^{ij}\right) = 0$. The t-statistics are based on robust standard errors that are constructed using Newey and West (1987) weight. The lag length for the weights is 10 for the one-day horizon and 20 for the 5-day horizon, and a negative or positive t-statistic indicates that the model produces a lower or higher loss on average than the benchmark model.

Table 4.13 Pairwise Comparison-Diebold Mariano test for stock indices

Model	DCC		CCC	
	1-Day	5-Day	1-Day	5-Day
DMSARV(1,3)	-0.84	-0.38	-0.95	-0.58
DMSARV(1,10)	-1.80	-1.61	-1.90	-1.70

4.5 Forecasting Performance

The forecasting ability of the models under study were also evaluated. In order to make forecasts, the full sample was divided into two parts, 1658 in-sample observations from 04/01/2010 to 05/08/2016 and 100 out-of-sample observations from 08/08/2016 to 30/12/2016. The forecast performance is shown in figure A below. Figure 4.7.1 shows that DCC model outperformed at volatility forecasting comparing the volatility of the original return series in Figure 4.6.

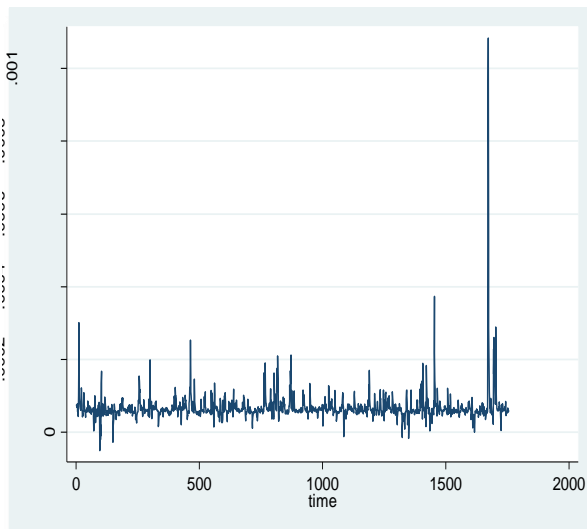


Figure4.7.1

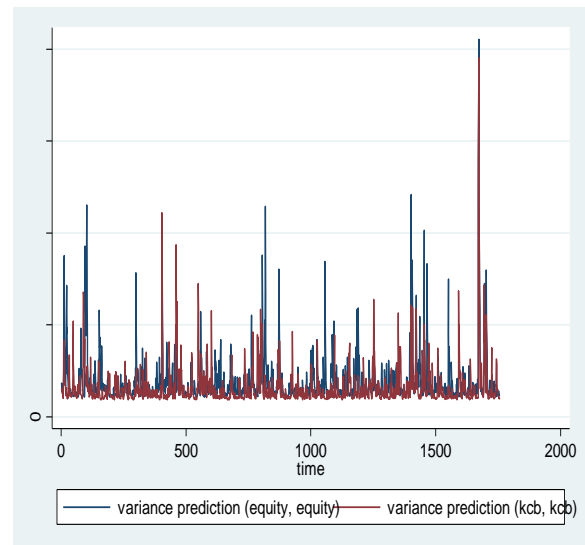


Figure4.7.2

Figure 4.7 Prediction of Conditional Covariance between Equity and KB and One-step prediction of volatility over the sample

Chapter 5

CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

The empirical analysis highlights the promise of our approach. Ultimately, volatilities and correlations among market returns are widely used in asset pricing. Although researchers have built many multivariate models, the multivariate stochastic volatility models remain the least mentioned in the literature. The multivariate volatility of NSE returns has been modelled and forecasted for a period of 1/01/2010 to 31/12/2016 using different GARCH-type model and by building on well-established techniques for constructing Markov chains with a specified autocorrelation function, we developed a multivariate stochastic volatility model in which volatility follows a low-order autoregressive process; the model specifications assume that volatility has discrete support. In this thesis, we also presented a theoretical and empirical modelling with some multivariate GARCH models and highlighted their features, we surveyed some MGARCH models and their basic constructions and we used maximum likelihood estimation procedures and SML estimation to estimate the models. Interestingly, however, the parameterization and estimation methods suggested here have already found use in other applications by Adriana and Kirby (2014), Mgr. Milan M. (2014), Choi (2011), ...among others. The results suggest that, the Equity market is influential in the pricing process of the KCB market and vice versa, and there is a close relationship between these two stock markets; therefore, investors may seize these stock markets as one investment opportunity instead of two separate classes of assets. One of the main findings is that volatility forecasts produced by first order discrete MSARV models outperform those produced by multivariate GARCH-type models and conditional correlations exhibit significant changes over time; these findings hold for the stock indices on the NSE. Therefore, there is opportunity to maximize portfolio returns through diversification.

5.2 Recommendations

Based on the results obtained from this research study, the following are the suggested recommendations for further study. Due to the evidence of high volatility persistence, it is recommended that DMSARV and DCC models are used in order to adequately describe the volatility process of the NSE returns.

Further research can be done to determine the effects of macroeconomic factors like exchange rate, inflation, taxes, exports on the stock prices using higher-order Discrete Multivariate Stochastic Autoregressive Volatility models.

We can conclude that, there is a number of interesting directions in which our analysis could be extended. One possibility is to investigate the performance of higher-order discrete MSARV models in assets returns context and this should be relatively straightforward.

References

- Adrianas, C., & Kirby, C. (2014). DISCRETE STOCHASTIC AUTOREGRESSIVE VOLATILITY. *Journal of Banking and Finance*, 43(1), 160-178.
- Akaike, H. (1974). A new look at the Statistical model identification . *IEEE Transactions on Automatic Control*, 19(6), 716-1974.
- AL-Najjar, D. (2016). Modelling and Estimation of Volatility Using ARCH/GARCH Models in Jordan's Stock Market. *Asian Journal o Finance and Accounting*, 8, 152-167.
- Anderson, T. (1994). Stochastic autoregressive volatility: A framework for volatility modelling. *Mathematical Finance*,4, 75-102.
- Baba, Y., Engle, R. F., Kraft, D., & Kroner, K. (1990). Multivariate Simultaneous Generalized ARCH. *mimeo, UCSD*.
- Bauwens, L., Laurent, S., & Rombouts, J. (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics*, 21, 79-109.
- Berkes, I., Horvath, L., & Kokoszka, P. S. (2003). GARCH Processes: Structure and estimation. *Bernoulli*, 9, 201-227.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity . *Journal of Econometrics*, 31, 307-327.
- Bollerslev, T. (1990). Modelling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH Model. *The Review of Economics and Statistics*, 498-505.
- Bollerslev, T., & Wooldridge, J. (1992). Quasi-Maximum likelihood estimation and Inference in dynamic models with time-varying covariances. *Econometrics Reviews*, 11, 143-172.
- Bollerslev, T., Chou, R. Y., & Kroner, F. K. (1992). ARCH modelling in finance* a review of the theory and empirical evidence. *Journal of Econometrics*, 52, 5-59.
- Bollerslev, T., Engle, R., & Wooldridge, J. (1988). A capital Asset Pricing model with time-varying covariances. *The Journal of Political Economy*, 96, 116-131.
- Calvet, & Fisher. (2004). How to Forecast Long-Run Volatility: Regime Swithing and the Estimation of Multifractal Processes. *Journal of Financial Econometrics*, 49-83.
- Chang, C.-Y., Xi-Yuan Q., & Sheng-Yuan J. (2011). Bayesian estimation and Comparison of MGARCH and MSV models via WinBUGS. *Journal of Statistical Computation and Simulation*, 82:1, 19-39.
- Choi, J. E. (2011). *Stochastic Volatility Models and Simulated Maximum Likelihood Estimation*. Ontario, Canada: University of Waterloo.
- Chong, C. W., Muhammad Ahmad, & Mat, Y. (1999). Performance of GARCH Models in Forecasting Stock Market Volatility. *Journal of Forecasting*, 18, 333-343.
- Comte, F., & Lieberman, O. (2003). Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis*, 84, 61-84.
- Danielsson, J. (1998). Multivariate Stochastic Volatility models: estimation and a comparison with VGARCH models. *Journal of empirical Finance*, 5(2), 155-173.

- Danielsson, J., & J.F. Richard. (1993). Accelerated Gaussian Importance Sampler with application to dynamic latent variable models. *Journal of Applied Econometrics*, 8, S153-S173.
- Dickey D. and Fuller, W. (1981). Distribution of the estimators for autoregressive time series with a unit root. *Econometrica*, 49, 1057-1072.
- Diebold, F., Lee, J.-H., & Weinbach, G. (1994). Regime Switching with Time-Varying Transition Probabilities. *Oxford University Press*, 283-302.
- Dielbold, F., & R. Mariano. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, 13, 253-265.
- Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with estimates of the variance of United Kingdom Inflation. *Econometrica*, 50, 122-150.
- Engle, R. F. (1982). The use of ARCH/GARCH Models in Applied Econometrics. *Journal of Economic Perspectives*, 15(4), 157-168.
- Engle, R. F. (2002). Dynamic Conditional Correlation. A Simple Class of Multivariate Generalized Autoregressive conditional Heteroscedasticity models. *Journal of Business and Economic statistics*, 20, 339-350.
- Engle, R. F., & Kroner, K. (1995). Multivariate Simultaneous Generalized ARCH. *Econometric Theory*, 11, 122-150.
- Engle, R. F., & Patton, J. A. (2001). What is a Good Volatility Model? *Quantitative Finance*, 1, 237-245.
- Engle, R. F., & Sheppard, K. (2001). Theoretical and Empirical properties of Dynamic Conditional Correlation Multivariate GARCH. *NBER Working Papers 8554, National Bureau of Economic Research, Inc.*
- Fleming, & Kirby. (2013). Component-Driven Regime-Switching Volatility. *Journal of Financial Econometrics*, 263-301.
- Francq, C., & Zakoian, J. (2004). Maximum Likelihood estimation of pure GARCH-ARMA processes. *Bernoulli*, 10(4), 605-637.
- Francq, C., & Zakoian, J. (2010). QML Estimation of a class of multivariate GARCH models without moment conditions on the observed process. *MPRA Paper No. 20779, Unpublished.*
- Giot, P., & Sebastien Laurent. (2001). Modelling Daily Value-At-Risk using Realized Volatility and ARCH type Models. *Journal of Empirical Finance*, 1-25.
- Glosten, L., Jagannathan, R., & Runkle, D. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Returns on Stocks. *Journal of Finance*, 48(5), 1779-1791.
- Hamilton, J. D. (1994). Time series Analysis. *Princeton University Press, Princeton New Jersey.*
- Harvey, A., & Genaro Sucarrat. (2013). EGARCH models with fat tails, skewness and leverage. *Computational Statistics and Data Analysis.*
- Harvey, Esther Ruiz, & Neil Shephard. (1994). Multivariate Stochastic variance Models. *Oxford University Press, The Review of Economic Studies*, 61, 247-264.

- Kobbacy, K., & Nicol, S. (1994). Sensitivity analysis of rent replacement models. *International Journal of Production Economics*, 36(3), 267-279.
- Koopman, S. J., & Eugenie Hol U. (2002). The Stochastic Volatility in Mean Model. Empirical Evidence from International Stock Markets. *Journal of Applied econometrics*, 667-689.
- Laurent E., C., & Adlai J., F. (2004). How to forecast Long-Run Volatility: Regime Switching and the Estimation of Multifractal Processes. *Journal of Financial Econometrics*, 49-83.
- Léon, N. K. (2015). Forecasting stock return Volatility: Evidence from the West African Regional stock market. *Journal of Applied Finance and Banking*, 5(6), 25-35.
- Lopes, M. H. (2004). *Modelling and Forecasting Stochastic Volatility*. Departement d'Economia i d'Historia Economica: International Doctorate in Economic analysis.
- McQueen, G., & Steven Thorley. ((March) 1991). Are Stock Returns Predictable? A test Using Markov Chains. *Journal of Finance* 66, 239-263.
- Mgr Milan, M. (2014). *Mulivariate GARCH*. RIGOROZNI PRACE.
- Nakatani, T., & Terasvirta, T. (2009). Testing for volatility interactions in the Constant Conditional Correlation GARCH model. *Econometrics Journal*, 147-163.
- Nelson, D. B. (1991). Conditional Heteroscedasticity in assets returns: a new approach. *Econometrica*, 59, 347-370.
- Newey, W., & K. West. (1987). A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703-708.
- Norris, J. (1997). *Markov Chains*. Cambridge University and Probabilistic Mathematics.
- Perrelli, R. (2001). Introduction to ARCH and GARCH models. *University of Illinois. Department of Economics. Econ 472*.
- Poon, S.-H. (2005). A practical Guide to Forecasting Financia Market Volatility. *John Wiley and Sons, Ltd*.
- Pye, G. (1966). A Markov model of the term Structure. *The Quarterly Journal of Economics*, 80, 60-72.
- Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics* 6(2), 461-464.
- Sheppard, K. (revised 25 Oct. 2003). Multi-step estimation of Multivariate GARCH models. *Proceedings of the International ICSC symposium: Advanced Computing in Financial Markets, University of California at San Diego*.
- Simeyo, O., Otumba, D. E., & Nyabwanga, R. N. (2015). Application of Markov Chain to model and Forecast stock market trend. *International Journal of current research*, 7(04), 14712-14721.
- Suliman, Z. S., & Winker, P. (2012). Modelling Stock Market Volatility Using Univariate GARCH models: Evidence from Sudan and Egypt . *International Journal of Economics and Finance*, 4(8), 161-176.
- Taylor, & Francis. (2016). Robust Parametric Tests of Constant Conditional Correlation in MGARCH model. *Econometric Reviews*.
- Taylor, S. (1986). Modeling Financial Time Series. *New York: Wiley*.

- Taylor, S. J. (2005). *Asset Price Dynamics, Volatility and Prediction*. Princeton University Press.
- Tsay, R. (2010). *Analysis of Financial Times Series*. John Wiley and Sons, Inc., New York, United States of America, 3rd edition.
- Tse, Y. K., & Tsui, K. C. (2002). A multivariate Generalized autoregressive conditional heteroscedasticity model with time-varying correlation. *Journal of Business and Economic statistics*, 20, 351-362.
- Zhang, D., & Zhang, X. (2009). *Study on Forecasting the Stock Market Trend Based on Stochastic Analysis Method* (Vol. 4). International Journal of Business and Management.

Appendix.

R. Codes

- `acf(Equity,lag=35,main="")`
- `acf(KCB,lag=35,main="")`
- `acf(Equity^2,lag=50,main="")`
- `acf(KCB^2,lag=50,main="")`
- `dcc.fcst=dccforecast(dcc.fit,n.ahead=100)`
- `dcc.fcst`
- `pacf(Equity,lag=35,main="")`
- `pacf(KCB,lag=35,main="")`
- `pacf(Equity^2,lag=50,main="")`
- `pacf(Equity^2,lag=50,main="")`
- `pacf(KCB^2,lag=50,main="")`
- `par(mfrow=c(2,1))`
- `plot.ts(Equity)`
- `plot.ts(KCB)`
- `plot(Equity.ret)`
- `plot(KCB.ret)`
- `plot(dcc.fit)`
- `plot(x=time(as.zoo(Equity.KCB.ret)),Y=Equity.KCB.cond.cov,type="L",xlab="Time",Ylab="covariance",lwd=2,col="blue",main="EWMA covariance between Equity and KCB")`
- `plot(garch_ccc,item="correlation")`
- `plot(garch_ccc,item="volatility")`
- `plot.ts(cbind(h[,2],hest[,2],type="L",main="h2",xlab="",Ylab="",plot.type=c("single"),col=c('red','blue')))`