The performance of MRS-GARCH models: A ML and Bayesian MCMC-based estimations

L.D. Xaba

orcid.org/0000-0002-1408-7595

Thesis submitted for the Degree *Doctor of Philosophy in Operations Research* at the North-West University

Promoter: Prof. N.D. Moroke

Co-promoter: Dr. L.D Metsileng

Graduation: May 2021

Student number: 18027709

DECLARATION

I, Lawrence Diteboho Xaba, hereby declare that this thesis: "The Performance of MRS-GARCH Models: A ML and Bayesian MCMC Based Estimation" is my own work towards the award of the Doctor of Philosophy degree and that, to the best of my knowledge, it contains no material previously published by another person nor material which had been accepted for the award of any other degree of a university, except where due acknowledgement had been made in the text.

LAWRENCE DITEBOHO XABA	………………	……………………
(18027709)	Signature	Date

DEDICATION

To my late friend, grandmother and mother, Mr.M.A Mahobe, Ms N. M. Xaba and Mrs M. F. Semelane for their care, support and love for the family.

ACKNOWLEDGMENT

Sir Isaac Newton, a great scientist once said: "if I have been able to see further than others it is because I was able to stand on the shoulders of giants." There are many "giants" whose efforts and contribution deserve to be mentioned in this study. Firstly, I would like to thank the Lord for His guidance in my life and in the preparation of this work. Special thanks go to Prof. N.D Moroke and Dr. L.D Metsileng for providing valuable guidelines and supporting me from the beginning till the end of the project. Secondly, I would like to extend my gratitude to Dr J. Arkaah and Mr V. Jain, for providing valuable comments during the process writing. Lastly, I thank Miss M. Semelane, Mr N.P. Xaba and my friends Dr J.T Tsoku and Dr. T.J Mosikari for the support and encouragement throughout my studies at the North-West University.

Finally, I would like to give my deepest gratitude to the North-West University for granting me financial assistance to do my post-graduate qualification.

PUBLICATIONS FROM THIS PHD STUDY

Book Chapter

Diteboho Xaba, Ntebogang D. Moroke, Dan Metsileng (2020). "**Performance of MRS-GARCH Models: A Bayesian MCMC Based Estimation investigation"** For the upcoming Book "Emerging Theories, Models and Applications of Financial Econometrics." Springer Publisher (Due Date: 13 December 2020)

Journal Paper

Diteboho Xaba and Ntebogang D. Moroke (2019). Modelling stock market returns of BRICS countries: Markov-switching Regression model Approach, Journal of Economics and Behavioral Studies. AMH International, vol. 11(3), pages 10-22.

Conference Paper

Diteboho Xaba and Ntebogang D. Moroke (2019). Modelling Exchange rate of South Africa using Bayesian MS-GARCH Models, Proceedings of the Finance and Economics Conference (December 20-21, 2019) Paris, France.

ABSTRACT

This study compared the in-sample and out-of-sample forecasting accuracy of nonlinear MRS-GARCH forecasting model using two estimation methods namely: The Maximum Likelihood and the Bayesian Markov Chain Monte Carlo. In addition, the study fitted an appropriate model that best describes the volatility of the exchange rates of BRICS countries. Data used was obtained from the Quantec Data Company. It covered the period from 1997 to 2017 with a total of 241 observations. Various tests of nonlinearity and nonlinear unit root were used to assess the validity of the assumptions of the study. Specifically, the study used a battery of linearity tests including Regression Specification Error Test (RESET), ARCH test, the McLeod and Li, Keenan's, Tsay, Brock- Dechert-Scheinkman (BDS), and CUSUM tests to determine whether or not exchange rates exhibit asymmetric behaviour. The Bierens Nonlinear Augmented Dickey-Fuller (B-NADF) and Kapetanois-Schmidt-Shin Nonlinear Augmented Dickey-Fuller (KSS-NADF) were applied to determine whether the time series properties of exchange rates of BRICS countries were nonlinear and have nonlinear unit root in nature.

To choose the best fitted model, Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), Log-likelihood (LOGL) and Deviance information criterion (DIC) were employed under different conditional distributions for innovation. Furthermore, the Mean Square Error (MSE), Mean absolute Error (MAE) and Root Mean Square Error (RMSE) served as the error measures in evaluating the forecasting ability of the models. The MRS-GARCH models proved to perform well with lower error measures when the Bayesian framework is used with student-t distribution for in-sample and for out-of-sample generalized error distribution. The decision by error measures were supported by Diebold and Mariano test of predictive accuracy. The results showed that the Bayesian method produces accurate and reliable forecasts. In this study, the Bayesian method is preferred over Maximum Likelihood estimation method.

The findings of the study proved both the Maximum Likelihood Estimation, and the Bayesian Markov Chain Monte Carlo estimation methods are good estimators, even though further improvement is evident. In the case of Maximum Likelihood Estimation method, where Expected Maximization (EM) algorithm was used, the study recommended that the Monte Carlo Maximum Likelihood estimator and/or Monte Carlo Expected Maximization estimator

can be used to improve Maximum Likelihood estimation method. The results from these two methods may be compared with those of the previous studies.

Table of Contents

DECLARATION i
DEDICATION ii
ACKNOWLEDGMENT iii
PUBLICATIONS FROM THIS PHD STUDY iv
ABSTRACT v
LIST OF FIGURES xi
LIST OF TABLES xii
ACRONYMS AND ABBREVIATIONS xiv
CHAPTER 1 1
OVERVIEW OF THE STUDY 1
1.1. INTRODUCTION 1
1.2. BACKGROUND AND CONTEXT 3
1.3. PROBLEM STATEMENT 4
1.4. AIM AND OBJECTIVES OF THE STUDY 5
1.5. RATIONALE OF THE STUDY 5
1.6. SIGNIFICANCE OF THE STUDY 6
1.7. DEFINITION OF TERMS 6
1.8. ORGANISATION OF THE STUDY 7
1.9. SUMMARY OF THE CHAPTER 8
CHAPTER 2 9
THEORY AND LITERATURE REVIEW 9
2.1. INTRODUCTION 9
2.2. THEORETICAL LITERATURE 9
2.2.1. OVERVIEW OF GARCH MODEL 9
2.2.2. THE MS-ARCH MODEL 11
2.2.3. THE MS-GARCH MODEL 12
2.2.4. PATH DEPENDENCE CHALLENGE 13
2.2.5. THE MS-GARCH MODEL APPROXIMATION 14
2.3. EMPIRICAL LITERATURE 18
2.4. SUMMARY OF THE CHAPTER 23
CHAPTER 3 25
RESEARCH METHODOLOGY 25

3.1. INTRODUCTION 25
3.2. ETHICAL CONSIDERATIONS 25
3.3. RESEARCH PROCESS 26
3.3.1. RESEARCH PHILOSOPHY 26
3.3.2. RESEARCH APPROACH 27
3.3.3. RESEARCH STRATEGY AND TIME HORIZONS 28
3.3.4. DATA COLLECTION AND ANALYSIS 29
3.3.4.1. THE PRIMARY DATA 29
3.3.4.2. SECONDARY DATA 29
3.4. PRELIMINARY DATA ANALYSIS 30
3.4.1. ASSESSING UNIVARIATE NORMALITY 30
3.5. ASSESSMENT OF DATA FOR LINEARITY 31
3.5.1. THE RESET TEST 32
3.5.2. KEENAN'S TEST 33
3.5.3. TSAY TEST (1986) 34
3.5.4. THE BROCK-DECKERT-SCHIENKMAN (BDS) TEST 34
3.5.5. CUSUM TEST 35
3.6. NONLINEAR UNIT ROOT TEST 37
3.6.1. THE B-NADF UNIT ROOT TEST 37
3.7. MODELLING AND ESTIMATION METHODS 40
3.7.1. MARKOV REGIME SWITCHING GARCH MODEL 40
3.7.2. ESTIMATION METHODS 41
3.7.2.1. MAXIMUM LIKELIHOOD ESTIMATION METHOD 41
3.7.2.2. BAYESIAN MCMC ESTIMATION METHOD 42
3.8. MODEL SELECTION CRITERIA 43
3.8.1. AKAIKE INFORMATION CRITERION (AIC) 44
3.8.2. SCHWARZ INFORMATION CRITERION (SIC) 44
3.8.3. HANNAN-QUINN INFORMATION CRITERION (HQC) 44
3.8.4. DEVIANCE INFORMATION CRITERION (DIC) 45
3.9. FORECASTING THE EXCHANGE RATE VOLATILITY 45
3.9.1. K-STEP AHEAD FORECASTING 48
3.10. FORECAST EVALUATION 48
3.11. MODEL FORECASTING ACCURACY 50

3.12. CHAPTER SUMMARY 51
CHAPTER 4 53
EMPIRICAL ANALYSIS AND RESULTS 53
4.1. INTRODUCTION 53
4.2. PRELIMINARY DATA ANALYSES RESULTS 53
4.2.1. DESCRIPTIVE STATISTICS 53
4.2.2. NONLINEAR MODEL SPECIFICATION/DIAGNOSTICS: NONLINEARITY TESTS 55
4.2.2.1. CUSUM TESTS RESULTS 56
4.2.2.2. BROCK- DECHERT-SCHEINKMAN (BDS) TEST RESULTS 58
4.2.2.3. LAGRANGE MULTIPLIER (LM) TEST RESULTS 59
4.2.3. NONLINEAR UNIT ROOT TEST RESULT 59
4.3. MODEL ESTIMATION RESULTS USING MAXIMUM LIKELIHOOD 61
4.3.1. BRAZIL RESULTS 62
4.3.2. RUSSIA RESULTS 63
4.3.3. INDIA 64
4.3.4. CHINA 65
4.3.5. SOUTH AFRICA 66
4.4. IN-SAMPLING EVALUATION USING MAXIMUM LIKELIHOOD 67
4.5. MODEL ESTIMATION RESULTS USING BAYESIAN MCMC 68
4.5.1. BRAZIL RESULTS 68
4.5.2. RUSSIA RESULTS 69
4.5.3. INDIA RESULTS 70
4.5.4. CHINA RESULTS 71
4.5.5. SOUTH AFRICA RESULTS 72
4.6. IN-SAMPLING EVALUATION USING BAYESIAN MCMC 73
4.7. MODEL COMPARISON RESULTS OF MRS-GARCH MODELS 73
4.8. PREDICTIVE ACCURACY OF MODELS (IN-SAMPLING) 74
4.9. OUT-OF-SAMPLING EVALUATION USING MAXIMUM LIKELIHOOD 75
4.9.1. THE 1-MONTH-AHEAD VOLATILITY FORECAST 75
4.9.2. THE 5-MONTH-AHEAD VOLATILITY FORECAST 76
4.9.3. THE 10-MONTHS-AHEAD VOLATILITY FORECAST 77
4.10. OUT-OF-SAMPLING VOLATILITY FORECASTS USING BAYESIAN MCMC 78

4.10.1. THE 1-MONTH-AHEAD VOLATILITY FORECAST 78
4.10.2. THE 5-MONTH-AHEAD VOLATILITY FORECAST 79
4.10.3. THE 10-MONTH-AHEAD VOLATILITY FORECAST 80
4.11. COMPARISON RESULTS OF MRS-GARCH MODELS 80
4.12. PREDICTIVE ACCURACY OF MODELS (OUT-OF-SAMPLING) 82
4.13. CHAPTER SUMMARY 84
CHAPTER 5 85
CONCLUSION AND RECOMMENDATIONS 85
5.1. INTRODUCTION 85
5.2. DISCUSSION OF THE RESULTS 85
5.3. CONTRIBUTIONS TO THE BODY OF KNOWLEDGE 91
5.4. LIMITATIONS 92
5.5. CONCLUSIONS 92
5.6. RECOMMENDATIONS FOR FUTURE RESEARCH 93
5.7. SUMMARY OF THE THESIS 93
REFERENCE 95
APPENDICES 107

LIST OF FIGURES

Figure 3.1: The research process onion (Saunders *et al*., 2012) 26

Figure 4.1: Graphical representation of the five exchange rates 55

Figure 4.2: CUSUM test for stability..57

LIST OF TABLES

Table 3.1: Critical Values for KSS Nonlinear Unit Root Tests 40

Table 4.1: Descriptive Statistics of Exchange Rates 54

Table 4.2: Test for Nonlinearity 56

Table 4.3: BDS tests results 58

Table 4.4: LM Test results 59

Table 4.5: Bierens Nonlinear Unit Root Test Results 60

Table 4. 6: KSS Nonlinear Unit Root Test 61

Table 4.7: Maximum likelihood estimates of MRS-GARCH models for Brazil 62

Table 4.8: Maximum likelihood estimates of MS-GARCH models for Russia 63

Table 4.9: Maximum likelihood estimates of MRS-GARCH models 64

Table 4.10: Maximum likelihood estimates of MRS-GARCH models 65

Table 4.11: Maximum likelihood estimates of MRS-GARCH models 66

Table 4.12: In-sample Results using MRS-GARCH models for BRICS countries 67

Table 4.13: Bayesian MCMC estimates of MRS-GARCH models for Brazil 68

Table 4.14: Bayesian MCMC Model estimates of MRS-GARCH models for Russia 69

Table 4.15: Bayesian MCMC Model estimates of MRS-GARCH models for India 70

Table 4.16: Bayesian MCMC Model estimates of MRS-GARCH models for China 71

Table 4.17: Bayesian MCMC Model estimates of MRS-GARCH models for South Africa .. 72

Table 4.18: MS-GARCH: In-sample Results using Bayesian MCMC estimator 73

Table 4.19: MRS-GARCH: In-sample Results for Model Comparison 74

Table 4.20: Comparison of Model Accuracy: Diebold-Mariano Absolute Error Test 75

Table 4.21: Out-of-sample evaluation of 1-month-ahead volatility Results 76

Table 4.22: Out-of-sample evaluation of 5-month-ahead volatility Results 77

Table 4.23: Out-of-sample evaluation of 10-month-ahead volatility Results 78

Table 4.24: Out of sample evaluation of 1-month-ahead volatility Results 79

Table 4.25: Out of sample evaluation of 5-month-ahead volatility Results 79

Table 4.26: Out of sample evaluation of 10-month-ahead volatility Results 80

Table 4.27: Out-of-sample Results for 1-month ahead.. 81

Table 4.28: Out-of-sample Results for 5-month ahead.. 81

Table 4.29: Out-of-sample Results for 10-month ahead.. 81

Table 4.30: Comparison of Model Accuracy: Diebold-Mariano Absolute Error Test (1-month-ahead volatility Results) 82

Table 4.31: Comparison of Model Accuracy: Diebold-Mariano Absolute Error Test (5-month-ahead volatility Results) 83

Table 4.32: Comparison of Model Accuracy: Diebold-Mariano Absolute Error Test (10-month-ahead volatility Results) 83

ACRONYMS AND ABBREVIATIONS

ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criterion
APGARCH	Asymmetric Power Generalized Autoregressive Conditional Heteroscedastic
ARCH	Autoregressive Conditional Heteroscedasticity
ARMA	Autoregressive Moving Average
AR-MS-GARCH	Autoregressive Markov Switching Generalized Autoregressive Conditional Heteroscedasticity
BDS	Brock-Dechert-Scheinkman
B-NADF	Bierens Nonlinear Augmented Dickey-Fuller
BRICS	Brazil, Russia, India, China, South Africa
CUSUM	Cumulative Sum test
CUSUMSQ	Cumulative Sum Square
DIC	Deviance Information Criterion
DM	Diebold Mariano test
FIEGARCH	Fractionally Integrated Exponential Generalized Autoregressive Conditional Heteroscedastic
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GED	Generalized Error Distribution
GMM	Generalized Method of Moments
HQC	Hannan-Quinn information criterion
IID	Independent Identically Distributed

IRAN	Islamic Republic of Iran
JB	Jarque-Bera
KSE	Karachi Stock Exchange
KSS	Kapetanois-Schmidt-Shin
KSS-NADF	Kapetanois-Schmidt-Shin Nonlinear Augmented Dickey-Fuller
LM	Lagrange Multiplier
MAE	Mean Absolute Error
MCEM	Monte Carlo Expected Maximization
MCMC	Markov Chain Monte Carlo
MCML	Monte Carlo Maximum Likelihood
ML	Maximum Likelihood
MRS-GARCH	Markov Regime Switching-GARCH
MS-ARCH	Markov Switching Autoregressive Conditional Heteroskedasticity
MS-ARMA	Markov-Switching Autoregressive Moving Average
MSE	Mean Squared Error
MS-GARCH	Markov Switching Generalized Autoregressive Conditional Heteroskedasticity
MS-GARCH-ARIMA	Markov Switching Generalized Autoregressive Conditional Heteroskedasticity Autoregressive Moving Average
MSM	Markov Switching Model
NORM	Normal distribution
OLS	Ordinary Least Square
PMF	Probability Mass Function

RESET	Regression Specification Error
RMSE	Root Mean Squared Error
SBC	Schwartz Bayesian criterion
STD	Student's–t distribution
VaR	Value-at-Risk
HMC	Hamiltonian Monte Carlo
GJR-GARCH	Glosten-Jagannathan-Runkle GARCH

CHAPTER 1

OVERVIEW OF THE STUDY

1.1. INTRODUCTION

The study investigated the performance of the hybrid models of Markov regime switching generalized autoregressive conditional heteroskedasticity (MRS-GARCH) using the two estimation methods of Maximum Likelihood estimation and Bayesian Markov Chain Monte Carlo (MCMC) estimation to provide better forecasts. The development of the generalized autoregressive conditional heteroskedastic (GARCH) model by Bollerslev (1986) and Taylor (1986) was intended to achieve the objective of forecasting volatility. According to Andersen & Bollerslev (1998) and Andersen *et al.* (2007), GARCH model can be used as a tool to measure risk and as well as providing accurate forecasting volatility of financial data. Furthermore, the GARCH model is used to estimate volatility characteristics of financial data such as dynamics of conditional heteroscedasticity (Epaphra, 2017). Matei (2009) concurred that GARCH models are the most suitable models to evaluate the volatility of the financial data with large amounts of observations. The suitability of the model is seen through the quality of volatility forecasts provided by GARCH when compared to other alternative models[1]. This class of models has been used to forecast volatility of financial data (such as commodities, stock market returns and exchange rates).

However, recent studies show that estimates of GARCH models can be biased because of structural breaks in the dynamics of volatility (Bauwens *et al*., 2014). These structural breaks happen during times of financial turnover. Furthermore, studies show that structural breaks in unconditional variance have potentially important implications for estimated GARCH models of exchange rate return volatility. Building on insights from Diebold (1986), Hendry (1986), and Lamoureux & Lastrapes (1990), recent research by Mikosch & Starica (2004) and Hillebrand (2005) demonstrates that neglected structural breaks in the parameters of GARCH processes induce upward biases in estimates of the persistence of GARCH processes. This is a relevant concern, as estimated GARCH models of exchange rate return volatility in the extant

[1] Autoregressive Conditional Heteroskedastic (ARCH) model compiled by Engle (1982), the Conditional Heteroskedastic Autoregressive Moving Average (CHARMA) model obtained by Tsay (1987), the random coefficient autoregressive (RCA) model of Nicholls and Quinn (1982), and the Stochastic Volatility (SV) models by Melino and Turnbull (1990), Taylor (1994), Harvey, Ruiz and Shephard (1994), and Jacquier, Polson, and Rossi (1994)

literature are typically highly persistent. More generally, failure to account for structural breaks can obviously lead to poor estimates of the unconditional volatility of exchange rate returns for GARCH processes. Klaassen (2002) adds that this persistence is not compatible with the poor forecasting results of the GARCH models. An approach to adapt to these drawbacks of GARCH models in modelling data with structural breaks gave birth to the MRS-GARCH models whose parameters fluctuate after some time as per an inert discrete Markov process (Ardia, 2008). These models can rapidly adjust to variations in the conditional volatility level which enhances hazard forecasts (Miron & Tudor, 2010).

This combination of Markov-Switching (MS) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models was proposed by Lamoureux and Lastrapes (1990), and Mikosch and Starica (2004). The MS-GARCH model can simply be understood as a GARCH model where parameters depend on the state of an unobserved Markov chain. Lamoureux & Lastrapes (1990), and Mikosch & Starica (2004) justify such a combination by noting that the high persistence observed in the variance of financial returns can be explained by time-varying GARCH parameters.

The high and low volatility probabilities of MS-GARCH models allow differentiating high and low volatility periods (Mikosch & Starica, 2004). By observing the periods in which volatility is high, it is possible to investigate the economic and political reasons that caused increased volatility. Furthermore, MS-GARCH models allow for a sudden change in the (uncondtional) volitility level which leads to significant improvements in volatility foreacasts (Gray, 1996; Dueker, 1997; Klaassen, 2002; Haas *et al*. 2004; Miron & Tudor, 2010; Gong & Lin, 2017). Since the flexibility of these models, the MS or regime-switching models have attracted a lot of attention in the econometric literature since the seminal paper of Hamilton (1989).

As such, the current study seeks to enquire more about the efficiency of these novel models. These models can rapidly adjust to variations in the conditional volatility level which enhances hazard forecasts. In the context of this study, the historical data of the exchange rates of BRICS countries was used as an experimental unit.

1.2. BACKGROUND AND CONTEXT

Modeling volatility of the exchange rate plays a vital role for a country's economy. Various authors such as Brunetti *et al*. (2007); Bala and Asemota (2013); May and Farrell (2017) made several attempts to study the characteristics of the foreign exchange markets rates' volatility. This has been inspired by the important role the exchange rate plays in a crisis, such as, the exogenous geopolitical events or financial crises, or disruptions due to weather related catastrophes (Vasanthi *et al*. 2011). An important ingredient during such crises is the possibility of modelling and forecasting the volatility under a reasonable level of accuracy (Arellano *et al*. 2017). Futhermore, the changes in exchange rate have important effects on a number of macroeconomic factors, such as inflation; change in the price of imported goods and services; change in the commodity prices; and change in the growth of exports. According to Arellano *et al*. (2017), these factors affect the competitiveness of sectors that produce and sell tradable goods and services; and on the valuation of assets and liabilities through currency mismatches. Therefore, modelling the returns and volatility of exchange rates would be useful for Portfolio Managers, Investment Analysts, Financial Engineers to know which distribution and estimation method to use.

According to Franses & van Dijk (2000) and Humala & Rodriguez (2013), time series of stock market returns and exchange rate have four typical stylized facts:

i. Large returns occur more often than expected (leptokurtosis or fat tails). This implies that the kurtosis is much larger than 3, or the tails of the distributions are fatter than the tails of the normal distribution;
ii. Large exchange rate and stock market returns are often negative (negative skewness). This implies that the left tail of the distribution is fatter than the right tail, or that large negative returns tend to occur more often than large positive ones;
iii. Large returns tend to occur in clusters. This implies that relatively volatile periods, characterized by large price changes (large returns) alternate with quieter periods in which prices remain more or less stable (small returns);
iv. High volatility often follows large negative stock and exchange rate. This implies that periods of high volatility tend to be triggered by a large negative return (this stylised fact is also called the leverage effect). These features of stock and exchange rate are modelled with nonlinear models, simply because linear models would not be able to generate data with these features.

The present study was motivated by the above mentioned stylised facts to apply nonlinear MRS-GARCH models to study the volatility of the exchange rates of BRICS countries. The study found MRS-GARCH models more relevant because of the abilities of the method. The MRS-GARCH can account for structural changes that can be found in the exchange rate data. Furthermore, the MRS-GARCH models is capable of describing the dynamics of exchange rate series that linear GARCH models failed to capture.

1.3. PROBLEM STATEMENT

Researchers show that there is an evidence of structural breaks in financial data which affect the financial indicators such as financial returns and volatility. These structural changes make it difficult to provide accurate forecasting volatility of exchange rates. The MRS-GARCH models are regarded as a promising way to capture nonlinearities in time series. The MRS-GARCH can account for sudden changes in the structure of the mean or variance of a process and give a straightforward interpretation of these shifts. Such shifts would cause a regular GARCH model to imply nonstationary processes. Hybrid Markov regime switching with GARCH model poses severe difficulties when it comes to understanding their dynamic properties and for the computation of parameter estimators.

Hamilton & Susmel (1994) were among the first authors to discuss the MS-GARCH model and they noted that the estimation of this path dependent model is a challenging task because exact computation of the likelihood is infeasible in practice. Furthermore, the computation method based on the maximum likelihood (ML) procedure has shown some convergence problems and it has a problem of local maxima (Ardia, 2008). Hamilton and Susmel (1994) and Cai (1994) reported that the disadvantage of MLE approach is that it becomes computationally infeasible for MS-GARCH models. This led to some authors such as Gray (1996), Dueker (1997), Klaassen (2002) and Haas *et al*. (2004) to propose modified estimation versions of the MS-GARCH model that avoid the path dependence problem which is often experienced when using the ML estimation method. Other authors suggested an alternative estimation method such as Bayesian Markov Chain Monte Carlo (MCMC) (Bauwens *et al*. 2007). Some authors (Bauwens *et al*. 2010, 2011, 2014, and Henneke *et al*., 2011) resorted to Bayesian framework because the MCMC methods are powerful tools for the numerical computation of integrals. Furthermore, Bayesian MCMC algorithm is able to circumvent the problem of path dependence by including the state variables in the parameter space and simulating them by Gibbs sampling.

Various studies attempted to model and forecast the volatility of exchange rate (Brunetti *et al.* 2008; Bala & Asemota, 2013; May & Farrell, 2017). In many instances, the findings obtained by these studies have been found to be inconclusive. Some of these studies found that MS-GARCH with a generalised error distribution (GED) fits exchange rates well (Satttayathem *et al.* 2012) while others have reported that MS-GARCH with a t-distribution fits the data better than competing distributions under MS-GARCH models (Abounoori *et al.* 2016). On the other hand, there are some studies that have found MS-GARCH with a normal distribution to fit the data better than all competing distributions (Augustsson, 2014). The studies mentioned above drew contradicting conclusions regarding the best distribution. In the context of the current study, the statement of the problem is how these three methods do apply under MRS-GARCH to model and forecast the exchange rate of BRICS countries, and how the estimated models compare in terms of efficiency and performance.

1.4. AIM AND OBJECTIVES OF THE STUDY

The study sought to explore the performance of the hybrid models of MRS-GARCH using the two estimation methods of Maximum Likelihood estimation and Bayesian MCMC estimation to provide better forecasts. The specific objectives of this study are as follows:

i. To estimate MLE and Bayesian MCMC MRS-GARCH models of the exchange rates of BRICS countries

ii. To compare the efficiency and performance of MLE and Bayesian MCMC MRS-GARCH models for in-sample and out-of-sample volatility forecasts

iii. To compare the predictive accuracy of MRS-GARCH models using MLE and Bayesian MCMC

1.5. RATIONALE OF THE STUDY

This study applies robust econometric techniques to investigate the dynamics of conditional heteroscedasticity of exchange rates. Despite the number of studies on the characteristics of exchange rate returns from BRICS countries, the MRS-GARCH models has not yet been exploited to the fullest in modelling and forecasting volatility of exchange rates of BRICS countries. The current study is trying to bridge this gap, specifically, as used in the investigation of the characteristics of the exchange rates of BRICS countries. By using the MRS-GARCH modelling techniques and various estimation methods, this study also provides a skilful guideline for modelling the different dynamic patterns of exchange rates in general as well as the art involved in capturing the nonlinear attributes of exchange rates. Overall, this study particularly intends to enhance the understanding of the patterns and movements in the exchange rates of BRICS countries as well as asymmetric regimes in exchange rates in general.

1.6. SIGNIFICANCE OF THE STUDY

This study provides a base for future researchers conducting studies on different economies, more specifically in the BRICS context. The findings of this study will provide a base for researchers in the field on the application of MRS-GARCH models. Furthermore, scholars who are interested in employing the ideal model (MRS-GARCH) for financial and economic data, more specifically when the intention is to predict the volatility of exchange rate will benefit from the findings of this study. The findings of the study will provide guidance to researchers or scholars to employ suitable distribution and estimation methods to model exchange rates of BRICS countries. The research findings have practical implications for traders, portfolio managers and policymakers.

More importantly, the findings may provide insight into the way the exchange rate volatility shocks are transmitted to stock markets and assess the degree and persistence of these innovations over time. From a financial stability perspective, the volatility transmission across the two markets is also an important consideration for policymakers.

1.7. DEFINITION OF TERMS

The following terms are used in the study:

- **Autoregressive Conditional Heteroskedasticity (ARCH):** This is a statistical model for time series data that describes the variance of the current error term or innovation as a function of the actual sizes of the previous time periods' error terms (Engle, 1982).
- **BRICS:** It is the acronym coined for an association of five major emerging national economies: Brazil, Russia, India, China and South Africa.
- **Exchange rate:** This is the price of one currency in terms of another currency (Pattichis, 2012).

- **Generalized Autoregressive Conditional Heteroskedasticity (GARCH):** This is a statistical model used in analysing financial time-series data. Heteroskedasticity describes the irregular pattern of variation of an error term, or variable, in a statistical model. Essentially, where there is heteroskedasticity, observations do not conform to a linear pattern. Instead, they tend to cluster (Bollerslev, 1986).
- **Gibbs Sampling:** This method requires all the conditional distributions of the target distribution to be sampled exactly. When drawing from the full-conditional distributions is not straightforward other samplers-within-Gibbs are used (Geman and Geman, 1984).
- **Maximum Likelihood Estimation (MLE):** This is a method of estimating the parameters of statistical model, given observations. The method obtains the parameter estimates by finding the parameter values that maximize the likelihood function (Rossi, 2018).
- **Markov chain Monte Carlo** (**MCMC**): s a simulation technique that can be used to find the posterior distribution and to sample from it (Korner-Nievergelt, *et al*., 2015)
- **Markov switching model or Regime switching model**: This is one of the most popular nonlinear time series models in literature. This model involves multiple structures (equations) that can characterize the time series behaviours in different regimes (Hamilton, 1990).
- **Stock Market Returns:** These are the returns that the investors generate out of the stock market. This return could be in the form of profit through trading or in the form of dividends given by the company to its shareholders from time-to-time (Restoy, and Rockinger, 1994).
- **Volatility**: It is a statistical measure of dispersion around a mean value; or is defined under a theoretical aspect as the changeability or randomness of the underlying asset (Schwert, 1990).

1.8. ORGANISATION OF THE STUDY

This study contains five chapters:

Chapter 1 presents the study overview using the introduction, background and context of the study, problem statement, aim and objectives of the study. The chapter gave a view on the research questions, rationale and significance of the study, and lastly the definition of terms were also presented.

Chapter 2 presents literature review and theory of the methods used in this study.

Chapter 3 provides a description of the research methodology. A detailed discussion of the research process, research design and methods used in this study is provided.

Chapter 4 gives a comprehensive report of the study results and an interpretation of the findings. The chapter also presents different methods using the BRICS data applied to each method.

Chapter 5 provides a summary of the findings of the study, discussion and recommendations for future research.

1.9. SUMMARY OF THE CHAPTER

This chapter provided the context and background to the research problem, the orientation to the study, and the structure and plan of the study. Research objectives were clearly aligned to the title and the research problem. Furthermore, the chapter highlighted some of the potential benefits of the study and the boundaries the study was confined to. The next chapter provides a critical review of the empirical literature related to the study and some of the important theories on MS-GARCH models.

CHAPTER 2

THEORY AND LITERATURE REVIEW

2.1.INTRODUCTION

This chapter discusses the theories and literature of the volatility methods found in time series. Furthermore, the chapter provides related theories, models and application of various models that has been used in the literature.

The rest of the chapter is organised as follows: Section 2.2 presents theoretical literature; Subsection 2.2.1 presents the overview of GARCH model; Subsection 2.2.2 provides the MS-ARCH; Section 2.2.3 presents the MS-GARCH model; Section 2.2.4 provides the explanation of the path dependence challenge under MS-GARCH model; 2.2.5 provides the various approximates of MS-GARCH models and Section 2.3 presents the empirical literature review on MS-GARCH. Finally, Section 2.4 presents the summary of the chapter.

2.2.THEORETICAL LITERATURE

This section presents the full description of the volatility models that are found in the theoretical literature.

2.2.1. OVERVIEW OF GARCH MODEL

GARCH is a statistical model used to analyse the financial and economic data. The GARCH model is an extension of autoregressive conditional Heteroscedasticity (ARCH). The model was extended by Bollerslev (1986) and invented by Engle (1982). This model was extended to bridge the shortcomings of the ARCH model which include the high number of unknown parameters and rapid decay of unconditional autocorrelation function of squared residuals, among others. Bollerslev (1986) in countering these shortcomings proposed the GARCH model in which conditional variance is also a linear function. This model is also a weighted average of past squared residuals, but it has declining weights that never go completely to zero as compared to ARCH model. According to Lama *et al*. (2015), GARCH model gives a flexible lag structure and it permits more prudent descriptions in most situations. Furthermore, the GARCH models are able to capture volatility which has been studied over time. The formulation of the GARCH model is as follows Engle (1982);

Let ϵ_t be an independent and identical distribution (*iid*) with a mean of 0 and variance of 1 and u_t is said to be an ARCH(k, g) process with respect to ϵ_t if:

$$u_t = \sqrt{h_t \epsilon_t}$$

$$u_t = \sigma_t^2 = \mathrm{Var}(u_t | u_s, s < t) = \alpha_0 + \sum_{i=1}^{g} \alpha_i u_{t-i}^2 \tag{2.1}$$

where h_t is the conditional variance of a series and a constant = α_i is strictly positive and defined for $i = 0, \dots \dots, g$. The volatility described by equation (2.1) is heteroskedastic implying that the conditional variance h_t is not constant. In the ARCH model, volatility is defined by weights of squared past observations of the financial time series data. The weights α_i's are estimated with various estimation methods.

The GARCH definition is as follows Bollerslev (1986):

Let ϵ_t be an *iid* sequence having mean 0 and variance of 1. A process u_t is said to be a GARCH(k, g) process with respect to ϵ_t if:

$$u_t = \sqrt{h_t \epsilon_t},$$

$$u_t = \alpha_0 + \sum_{i=1}^{g} \alpha_i u_{t-i}^2 + \sum_{j=1}^{k} \beta_j h_{t-j,} \tag{2.2}$$

where the constants α_i and β_j defined for $i = 0, \dots \dots, g$ and $j = 1, \dots \dots, k$ are positive. The condition for equation (2.2) defining a unique strictly stationary process u_t with $E(u_t^2) < \infty$ is:

$$\sum_{i=1}^{g} \alpha_i + \sum_{j=1}^{k} \beta_j < 1. \tag{2.3}$$

Furthermore, $E(u_t) = 0$, $\mathrm{Cov}(u_t, u_s) = 0$ for $s < t$ and $\mathrm{Var}(u_t) = \frac{\alpha_0}{1 - \sum_{i=1}^{g} \alpha_i - \sum_{j=1}^{k} \beta_j}$.

In addition, $E(log(u_t^2)) < \infty$, provided max $(1, E(\epsilon_t^4)^{0.5}). \frac{\sum_{i=1}^{g} \alpha_i}{1 - \sum_{j=1}^{k} \beta_j} < 1$.

The GARCH model in equation (2.2) is different to ARCH model since GARCH has an extra term known as, $\sum_{j=1}^{k} \beta_j h_{t-j}$. This extra term ensures a parsimonious model and often it is found that only one GARCH term is appropriate (Mazviona, 2012). The value of α_0 is the average conditional variance in the long run. The GARCH model introduced can improve the modelling of the evolution of the financial time series data.

2.2.2. THE MS-ARCH MODEL

The MS-ARCH is a time series model that describes the variance of the current error term or innovation as a function of the actual sizes of the previous time periods' error terms. The model was developed by Cai (1994), with the combination ARCH of Engle (1982) and MS model of Hamilton (1988). The combination of these models was done after several authors identified the shortcomings of the standard GARCH models. Standard GARCH models usually indicate a high persistence of the conditional variance (i.e. a nearly integrated GARCH process). Diebold (1986) and Lamoureux & Lastrapes (1990), among others, argued that the nearly integrated conduct of the conditional variance may originate from structural changes in the variance process which are not accounted for by standard GARCH models.

Furthermore, Mikosch & Starica (2004) showed that estimating a GARCH model on a sample displaying structural changes in the unconditional variance does indeed create an integrated GARCH effect. These findings clearly indicated a potential source of misspecification, to the extent that the form of the conditional variance is relatively inflexible and held fixed throughout the entire sample period. Hence, the estimates of a GARCH model may suffer from a substantial upward bias in the persistence parameter. Therefore, models in which the parameters are allowed to change over time may be more appropriate for modelling volatility. MS-ARCH(g) process is followed by a process u_t if Cai (1994):

$$u_t = \sqrt{h_t \epsilon_t},$$

$$h_t = k_{(s_t)} + \sum_{i=1}^{g} \alpha_i u_{t-i}^2, \tag{2.4}$$

where ϵ_t is an iid of having a zero mean and unit variance, $k_{(s_t)}$=$c_0 + c_1 s_t$, $c_0 > 0$ and $c_1 > 0$. $s_t = 0$ or 1 denotes the unobserved states. The latent variable s_t is assumed to be stationary and of first-order Markov process with transition probabilities:

$$f(s_t = j | s_{t-1} = i) = \lambda_{ij} \tag{2.5}$$

where $\lambda_{11} + \lambda_{12} = 1, \lambda_{21} + \lambda_{22} = 1$ and $0 < \lambda_{ij} < 1$

The term $k_{(s_t)}$ is the main feature in the MS-ARCH(g) model that makes the volatility to switch between the states. This term explains the business cycle of the economy (e.g. high and low volatility).

2.2.3. THE MS-GARCH MODEL

A Markov switching GARCH model is a nonlinear specification of the evolution of a time series which is affected by different states of the world and for which the conditional variance in each state follows a GARCH process (Francq *et al.*, 2001). This model involves multiple structures (equations) that can characterize the time series behaviors in different regimes. The MS-GARCH model can be simply understood as a GARCH model where parameters depend on the state of an unobserved Markov chain. Mikosch and Starica (2004) justify such a combination by noting that the high persistence observed in the variance of financial returns can be explained by time-varying GARCH parameters. To define MS-GARCH model, let y_t be the observed variable and s_t a discrete, unobserved state variable which could be interpreted as the state of the world at time *t*. Define $(y_t \dots\dots, y_u)$ and $(s_t \dots\dots, s_u)$ as $y_{t:u}$ and $s_{t:u}$ respectively, whenever $s \leq t$, then an MS-GARCH model can be defined as follows according to Bauwens *et al.* (2010) and Francq *et al.* (2001):

$$y_t = \mu_t\big(y_{1:t-1}, \theta_t(s_t)\big) + \sigma_t\big(y_{1:t-1}, \theta_t(s_t)\big)\eta_t, \quad \eta_t \sim N(0,1), \tag{2.6}$$

$$\sigma_t^2 = \big(y_{1:t-1}, \theta_\sigma(s_t)\big) = \gamma(s_t) + \alpha(s_t)\epsilon_{t-1}^2 + \beta(s_t)\sigma_{t-1}^2\big(y_{1:t-2}, \theta_t(s_{t-1})\big), \tag{2.7}$$

where, $\epsilon_t = \sigma_t\big(y_{1:t-1}, \theta_\sigma(s_t)\big)\eta_t, \theta_\sigma(s_t) = \big(\gamma(s_t), \alpha(s_t)\beta(s_t)\big), \gamma(s_t) > 0, \alpha(s_t) \geq 0, \beta(s_t) \geq 0,$ and $s_t \in \{1, \dots\dots.., M\}, t = 1, \dots\dots\dots.., T$, is assumed to follow an *M*-state first order Markov chain with transition probabilities $\{\pi_{ij,t}\}i, j = 1,2, \dots\dots.M$:

$$\pi_{ij,t} = p(s_t = i | s_{t-1} = j, y_{1:t-1}, \theta_\pi), \textstyle\sum_{i=1}^{M} \pi_{ij,t} = 1 \qquad \forall\ j = 1,2, \dots\dots\dots., M. \tag{2.8}$$

The parameter shift functions $\gamma(s_t), \alpha(s_t)$ and $\beta(s_t)$, describe the dependence of parameters on the realized regime s_t, i.e.,

$$\gamma(s_t) = \textstyle\sum_{m=1}^{M} \gamma_m\ \mathbb{I}_{st=m}, \alpha(s_t) = \sum_{m=1}^{M} \gamma_m\ \mathbb{I}_{st=m}, \quad \text{and}\ \beta(s_t) = \sum_{m=1}^{M} \gamma_m\ \mathbb{I}_{st=m}, \tag{2.9}$$

where

$$\mathbb{I}_{st=m} = \begin{cases} 1, & \text{if}\ \ s_t = m \\ 0, & \text{otherwise} \end{cases}. \tag{2.10}$$

By defining the allocation variable, s_t, as a M-dimensional discrete vector, $\xi_t = (\xi_{1t}', \dots\dots\dots, \xi_{Mt}')'$, where $\xi_{mt} = \mathbb{I}_{st=m}, m = 1, \dots\dots\dots..M$, the system of equation (2.6) and equation (2.7) can be rewritten compactly as:

$$y_t = \mu_t(y_{1:t-1}, \xi_t'\theta_t) + \sigma_t(y_{1:t-1}, \xi_t'\theta_t)\eta_t, \quad \eta_t \sim N(0,1), \tag{2.11}$$

$$\sigma_t^2 = (y_{1:t-1}, \xi_t'\theta_\sigma) = (\xi_t'\gamma) + (\xi_t'\alpha)\epsilon_{t-1}^2 + (\xi_t'\beta)\sigma_{t-1}^2(y_{1:t-2}, \xi_{t-1}'\theta_\sigma), \tag{2.12}$$

where e_i is the $i-$th column of a M-by-M identity matrix. The conditional probability of ξ_t given ξ_{t-1}, θ_π and $y_{1:t-1}$ is given by:

$$p(\xi_t'|\xi_{t-1}', \theta_\pi) = \prod_{m=1}^{M}(\pi_{mt}\xi_{t-1})^{\xi mt}. \qquad (2.13)$$

Equation (2.13) implies that the probability with which event m occurs at time t is $\pi_{mt}\xi_{t-1}$. Next Subsection 2.2.4 presents the path dependence challenge faced by MS-GARCH models.

2.2.4. PATH DEPENDENCE CHALLENGE

Estimating the MS-GARCH model is a challenge since the likelihood of y_t depends on the entire sequence of past states up to time t due to the recursive structure of its volatility (Grey, 1996). To elaborate on this, the likelihood function of the MS-GARCH model is given by:

$$\mathcal{L}(\theta|y_{1:T}) = f(y_{1:T}|\theta) = \sum_{i=1}^{M} \dots\dots.. \sum_{j=1}^{M} f(y_{1:T}, \xi_1' = e_T'|\theta), \qquad (2.14)$$

where $\theta = \left(\{\theta_{m\mu}, \theta_{m\sigma}\}m = 1 \dots\dots.., M, \theta_\pi\right)$. Setting $\xi_{s:t} = (\xi_s', \dots\dots\dots, \xi_t')$ whenever $s \le t$, the joint density functions of $y_{1:t}$ and $\xi_{1:t}$ on the right hand side of equation (2.14) is:

$$\begin{aligned} f(y_{1:T}, \xi_{1:T}|\theta) &= f\left(y_1\middle|\xi_{1:1}, \theta_\mu, \theta_\sigma\right) \prod_{t-2}^{T} f\left(y_t\middle|y_{1:t-1}, \xi_{1:1}, \theta_\mu, \theta_\sigma\right) p(\xi_t|y_{1:t-1}, \xi_{1:t-1}, \theta_\pi) \\ &= f\left(y_1\middle|\xi_{1:1}, \theta_\mu, \theta_\sigma\right) \prod_{t-2}^{T} f\left(y_t\middle|y_{1:t-1}, \xi_{1:1}, \theta_\mu, \theta_\sigma\right) \left(\prod_{i=1}^{M}(\pi_{it}\xi_{t-1})^{\xi it}\right) \qquad (2.15) \end{aligned}$$

with,

$$f\left(y_t\middle|y_{1:t-1}, \xi_{1:1}, \theta_\mu, \theta_\sigma\right) \infty \frac{1}{\sigma_t(y_{1:t-1}, \xi_t'\theta_\sigma)} \exp\left(-\frac{1}{2}\left(\frac{y_t - \mu_t(y_{1:t-1}, \xi_t'\theta_\mu)}{\mu_t(y_{1:t-1}, \xi_t'\theta_\mu)}\right)^2\right).$$

Given σ_1, recursive substitution in equation (2.12) yields:

$$\sigma_t^2 = \sum_{i=0}^{t-2}[\xi_{t-i}'\gamma + (\xi_{t-1-i}'\sigma)\epsilon_{t-1-i}^2] \prod_{j=0}^{i=1} \xi_{t-j}'\beta + \sigma_1^2 \prod_{i=0}^{t-2} \xi_{t-i}'\beta. \qquad (2.16)$$

Equation (2.14) shows the dependence of conditional variance at time t on the entire history of the regimes, and by inference, the dependence of the likelihood function on the entire history of the regimes (Billio *et al*., 2013). The evaluation of the likelihood function over a sample of length T, as noticed in equation (2.9), involves integration (summation) overall M^T unobserved states *i.e.* integration over all M^T possible (unobserved) regime paths. This requirement makes the maximum likelihood estimation of θ in equation (2.9) infeasible in practice. Maximum likelihood (ML) procedure has shown some convergence problems and it has a problem of local maxima (Ardia, 2008). Cai (1994) and Hamilton & Susmel (1994) reported that the disadvantage of MLE approach becomes computationally unfeasible for MS-GARCH models. This led to some authors such Dueker (1997), Gray, (1996), Haas *et al*., (2004) and Klaassen, (2002) to propose modified estimation versions of the MS-GARCH model that avoids the path dependence problem which is often experienced when using the ML estimation method. In order to solve this estimation problem using maximum likelihood procedure, different procedures were suggested in the literature based on model approximation. Section 2.5 presents the various MS-GARCH estimation methods.

2.2.5. THE MS-GARCH MODEL APPROXIMATION

Billio *et al*. (2013) introduced different approximation approaches to address the path dependence problem. The various approximates of the MS-GARCH models are the following: Basic Approximation; Gray (1996) approximation; Dueker (1997) approximation; simplified version of Klaassen (2002) approximation; and Klaassen (2002) approximation. According to Billio *et al*. (2013) the first step in approximating the MS-GARCH model is by eliminating the problem of path dependence and then deriving a proposed distribution for state variables from the auxiliary model. A possible way for avoiding the path dependence problem inherent in the MS-GARCH model is to replace the lagged conditional variance appearing in the definition of the GARCH model with a proxy. Literature reported on different auxiliary models which differ only by the content of the information in defining the proxy used in each case. In general, various MS-GARCH (as available in the literature) can be obtained by approximating the conditional variance,

$$\sigma_t^2(y_{1:t-1}, \theta_\sigma(s_t)) = V\left[y_t \middle| y_{1:y-1}, s_{1:t}\right] = V\left[\epsilon_t \middle| y_{1:y-1}, s_{1:t}\right] \tag{2.17}$$

of the GARCH process as

$$\sigma_t^2(y_{1:t-1}, \xi_t'\theta_\sigma) \approx \xi_t'\gamma + (\xi_t'\alpha)\epsilon_{(X)t-1}^2 + (\xi_t'\beta)\epsilon_{(X)t-1}^2. \tag{2.18}$$

In this section, the study presents alternative specifications of $\epsilon_{(x)t-1}$ and $\sigma_{(x)t-1}^2$ that define different approximations of the MS-GARCH model. The following are the five (5) different approximate models:

Model 1: Basic approximation (Model B)

A first attempt to eliminate the path dependence challenge was by Billio *et al*. (2013). The authors also mention that the conditional density of ϵ_t is a mixture of normal distributions with zero means and time varying variances. Hence, they estimated the switching GARCH model by replacing the lagged conditional variance, σ_{t-1}^2 , with the variance σ_{t-1}^2 of the conditional density of ϵ_t, *i.e.*;

$$\epsilon_{(B)t-1} = y_{t-1} - \mu_{(B)t-1},$$

$$\mu_{(B)t-1} = E[\mu_{t-1}(y_{1:t-2}, \xi'_{t-1})|y_{1:t-2}|] = E[y_{t-1}|y_{1:t-2}]$$

$$= \sum_{m=1}^{M} \mu_{t-1}(\mu_{t-2}, e'_m\theta_\mu)q(\xi'_{t-1} = e'_m|y_{1:t-2}),$$

$$\sigma_{(B)t-1}^2 = E[\sigma_{t-1}^2(y_{1:t-2}, \xi'_{t-1})|y_{1:t-2}] = E[\epsilon_{t-1}^2|y_{1:t-2}] = V(\epsilon_{t-1}|y_{1:t-2})$$

$$= \textstyle\sum_{m=1}^{M} \mu_{t-1}(\mu_{t-2}, e'_m\theta_\mu)q(\xi'_{t-1} = e'_m|y_{1:t-2}). \qquad (2.19)$$

Observe that in this approximation scheme, $\mu_{(B)t-1}$ and $\sigma_{(B)t-1}^2$ are the functions of $y_{1:t-2}$ and the information coming from $y_{1:t-1}$ is lost. With $q(\xi'_{t-1} = e'_m|y_{1:t-2})$ known from $m = 1, \ldots\ldots\ldots, M, \mu_{(B)t-1}$ can be computed while $\sigma_{(B)t-1}^2$ can be computed recursively since $\sigma_{t-1}^2(y_{1:t-2}, \xi'_{t-1})|y_{1:t-2}$ depends on $\sigma_{(B)t-1}^2$. Note that in this approximation, the conditioning is on $y_{1:t-2}$. This approach represents a starting point for other approximations hence Billio *et al*. (2013) tagged it basic model approximation.

Model 2: Gray's approximation (Model G)

Gray (1996) notes that the conditional density of the return process, y_t, of the MS-GARCH model is a mixture of normal distribution with time-varying parameters. Hence, the author suggested the use of the variance of conditional density $\sigma_{(G)t-1}^2$ of y_t as a proxy for the lag of the conditional variance σ_{t-1}^2 switching GARCH process, *i.e.*:

$\epsilon_{(G)t-1} = y_{t-1} - \mu_{(G)t-1,}$ where

$\mu_{(G)t-1} = \mu_{(B)t-1}.$ and

$$\sigma_{(G)t-1}^2 = V(y_{t-1}|y_{1:t-2}) = V(E[y_{t-1}|y_{t-2}, \xi'_{t-1}]|y_{1:t-2}) + E[V(y_{t-1}|y_{t-2}, \xi'_{t-1})|y_{1:t-2}]$$

$$= V(\mu_{t-1}(y_{1:t-2}, \xi'_{t-1}\theta_\mu|y_{1:t-2}) + E[(\sigma_{t-1}^2(y_{1:t-2}, \xi'_{t-1}\theta_\mu)|y_{1:t-2})]$$

$$= E(\mu_{t-1}(y_{1:t-2}, \xi'_{t-1}\theta_\mu|y_{1:t-2}) + E[(\sigma_{t-1}^2(y_{1:t-2}, \xi'_{t-1}\theta_\mu)|y_{1:t-2})]$$

$$= \sum_{m=1}^{M} \left(\mu_{t-1}\left(\mu_{t-2}, e'_m\theta_\mu\right)q(\xi'_{t-1} = e'_m|y_{1:t-2})\right)q(\xi'_{t-1} = e'_m|y_{1:t-2}) - \left(\mu_{(B)t-1}\right) + \sigma^2_{(B)t-1}. \quad (2.20)$$

Similarly, as in the basic approximation, information on y_{t-1} is lost in this approximation scheme as $\mu_{(B)t-1}$ and $\sigma^2_{(G)t-1}$ are functions of $y_{1:t-2}$. By recursion, $\sigma^2_{(G)t-1}$ can be computed since $\sigma^2_{(B)t-1}$ depends on $\sigma^2_{(B)t-2}$ through $\sigma^2_{t-1}\left(\mu_{t-2}, e'_m\theta_\mu\right)$. Within this framework, the conditioning is also on $y_{1:t-2}$. The major difference between Model 1 and 2 can be seen from the development of the proxy, *i.e,* $V(y_{t-1}|y_{1:t-2})$ is replaced with $V(\epsilon_{t-1}|y_{1:t-2})$ in Gray's approximation.

Model 3: Dueker's approximation (Model D)

In the previous approximation schemes, the information coming from y_{t-1} is not used. Dueker (1997) suggests that y_{t-1} be included in the conditioning set of the proxy while assuming that μ_{t-1} and σ^2_{t-1} are functions of $(y_{1:t-2}, \xi'_{t-2})$. The following relation can thus be credited to Dueker (1997):

$$\epsilon_{(D)t-1} = y_{t-1} - \mu_{(D)t-1},$$

$$\mu_{(D)t-1} = E\left[\mu_{t-1}\left(y_{1:t-2}, \xi'_{t-1}\theta_\mu\right)|y_{1:t-1}\right] = \sum_{m=1}^{M} \mu_{t-1}\left(\mu_{t-2}, e'_m\theta_\mu\right)q(\xi'_{t-1} = e'_m|y_{1:t-1}),$$

$$\sigma^2_{(D)t-1} = E\left[\sigma^2_{t-1}\left(y_{1:t-2}, \xi'_{t-1}\theta_\mu\right)|y_{1:t-1}\right] = \sum_{m=1}^{M} \sigma^2_{t-1}\left(\mu_{t-2}, e'_m\theta_\mu\right)q(\xi'_{t-1} = e'_m|y_{1:t-1}). \quad (2.21)$$

The probability $q(\xi'_{t-1} = e'_m|y_{1:t})$ is a one period ahead smoothed probability which can be computed as:

$$q(\xi'_{t-1} = e'_m|y_{1:t}) = \sum_{m=1}^{M} q(\xi'_{t-1} = e'_m, \xi'_{t-1} = e'_m|y_{1:t})$$

$$= \sum_{i=1}^{M} q(\xi'_{t-1} = e'_m|\xi'_t = e'_i, y_{1:t})q(\xi'_i = e'_i|y_{1:t})$$

$$= \sum_{i=1}^{M} q(\xi'_{t-1} = e'_m|\xi'_t = e'_i, y_{1:t-1})q(\xi'_i = e'_i|y_{1:t})$$

$$= \sum_{i=1}^{M} \frac{q(\xi'_{t-1} = e'_m | \xi'_t = e'_i, y_{1:t}) q(\xi'_i = e'_i | y_{1:t})}{q(\xi'_t = e'_i | y_{1:t-1})}$$

$$= q(\xi'_{t-1} = e'_m | y_{1:t-1}) \sum_{i=1}^{M} \frac{q(\xi'_{t-1} = e'_m | \xi'_t = e'_i, y_{1:t}) q(\xi'_i = e'_i | y_{1:t})}{q(\xi'_t = e'_i | y_{1:t-1})}. \quad (2.22)$$

Within this framework, Dueker (1997) notes that the conditioning is on $y_{1:t-1}$ while the functional form depends on $(y_{1:t-2}, \xi'_{t-2})$. Dueker (1997) equally notes that at every time step t, the value of $q(\xi'_{t-2} = e'_m | y_{1:t-1})$ for all m is required for computation.

Model 4: Klaassen's approximation (Model SK)

The following approximation is similar to model 3. As opposed to Dueker's approximation, Billio *et al*. (2013) assume that μ_{t-1} and σ^2_{t-1} are functions of $(y_{1:t-2}, \xi'_{t-2})$. This modification leads to the following approximation Klaassen (2002) model:

$$\epsilon_{(K)t-1} = y_{t-1} - \mu_{i,(K)t-1}$$

$$\mu_{(SK)t-1} = E\left[\mu_{t-1}\left(y_{1:t-2}, \xi'_{t-1}\theta_\mu\right) | y_{1:t-1}\right]$$

$$= \sum_{m=1}^{M} \mu_{t-1}\left(y_{1:t-2}, e'_{t-1}\theta_\mu\right) q\left(\xi'_{t-1} = e'_m | y_{1:t-1}\right)$$

$$\sigma^2_{(SK)t-1} = E[\sigma^2_{t-1}(y_{1:t-2}, \xi'_{t-1}\theta_\sigma) | y_{1:t-1}]$$

$$= \sum_{m=1}^{M} \sigma^2_{t-1}(y_{1:t-2}, e'_{t-1}\theta_\sigma) q\left(\xi'_{t-1} = e'_m | y_{1:t-1}\right). \quad (2.23)$$

In the next approximation, the current regime will be added to the conditioning set of this version of the auxiliary model. Hence, this approximation will be identified as the simplified version of Klaassen (2002) model. In order to implement this approximation scheme, the value of $q(\xi'_{t-1} = e'_m | y_{1:t-1})$ for all m is required at each point in time t.

Model 5: Klaassen's approximation (Model K)

In each of the approximations described above, information relating to the current regime is ignored in the conditioning set. On observing this, Klaassen (2002) suggested the following approximation:

$$\epsilon_{(K)t-1} = y_{t-1} - \mu_{(K)t-1},$$

$$\mu_{(K)t-1} = E\left[\mu_{t-1}\left(y_{1:t-2}, \xi'_{t-1}\theta_\mu\right) | y_{1:t-1}, \xi'_{t-1} = e'_i\right]$$

$$= \sum_{m=1}^{M} \mu_{t-1}\left(y_{1:t-2}, \xi'_{t-1}\theta_{\mu}\right) q\left(\xi'_{t-1} = e'_m | y_{1:t-1}, \xi'_{t-1} = e'_i\right)$$

$$\sigma^2_{i,(K)t-1} = E[\sigma^2_{t-1}(y_{1:t-2}, \xi'_{t-1}\theta_{\sigma})|y_{1:t-1}]$$

$$= \sum_{m=1}^{M} \left(\mu_{t-1}\left(y_{1:t-2}, e'_{t-1}\theta_{\mu}\right) + \sigma^2_{t-1}(y_{1:t-2}, \xi'_{t-1}\theta_{\sigma})\right) q(\xi'_{t-1} = e'_m | y_{1:t-1}, \xi'_{t-1} = e'_i)$$

$$- \left(\sum_{m=1}^{M} \mu_{t-1}\left(y_{1:t-2}, e'_{t-1}\theta_{\mu}\right) q(\xi'_{t-1} = e'_m | y_{1:t-1}, \xi'_{t-1} = e'_i)\right)$$

where

$$q(\xi'_{t-1} = e'_m | y_{1:t-1}, \xi'_{t-1} = e'_m) = \frac{q(\xi'_{t-1} = e'_m | \xi'_t = e'_i, y_{1:t}) q(\xi'_i = e'_i | y_{1:t})}{q(\xi'_t = e'_i | y_{1:t-1})}$$

$$= \frac{q\left(\xi'_{t-1} = e'_m \middle| \xi'_t = e'_i, y_{1:t}\right) q\left(\xi'_i = e'_i \middle| y_{1:t}\right)}{q\left(\xi'_t = e'_i \middle| y_{1:t-1}\right)}. \tag{2.24}$$

Note that this approximation requires the computation of $q(\xi'_{t-1} = e'_m | y_{1:t-1}, \xi'_{t-1} = e'_i)$ for all m and i at time t. Several authors (Gray, 1996; Dueker, 1997; Klaassen, 2002) made attempts to tackle the path dependence problem, while other provided an alternative method to tackle to path dependence problem using different estimation methods. The next Section 2.3 presents the empirical literature.

2.3. EMPIRICAL LITERATURE

There is a developing interest to Markov Switching models (MSM) in economic and financial time series data works. However, in most cases, literature available is on studies around the interest rate, stock market returns and exchange rates. Mostly, studies in the literature found MSM superior and relevant to explain the business cycles of the economy. Hamilton (1989) suggested the applications of Hamiltonian Monte Carlo (HMC) models to be incorparated in Markov switching framework. The parameters of the suggested models were estimated by MLE, where the regime probabilities are obtained by the proposed Hamilton-filter (1994). Maximum Likehood estimation method of the models is centered on the Expectation Maximization (EM) algorithm (Hamilton, 1990).

In the MS models, regime changes are unobserved and are a discrete state of a Markov chain which governs the endogenous switches between different AR processes throughout time.

Following the seminar study of Hamilton (1989, 1990), many studies used MS framework to analyse the behaviour of different markets. Later, various modifications of MS models were presented and models shown improvements. Hamilton & Susmel (1994) and Cai (1994) suggested the use of MS-ARCH model to study the behaviour of financial time series data. Lamoureux & Lastrapes (1990) and Mikosch & Starica (2004) also made a suggestion that MS-GARCH model can be useful in modelling financial time series data. These models were found revelent since they are able to capture the effects of the structural changes (or sudden changes) in the conditional variance.

Since the modifications and the improvements of MS models, the number of studies has significantly increased in modelling the volatility of financial time series data using MS-extension models such as MS-ARCH and MS-GARCH. However, the estimation of GARCH-type models is often carried out using the MLE and Bayesian MCMC estimation methods. These estimation methods are used to estimate the parameters of the models and provide the better accuracy forecasts of the GARCH-type models. Furthermore, the estimation methods are valuable to these GARCH-type in studying the volatility of the financial time series data and its behaviour.

Many experiments or demonstrations were conducted across the globe to show the abilities of these models in financial time series data. A study on forecasting the volatility for crude oil prices was conducted by Zhang et al. (2015) of which the single-regime GARCH models as well as the two-regime Markov Regime Switching GARCH (MRS-GARCH) models were adopted in the study. The authors believe that for the study to attain reliable and accurate forecast approach a performance assessment on the above mentioned models must be conducted at different data occurences and distances in time. The research study resulted that: (i) the two-regime MRS-GARCH model performs better than the other three single-regime GARCH type models in in-sample data estimation, (ii) the linear single-regime GARCH model overall performs better than other three nonlinear GARCH type models in VaR forecast.

Naeini & Fatahi (2012) compared the standard GARCH model and Markov switching GARCH models in forecasting the volatility of the Tehran stock market. The study provided an analysis of regime switching in volatility and out-of-sample forecasting of the Tehran stock market. The study modelled volatility regime switching within a univariate Markov-Switching framework. Then, the study provided the results of the in-sample and out-of-sample forecasts of the Tehran stock market. The results of the study showed that MS-GARCH models performed better than standard GARCH model.

Chen *et al*. (2009) proposed the use of Double Markov-Switching GARCH (DMS-GARCH) to model mean and volatility asymmetry of financial time series data. The advantage of the model is the ability to capture the characteristics of the stock returns, such as as leptokurtosis, volatility clustering, and an asymmetric relationship between stock returns and conditional variance. The authors applied the models in the eight international stock market returns. The parameters of the models were estimated using Bayesian MCMC estimation method and the Metropolis-Hastings (MH) method. The results of the study showed that Markov switching model with t-distribution and the GARCH-t performed well in capturing the characteristics of stock returns. The findings showed that the eight market indices have the fat-tail characteristics when they are modelled with DMS-GARCH model. Furthermore, the study compared the model based on Value-at-risk (VaR) forecasting and results showed that DMS-GARCH model outperformed other computing models.

Later, Bauwens *et al. (*2010) developed an MS-GARCH model that has a finite number of regimes in each of which the conditional mean is constant and the conditional variance takes the form of a GARCH(1,1) process. The switching is governed by a hidden Markov chain. The study provided sufficient conditions for geometric periodicity and the existence of moments of the process. Beuwens *et al*. (2010) developed this model to overcome the infeasible path dependence of MLE using S&P500 daily returns. The study incorporated a Bayesian MCMC estimation that is feasible. Based on residual diagnostics and the Bayesian information criterion, the study found that the MS-GARCH model with two regimes fits the data much better than the corresponding ARCH model.

Henneke *et al*. (2011) demonstrated the ability of MS-ARIMA-GARCH using MLE and Bayesian MCMC estimation methods. The study used New York Stock exchange rate and simulated data. The results showed that when MS-ARIMA-GARCH model is estimated with MLE, it poses severe difficulties when it comes to computation of parameter estimates. Furthermore, the MLE becomes unfeasible due to the path dependence of GARCH-type models. But when the MS-ARIMA-GARCH which estimated with Bayesian MCMC framework provided better estimate the parameters of the models. The results of study revealed that models estimated using Bayesian MCMC provided better accuracy compared to the models that used MLE method.

Sattayathama *et al.* (2012) applied Markov Regime Switching-GARCH models that allow volatility to have different dynamics according to unobserved regime variables. The main objective of this study was determining whether or not the MRS-GARCH models are an improvement on the GARCH type models in terms of modelling and forecasting the SET50 Index closing price volatility. Authors compared the MRS-GARCH (1,1) models with GARCH, EGARCH and GJRGARCH models. These models are estimated with MLE under three distributional assumptions that are Normal, Student-t and GED. The study used in-sample and out-of-sample technique to measure the ability of the models. The results of the study revealed that the MRS-GARCH with t-distribution (two degree of freedoms) model outperformed the GARCH, EGARCH and GJRGARCH models.

Bildirici & Ersin (2014) proposed asymmetric power properties to MS-GARCH processes and a group of nonlinear GARCH models, which encompasses fractional assimilation. The study also aimed at enhancing the MS-GARCH type models with artificial neural networks that will benefit from the general estimate properties to attain enhanced forecasting accuracy. The study used the ISE100 daily returns to demonstrate the ability of the models. The study proposed the MS-ARMA-FIGARCH, APGARCH, and FIAPGARCH procedures that are then improved with MLP, Recurrent NN, and Hybrid NN type neural networks. The study outcome suggests that the fractional assimilation and asymmetric equivalents of the MS-GARCH model yielded encouraging outcomes whereas the best results are attained for their neural network counterparts. The study compared the MS-ARMA-FIGARCH, APGARCH, and FIAPGARCH using error metrices (MAE, MSE, and RMSE) and Diebold-Mariano (DM) equal forecast accuracy tests. The results of the study showed that neural network outperformed the MS-GARCH models.

Günay (2015) demonstrated the ability of Markov regime-switching GARCH (MRS-GARCH) to compare GARCH, exponential GARCH, and Glosten-Jagannathan-Runkle GARCH models. Brent crude oil returns was used in this study. The study used the MLE method through Expectation Maximization to estimate the parameters of the models. Furthermore, the study used cumulative sum of squares test to assess the existence of structural breaks in the return series. The volatility models were formed under normal, generalized error distribution and

Student's t distributions. The best model was selected based on the values of Akaike Information Criterion (AIC) and a Bayesian information criterion (BIC) value. MRS-GARCH model outperformed all other alternative models using Student’s t distributions.

Abdullah *et al.* (2017) applied GARCH-type models to estimate and forecast exchange rate volatility of Bangladeshi taka (BDT) and the US dollar. The objective of the study was to address the issue of error distribution assumption in modelling and forecasting exchange rate volatility between the Bangladeshi taka and the US dollar. The study compared generalized autoregressive conditional heteroscedastic (GARCH), asymmetric power ARCH (APARCH), exponential generalized autoregressive conditional heteroscedastic (EGARCH), threshold generalized autoregressive conditional heteroscedastic (TGARCH), and integrated generalized autoregressive conditional heteroscedastic (IGARCH) processes under both normal and Student’s t-distribution assumptions for errors. The results of the study showed that, the application of Student-t distribution for errors induced the models to meet the diagnostic tests and improve forecasting accuracy. With such error distribution for out-of-sample volatility forecasting, AR(2)–GARCH(1,1) was considered the best.

Iqbal (2016) compared and evaluated various GARCH models’ abilities to forecast volatility and risk of Karachi Stock Exchange (KSE). Linear and nonlinear GARCH and Markov Regime-Switching GARCH (MRS-GARCH) models with normal and fat tail errors were employed to predict 1-day to 1-month ahead forecast of volatility and Value-at-Risk (VaR). The MRS-GARCH model was estimated with two regimes, representing periods of low and high volatility of stock returns. The MRS-GARCH-*t* model for the short horizon and EGARCH-t model for long-horizon were found fit to predict the volatility and risk of KSE better than the competing models.

Alemohammad *et al*. (2017) proposed a Markov switching asymmetric GARCH model that follows a logistic smooth transition structure between the effects of positive and negative shocks. The study applied the MCMC method through Gibbs and griddy Gibbs sampling to parameters of the model. The authors made a demonstration by using S&P 500 indices to show the competing performance of in-sample fit and out-of-sample forecast volatility and value at risk of the proposed model. The Diebold-Mariano test showed that the proposed model outperforms all competing models in forecast volatility. Shiferaw (2018) objective of this study was to examine the relationship between the prices of agricultural commodities with the oil price, gas price, coal price and exchange rate (USD/Rand).

In addition, the study tried to fit an appropriate model that best describes the log return price volatility and estimate Value-at-Risk (VaR). The data used in this study were the daily returns of agricultural commodity prices from 02 January 2007 to 31st October 2016. The study applied the three-state Markov-switching (MS) regression, the standard single-regime GARCH, and the Markov-switching GARCH (MS-GARCH) models. To choose the best fit model, the log-likelihood function, Akaike information criterion (AIC), Bayesian information criterion (BIC) and deviance information criterion (DIC) were employed under different distributions for innovations. The results indicated that the price of agricultural commodities was significantly associated with the price of coal, the price of natural gas, price of oil and exchange rate. Moreover, for most of the agricultural commodities considered in this study, the MS-GARCH models under the MCMC approach outperformed the standard single regime GARCH models in measuring VaR. In conclusion, this study provided a practical guide for modelling agricultural commodity prices by MS regression and MS-GARCH processes.

Ardia *et al.* (2018) conducted a large-scale empirical study to compare the forecasting performances of single-regime and MS-GARCH models from a risk management outlook. The MS-GARCH models produced the most accurate VaR, expected shortfall, and left-tail distribution forecasts than their single-regime counterparts for daily, weekly, and ten-day equity log-returns. These authors also suggested that the MS-GARCH models whose parameters can change over time according to a separate hidden variable are provided to contend the switch in the return process. These models can easily attune to variations in the unconditional volatility level, which then progresses the risk predictions. Furthermore, the results indicated that accounting for parameter uncertainty also progresses the left-tail predictions, autonomously of the presence of the Markov-switching mechanism.

2.4.SUMMARY OF THE CHAPTER

The chapter reviewed the theoretical and empirical literature of volatility models. Literature showed that GARCH-type models are good in terms of model economic and financial data. Furthermore, the GARCH-type models identified as the better models to capture the volatility of exchange rates, stock market returns and interest rates especially MS-GARCH models. Literature showed that GARCH-type models that were estimated using MLE procedure are not producing the best results of GARCH-type models. The studies showed that MLE procedure suffers from convergence problems and it has a problem of local maxima.

Further, the studies showed that MLE procedure becomes computationally unfeasible for GARCH-type such as MS-GARCH and MS-ARIMA-GARCH models. Alternative models were developed to tackle the path dependence problem. Bayesian MCMC models were developed as an alternative estimation method to circumvent the path dependence problem and several studies from the reviewed literature showed this estimation method provide promising results. Even though Bayesian MCMC is able to avoid the path dependence problem, it comes with a cost. The disadvantage is that the Bayesian MCMC does not have criterion to select a prior and there is no correct way to choose a prior. Bayesian inferences require skills to translate subjective prior beliefs into a mathematically formulated prior. If the study does not proceed with caution, the study might generate misleading results.

Based on the benefits and shortcomings of the two procedures, the current study finds it fit to assess the abilities of the methods using MRS-GARCH to model the BRICS countries exchange rates. Furthermore, the study uses various distributions in these BRICS countries exchange rates. To the best knowledge of the researcher, there is no study conducted where the two procedures (MLE and Bayesian MCMC) and various distributions were used to model BRICS countries exchange rates data to do “in-sample and out-of-sample evaluation” volatility forecasts. Furthermore, there is no study conducted using Diebold-Mariano test of predictive accuracy using exchange rates of BRICS countries.

CHAPTER 3
RESEARCH METHODOLOGY

3.1. INTRODUCTION

This chapter presents a detailed description of the research methods that were used to conduct the research. The current study investigates the performance of MS-GARCH estimation methods and distribution with application of the exchange rates of BRICS countries. For the purpose of this study, the researcher aims to achieve the following objectives:

i. To estimate MLE and Bayesian MCMC MRS-GARCH models of the exchange rates of BRICS countries
ii. To compare the efficiency and performance of MLE and Bayesian MCMC MRS-GARCH models for in-sample and out-of-sample volatility forecasts
iii. To compare the predictive accuracy of MRS-GARCH models using MLE and Bayesian MCMC

The rest of the chapter is organised as follows: Section 3.2 discusses the study ethical considerations, Section 3.3 provides this study research process and Section 3.4 discusses the preliminary data analysis. Section 3.5 presents the assessment of data for linearity and Section 3.6 presents nonlinear unit root test. Section 3.7 presents modelling and estimation methods. Section 3.8 presents model selection criteria. Section 3.9 discusses the forecasting exchange rate volatility (in-sample and out-of-sample). Section 3.10 presents the forecast evaluation. Section 3.11 discusses the model forecasting accuracy and the last Section presents concluding remarks on the chapter.

3.2. ETHICAL CONSIDERATIONS

There is no ethical consideration related to environment, animals or human subjects in this study. The study only uses secondary data. However, the researcher remains truthful to the recommendations of the North-West University institutional manual of postgraduate studies for the handling and manipulation of data in the execution of the study. Permission to conduct the study was sought from the University by applying for ethical clearance through the supervisor after the defence and approval of the proposal.

3.3. RESEARCH PROCESS

In order to give direction to this study, the research process as described by Saunders *et al*. (2012) was followed. The illustration by Saunders *et al*. (2012) uses the onion metaphor. This onion illustrates a range of choices, paradigms, strategies and steps followed by researchers during the research process (see Figure 3.1).

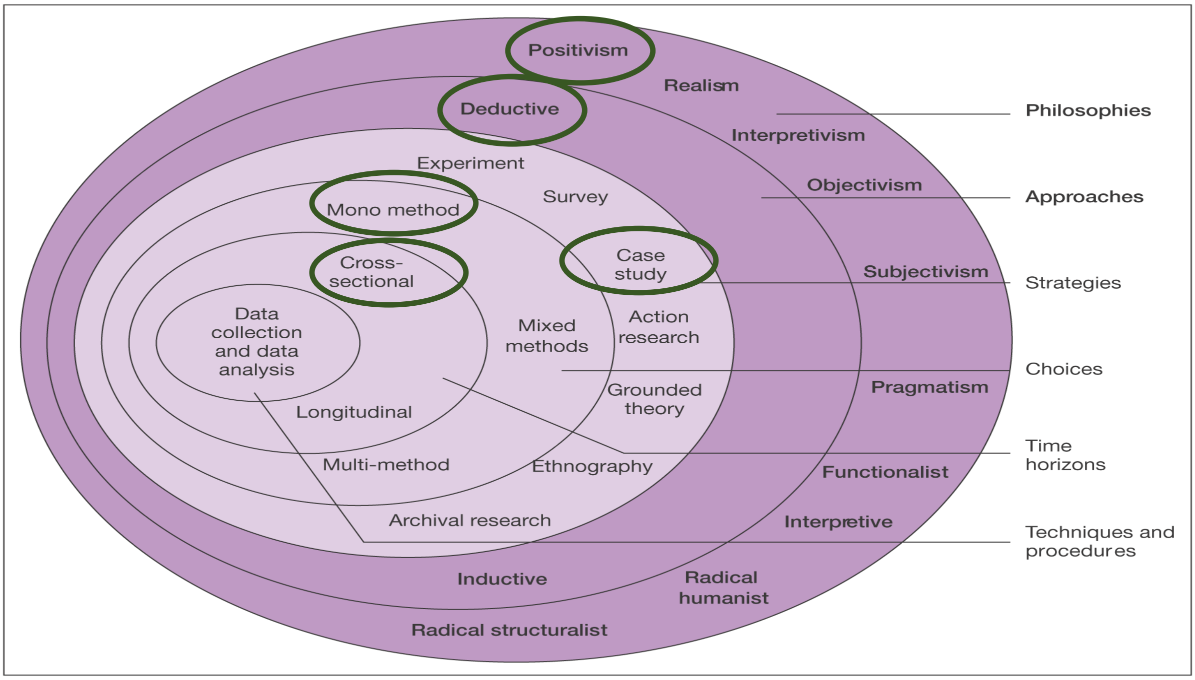


Figure 3.1: The research process onion (Saunders *et al*., 2012)

3.3.1. RESEARCH PHILOSOPHY

A research philosophy is the first and most important layer of the research onion. It discusses the number of philosophies that are available to the researchers. According to Saunders *et al*. (2012), research philosophy is important as it is what a researcher does when conducting research in a topic: Research is increasing the understanding in a certain field effectively. It is also a belief or an idea about the collection, interpretation, and analysis of data collected. One can further refer to a research philosophy as the development of knowledge in a particular field. Moreover, a research philosophy refers to the set of beliefs concerning the nature of the reality being investigated (Bryman, 2012). It is the underlying definition of the nature of knowledge. The assumptions created by a research philosophy provide the justification for how the research will be undertaken (Flick, 2011). These are not necessarily at odds with each other, but the choice of research philosophy is defined by the type of knowledge being investigated in the research project (May, 2011). All the authors mentioned above agree that the research philosophy is mainly the how aspect of the research (how knowledge is generated, and how we know it).

There are various philosophies explained in the research onion; the major standpoints relate to epistemology, ontology, and axiology (Saunders & Thornhill, 2012). Collis and Hussey (2013) define epistemology as to what makes the acceptable knowledge. Epistemology can be divided into two areas namely: positivism and interpretivism. According to Hwang (1996), positivism is a broader variety of theories and practices. Gephart (1999) described interpretivism as directed to different meanings and understandings of the social interactions between humans. Positivism depends on the scientific knowledge while interpretivism mainly depends on the individual viewpoint. Therefore, the most appropriate philosophical stance for this study is positivistic epistemology. The study seeks answers to theory-driven questions or objectives (Creswell, 2013; and Creswell & Tashakkori, 2007) as outlined in Chapter 1 hence the applicability of this philosophy.

3.3.2. RESEARCH APPROACH

Research approach is the second layer of the research onion as illustrated in Figure 3.1. There are two types of approaches outlined: the deductive approach and the inductive approach. The deductive approach concentrates on using the literature to identify theories and ideas that the researcher will test using data. This approach develops the hypothesis upon a pre-existing theory and then formulates the research approach to test it (Silverman, 2013). The deductive approach is best suited to contexts where the research project is concerned with examining whether the observed phenomena fit with expectation based upon previous research (Wiles *et al*., 2011). The deductive approach, thus, might be considered particularly suited to the positivist approach, which permits the formulation of hypotheses and the statistical testing of expected results to an accepted level of probability (Snieder & Larner, 2009).

On the other hand, the inductive approach involves collecting data and developing a theory based on the results of data analysis (Saunders *et al*., 2009). The inductive approach is characterised as a move from the specific to the general (Bryman & Bell, 2011). In this approach, the observations are the starting point for the researcher, and patterns are looked for in the data (Beiske, 2007). There is no framework that initially informs the data collection and the research focus can thus be formed after the data has been collected (Flick, 2011). This method is more commonly used in qualitative research, where the absence of a theory informing the research process may be of benefit by reducing the potential for researcher bias in the data collection stage (Bryman and Bell, 2011). Interviews are carried out concerning specific phenomena and then the data may be examined for patterns between respondents (Flick, 2011).

The current study uses secondary data and also uses methods that follow a quantitative approach, hence a deductive approach was followed to help achieve the set objectives.

3.3.3. RESEARCH STRATEGY AND TIME HORIZONS

The research strategy refers to how the researcher intends to carry out the work (Saunders *et al*., 2007). This enables a researcher to answer research objectives or questions which shape the flow and structure of the study. The necessity of deploying a strategy is based on the aims and objectives of a study. According to Saunders *et al*. (2012) emphasised that the choice of a strategy is guided by research questions and objectives. This study guided by research objectives.

Primarily, in existing knowledge, the amount of time available, as well as philosophical underpinnings are also important. Furthermore, there are five steps to be followed as part of research strategy. According to Verschuren (2010) these are (1) Deciding on opting for breadth or depth in view of the research objective and researcher's own expertise and interests. (2) Deciding on a quantitative or qualitative approach, by using the same type of arguments as in Step 1. (3) Determining empirical or nonempirical type of research. (4) Selecting one of the five research strategies to use in the research project, not necessarily in its exact form, based on the decisions taken in the above steps. (5) Selecting one variant of a strategy and its detailed characteristics on the basis of research objectives. For the current study, the research strategy used is a case study since the study is conducted in BRICS countries using the exchange rates of those countries. The countries referred to are Brazil, Russia, India, China and South Africa.

Collis & Hussey (2013) defined a case study as a strategy for conducting a research that involves empirically investigating certain contemporary phenomenon in the context of its real life and using multiple sources of evidence. Case study is the assessment of a single unit in order to establish its key features and draw generalisations (Bryman, 2012). This study uses secondary data for BRICS as a case study. The fact is that this study conducts an empirical investigation into BRICS countries' exchange rates as a research strategy. The study further develops conclusions based on the data analysis; Shiferaw (2018) and Ardia *et al*. (2018) used a strategy that is relevant for the current study.

Another important part of a research is the time horizon. Time horizon has two elements, which are cross-sectional and longitudinal studies. Cross-sectional studies are described by Saunders *et al*. (2012) as a type of observational study that analyses data collected from a population, or a representative subset, at a specific point in time, while longitudinal studies represent events over a long period of time. Schmidt & Kohlmann (2008) emphasised that cross-sectional studies can be used to describe, not only the odds ratio but also absolute risks and relative risks from prevalence (sometimes called prevalence risk ratio). This study is undertaken to answer the predetermined research objectives and address a specific problem at a particular time, therefore a cross-sectional strategy is thought appropriate.

3.3.4. DATA COLLECTION AND ANALYSIS

Data collection and analysis are dependent on the methodological approach used (Bryman, 2012). The process used at this stage of the research contributes significantly to the study and its overall reliability and validity (Saunders *et al*., 2007). Regardless of the approach used in the project, the type of data collected can be separated into two types: the primary and the secondary data.

3.3.4.1. THE PRIMARY DATA

Primary data is that which is derived from first-hand sources. This can be historical first-hand sources, or the data derived from the respondents in the survey or interview data (Bryman, 2012). However, it is not necessarily data that has been produced by the research being undertaken. For example, data derived from statistical collections such as the census can constitute primary data. Likewise, data that is derived from other researchers may also be used as primary data, or it may be represented by a text being analysed (Flick, 2011).

3.3.4.2. SECONDARY DATA

Secondary data is that which is derived from the work or opinions of other researchers (Newman, 1998). For example, the conclusions of a research article can constitute secondary data because it is information that has already been processed by another. Likewise, analyses conducted on statistical surveys can constitute secondary data (Kothari, 2004). However, there is an extent to which the data is defined by its use, rather than its inherent nature (Flick, 2011). Newspapers may prove both a primary and secondary source for data, depending on whether the reporter was actually present. Therefore, the most effective distinction of the two types of data is perhaps established by the use to which it is put in a study, rather than to an inherent characteristic of the data itself.

The current study used secondary data which was obtained from www.quantec.com. To be specific, the study employed monthly exchange rates (US dollars) from BRICS countries for the period from January 1997 to November 2017, a total of 241 observations. The choice of the data and time period was based on the size required to fit the MRS-GARCH model and the frequency (monthly) enabled the study to reveal proper structural breaks in the data. This data is also relevant and suited for the set objectives and study significance. Researchers from other countries such as Kutu & Ngalawa (2017) and Jandoo & Gonpot (2018) have also used similar data on studies involving the application of MRS-GARCH models. Statistical software package used was R program, Ox-Metrics and E-views. These softwares provide results relevant to achieve the study objectives.

3.4. PRELIMINARY DATA ANALYSIS

In this section, the study reviews the preliminary data analysis methods used. To assess the characteristics of the data, tests such as the normality, nonlinearity and nonlinear unit root were conducted as suggested by Thorsten & Herbert (2004) and Schumacker & Lomax (2004).

3.4.1. ASSESSING UNIVARIATE NORMALITY

Financial time series data is non-normal in nature. Financial time series is characterised by stylised facts. According to Franses & van Dijk (2000) and Humala & Rodriguez (2013), large returns occur more often than expected (leptokurtosis or fat tails). This implies that the kurtosis is much larger than 3, or the tails of the distributions are fatter than the tails of the normal distribution. Therefore, it is important to confirm whether or not the data is normally distributed before conducting further analysis. To confirm the assertion, Tsay (2005) highlighted that financial time series are high-peaked and fat-tailed distributions of financial returns. These are common knowledge amongst practitioners.

This study used Jarque-Bera (JB) test which was developed by Jarque & Bera (1980). This test is based on skewness which is supposed to be equivalent to or around zero, reflecting a symmetric distribution where the mean, median and mode are equal. It is also based on kurtosis which should be equal to or around three, indicating perfect tallness of the data. Skewness and kurtosis of the data are expressed as follows respectively:

$$s = \frac{\sum_{i=1}^{N}(Y_1 - \overline{Y})^3}{(N-1)S^3} \tag{3.1}$$

$$k = \frac{\sum_{i=}^{N}(Y - \overline{Y})^4}{s^4} \tag{3.2}$$

where S is the standard deviation, Y is the actual data, $\overline{Y}$ is the mean and N is the total of the whole data set in used. After computing the prerequisites of the normality test, the last step is to calculate the JB test which is calculated using the following equation:

$$JB = \frac{N-k}{6}[s^2 + \frac{(k-3)^2}{4} \tag{3.3}$$

where

N = number of observations

S = skewness of a variable

k = kurtosis of a variable

When computing the JB test, the results are expected to be insignificant because the null hypothesis is that, the data is normally distributed while the alternative is that the data is not normally distributed. The conclusion is based on the benchmark that if p-value is less than 0.05, then the null hypothesis is rejected. The subsequent section presents the nonlinearity tests.

3.5. ASSESSMENT OF DATA FOR LINEARITY

In this section, the study reviews the nonlinearity methods for data assessment. In order to apply the nonlinear methods, the data was first tested for nonlinearity. Since nonlinearity in time series may occur in several ways, it is imprudent to rely solely on a single test (Ismail & Isa, 2006). In that case, to test for nonlinearity in the data sets, the Regression Specification Error Test (RESET), Keenan’s test, Tsay test, Brock- Dechert-Scheinkman (BDS), and CUSUM test were used.

3.5.1. THE RESET TEST

This section discusses the RESET test for assessing linearity in the data. According to Ramsey (1969), the RESET test is a specification test for linear regression analysis. In the context of this study, the commonly used linear regression model is the univariate autoregressive model of order p, denoted by AR (p) and is defined as follows:

$$Y_t = \beta_0 + \sum_{j=1}^{p} \beta_j Y_{t-p} + \varepsilon_t, \tag{3.4}$$

where $\beta_0, \beta_1, \beta_2, \dots\dots.., \beta_p$ are parameters and ε_t is *iid* random variable with mean 0 and variance σ_ε^2. The AR order, p, is selected to minimize the error, ε_t. This is practically accomplished by selecting a value for p that minimizes an information criterion, such as the SBC (Franses and Dijk, 2000). If $Y_t = \left(Y_{t-1} Y_{t-2} \dots\dots.. Y_{t-p}\right)'$, equation (3.4) becomes:

$$Y_t = Y_{t-1}' \beta_0 + \varepsilon_t\ . \tag{3.5}$$

The RESET test involves, first, obtaining the OLS estimate, β, in equation (3.5), the residual $\varepsilon_t = Y_t - \widehat{Y}_t$, and the sum squared residuals:

$$SSR_0 = \sum_{t=p+1}^{n} \hat{\varepsilon}_t^2. \tag{3.6}$$

The second step involves estimating the regression

$$\hat{\varepsilon}_t = Y_{t-1}' + M_{t-1}' \lambda_2 + e_t, \tag{3.7}$$

where $M_{t-1} = \left(\widehat{Y}_t^2\ \widehat{Y}_t^3 \dots\dots \widehat{Y}_t^{s+1}\right)$ for some $s \geq 1$, $\hat{e}_t$ is an independent and identically distributed random variable with mean 0 and variance σ_e^2. From the estimated residuals $\hat{e}_t = \hat{\varepsilon}_t - \hat{\varepsilon}_t$, the sum of squared residuals is computed as:

$$SSR_1 = \sum_{t=p+1}^{n} \hat{\varepsilon}_t^2. \tag{3.8}$$

If the underlying AR (p) is adequate, the RESET test asserts that λ_1 and λ_2 are zero. Thus, the hypotheses to test are:

$H_0 = \lambda_1 = \lambda_2 = 0$ (Specification is indeed linear)

vs.

$H_1: \lambda_j \neq 0$ for at least one j (Specification is nonlinear)

The test statistic is the usual F-statistic of the equation given by

$$F^* = \frac{(SSR_0 - SSR_1)/r}{SSR_1/(n-p-r)} \sim F_\alpha(r, n-p-r) \tag{3.9}$$

where $(r = s + p + 1)$. At the α level, the null hypothesis of linearity is rejected in favour of the alternative hypothesis if

$$F^* = F_\alpha(r, n - p - r) \quad \text{or} \quad \text{pro}(F^*) < \alpha. \tag{3.10}$$

This means that the F test statistic is greater than the F critical value, and the study rejects the null hypothesis that the true specification is linear (which implies that the true specification is non-linear).

3.5.2. KEENAN'S TEST

The section discusses the Keenan's test for nonlinearity. Keenan (1985) proposed a nonlinearity test for time series that uses $\widehat{Y}_t^2$ only and modified the second step of the RESET test to avoid multicollinearity between $\widehat{Y}_t^2$ and Y_{t-1}. In particular, Keenan assumed that the series can be approximated (Volterra expansion) as follows:

$$Y_t = \mu + \sum_{u=-\infty}^{\infty}\sum_{v=-\infty}^{\infty}\theta_u a_{t-u} + \sum_{u=-\infty}^{\infty}\sum_{v=-\infty}^{\infty}\theta_{uv} a_{t-u}\, a_{t-v} \tag{3.20}$$

Clearly, if $\sum_{u=-\infty}^{\infty}\sum_{v=-\infty}^{\infty}\theta_u a_{t-u} + \sum_{u=-\infty}^{\infty}\sum_{v=-\infty}^{\infty}\theta_{uv} a_{t-u}\, a_{t-v}$ is zero, the approximation is linear, so Keenan's idea shares the principle of an *F* test. The procedure is in the same steps as Ramsey's test. Firstly, select (with a selection criterion, e.g. AIC) the value p of the number of lags involved in the regression, then fit $(\widehat{Y}_t)$ on $(1, Y_{t-1}, \ldots, Y_{t-p})$ to obtain the fitted values $(\widehat{Y}_t)$, the residuals set (at) and the residual sum of squares SSR. Then regress $\widehat{Y}_t^2$ on $(1, Y_{t-1}, \ldots, Y_{t-p})$ to obtain the residuals set$(\hat{\zeta}_t)$. Finally calculate

$$\hat{\eta}_t = \frac{\sum_{t=p+1}^{n}\hat{a}_t\hat{\zeta}_t}{\sum_{t=p+1}^{n}\hat{\zeta}_t^2} \tag{3.21}$$

and the test statistic equals

$$\hat{F} = \frac{(n-2p-2)\hat{\eta}^2}{(SSR-)\hat{\eta}^2}. \tag{3.23}$$

Under the null hypothesis of linearity, i.e.,

$$H_0 = \sum_{u=-\infty}^{\infty}\sum_{-\infty}^{\infty}\theta_{uv}\, a_{t-u}a_{t-v} = 0$$

and the assumption that (at) are i.i.d. Gaussian, asymptotically $\hat{F} \sim F_{1,n-2p-2}$. Tsay (1986) improved on the power of the Keenan (1985) test by allowing for disaggregated nonlinear variables (all cross products $Y_{t-1}, Y_{t-p}, i, j = 1, \dots, p$), thus generalizing Keenan test by explicitly looking for quadratic serial dependence in the data. While the first step of Keenan test is unchanged, in the second step of Tsay test, instead of $(Y_t)^2$, the products $Y_{t-1}, Y_{t-p}, i, j = 1, \dots, p$ are regressed on $\left(1, Y_{t-1}, \dots, Y_{t-p}\right)$. Hence, the corresponding test statistic ~ F is asymptotically distributed as $F_{m,n-m-p-1}$, where $m = \frac{p(p-1)}{2}$.

3.5.3. TSAY TEST (1986)

Tsay (1986) developed a test that helps improve the power of the Keenan test. In order to improve the power of the nonlinearity tests developed by Keenan (1985) and Ramsey (1969), Tsay (1986) proposed to use a different set of explanatory variables for the test. The test is based on running an auxiliary equation in the form:

$$\hat{a}_t = \beta'^{Z_t} + u_t, \tag{3.24}$$

where $Z_t = \mathrm{vech}(X_t X_t')$ is a vector of predetermined variables, their squares and cross products, and vech(.) denotes a half-stacking operator. The LM version of the test statistic is defined as:

$$\mathrm{TSAY}(p) = TR^2 \xrightarrow{d} \chi^2\left(\frac{p(p+1)}{2}\right), \tag{3.25}$$

where T denotes the sample size, and R^2 is the coefficient of determination from an auxiliary model. In a special case when p = 1, the Tsay test coincides with the Keenan test proposed by Keenan (1985).

3.5.4. THE BROCK-DECKERT-SCHIENKMAN (BDS) TEST

The BDS test was introduced by Brock-Deckert-Schienkman to test if the data is nonlinear in nature (Brock *et al.*, 1996). If equation (3.4) is correctly specified, then under the null hypothesis of linearity, the residuals should be serially independent. This forms the basic idea behind various tests of nonlinearity. In practice, diagnostic tests of serial independence typically are based on certain aspects of the data such as the serial correlations or ARCH-type dependence, while other tests explore dependence by testing the iid condition of the residual term, which is sufficient for serial independence (Kuan, 2002). One such test is the so-called BDS test - a form of portmanteau test. Portmanteau tests are residual-based tests in which the null hypotheses are well-stated, but do not necessary have well-stated alternative hypotheses.

The BDS test can be applied to the estimated residuals from any time series process provided the time series process can be transformed into a form with *iid* errors. The BDS test, which focuses on the residual obtained after a linear structure has been removed from a process, tests the null hypothesis of linearity against a variety of alternative hypotheses. Under the null hypothesis of the BDS test, if the residuals are iid or follow a white noise process, then its m-lagged (also referred to as embedding dimension) correlation integral (also referred to as correlation function) is equal to the correlation integral of the (m-1)-lagged residuals. BDS test statistic is given by Brock *et al*. (1996) as:

$$\mathrm{BDS}_{m,n} = \frac{\sqrt{n}[C_{m,n}(\epsilon) - C_{1,n}(\epsilon)^m]}{\sigma_m(\epsilon)}, \tag{3.26}$$

where $C_{m,n}(\epsilon)$ the correlation is integral, $\sigma_m(\epsilon)$ is the asymptotic standard deviation of the numerator, and ϵ is the maximum difference between pairs of observations used in calculating the correlation integral. Brock *et al*. (1996) showed that under the null hypothesis of the residuals being iid or following a white noise process,

$$\mathrm{BDS}_{m,n} \sim N(0,1) \tag{3.27}$$

The null hypothesis of iid residuals (whiteness or linearity) is rejected if the test statistic exceeds the critical value at the α-level of significance or if the p-value of $\mathrm{BDS}_{m,n}$ is lower than α. Rejection of the null hypothesis is indicative of nonlinear dependence in time series data.

3.5.5. CUSUM TEST

Stability is another aspect of nonlinearity in data (Brown *et al*., 1975). CUSUM examines data stability by testing for possible structural change in the data. On the one hand, if the model is stable, then β and the variance of the residuals do not change over time. In that case, the coefficients, $\hat{\beta} = (1\ \beta_1, \beta_2, \dots\dots, \beta_p)$, in equation (3.5) can be obtained from the matrix (Brown *et al*.,1975):

$$\hat{\beta} = (Y'_{t-1}Y_{t-1})^{-1}Y'_{t-1}Y_t \tag{3.28}$$

where Y_t is the dependent variable in equation (3.4) and $Y_{t-1} = \left(1\ Y_{t-1}\ Y_{t-2} \dots\dots Y_{t-p}\right)'$ and $\varepsilon_t \sim \mathrm{iid}(0, \sigma_\varepsilon^2)$. On the other hand, if the model is unstable, then β and the variance of the residuals possibly change over time. In that case, then β is replaced by b_t, say, and so

$$b_{t-1} = (Y'_{t-1}Y_{t-1})^{-1}Y'_{t-1}Y_{t-1} \tag{3.29}$$

where $Y_{t-1} = \left(1\; Y_{t-1}\, Y_{t-2} \ldots \ldots Y_{t-p}\right)'$ and $\varepsilon_t \sim iid\left(0, \sigma_{t,s}^2\right)$. If the AR(p) is stable, the parameters remain constant over time, suggesting the absence of any structural change in the data. Thus, the hypotheses to test are:

$H_0 = \beta_1 = \beta_2 = \cdots = \beta_2 = \beta$

vs.

$H_1 = \sigma_{1,s}^2 = \sigma_{2,s}^2 = \cdots = \sigma_{n,s}^2 = \sigma_s^2$.

Then, the variance of the recursive residuals is computed as

$$\mathrm{var}(\varepsilon_t) = E(\varepsilon_t'\varepsilon_t) = \mathrm{var}\{\varepsilon_t - Y_t'(b_{t-1} - \beta)\}$$
$$= \sigma_s^2 + Y_t'(Y_{t-1}'Y_{t-1})^{-1}Y_t. \tag{3.30}$$

Define the scaled recursive residuals, ω_t, as

$$\omega_t = \frac{\varepsilon_t}{\sqrt{\left\{1+Y_t'\left(Y_{t-1}'Y_{t-1}\right)^{-1}Y_t\right\}}} = \frac{\varepsilon_t}{\sqrt{\mathrm{var}(\varepsilon_t)}} = \frac{\varepsilon_t}{\mathrm{s.e}(\varepsilon_t)} = \frac{\varepsilon_t}{\sigma_\omega}, \tag{3.31}$$

then under constant parameters, $\varepsilon_t \sim iid\left(0, \sigma_{t,s}^2\right)$. Then the CUSUM test statistic is given by:

$$W_t = \sum_{j=t-1}^{t} w_j \quad t = k+1, k+2, \ldots \ldots \ldots, n. \tag{3.32}$$

The cumulative sum of the square (CUSUMSQ) test statistic is given by:

$$S_t = \sum_{j=k+1}^{t} \omega_j^2 / \sum_{j=k+1}^{n} \omega_j^2 \qquad t = k+1, k+2, \ldots \ldots \ldots, n. \tag{3.33}$$

These tests are performed by plotting W_t or S_t against time t. The confidence bounds are obtained by plotting the two lines that connect the points $\left[k, \pm a\sqrt{n-k}\right]$ and $\left[n, \pm a\sqrt{n-k}\right]$. At the 5% level (that is, 95% confidence interval) $a = 0.948$, while at the 1% level (that is, 99% confidence interval) $a = 1.143$. A test statistic meandering outside the confidence interval is indicative of a possible a structural change, non-constancy in the parameters, and hence instability in the data, leading to the rejection of the null hypothesis of model stability.

Under the null hypothesis of model stability, Harvey & Collier (1977) developed a test with test statistic given by:

$$T^* = (n-1)^{1/2} S^{-1}.\bar{W} \tag{3.34}$$

where $\bar{W} = \frac{1}{n-k}\sum_{j=k+1}^{n}\omega_j$ and $S^2 = \frac{1}{n-k}\sum_{j=k+1}^{n}\left(\omega_j - \bar{W}\right)$.

The estimated T^* has a t-distribution with (n-k-1) degree of freedom. The null hypothesis of model stability is rejected if T^* is greater than a critical value at the α-level or if the p-value of T^* is less than α, often 0.05.

3.6. NONLINEAR UNIT ROOT TEST

In this section, nonlinear test for stationarity is conducted to determine the existence of nonlinear dependence in the data. This section presents nonlinear test for stationarity. In order to apply the nonlinear GARCH for modelling and forecasting, the five exchange rates used in the study must be investigated for the presence of stationarity. For the current study, nonlinear stationarity is tested using the Bierens Nonlinear ADF (B-NLADF) and Kapetanois-Schmidt-Shin Nonlinear Augmented Dickey-Fuller (KSS-NADF).

3.6.1. THE B-NADF UNIT ROOT TEST

The Bierens Nonlinear ADF (B-NLADF) test is a modification of the Augmented Dickey-Fuller (ADF) unit root test in which ADF auxiliary regression is further augmented with orthogonal Chebyshev polynomial. In the B-NLADF test, the null hypothesis of a unit root is tested against the alternative of nonlinear trend stationarity. The B-NLADF auxiliary regression is given by the following equation which is adopted from (Bierens, 1997):

$$\Delta Y_t = \xi Y_{t-1} + \sum_{j=1}^{p} \delta_j \Delta Y_{t-j} + \sum_{k=0}^{m} \theta_k C_{k,t} + \varepsilon_t, \qquad (3.35)$$

where $C_{1,t}, C_{2,t}, \ldots\ldots, C_{m,t}$ are Chebychev polynomials with $C_{0,t} = 1$,, $C_{1,t}$ capturing a linear trend in the data, and are cosine functions. The hypotheses to test are:

$H_0: \xi = 0$ (Nonlinear unit root)

vs. $H_1: \xi < 0.$ (No nonlinear unit root)

Three important tests - $t(m), A(m)$ and $F(m)$ – are proposed by Bierens (1997) under the B-NLADF test. The t(m) test is a t-test on the significance of the estimated coefficient, ξ. Under the null hypothesis, the F(m) test is a joint F-test that the estimated coefficient, ξ, and the coefficients of the nonlinear Chebyshev polynomials are zero (not significant). The A(m) test statistic given by the following equation and is adopted from Bierens (1997):

$$A(m) = \frac{n.\xi}{\left|1-\sum_{j=1}^{p} \hat{\delta}_j\right|}, \qquad (3.36)$$

where n is the sample size, is an alternative test for t(m) test, and therefore can be used to check the robustness of results of the t(m) test. As the tests are prone to size distortions, critical values used are based on the Monte Carlo simulations with 1000, 2000, 5000 or 10000 replications of a Gaussian AR(p) process for the first-difference data, ΔY_t, where p is determined by an information criterion such as AIC, SBC and HQC. As argued in Bierens (1997), there is not a unique way of choosing m, as a low value may be insufficient to detect nonlinearity under the alternative hypothesis and a too high value may result in lack of power. For this reason, this study considers different values of m with the optimal order, p, determined by the SBC criterion. Decisions about nonlinear stationarity or otherwise, would be made based on comparisons of test statistics with asymptotic critical values. The null hypothesis of a nonlinear unit root is rejected if a test statistic is less than its associated critical value at the 1%, 5% or 10% level. This concludes that the data in use is nonlinear unit root.

3.5.2 The KSS-NADF Unit Root Test

In the presence of nonlinearity in a time series data, the conventional unit root test, such as the ADF-GLS discussed above, may be inadequate for detection of stationarity in the data. In such a scenario it is therefore important that stationarity tests that accommodate nonlinearity, such as the KSS test, be used. The KSS test is a modification of the Augmented Dickey-Fuller (ADF), based on the following nonlinear model specification (Kapetanios *et al.,* 2003):

$$Y_t = \beta Y_{t-1} + \gamma Y_{t-1}\left[1 - \exp\left(-\theta Y_{t-d}^2\right)\right] + \varepsilon_t \qquad (3.37)$$

which when parameterised yields:

$$\Delta Y_t = \delta Y_{t-1} + \gamma Y_{t-1}\left[1 - \exp\left(-\theta Y_{t-d}^2\right)\right] + \varepsilon_t \qquad (3.38)$$

where $\delta = \beta - 1$, γ, θ are parameters that must be estimated and ε_t is the residual term. The KSS test sets $\delta = 0$ and the decay parameter, $d = 1$, so that the test is formally based on the following specification:

$$\Delta Y_t = \gamma Y_{t-1}\left[1 - \exp\left(-\theta Y_{t-d}^2\right)\right] + \varepsilon_t. \qquad (3.39)$$

The KSS tests the null hypothesis of linear stationarity by setting $\theta = 0$ against the alternative that $\theta > 0$. However, Katepenios *et al.,* (2003) argue that it is impossible to directly test the null hypothesis since the speed of reversion, γ, is unknown. Using a first-order Taylor series approximation, Luukkonen *et al.* (1998) reformulated an estimable nonlinear specification for testing nonlinear stationarity in X_t as:

$$\Delta Y_t = \xi Y_{t-1}^3 + \varepsilon_t. \tag{3.40}$$

To account for the possibility of serial correlation in the error term, equation (3.40) is augmented with lags of the first-difference of X_t as:

$$\Delta Y_t = \xi Y_{t-1}^3 + \sum_{j=1}^{p} \delta_j \Delta Y_{t-j} + \varepsilon_t \tag{3.41}$$

where ξ is the coefficient used to test the presence of a unit root. From the nonlinear stationarity specification, the KSS-NADF unit root test is based on the t-statistic:

$$\tau_{NL} = \frac{\xi}{s.e(\xi)} \tag{3.42}$$

where s.e.(ξ) is the standard error of ξ. The following hypotheses are tested:

$H_0 : \xi = 0$ (Nonlinear nonstationarity)

vs. $H_1 : \xi < 0$. (Nonlinear stationarity)

Three different asymptotic critical values are constructed with three different nonlinear model specifications - raw data, demeaned data, and de-trended data (Kapetanios *et al.,* 2003). The following scenarios prevail:

- If X_t has a zero mean, then the appropriate data to use is $Y_t = X_t$, the raw data.
- If X_t has a non-zero mean and zero trend, then the appropriate data to use is $Y_t = X_t - \bar{X}$, the demeaned data, where $\bar{X}$ is the mean of the data.
- If X_t has a non-zero mean and non-zero trend, then the appropriate data to use is $Y_t = X_t - (\alpha_0 + \alpha_1 t)$, the de-trended data, where $\alpha_0 + \alpha_1 t$ is the trendline obtained by regressing X_t on time point t=1, 2, 3,.., n with an intercept term.

The KSS-NLADF test is sensitive to the choice of lag length, p. One prominent approach to select p is use of general-to-specific method suggested by Hall (1994). This involves setting up an upper bound, p_{max}, suggested by Schwert (1989):

$$P_{max} = \text{integer}\left[12.\left(\frac{n}{100}\right)^{1/4}\right] \tag{3.43}$$

where n is the sample size, estimating the test regression with p= p_{max}. This study will, however, appeal to the lag length of 8 as recommended by Liew *et al.* (2004). At the 1%, 5% or 10% level, if the last included lag is significant, it is retained as the optimal lag and used in the KSS-NLADF unit root test. However, if the last included lag is not significant, p is reduced by one lag until last included lag is significant, and used as the optimal lag for the KSS-NLADF unit root test. The null hypothesis of the nonlinear unit root is rejected in favour of the alternative if the t-statistic is greater than the critical value at some α-level of significance. Table 3.1 presents asymptotic critical values at the 1%, 5% and 10% levels.

Table 3.1: Critical Values for KSS Nonlinear Unit Root Tests

Significance Level	Raw Data	De-Meaned Data	De-Trended Data
1	-2.82	-3.48	-3.93
5	-2.22	-2.93	-3.40
10	-1.92	-2.66	-3.13

3.7. MODELLING AND ESTIMATION METHODS

This section presents an overview of the nonlinear Markov regime switching GARCH model and MLE and Bayesian MCMC methods.

3.7.1. MARKOV REGIME SWITCHING GARCH MODEL

The Markov switching regime model adopted relies on two assumptions: i) the volatility process is characterized by two regimes, high volatility and low volatility; ii) the high volatility regime is associated with exchange rate (high values of the mean process) and the low volatility regime is associated with small exchange rate movements (low values of the mean process). The presence of two regimes of low volatility versus high volatility motivates the Markov switching approach adopted. Brunetti *et al*. (2008) documented that within the two regimes, volatility is not constant, and hence the need for a GARCH approaches. Following the study of Sajjad *et al*. (2008), Sopipan *et al*. (2012) and Iqbal (2016), the $\mathrm{MRS}-\mathrm{GARCH}(\mathrm{p},\mathrm{q})$ is presented as:

$$r_t = \delta_{s_t} + \varepsilon_t = \delta_{s_t} + \eta_t\sqrt{h_{t,s_t}}, \tag{3.44}$$

and $h_{t,s_t} = \alpha_{0,s_t} + \alpha_{1,s_t}\varepsilon_{t-1}^2 + \beta_{1,s_t}h_{t-1}$,

where S_t=1or 2, δ_{s_t} is the mean and, h_{t,s_t} is the volatility under regime S_t on F_{t-1}, both are measurable functions of $F_{t-\tau}$ for $\tau \le t-1$. In order to ensure the positive of conditional variance, the study imposes the restrictions $\alpha_0 > \alpha_1 \ge 0$ and $\beta_1 \ge 0$. The sum $\alpha_{1,s_t} + \beta_{1,s_t}$ measures the persistence of a shock to the conditional variance. The unobserved regime variable S_t is governed by a first order Markov Chain with constant transition probabilities given by:

$$\Pr(S_t = i \mid S_{t-1} = j) = p_{ji} \text{ for } i,j = 1,2. \tag{3.45}$$

In matrix notation,

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \tag{3.46}$$

so that $p_{11} + p_{12} = 1$ and $p_{21} + p_{22} = 1$. P represents the probability of change in regime.

3.7.2. ESTIMATION METHODS

This section presents the estimation methods that used to estimate the parameters of the linear or nonlinear models. For the present study, to estimate the parameters of MRS-GARCH model, the study used both ML and Bayesian MCMC estimation methods. The estimation methods are discussed in the following subsections.

3.7.2.1. MAXIMUM LIKELIHOOD ESTIMATION METHOD

The maximum likelihood method is the most popular procedure used to estimate one or more parameters of the given model (Harris & Stocker, 1998, Fisher, 1938). The specification (3.25) causes difficulties in estimation since the Conditional variance at time t depends on the entire regime path S_t. To underline this dependence, Augustyniak (2014), denoted that the notation $(s_t)\varepsilon_{t-i}^2$ is used, but for simplicity, ε_t^2 is used to represent $(s_t)\varepsilon_{t-i}^2$. Furthermore, letting y and s to represent Y_t and (S_t) and the probability mass function (PMF) by $f(p)$, then the computation of the likelihood of the observations denoted by $f(y|\theta)$ can be proficient by integrating out all possible regime paths:

$$f(y|\theta) = \sum_{s} f(y|\theta) \sum_{s} f(y|\theta)\, p(s|\theta)$$

$$= \sum_{s} \left[\prod_{t-1}^{T} \varepsilon_t^{-1} \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{y_t - \mu_{s_t}}{2\varepsilon_t^2}\right)} \right] p(s|\theta), \tag{3.47}$$

when T is large, the integration in equation (3.38) is numerically infeasible because it contains N^t terms that quickly increases exponentially. Therefore, Bauwens *et al*. (2014) had shown that it is possible to accurately estimate the log-likelihood by:

$$\log f(y|\theta) = \log f(y|\theta) + \sum_{t=1}^{T-1} \log(y_{t+1}|y_{1:t}\theta), \tag{3.48}$$

and also estimating $f(y_{t+1}|y_{1:t}, \theta), t = 1, \ldots, T-1$ successively with the support of particle filters. For a parsimonious expression of the likelihood function, we define the vector $y = (y_1, \ldots, y_T)^{`}, \varpi = (\varpi_1, \ldots, \varpi_T)^{`}$, $s = (s_1, \ldots, s_T)^{`}$ and $\alpha = \left(\alpha_0^{`}, \alpha_1^{`}, \alpha_2^{`}\right)^{`}$. Finally, the diagonal matrix is $\sum = \sum \Theta = \operatorname{diag}\left(\{\varpi_t \varrho e_t^{`} \varepsilon^t\}_{t=1}^{T}\right)$. Note that ϱ, e_t and ε_t are the model parameter functions respectively. Therefore, the likelihood function of Θ is as follows (Bauwens *et al*., 2014):

$$l(\theta|y) \propto \left(\det \sum\right)^{\frac{-1}{2}} e^{\left[\frac{-1}{2} y \sum_{y}^{-1}\right]^{`}}. \tag{3.49}$$

3.7.2.2. BAYESIAN MCMC ESTIMATION METHOD

Estimating the MRS-GARCH model with Bayesian methods have more advantage over the classical approach. Firstly, proper computation methods are based on MCMC procedures and circumvents the common problem of local maxima encountered in MLE of this class of models (Ardia, 2008). Secondly, exploring the posterior joint distribution gives a comprehensive picture of the uncertainty of parameters that is not achievable with classical approach. For the model considered in this study, the asymptotic distribution theory of the MLE is inoperable because of the threshold parameters. Third, constraints on the model parameters can be incorporated through prior specifications. Finally, the likelihoods and Bayes factors can be used to achieve the discrimination between models. In the classical framework, testing the null of K versus K' states is not possible (Frühwirth-Schnatter, 2006).

Following Frühwirth-Schnatter (2006), Bauwens *et al.*(2014) and Ardia *et al.* (2016), the Normal, Student-t, and Generalised Error distribution (GED) specification (3.45) is nested in Bayes factor and the following posterior distribution is constructed using Baye's rule:

$$\begin{aligned} P(\Theta, S|Y) &\propto P(\Theta, S)P(Y|\Theta, S) \\ &\propto P(\Theta)P(S|\Theta)P(Y|\Theta, S), \end{aligned} \tag{3.50}$$

where, $\Theta = (\Upsilon, \alpha, \beta, \mu_1, \mu_2, P_{11}, P_{22})$, $S = (S_1, S_T)$ and $Y = (y_1, y_T)$. The Bayes' rule is applied in the first line. The second line comes from the definition of conditional probability. The posterior parameters are denoted by $P(\Theta)$. Under the assumption of independence, the prior density is chosen as:

$$P(\Theta) = P(\Upsilon)P(\mu_1)P(\mu_2)P(\alpha)P(\beta)P(P_{11})P(P_{22}). \tag{3.51}$$

Furthermore, the likelihood function of S is given by:

$$\begin{aligned} P(S|\Theta) &= P(S|P_{11}, P_{22}) \\ &= \prod_{t=1}^{T} P(S_{t+1}|S_t, P_{11}, P_{22}) \\ &= P_{11}^{\eta_{11}}(1 - P_{11})^{\eta_{12}} P_{22}^{\eta_{22}}(1 - P_{22})^{\eta_{21}} \end{aligned} \tag{3.52}$$

where the number of the transitions from state i to j are denoted by η_{ij}. The second line is due to the Markov property of S. The last term, $P(Y|\Theta, S)$ is the likelihood function Θ and S.

$$P(Y|\Theta, S) = \begin{aligned} &\prod_{t=1}^{T} P(y_t|Y_{t-1}, S_t, S_1, \Theta) \\ &= \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{\left[-\frac{\varepsilon_t^2}{2\sigma_t^2}\right]}. \end{aligned} \tag{3.53}$$

3.8. MODEL SELECTION CRITERIA

The section discusses the model selection by showing how the best fit model was arrived at. The study uses a few criteria to determine the best model and those criteria are listed below. To identify the best fitted model among the estimated nonlinear time series models, the study used the Akaike information criterion (AIC) by Akaike (1973), Schwartz Bayesian criterion (SBC) by Schwartz (1978) and Hannan-Quinn criterion by Hannan & Quinn (1979) and Hannan (1980). These model selections help to evaluate the ability of the models to describe the data (Xekalaki & Degiannakis, 2010). Brooks and Burke (2003) also applied these model selection criterions to select the best ARCH model in their study. The subsections below show the formulae to compute the AIC, SBC, HQC and DIC.

3.8.1. AKAIKE INFORMATION CRITERION (AIC)

Akaike (1973) derived the AIC as an approximately unbiased estimator of the Kullback-Leibler information. This criterion has been widely used by researchers. For instance, Hurvich & Tsai (1989) illustrated that AIC over-fit when the sample sizes are small or when the number of fitted parameters are a moderate to large fraction of the sample size. The study uses the AIC for the purpose of model selection.

$$\mathrm{AIC} = -2\mathrm{L}(\beta; \mathrm{D}) + 2\mathrm{p}, \tag{3.54}$$

with p being the dimension of β and D is data series. Model selection is accomplished by selecting from M models the one that minimizes AIC.

3.8.2. SCHWARZ INFORMATION CRITERION (SIC)

Schwarz (1978) derived SIC by approaching model selection from a Bayesian perspective. The criterion developed by Schwarz (1978) is referred to in literature in three different ways. Some researchers refer to the Schwarz information criteria as SIC, others call it the BIC, whilst there are some who combine the terms and refer to the criterion as the Schwarz Bayesian criterion (SBC). Throughout this study, the study standardizes on the notation and uses the SIC. Schwarz (1978) derived the original criterion in the linear models framework. The IC used is defined as:

$$\mathrm{SIC} = -2[\ln \widehat{\varphi} + \mathrm{k} * \ln(\mathrm{n})], \tag{3.55}$$

where n is the sample size, k is the number of parameters to be estimated, $\widehat{\varphi}$ is the likelihood function of the estimated model (M) which is $\widehat{\varphi} = \mathrm{p}\left(\mathrm{x}|\widehat{\theta}, \mathrm{M}\right)$,X is the observed data and θ is the parameter of the estimated model.

3.8.3. HANNAN-QUINN INFORMATION CRITERION (HQC)

Hannan & Quinn (1979) based their derivation of HQ on the approach adopted by Shibata for developing the Shibata's information criterion. They provided a procedure where the method was strongly consistent for estimating the order of an AR model. The IC derived is often referred to in literature as HQIC. For this study, the notation used is HQ. The IC function is given as:

$$HQ = -2\log\tilde{L}_n(\theta_n) + 2ck_0\log\log n, c > 1, \tag{3.56}$$

where k_0 number of is estimated parameters of the model, n is the sample size and $\tilde{L}_n(\theta_n)$ is the likelihood function.

3.8.4. DEVIANCE INFORMATION CRITERION (DIC)

Spiegelhalter *et al*. (2002) proposed DIC for Bayesian model selection. Deviance information criterion (DIC) has been extensively used for making Bayesian model selection. It is a Bayesian version of AIC and chooses a model that gives the smallest expected Kullback-Leibler divergence between the data generating process and a predictive distribution asymptotically. The DIC function is given as:

$$DIC = \overline{D} + p_D, \tag{3.57}$$

where the measure of the goodness of fit (i.e., deviance) is given by $\overline{D} = E_{\Theta|y}[D(\Theta)] = E_{\Theta|y}[-2\ln f(\Theta|y)]$ and the term which penalizes the complex models is given as $\overline{D} - D(\Theta), \Theta$ is the posterior mean of parameters.

3.9. FORECASTING THE EXCHANGE RATE VOLATILITY

The section looks to forecast ahead the certain time period using a k-step-ahead approach. There are three periods of interest given as 1-month, 5-months and 10-months. In the MRS-GARCH model with two regimes, Klaassen (2002) forecasted volatility for k-step-ahead by using the recursive as in the standard GARCH model where k is a positive integer. In order to compute the multi-step-ahead volatility forecast, the study firstly computes a weighted average of the multi-step-ahead volatility forecasts in each regime where the weights are the prediction probability $\left(\Pr(S_{l+\tau} = i|F_{t-1})\right)$. Since there is no serial correlation in the returns, the k-step-ahead volatility forecast at time t depends on information at times $t-1$. Let $h_{t,t+k}$ denote the time t aggregated volatility forecast for the next k steps. $h_{t,t+k}$ can be calculated as follows (Marcucci, 2005):

$$\hat{h}_{t,t+k} = \sum_{\tau=1}^{k} h_{t,t+\tau} = \sum_{\tau=1}^{k}\left[\sum_{i=1}^{2} \Pr S_{i+\tau} = i|\, F_{i-1}\, h_{t,t+\tau,S_{i+\tau=i}}\right], \tag{3.58}$$

where $h_{t,t+\tau,S_{i+\tau=i}}$ indicates the $\tau - \text{step} - \text{ahead}$ volatility forecast in regime i made at time t and can be calculated recursively as follows:

$$\hat{h}_{t,t+\tau,S_{i+\tau=i}} = E_{t-1}\, h_{t+\tau-1}\big|S_{t+\tau} = i = \alpha_{0..S_{t+\tau=i}} + \alpha_{1..S_{t+\tau=i}} + \beta_{1..S_{t+\tau=i}}\,.\,E_{t-1}h_{t+\tau-1}\big|S_{t+\tau} = i. \tag{3.59}$$

Also, in general the prediction probability (equation 3.58) is expressed as

$$\begin{bmatrix} \Pr S_{t+\tau} = 1| & F_{t-1} \\ \Pr S_{t+\tau} = 2| & F_{t-1} \end{bmatrix} P^{\tau+1}.\begin{bmatrix} \Pr S_{t+\tau} = 1| & F_{t-1} \\ \Pr S_{t+\tau} = 2| & F_{t-1} \end{bmatrix},$$

where P defined in (equation 3.44) as $\Pr(S_{l+\tau} = i|F_{t-1})$ will be discussed in (equation 3.41). Lastly, the study computes expectation part $E_{t-1}(h_{t,l+\tau-1}|S_{i+\tau} = i)$ as follows:

$$E_{t-1}(h_{t,l+\tau-1}|S_{i+\tau} = i) = E\left[h_{t,l+\tau-1}|S_{i+\tau} = i, F_{t-1}\right] = E\,\lceil E[{r^2}_{t+\tau-1}|S_{t+\tau-1} = j, F_{t-1}]|S_{t+\tau} = i, F_{t-1}\rceil - E\left[\left[E[r_{t+\tau-1}|S_{i+\tau-1} = j, F_{t-1}]\right]^2|S_{i+\tau} = i, F_{i-1}\right], \tag{3.60}$$

where

$$E[E[{r^2}_{t+\tau-1}|S_{t+\tau-1} = j, F_{t-1}]|S_{t+\tau} = i, F_{t-1}]$$
$$= \sum_{j=1}^{2}\left\{\begin{matrix} E[{\sigma^2}_{t+\tau-1} + 2\sigma_{t+\tau-1}\varepsilon_{t+\tau-1} + {\varepsilon^2}_{t+\tau-1}|S_{t+\tau-1} = j, F_{t-1}] \\ .\Pr[S_{t+\tau-1} = j|S_{t+\tau} = i, F_{t-1}] \end{matrix}\right\}$$
$$= \textstyle\sum_{j=1}^{2} \overline{P}_{ji,t-1}\left[{\sigma^2}_{t+\tau-1,S_{i+\tau-1=j}} + h_{t+\tau-1,S_{t+\tau-1=j}}\right], \tag{3.61}$$

and

$$\overline{P}_{ji,t-1} = \frac{P_{ji}\Pr S_{t+\tau-1}=j|F_{t-1}}{\Pr S_{t+\tau}=i|F_{t-1}} \tag{3.62}$$

Similarly, the study showed in the second term on the right hand side such that

$$E\left[\left[E[r_{t+\tau-1}|S_{t+\tau-1} = j, F_{t-1}]\right]^2|S_{t+\tau} = i, F_{t-1}\right] = \textstyle\sum_{j=1}^{2} \overline{P}_{ji,t-1}\left[\sigma_{t+\tau-1,S_{i+\tau-1=j}}\right]^2. \tag{3.63}$$

Substitute both equations (3.54) and (3.56) to equation (3.53) such that

$$E_{T-1}(h_{t,t+\tau-1}|S_{t+\tau} = i) = \textstyle\sum_{j=1}^{2} \overline{P}_{ji,t-1}\left[{\sigma^2}_{t+\tau-1,S_{i+\tau-1=j}} + h_{t+\tau-1,S_{t+\tau-1=j}}\right] - \sum_{j=1}^{2} \overline{P}_{ji,t-1}\left[\sigma_{t+\tau-1,S_{i+\tau-1=j}}\right]^2. \tag{3.64}$$

Next, the study will compute those regime probabilities

$p_{it} = \Pr(S_t = i|F_{t-1})$ for $i = 1,2$ in equation (3.45). (3.65)

Note that when the regime probabilities are on information up to time t, the study describes this as filtered probability $\left(\Pr(S_t = i|F_t)\right)$. In order to compute the regime probabilities, the study denotes $f_{1t} := f(r_t|S_t = 1, F_{t-1}), f_{2t} := f(r_t|S_t = 2, F_{t-1})$. Then the conditional distribution of return series r_t becomes a mixture-of-distribution models in which mixing variables are regime probability p_{it}. That is:

$$r_t|F_{t-1}] \begin{cases} f(r_t|S_t = 1, F_{t-1}) \text{ with probability } p_{1t} \\ f(r_t|S_t = 2, F_{t-1}) \text{ with probability } p_{2t} = 1 - p_{1t}, \end{cases} \tag{3.66}$$

where $f(r_t|S_t, F_{t-1})$ denotes one of the assumed conditional distributions for errors: Normal distribution (N), student-t distribution with only single degree of freedom (t) and generalized error distributions (GED) (Appendix A). The study shall compute regime probabilities recursively by following two steps (Kim and Nelson, 1999):

Step 1: given $\Pr\left((S_{t-1} = j|F_{t-1})\right)$ at the end of the time $t-1$, the regime probabilities $p_{it} = \Pr(S_t = i|F_{t-1})$ are computed as $\Pr(S_t = i|F_{t-1}) = \sum_{j=1}^{2} \Pr(S_t = i, S_{t-1} = j|F_{t-1})$. Since the current regime (S_t) only depends on the regime one period ago (S_{t-1}), then

$$\Pr(S_t = i|F_{t-1}) = \sum_{j=1}^{2} \Pr(S_t = i|S_{t-1} = j)\Pr(S_{t-1} = j|F_{t-1}) = \sum_{j=1}^{2} p_{ji}\Pr(S_{t-1} = j|F_{t-1}). \tag{3.67}$$

Step 2: At the end of the time t, when observed return at time $t(r_t)$ the information at time t set $F_t = \{F_{t-1}, r_t\}$, the $\Pr(S_t = i|F_t)$ is calculated as follows:

$$\Pr(S_t = i|F_t) = \Pr(S_t = i|r_t, F_{t-1}) = \frac{f(r_t, S_t = i|F_{t-1})}{f(r_t|F_{t-1})} \tag{3.68}$$

where $f(r_t, S_t = i|F_{t-1})$ is joint density of returns and unobserved regime at state i for $i = 1,2$ variables can be written as follows:

$$f(r_t, S_t = i|F_{t-1}) = f(r_t|S_t = I, F_{t-1})f(S_t = i|F_{t-1}) = f(r_t, S_t = i|F_{t-1})\Pr(S_t = i|F_{t-1}), \tag{3.69}$$

and $f(r_t|F_{t-1})$ is marginal density function of returns which can be constructed as follows:

$$f(r_t|F_{t-1}) = \sum_{i=1}^{2} f(r_t|S_t = i, F_{t-1})\Pr(S_t = i|F_{t-1}), \tag{3.70}$$

where Bayesian argument is used

$$\Pr(S_t = i|F_t) = \frac{f(r_t, S_t = i|F_{t-1})}{f(r_t|F_{t-1})} = \frac{f(r_t, S_t = i|F_{t-1})\Pr(S_t = i|F_{t-1})}{\sum_{i=1}^{2} f(r_t, S_t = i|F_{t-1})\Pr(S_t = i|F_{t-1})} = \frac{f_{it}p_{it}}{\sum_{i=1}^{2} f_{it}p_{it}}. \tag{3.71}$$

Then, all regime probabilities p_{it} can be computed by iterating these two steps. However, at the beginning of iterations $\Pr(S_0 = i|F_0)$ for $i = 1,2$ are necessary to start iterations. Hamilton (1989, 1990) suggests the use of unconditional regime probabilities instead of $\Pr(S_0 = i|F_0)$. These are given by:

$$\pi_1 = \Pr(S_0 = 1|F_0) = \frac{1-q}{2-p-q}, \pi_2 = \Pr(S_0 = 2|F_0) = \frac{1-q}{2-p-q}. \tag{3.72}$$

Given initial values for regime probabilities, conditional mean and conditional variance in each regime, the parameters of the MRS-GARCH model are obtained by maximizing numerically the log-likelihood function. The log-likelihood function is constructed recursively similar to that in GARCH models.

3.9.1. K-STEP AHEAD FORECASTING

This study forecasts exchange rate at k-step-ahead with MRS-GARCH models. Denote $\hat{r}_{t,t+k}$ is k-step-ahead forecasting logarithm return of exchange rate at time t depend on F_{t-1}. This study computes $\hat{r}_{t,t+k}$ as:

$$\hat{r}_{t,t+k} = E_{t-1}[r_{t+k}] = \sum_{i=1}^{2} \Pr(S_{t+k} = i|F_{t-1}).\hat{r}_{t,t+k}.S_{t+k=i}, \tag{3.73}$$

where $\hat{r}_{t,t+k}.S_{t+k=i} = E_{t-1}(r_{t+k}|S_{t+k} = i) = E_{t-1}(\sigma + \varepsilon_{t+k}|S_{t+k} = i) = E_{t-1}(\sigma|S_{t+k} = i) +$

$$E_{t-1}(\varepsilon_{t+k}|S_{t+k} = i) = \sum_{j=1}^{2} \begin{matrix} \Pr(S_{t+k-1} = j|S_{t+k} = i, F_{t-1}).\sigma_{S_{t+k=i}} \\ = \sum_{j=1}^{2} \bar{p}_{ji,t-1}.\sigma_{S_{t+k=i}} \end{matrix}.$$

For forecasting exchange rate return one, five, ten-step-ahead, the study used equation (3.66) and log-return of exchange rate is

$$P_{t+1} = P_t.\exp\left[\frac{\sum_{i=1}^{2} \Pr(S_{t+1} = i|F_{t-1}).\sum_{j=1}^{2} \bar{p}_{ji,t-1}.\sigma_{S_{t+k=i}}}{100}\right]. \tag{3.74}$$

3.10. FORECAST EVALUATION

This section discusses the evaluation based forecasting which depends on the three error metrices. According to Marcucci (2005), the commonly used error metrices for forecast evaluation are Mean Squared Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE). The three error metrices are used in the current study in order to select the appropriate models for BRICS exchange rates using MS-GARCH model, where n represents the number of total observations both in-sample and out-of-sample and n_1 is the first out-of-sample forecast observation. The observation that the model is estimated for start from 1 to $n_t - 1$ observations and n_1 to n are used for the out-of-sample forecasting. Y_t and $\widehat{Y_t}$ denote the actual and the estimated series. The three error metrices are defined below (Brooks, 2008). The MSE is computed using the following:

$$MSE = \frac{1}{n-(n_1-1)} \sum_{i=n_1}^{n} (Y_t - \widehat{Y}_t)^2. \tag{3.75}$$

MSE is commonly used measure for forecasting errors and its variance. MSE is related to standard deviation of forecast errors and is therefore an appropriate error for mathematical operations (Montgomery *et al*., 1990). The MSE measures the variability in forecast errors. Obviously, the variability in forecast errors must be small. MAE and MSE are all scale dependent measures of forecast accuracy, that is, their values are expressed in terms of the original units of measurement. The RMSE is computed using the following:

$$RMSE = \sqrt{\frac{1}{n-(n_1-1)} \sum_{i=n_1}^{n} (Y_t - \widehat{Y}_t)^2}. \tag{3.76}$$

RMSE and MAE measure how close the forecast is from the actual data, thus the forecast which has the smallest RMSE and MAE is probably the most accurate forecast (Brooks, 2008). RMSE is calculated based on the quadratic loss function (Brooks, 2008). It is advantageous in the sense that it is most sensitive to large errors of the four criterions because of the square of the errors. The most advantageous thing about it is that large estimate errors could lead to serious problems. RMSE can lead to a tailback if the larger errors cannot lead to serious problems. The MAE is computed using the following:

$$MAE = \frac{1}{n-(n_1-1)} \sum_{i=n_1}^{n} \left| Y_t - \widehat{Y}_t \right|. \tag{3.77}$$

MAE measures the average absolute forecast error (Brooks, 2008). The advantage of these loss functions is that it penalizes large errors less than the RMSE because of non-squaring errors. RMSE and MAE are very simple loss function, and by that they are inconsistent to scalar transformation and symmetric which implies that they are not realistic in some cases (Brooks, 2008).

MSE is commonly used measure for forecasting errors and its variance. MSE is related to standard deviation of forecast errors and is therefore an appropriate error for mathematical operations (Montgomery *et al*., 1990). The MSE measures the variability in forecast errors. Obviously, the variability in forecast errors must be small. MAE and MSE are all scale

dependent measures of forecast accuracy, that is, their values are expressed in terms of the original units of measurement. The RMSE is computed using the following:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n-(n_1-1)}\sum_{i=n_1}^{n}(Y_t - \widehat{Y}_t)^2}. \quad (3.78)$$

RMSE and MAE measure how close the forecast is from the actual data, thus the forecast which has the smallest RMSE and MAE is probably the most accurate forecast (Brooks, 2008). The RMSE is calculated based on the quadratic loss function (Brooks, 2008). It is advantageous in the sense that it is most sensitive to large errors of the four criterions because of the square of the errors. The most advantageous thing about it is that large estimate errors could lead to serious problems. RMSE can lead to a tailback if the larger errors cannot lead to serious problems. The MAE is computed using the following:

$$\mathrm{MAE} = \frac{1}{n-(n_1-1)}\sum_{i=n_1}^{n}\left|Y_t - \widehat{Y}_t\right|. \quad (3.79)$$

MAE measures the average absolute forecast error (Brooks, 2008). The advantage of these loss functions is that it penalizes large errors less than the RMSE because of non-squaring errors. RMSE and MAE are very simple loss function, and by that they are inconsistent to scalar transformation and symmetric which implies that they are not realistic in some cases (Brooks, 2008).

3.11. MODEL FORECASTING ACCURACY

This section discusses the test used to determine the accuracy of the forecasts. Diebold-Mariano's (1995) (DM) test is used to check and compare the forecasting ability of two or more competing models. According to the research, the DM test is the only test used to determine the accuracy of the forecast. Diebold and Mariano (1995) are of the view that comparing forecasting ability of competing models built for a given data is very important. Although the lower the forecast measure the better the forecasting performance, it does not necessarily imply superior forecasting ability since the difference between error measures from competing models may not be significantly different from zero (Diebold & Mariano, 1995). One simple approach to compare the forecast accuracy of competing models is the DM test, which evaluates the fairness of forecast accuracy of pairs of competing models built for a particular data. If the pair of forecast errors, $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ have equal forecasting accuracy, then the expected square differences (mean of the square differences),

$$\bar{d} = E(d_t) = E\left(\varepsilon_{1,t}^2 - \varepsilon_{2,t}^2\right), \quad (3.83)$$

should be approximately zero, that is $\bar{d} = \frac{1}{n}\sum_{t=1}^{n} = 0$. The DM test, thus, formulates the following hypotheses to be tested:

$H_0: \bar{d} = 0$ (Forecasting accuracy of pair of models is equal)

vs.

$H_1: \bar{d} \neq 0$ (Forecasting accuracy of pair of models is not equal)

With h-step ahead, forecasts are generated from a model, the Diebold-Mariano test statistic, S, is given by:

$$S = \frac{\bar{d}}{\sqrt{\mathrm{var}(\bar{d})}} \sim N(0,1) \quad (3.84)$$

where $\mathrm{var}(\bar{d}) = \frac{1}{n}\left[\gamma_0 + 2\sum_{k-1}^{h-1}\gamma_k\right]$ is the asymptotic variance of $\bar{d}$. The term, γ_0, is the variance of the loss differential, d_t, while γ_k, the k-th auto-covariance of d_t at time t, is given by:

$$\hat{\gamma}_k = \frac{1}{n}\sum_{t=k+1}^{n}\left(d_t - \bar{d}\right)\left(d_{t-k} - \bar{d}\right) \quad (3.85)$$

The null hypothesis of Model 1 (MS-GARCH-ML) having equal forecasting accuracy as Model 2 (MS-GARCH-MCMC) is rejected if the test statistic, S, is greater than the critical value or the p-value of S is less than the 1%, 5% or 10% level. In case the null hypothesis is rejected, a higher positive t-statistic for Model 1 than Model 2 suggests that Model 2 is more accurate than Model 1.

3.12. CHAPTER SUMMARY

This chapter presented the methodology applied to the exchange rates of BRICS countries. It presented the process that the methodology followed in the study process. The chapter presented the ethical considerations the study abided by. The research process picture that the study followed was also painted for ease of reference. This provided the steps that the study followed in its entirety. This chapter has presented an overview of various statistical tests and forecasting methods used in analysing the exchange rates of BRICS countries. The MRS-

GARCH models were discussed followed by various estimation methods, then forecasting evaluation and lastly model forecasting accuracy procedures were illustrated. Criteria used to identify appropriate forecasting models for each of the closing stock prices were also discussed. The next chapter presents empirical analysis and results.

CHAPTER 4
EMPIRICAL ANALYSIS AND RESULTS

4.1.INTRODUCTION

The chapter presents the empirical analysis of data and discusses this in light of the study aim and objectives as outlined in Chapter 3. The results are summarised using tables and figures. The graphical presentations assist in displaying the performance of the data.

The rest of the chapter is organised as follows: Section 4.2 presents the preliminary data analysis results. Section 4.3 presents the model estimation results using maximum likelihood and Section 4.4 presents the in-sampling evaluation using maximum likelihood. Section 4.5 presents model estimation results using Bayesian MCMC and Section 4.6 presents the in-sampling evaluation using Bayesian MCMC. Section 4.7 presents the model comparison results of MRS-GARCH models. Section 4.8 presents predictive accuracy of the models (in-sampling). Section 4.9 presents the out-of-sampling evaluation using maximum likelihood. Section 4.10 presents the out-of-sampling volatility forecasts using Bayesian MCMC. Section 4.11 presents the comparison results of MRS-GARCH models. Section 4.12 presents the predictive accuracy of models (out-of-sampling). Lastly, the summary of the chapter was presented.

4.2.PRELIMINARY DATA ANALYSES RESULTS

This section presents the preliminary analysis of data with the aim to identify the broad behaviour of the measured exchange rates of BRICS countries used in the study. The study conducted descriptive statistics, nonlinearity and nonstationary tests to assess the characteristics of exchange rates of BRICS countries. The next Section 4.2.1 presents the descriptive analysis of the exchange rates of BRICS countries.

4.2.1. DESCRIPTIVE STATISTICS

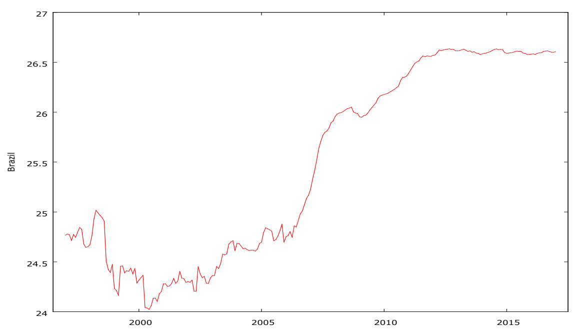
The descriptive statistics used in the study are Skewness, Kurtosis and Jarque-Bera test to assess the characteristics of the exchange rates of BRICS countries. Prior to executing the primary analysis of data, the study also addresses a few very important issues such as the normality of the actual data as suggested by St-Hilaire (2015) as well as Trencaa (2015). Normality test gives an indication of which appropriate methods (linear or nonlinear) to be used. Furthermore, the test gives an indication of whether it is appropriate to use normal distribution to estimate MRS-GARCH model or not. Table 4.1 gives summary results for descriptive statistics of the data.

Table 4.1: Descriptive Statistics of Exchange Rates

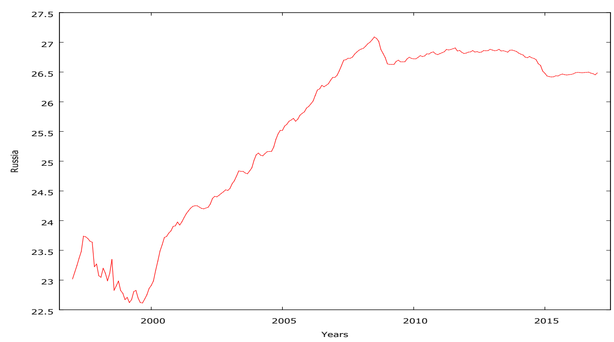
	Brazil	China	India	Russia	South Africa
Mean	25.45582	25.83268	25.53752	25.52313	23.48393
Std. Dev.	0.951398	0.512947	0.936779	1.450933	0.929245
Skewness	0.032629	0.222667	-0.599311	-0.692757	-0.439318
Kurtosis	1.288816	1.581710	1.769979	1.957169	1.675854
Jarque-Bera	29.44629	22.19075	29.61936	30.19678	25.35885
Probability	0.000000	0.000015	0.000000	0.000000	0.000003

Table 4.1 presents the descriptive statistics of BRICS exchange rate. The exchange rates of India, Russia, and RSA were negatively skewed. It can also be noted that the series have excess kurtosis of less than 3. These values indicate that the exchange rates of BRICS countries are not normally distributed. The assertion was also confirmed by the Jarque-Bera (JB) test. This indicates that exchange rates of BRICS can only be estimated and forecasted using nonlinear models. These results are consistent with those obtained by Sopipan *et al*. (2012) as well as Naeini & Fatahi (2012).

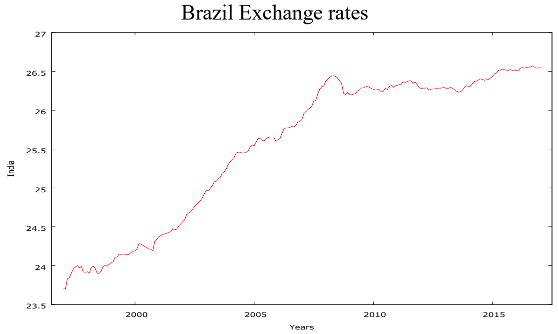
Figure 4.1 presents the results of the time series plots of BRICS exchange rates.



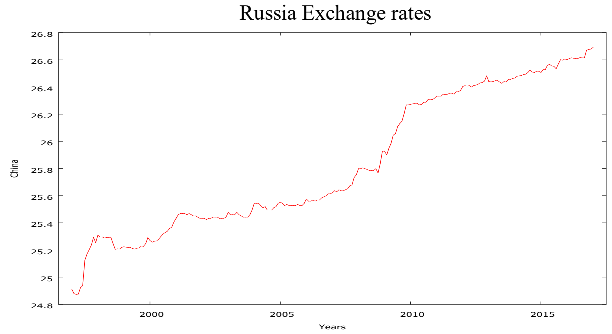
Brazil Exchange rates



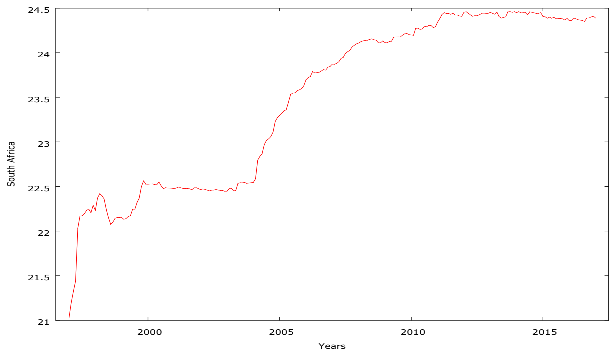
Russia Exchange rates



India Exchange rates



China Exchange rates



South Africa Exchange rates

Figure 4.1: **Graphical representation of the five exchange rates**

A graphical examination of the plot reveals different kinds of patterns for BRICS countries. Figure 4.1 indicates that there was a high and consistent increase in the exchange rates of the five countries after the period 2007-2008 global financial crises. India and China appear to be the two countries with strong currencies as evidenced in an upward sloping trend after recovery from the crises. On the other hand, Brazil, Russia, and South Africa experienced the increasing pattern in the exchange rates until the period 2015 to 2018 which is evidenced by the data showing a plateau or constant slow increase. In general, the exchange rates of the five countries is nonstationary by eye inspection. Further empirical analysis was conducted by employing the nonlinear tests to confirm the graphical presentation and the results are presented in Section 4.2.2.

4.2.2. NONLINEAR MODEL SPECIFICATION/DIAGNOSTICS: NONLINEARITY TESTS

This section presents the nonlinear test to confirm the use of MRS-GARCH model. The study used several nonlinearity tests to detect the presence of nonlinearity. According to Ismail and Isa (2006), it is unwise to rely solely on a single test, since nonlinear behaviour can be detected differently. Based on the advice by Ismail and Isa (2006), the study computed several nonlinearity tests to detect nonlinear behaviour in the exchange rates series. The results of the test are summarised in Table 4.2.

Table 4.2: Test for Nonlinearity

Test/Variable	Brazil	Russia	India	China	South Africa
RESET Test for Specification	3.328 [0.0377]	2.694 [0.0696]	3.805 [0.0236]	7.3397 [0.0008]	25.9658 [0.0000]
McLeod and Li Test	203.629 [0.0000]	219.993 [0.0000]	219.760 [0.0000]	161.036 [0.0000]	957.84 [0.0000]
Keenan's test	2.573 [0.0981]	22.347 [0.0000]	17.436 [0.0000]	17.436 [0.0000]	18.743 [0.0000]
Tsay test	9.078 [0.0028]	15.92 [0.0000]	11.650 [0.0008]	11.650 [0.0008]	3.625 [0.0581]

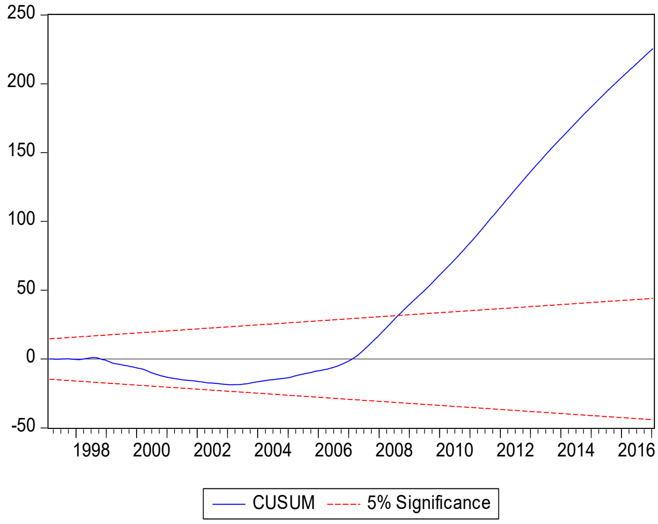
Note: Parentheses indicate p-values

The results from the five tests revealed that the linear model is inadequate in capturing the stochastic properties of the BRICS exchange rates. The corresponding observed probabilities (the p-values are less than 1%, 5% and 10%) are also significant according to the results. This implies that the exchange rates of BRICS are nonlinear in nature and could be better characterised by nonlinear models. Similar results were reported in the studies of Ashley and Patterson (2006), Ismail and Isa (2006), Chen (2011), and Chen (2014).

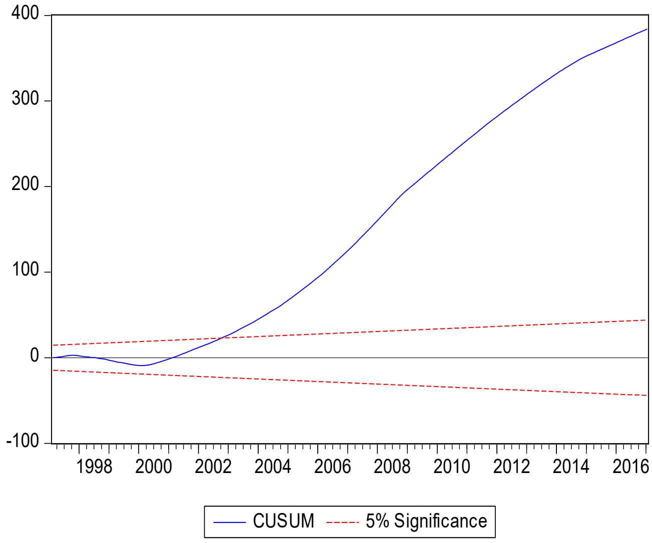
The next subsection 4.2.2.1 presents the CUSUM test results.

4.2.2.1. CUSUM TESTS RESULTS

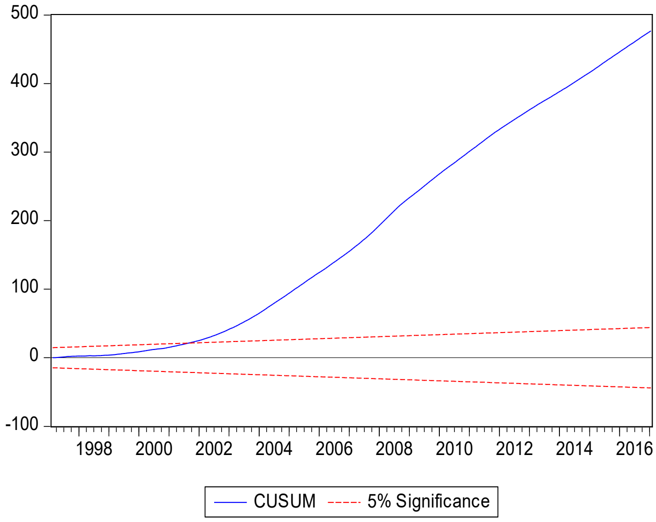
The CUSUM test was conducted to demonstrate the existence of parameters instability. This test assists the study to find the best suitable model for exchange rates of BRICS. Figure 4.2 below presents the graphical presentation of the CUSUM test results.



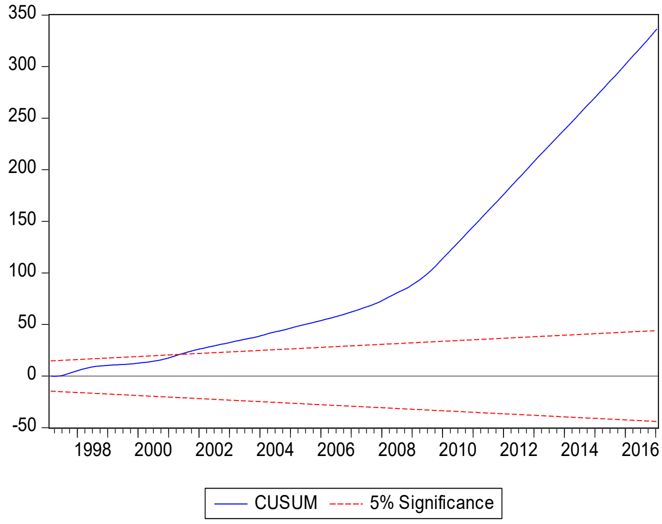
Brazil



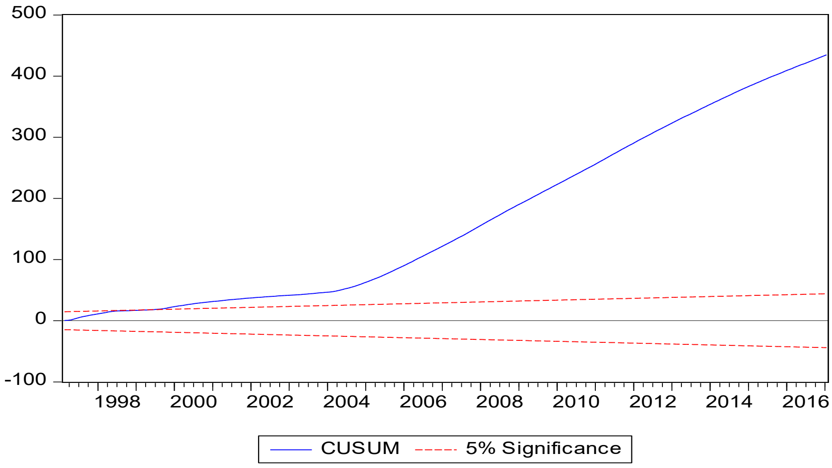
Russia



India



China



South Africa

Figure 4.2: CUSUM test for stability

The results of the CUSUM test presented in Figure 4.2 reveal that the exchange rates of the BRICS countries are outside the bends. The study concluded that there is a nonlinear behaviour in exchange rates of the BRICS countries. These results agree with studies by Kirch and Kamgaing (2011), Farhani (2012), as well as Ardia et al. (2016). The next subsection 4.2.2.2 presents the results of the BDS test for iid.

4.2.2.2. BROCK- DECHERT-SCHEINKMAN (BDS) TEST RESULTS

The study conducted nonlinearity test on exchange rates using a BDS test. The test was conducted to assess if data is iid or not. The results of the BDS test are summarised below in Table 4.3.

Table 4.3: BDS tests results

Variable	Dimension	BDS Statistic	Std. Error	z-Statistic	Normal Prob.
Brazil	2	0.193023	0.002849	67.74017	0.0000
	3	0.328738	0.004516	72.78659	0.0000
	4	0.422557	0.005361	78.82555	0.0000
	5	0.486900	0.005568	87.45312	0.0000
	6	0.530500	0.005349	99.17237	0.0000
China	2	0.202102	0.002842	71.10188	0.0000
	3	0.342895	0.004484	76.46538	0.0000
	4	0.440934	0.005298	83.23104	0.0000
	5	0.509493	0.005477	93.03174	0.0000
	6	0.557704	0.005237	106.4861	0.0000
India	2	0.206883	0.003836	53.92533	0.0000
	3	0.352048	0.006086	57.84464	0.0000
	4	0.453693	0.007232	62.73832	0.0000
	5	0.524718	0.007520	69.78090	0.0000
	6	0.574294	0.007234	79.39128	0.0000
Russia	2	0.206638	0.003897	53.02568	0.0000
	3	0.351131	0.006161	56.99181	0.0000
	4	0.451683	0.007296	61.90860	0.0000
	5	0.521265	0.007561	68.94071	0.0000
	6	0.568886	0.007249	78.47444	0.0000
South Africa	2	0.203467	0.003035	67.03164	0.0000
	3	0.346174	0.004819	71.84096	0.0000
	4	0.446159	0.005728	77.88567	0.0000
	5	0.516095	0.005959	86.60717	0.0000
	6	0.564821	0.005735	98.49233	0.0000

Following the work of Gooijer and Yuan (2016), a dimension m from 2 to 6 and distance (ϵ) of 0.7 was selected to complete the BDS test. The quest is to test the null hypothesis that the exchange rates are iid against the alternative hypothesis of nonlinearly. The BDS test indicates that exchange rates with p-values less than 5% are significant for all countries. Thus, the study strongly rejects the null hypothesis of iid. The results clearly suggest that BRICS exchange rates are nonlinearly dependent. These results agree with those reported by Panagiotidis (2005), Winker and Jeleskovic (2007) as well as Cheteni (2017). The next subsection 4.2.2.3 presents the LM test results.

4.2.2.3. LAGRANGE MULTIPLIER (LM) TEST RESULTS

In this section, the study assessed the presence of nonlinear using LM test. The LM test is utilised to detect the presence of ARCH effects in exchange rates of BRICS countries. The results of the test are summarized in Table 4.4 below.

Table 4.4: LM Test results

Variable	**Engle Test**	**F-statistic**	**Prob.**
Brazil	Using up to lag 1	3936.794	0.0000
	Using up to lag 2	1947.370	0.0000
	Using up to lag 3	1283.462	0.0000
	Using up to lag 4	962.6927	0.0000
Russia	Using up to lag 1	10451.90	0.0000
	Using up to lag 2	5173.446	0.0000
	Using up to lag 3	3417.875	0.0000
	Using up to lag 4	2845.611	0.0000
India	Using up to lag 1	37712.18	0.0000
	Using up to lag 2	18585.87	0.0000
	Using up to lag 3	15378.97	0.0000
	Using up to lag 4	11137.86	0.0000
China	Using up to lag 1	12515.39	0.0000
	Using up to lag 2	6482.443	0.0000
	Using up to lag 3	4412.339	0.0000
	Using up to lag 4	3115.155	0.0000
South Africa	Using up to lag 1	5153.437	0.0000
	Using up to lag 2	2176.205	0.0000
	Using up to lag 3	1148.975	0.0000
	Using up to lag 4	702.1693	0.0000

The results in Table 4.4 above show that there is evidence of ARCH effects in the exchange rates of BRICS countries as p-values are less than 5% significant level. Therefore, the study strongly rejects the null hypothesis that the exchange rates are *iid* and then the study concludes that the exchange rates are nonlinearly dependent. Similar results were reported in the studies by Zhang (2002) and Cheteni (2017). The subsequent section 4.2.3 of the study presents the results of nonlinear unit root test.

4.2.3. NONLINEAR UNIT ROOT TEST RESULT

This section presents nonlinear unit root test results. The tests are used to assess the presence of the nonlinearity in the exchange rates of BRICS countries. The test is used to confirm if it is appropriate to apply MRS-GARCH model in the study. For the current study, nonlinear unit root was tested using the Bierens Nonlinear Augmented Dickey-Fuller (B-NADF) and KSS Nonlinear unit root tests. The results of the B-NADF test are summarized in Table 4.5.

Table 4.5: Bierens Nonlinear Unit Root Test Results

Variables	M	t-Test Statistic	t-Test 5% Critical Value	Am Statistic	Am 5% Critical Value	F-Test Statistic	F-Test 5% Critical Value
Brazil	2	-1.608	-3.97	-6.117	-27.2	1.744	1.08
	5	-4.496	-5.16	-41.408	-48.7	3.757	1.83
	10	-4.605	-6.67	-43.744	-80.3	2.727	2.15
	15	-6.452	-7.89	-90.703	-112.4	3.783	2.36
	20	-7.609	-9.00	-131.096	-145.7	4.214	2.52
Russia	2	-6.14	-3.97	0.66	-27.2	3.1916	1.08
	5	-4.4372	-5.16	-43.3766	-48.7	4.0562	1.83
	10	-6.8818	-6.67	-89.6121	-80.3	4.8848	2.15
	15	-7.3307	-7.89	-103.0385	-112.4	3.9183	2.36
	20	-8.6817	-9.00	-137.5241	-145.7	4.0834	2.52
India	2	-2.2387	-3.97	-12.6037	-27.2	2.2546	1.08
	5	-4.1939	-5.16	-33.9801	-48.7	3.6163	1.83
	10	-4.4909	-6.67	-41.3511	-80.3	2.5293	2.15
	15	-7.2159	-7.89	-97.3750	-112.4	4.1744	2.36
	20	-7.4247	-9.00	-104.5964	-145.7	3.3754	2.52
China	2	-2.6593	-3.97	-12.7934	-27.2	2.9569	1.08
	5	-5.2442	-5.16	-54.5963	-48.7	4.9939	1.83
	10	-6.2399	-6.67	75.2555	-80.3	3.8577	2.15
	15	-7.4658	-7.89	-104.1044	-112.4	3.7836	2.36
	20	-7.8911	-9.00	-113.3282	-145.7	3.3911	2.52
South Africa	2	-4.2306	-3.97	-30.0178	-27.2	6.3504	1.08
	5	-6.0093	-5.16	-67.1012	-48.7	79.0804	1.83
	10	-7.4250	-6.67	-103.2396	-80.3	5.5080	2.15
	15	-8.4151	-7.89	-130.5011	-112.4	4.5205	2.36
	20	-9.3462	-9.00	-158.2898	-145.7	4.5205	2.52

The results of the B-NADF unit root tests for various Chebyshev polynomial orders were computed. The optimal lag length for each variable was chosen using the AIC, while the test statistics were obtained using Gaussian AR (m) process. The results in Table 4.5 showed that the null hypothesis of nonlinear unit root cannot be rejected at 5% significance level. The t-test, Am-test and F-test statistics are all greater than their corresponding critical values. This indicates that the exchange rates of BRICS countries are nonstationary also confirming the depiction in Figure 4.1. Similar studies conducted by Anoruo & Elike (2008) agrees with the findings. Since various tests confirmed the presence of nonlinearity and nonlinear unit root in BRICS exchange rates, the study proceeded to estimate the MRS-GARCH models using different estimation methods. The Table 4.6 presents the KSS Nonlinear Unit Root Test results.

Table 4. 6: KSS Nonlinear Unit Root Test

Country	Optimal Lag, P	Estimate	Std. Err	t-Value	p-Value	KSS
Brazil	2	-.000000421	7.1592E-9	-2.96	0.0063	-2.7713
Russia	3	-.000000246	2.2856E-8	-2.82	0.0011	-2.6124
India	1	-.000000040	5.6322E-8	-3.56	0.0020	-2.5321
China	2	-.000000368	6.5941E-9	-2.42	0.0021	-3.4324
South Africa	2	-.000000726	3.3276E-9	-1.89	0.0031	-2.7364

Critical values of the KSS-NLADF test with constant and trend at the 10%, 5% and 1% significant levels are -3.13, -3.40 and -3.93, respectively.

As a starting point for the KSS-NADF, various regression analyses were conducted to examine whether the raw (zero-mean), de-meaned (single-mean) or de-trended data be used in conducting the KSS-NADF test. The results, reported in Appendix C and D, show the intercept term, c, and trend variable, t, being significant at the 5% significance level for all five variables, thus suggesting that de-trended variables be used in the KSS-NADF unit root tests.

Upon using the SBC for selection of the optimal lags associated with exchange rate for the five countries, the KSS-NADF unit root test was computed and Table 4.6 gives summary results of this test. At the 0.05 significance level, the null hypothesis of a nonlinear unit root in Brazil, Russia, India and China is rejected since the KSS test statistics are all greater than the critical values of -3.13, -3.40 and -3.93 at 1%, 5% and 10% significance levels, respectively. This means that exchange rates of Brazil, Russia, India and China exhibit asymmetric and nonlinear stationarity. The results, however, indicate that the null hypothesis of nonlinear unit root in South Africa's exchange rate cannot be rejected, indicating the presence of unit root in this respect.

4.3.MODEL ESTIMATION RESULTS USING MAXIMUM LIKELIHOOD

This section presents the results of the MRS-GARCH model using maximum likelihood estimation method. MRS-GARCH models were estimated using three various error distributions (normal, student-t and GED). The subsequent section 4.3.1 presents the results of MRS-GARCH(1,1) where MLE is used to estimate the best fitted models of the exchange rates of BRICS countries.

4.3.1. BRAZIL RESULTS

The parameter estimates and unconditional probability of these models are presented in Table 4.7 below for Brazil exchange rate.

Table 4.7: Maximum likelihood estimates of MRS-GARCH models for Brazil

Brazil						
	Normal		Student t		GED	
Model	Low Volatility	High volatility	Low volatility	High volatility	Low volatility	High volatility
μ			5.2817	2.9103	1.3396	0.9521
P-value			0.0000	0.0000	0.0000	0.0000
α_0		0.2051	0.0000	0.0991	0.0000	0.0947
P-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
α_1	0.0000	0.0415	0.0026	0.0001	0.0010	0.0010
P-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
β_1	0.7948	0.9584	0.9007	0.6969	0.9051	0.8593
P-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
P_{00}	0.8708		0.9864		0.9864	
P-value	0.0000		0.0000		0.0000	
P_{11}	1.0000		0.0188		0.0188	
P-value	0.0000		0.0000		0.0000	
Persistence	0.7948	0.9999	0.9033	0.6970	0.9061	0.8603
π_1 (Low volatility)	0.89		0.58		0.59	
π_2 (High volatility)	0.11		0.42		0.42	

According to the results, all the parameter estimates for Brazil are significant at 1%, 5% and 10% of levels of significant as presented in Table 4.6. The two regimes (from the columns), first with low volatility and second with high volatility can be identified from the estimates. The transition probabilities are all significant, implying that all regimes are particularly persistent in MRS-GARCH models. The results also reported the unconditional probabilities for each MRS-GARCH model. The persistence[2] is computed by adding α_1 and β_1. All models show a strong persistence in volatility since the persistence values ranges from 0.6970 to 0.9999. This implies that the volatility is likely to remain high over several exchange rate return periods once it increases. The unconditional probability π_1 for being in the first regime representing low volatility and it ranges between 58% (Student-t) and 89% (GED) for different MRS-GARCH models. Similarly, the unconditional probability π_2 for being in the second regime representing high volatility and it ranges between 11% (normal) and 42% (student-t and

[2] The study chose the cut-off point as 0.7 indicating strong persistence

GED) for different MRS-GARCH models. The subsection 4.3.2 below presents the results for Russia.

4.3.2. RUSSIA RESULTS

The parameter estimates and unconditional probability of these models are presented in Table 4.8 below for Russia exchange rate.

Table 4.8: Maximum likelihood estimates of MS-GARCH models for Russia

Russia						
	Normal		Student t		GED	
Model	Low Volatility	High volatility	Low volatility	High volatility	Low volatility	High volatility
μ			99.2939	11.8796	3.5139	1.5859
P-value			0.0000	0.0000	0.0000	0.0000
α_0	0.0000	0.0087	0.0000	0.0086		0.0089
P-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
α_1	0.0001	0.0663	0.0001	0.0812	0.0001	0.2650
P-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
β_1	0.9679	0.9322	0.9667	0.9174	0.9668	0.7347
P-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
P_{00}	0.9276		0.9397		0.7848	
P-value	0.0000		0.0000		0.0000	
P_{11}	0.1393		0.0970		0.1550	
P-value	0.0000		0.0000		0.0000	
Persistence	0.9680	0.9985	0.9668	0.9986	0.9669	0.9997
π_1 (Low volatility)	0.66		0.62		0.42	
π_2 (High volatility)	0.34		0.38		0.58	

The results in Table 4.8 above shows that all the estimated parameters for Russia are significant at all levels of significant. The estimates confirmed the existence of two regimes; the first regime is regarded as low volatility and second as high volatility. The transition probabilities are found significant in all cases showing the persistence of both regimes. The results revealed that there is a strong persistence in volatility ranging from 0.9668 to 0.9997. This implies that the volatility is likely to remain high over several exchange rate return periods once it increases. Table 4.8 also reported unconditional probabilities for each model. The unconditional probability of being in the first regime (low volatility/ π_1) ranges between 42% (GED) to 66% (normal), whereas the unconditional probability of being in the second regime (high volatility/ π_2) ranges between 34% (normal) to 58% (GED). The subsequent subsection presents the results for India.

4.3.3. INDIA

The parameter estimates and unconditional probability of these models are presented in Table 4.9 below for India exchange rate.

Table 4.9: Maximum likelihood estimates of MRS-GARCH models

India						
	Normal		Student t		GED	
Model	Low Volatility	High volatility	Low volatility	High volatility	Low volatility	High volatility
μ	-	-	89.9671	2.7826	1.4146	1.2667
P-value			0.0000	0.0000	0.0000	0.0000
α_0	0.0006	0.0001	0.0002	0.0001		
P-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
α_1	0.0000	0.1179	0.0000	0.0315	0.0000	0.1005
P-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
β_1	0.5342	0.8820	0.8096	0.9684	0.9952	0.8994
P-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
P_{00}	0.9872		0.9771		0.9874	
P-value	0.0000		0.0000		0.0000	
P_{11}	0.0085		0.0079		0.0070	
P-value	0.0000		0.0000		0.0000	
Persistence	0.5342	0.9999	0.8096	0.9684	0.9952	0.9999
π_1 (Low volatility)	0.40		0.26		0.36	
π_2 (High volatility)	0.60		0.74		0.64	

The estimated parameters and their respective probability values for India are presented in Table 4.9 above. All the parameters estimate for India in MRS-GARCH are significant at 1%, 5% and 10% of levels of significant. These estimates can be identified from both the first regime (low volatility) and the second regime (high volatility). Reported results show that for both regimes in all three error distributions, the probabilities are found to be significant for all the distributions. MRS-GARCH models with student-t and GED show a strong persistence in volatility since the persistence values ranges from 0.5342 to 0.9999. This implies that the volatility is likely to remain high over several exchange rate return periods once it increases. Furthermore, the study reported unconditional probabilities for each model which ranges between 26% (student t) to 40% (normal) for being in the first regime, whereas the unconditional probability of being in the second regime ranges between 60% (GED) to 74% (normal). The subsequent subsection 4.3.4 presents the results for China.

4.3.4. CHINA

The results presented in Table 4.10 show the parameter estimates and unconditional probability of these models for China exchange rate.

Table 4.10: Maximum likelihood estimates of MRS-GARCH models

China						
	Normal		Student t		GED	
Model	Low Volatility	High volatility	Low volatility	High volatility	Low volatility	High volatility
μ			99.6867	67.0870	1.7261	0.7000
P-value			0.0000	0.0000	0.0000	0.0000
α_0	0.0001	0.0000	0.0001	0.0000	0.0001	0.8438
P-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
α_1	0.0003	0.6593	0.0002	0.5583	0.0001	0.2548
P-value	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000
β_1	0.0008	0.3406	0.0003	0.4413	0.0004	0.7441
P-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
P_{00}	0.8217		0.8263		0.7916	
P-value	0.0000		0.0000		0.0000	
P_{11}	0.2766		0.2717		0.3652	
P-value	0.0000		0.0000		0.0000	
Persistence	0.0011	0.9999	0.0005	0.9996	0.0005	0.9989
π_1 (Low volatility)	0.61		0.61		0.64	
π_2 (High volatility)	0.39		0.39		0.36	

The parameter estimates for MS-GARCH models for China are presented in Tables 4.10 above. According to the results, all the parameter estimates are significant at three levels of significance. The estimates confirmed the existence of two regimes. The transition probabilities are all significant, showing that all regimes are particularly persistent in MRS-GARCH models. The results of persistence in volatility for low regime shows a weak persistence that ranges from 0.0005 to 0.0011. This implies that the volatility is likely to remain low for shorter exchange rate periods. On the other hand, the high regime shows a strong persistence ranging from 0.9996 to 0.9999. This implies that the volatility is likely to stay on the high for several exchange rate return periods once it increases. The unconditional probability π_1 of being in the first regime ranges between 61% (normal and student-t) and 64% (GED) for different MRS-GARCH models. The unconditional probability π_2 of being in the second regime ranges between 36% (GED) and 39 % (normal and student-t) for different MS-GARCH models. The subsequent subsection 4.3.5 presents the results for South Africa.

4.3.5. SOUTH AFRICA

The parameter estimates and unconditional probability of these models are presented in Table 4.11 below for South African exchange rate.

Table 4.11: Maximum likelihood estimates of MRS-GARCH models

South Africa						
	Normal		Student t		GED	
Model	Low Volatility	High volatility	Low volatility	High volatility	Low volatility	High volatility
μ			3.1816	2.6406	4.0017	0.7000
P-value			0.0000	0.0000	0.0000	0.0000
α_0	0.0001	0.1381	0.0000	0.0048		0.3051
P-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
α_1	0.0012	0.7569	0.0001	0.1312	0.0001	0.2134
P-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
β_1	0.2511	0.2430	0.8321	0.8681	0.1356	0.7861
P-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
P_{00}	0.6884		0.7617		0.5507	
P-value	0.0000		0.0000		0.0000	
P_{11}	0.6434		0.1064		0.2462	
P-value	0.0000		0.0000		0.0000	
Persistence	0.2523	0.9999	0.8322	0.9993	0.1357	0.9995
π_1 (Low volatility)	0.67		0.31		0.35	
π_2 (High volatility)	0.33		0.69		0.65	

The estimated results and summary statistics of MRS-GARCH models for South Africa are presented in Table 4.11 above. All parameter estimates are found to be significant at 1%, 5% and 10% of significant levels. The estimates confirmed the existence of two regimes. The transition probabilities are found significant in all cases showing the persistence of both regimes. The results of persistence for student-t distribution is high (0.8322 and 0.9993) for both regimes. It follows that student-t has high persistence as compared to the other two distributions. The results of persistence in volatility for normal and GED distribution in low regime shows a weak persistence that ranges from 0.1357 to 0.2523 and for high regime it ranges from 0.9993 to 0.9999. The unconditional probability of being in the first regime ranges between 42% (GED) to 66% (normal), whereas the unconditional probability of being in the second regime ranges between 34% (normal) to 58% (GED). The next Section 4.4 presents the in-sampling evaluation using maximum likelihood estimation method.

4.4. IN-SAMPLING EVALUATION USING MAXIMUM LIKELIHOOD

The section presents the in-sampling results of MRS-GARCH using MLE method. The study used various model section criteria and error metrics to compare in-sampling evaluation between MRS-GARCH-MLE using normal, student-t and GED innovation distributions. The results are presented in Table 4.12 below.

Table 4.12: In-sample Results using MRS-GARCH models for BRICS countries

Model	Maximum Likelihood Estimation Method							
	Brazil							
	No. of Par.	BIC	AIC	Log(L)	MSE	RMSE	MAE	Best Model
MRS-GARCH-N	5	-865.614	-893.459	454.73	0.80	0.90	0.62	MRS-GARCH-t
MRS-GARCH-t	6	-872.625	-907.431	463.72	0.66	0.81	0.55	
MRS-GARCH-GED	6	-868.152	-902.959	461.479	0.77	0.88	0.59	
Russia								
MS-GARCH-N	5	-753.534	-781.38	398.69	1.03	1.02	0.76	MRS-GARCH-N
MS-GARCH-t	6	-741.648	-776.454	398.23	1.09	1.04	0.77	
MS-GARCH-GED	6	-741.547	-776.353	398.12	1.15	1.07	0.81	
India								
MS-GARCH-N	5	-993.667	-1021.513	518.756	0.93	0.96	0.72	MRS-GARCH-t
MS-GARCH-t	6	-995.875	-1030.682	525.341	0.73	0.85	0.64	
MS-GARCH-GED	6	-991.770	-1026.577	523.289	0.91	0.95	0.71	
China								
MS-GARCH-N	5	-1255.931	-1290.737	655.474	0.97	0.98	068	MRS-GARCH-GED
MS-GARCH-t	6	-1256.427	-1291.233	655.369	0.96	0.98	0.68	
MS-GARCH-GED	6	-1267.103	-1294.948	655.617	0.87	0.93	0.65	
South Africa								
MS-GARCH-N	5	-1016.372	-1044.217	530.108	1.05	1.02	0.66	MRS-GARCH-t
MS-GARCH-t	6	-1020.007	-1054.813	537.406	0.78	0.89	0.55	
MS-GARCH-GED	6	-1018.835	-1053.641	536.821	0.87	0.93	0.59	

In Table 4.12 above, the results of model selection criterions and error metrics for all volatility models are presented. According to the BIC, AIC, LOGL, MSE, MAE and RMSE, the MRS-GARCH model with the t-distribution performed better in modelling the exchange rates of Brazil, India, and South Africa. However, Russia and China were modelled using different distributions. In the case of Russia, model selection criterions and error metrics chose the MRS-GARCH with normal distribution while China's exchange rate was modelled with MRS-GARCH using GED. The next Section 4.5 presents the model estimation results using Bayesian MCMC.

4.5. MODEL ESTIMATION RESULTS USING BAYESIAN MCMC

This section provides the results of the MRS-GARCH using Bayesian MCMC estimation method. The parameter estimates for MRS-GARCH(1,1) models are presented in Tables 4.13 to 4.17.

4.5.1. BRAZIL RESULTS

This section presents the summarized results of MRS-GARCH models for Brazil exchange rate. The study used various distributions under MRS-GARCH model to estimate the best model.

Table 4.13: Bayesian MCMC estimates of MRS-GARCH models for Brazil

Brazil						
	Normal		Student t		GED	
Model	Low Volatility	High volatility	Low volatility	High volatility	Low volatility	High volatility
μ			6.9680	3.1274	1.1603	1.2283
Std.err.			0.2958	0.2488	0.0439	0.0521
α_0	0.0000	0.0122	0.0057	0.0135		0.3255
Std.err.	0.0000	0.0135	0.0004	0.0035		0.0350
α_1	0.0060	0.0417	0.0011	0.0518	0.0041	0.2209
Std.err.	0.0005	0.0120	0.0004	0.0035	0.0002	0.0219
β_1	0.9461	0.5148	0.3975	0.8382	0.5885	0.2384
Std.err.	0.0021	0.013	0.0335	0.0331	0.0439	0.0332
P_{00}	0.9751		0.9485		0.9864	
Std.err.	0.0013		0.0032		0.0011	
P_{11}	0.0664		0.0105		0.0347	
Std.err.	0.0037		0.0031		0.0028	
Persistence	0.9521	0.5565	0.3986	0.8900	0.5926	0.4593
π_1 (Low volatility)	0.84		0.52		0.66	
π_2 (High volatility)	0.16		0.48		0.34	

The results in Table 4.13 show that the parameters estimated for Brazil are significant at 1%, 5% and 10% level of significant. The first regime of the two existing regimes is characterized by low volatility, whereas the second one reveals high volatility. The transition probabilities are all significant, showing that all regimes are particularly persistent in MRS-GARCH models. The unconditional probability of the first regime ranges between 52 % (normal) and 84% (student-t) for different MRS-GARCH models. The unconditional probability of the second regime ranges between 16% (norm) and 48% (student-t) for different MRS-GARCH models. Section 4.5.2 presents the Bayesian estimates of MRS-GARCH models for Russia.

4.5.2. RUSSIA RESULTS

Section 4.5.2 presents the results of Russian exchange rate. The parameter estimates for MRS-GARCH models are presented in Tables 4.14 below.

Table 4.14: Bayesian MCMC Model estimates of MRS-GARCH models for Russia

Russia						
	Normal		Student t		GED	
Model	Low Volatility	High volatility	Low volatility	High volatility	Low volatility	High volatility
μ	-	-	19.8070	88.2912	2.6101	1.7385
Std.err.	-	-	4.1827	4.1781	0.1302	0.0427
α_0	0.0001	0.0807	0.0001	0.1946	-	0.0282
Std.err.	0.0000	0.0017	0.0000	0.0054	-	0.0018
α_1	0.0002	0.1187	-	0.1374	0.0003	0.1491
Std.err.	0.0000	0.0334	-	0.0058	0.0000	0.0041
β_1	0.7758	0.8765	0.6860	0.8564	0.8461	0.8308
Std.err.	0.0062	0.0044	0.0115	0.0078	0.0062	0.0053
$\mathbf{P_{00}}$	0.8559		0.8639		0.9594	
Std.err.	0.0106		0.0034		0.0022	
$\mathbf{P_{11}}$	0.2052		0.0947		0.0387	
Std.err.	0.0085		0.0037		0.0014	
Persistence	0.7760	0.9952	0.6860	0.9938	0.8464	0.9799
π_1 (Low volatility)	0.73		0.48		0.29	
π_2 (High volatility)	0.27		0.52		0.71	

The above Table 4.14 presents the parameter estimates for MRS-GARCH models. All the parameter estimates for Russian are significant at 1%, 5% and 10% level of significant. The estimates confirmed the existence of two regimes: the first regime is regarded as low volatility, whereas the second one reveals high volatility. The transition probabilities are all significant at 1%, 5% and 10% level of significant, showing that all regimes are persistent in MRS-GARCH models. All models show a strong persistence in volatility since the persistence values ranges from 0.6860 to 0.9952. This implies that the volatility is likely to remain high over several exchange rate return periods once it increases. The unconditional probability of the first regime ranges between 29% (GED) and 73% (norm) for different MS-GARCH models. The unconditional probability of the second regime ranges between 27% (norm) and 71% (GED) for different MRS-GARCH models. The next section 4.5.3 presents the Bayesian estimates of MRS-GARCH models for India.

4.5.3. INDIA RESULTS

Section 4.5.3 presents the Indian exchange rate results. The estimation results and summary statistics of MRS-GARCH models are presented in Tables 4.15 below.

Table 4.15: Bayesian MCMC Model estimates of MRS-GARCH models for India

India						
	Normal		Student t		GED	
Model	Low Volatility	High volatility	Low volatility	High volatility	Low volatility	High volatility
μ			62.6217	29.9564	7.4348	3.1757
Std.err.			3.4992	5.0842	1.0157	0.7916
α_0	0.0002	0.2154	0.0002	0.2434	0.0013	0.0788
Std.err.	0.0000	0.0225	0.0000	0.0207	0.0001	0.0192
α_1	0.0009	0.0001	0.0002	0.1377	0.0005	0.2592
Std.err.	0.0001	0.0000	0.0001	0.0193	0.0001	0.0230
β_1	0.0424	0.4234	0.4029	0.7435	0.1099	0.5228
Std.err.	0.0057	0.0313	0.0205	0.0324	0.0273	0.0366
P_{00}	0.9481		0.3020		0.8363	
Std.err.	0.0096		0.0364		0.0229	
P_{11}	0.0716		0.5082		0.1292	
Std.err.	0.0143		0.0439		0.0189	
Persistence	0.0433	0.4235	0.4031	0.8812	0.1104	0.7820
π_1 (Low volatility)	0.20		0.90		0.51	
π_2 (High volatility)	0.80		0.10		0.49	

The parameter estimates for MRS-GARCH models are presented in Tables 4.15 above. All the parameter estimates for India are significant at 1%, 5% and 10% level of significance. The two regimes were identified from the estimated models. All the transition probabilities were found to be significant at 1%, 5% and 10% level of significance, showing that regimes are persistent in MRS-GARCH models. All models show a strong persistence in volatility since the persistence values ranges from 0.0433 to 0.8812. This implies that the volatility is likely to remain high over several exchange rate return periods once it increases. The unconditional probability of the first regime ranges between 20% (norm) and 90% (student-t) for different MRS-GARCH models. The unconditional probability of the second regime ranges between 10% (student-t) and 80% (norm) for different MRS-GARCH models. The subsequent Section 4.5.4 presents the Bayesian estimates of MRS-GARCH models for China.

4.5.4. CHINA RESULTS

This section presents the results of China's exchange rate using Bayesian MRS-GARCH models. Table 4.16 below shows the summarized results of estimated MRS-GARCH models.

Table 4.16: Bayesian MCMC Model estimates of MRS-GARCH models for China

China						
	Normal		Student t		GED	
Model	Low Volatility	High volatility	Low volatility	High volatility	Low volatility	High volatility
μ			95.7979	3.5539	1.5416	1.0514
Std.err.			0.3569	0.0964	0.0391	0.0376
α_0	0.0001	0.1610	0.0001	0.0335	0.0001	0.4072
Std.err.	0.0000	0.0046	0.0000	0.0047	0.0000	0.0297
α_1	0.0003	0.5330	0.0014	0.5943	0.0018	0.2415
Std.err.	0.0000	0.0099	0.0001	0.0197	0.0002	0.0238
β_1	0.0005	0.4638	0.0758	0.1900	0.0063	0.3436
Std.err.	0.0000	0.0096	0.0122	0.0126	0.0005	0.0201
$\mathbf{P_{00}}$	0.8777		0.8725		0.9216	
Std.err.	0.0016		0.0080		0.0063	
$\mathbf{P_{11}}$	0.3182		0.2461		0.2865	
Std.err.	0.0073		0.0086		0.0145	
Persistence	0.0008	0.9968	0.0136	0.7843	0.0081	0.5851
π_1 (Low volatility)	0.75		0.79		0.97	
π_2 (High volatility)	0.25		0.21		0.03	

The parameter estimates for MRS-GARCH models are presented in Tables 4.16 above. The results revealed that the estimated parameters for China are significant at 1%, 5% and 10% level of significance. The estimates confirmed the existence of two regimes: The first regime is categorized by low volatility, whereas the second one reveals high volatility. The transition probabilities are all significant. This means that all regimes are persistent in MRS-GARCH models for China's exchange rate. All models show a strong persistence in volatility since the persistence values ranges from 0.0008 to 0.9968. This implies that the volatility is likely to remain high over several exchange rate return periods once it increases. The unconditional probability of the first regime ranges between 75% (norm) and 97% (GED) for different MRS-GARCH models. The unconditional probability of the second regime ranges between 3% (GED) and 25% (norm) for different MRS-GARCH models. The next section 4.5.5 presents the Bayesian estimates of MRS-GARCH models for South Africa.

4.5.5. SOUTH AFRICA RESULTS

This section presents the results of South African's exchange rate. The summary of the results are presented in Table 4.17 below.

Table 4.17: Bayesian MCMC Model estimates of MRS-GARCH models for South Africa

South Africa						
	Normal		Student t		GED	
Model	Low Volatility	High volatility	Low volatility	High volatility	Low volatility	High volatility
μ			4.1866	3.1138	0.9004	15.5730
Std.err.			0.2251	0.2214	0.0112	0.3731
α_0	0.0001	0.0189	0.0001	0.3226	0.0003	0.6930
Std.err.	0.0000	0.0010	0.0000	0.0251	0.0000	0.0207
α_1	0.0016	0.6398	0.0028	0.3386	0.3243	0.8092
Std.err.	0.0001	0.0154	0.0003	0.0373	0.0414	0.0197
β_1	0.3633	0.3568	0.3682	0.5500	0.2104	0.1061
Std.err.	0.0102	0.0152	0.0188	0.0448	0.0191	0.0128
P_{00}	0.8685		0.8785		0.9877	
Std.err.	0.0050		0.0070		0.0007	
P_{11}	0.3466		0.4181		0.3800	
Std.err.	0.0114		0.0308		0.0179	
Persistence	0.0118	0.9966	0.3710	0.8886	0.2107	0.9153
π_1 (Low volatility)	0.60		0.69		0.41	
π_2 (High volatility)	0.40		0.31		0.59	

Table 4.17 presents the parameter estimates for MRS-GARCH models. The results revealed that the parameters estimated for South Africa are significant at 1%, 5%, and 10% levels of significance. The estimates showed that the two regime existence, the first regime is regarded as the low volatility, and the second regime as the high volatility. The transition probabilities are all significant at 1%, 5%, and 10% levels of significance. This means that the two regimes are persistent in MRS-GARCH models. All models show a strong persistence in volatility since the persistence values ranges from 0.0118 to 0.9966. This implies that the volatility is likely to remain high over several exchange rate return periods once it increases. The unconditional probability of the first regime ranges between 41% (GED) and 69% (student-t) for different MRS-GARCH models, while second regime ranges between 31% (student-t) and 59% (GED) for different MRS-GARCH models. The next Section 4.6 presents the in-sampling results of Bayesian MCMC estimates of MRS-GARCH models.

4.6.IN-SAMPLING EVALUATION USING BAYESIAN MCMC

In Table 4.18, the results of model selection and error metrics for all volatility models are presented.

Table 4.18: MS-GARCH: In-sample Results using Bayesian MCMC estimator

Model	Bayesian MCMC Estimation Method					
	Brazil					
	No. of Par.	DIC	MSE	RMSE	MAE	Best Model
MRS-GARCH-N	5	-888.79	0.94	0.97	0.61	MRS-GARCH-t
MRS-GARCH-t	6	-927.51	0.46	0.68	0.47	
MRS-GARCH-GED	6	-901.80	0.77	0.88	0.58	
Russia						
MS-GARCH-N	5	-771.63	1.10	1.05	0.80	MRS-GARCH-t
MS-GARCH-t	6	-776.11	0.97	0.98	0.74	
MS-GARCH-GED	6	-769.26	1.07	1.04	0.78	
India						
MS-GARCH-N	5	-1016.62	0.97	0.98	0.72	MRS-GARCH-t
MS-GARCH-t	6	-1028.78	0.82	0.91	0.66	
MS-GARCH-GED	6	-998.63	0.86	0.93	0.67	
China						
MS-GARCH-N	5	-1291.22	0.88	0.94	0.65	MRS-GARCH-t
MS-GARCH-t	6	-1300.67	0.77	0.88	0.57	
MS-GARCH-GED	6	-1293.39	1.10	1.05	0.67	
South Africa						
MS-GARCH-N	5	-1038.47	0.96	0.98	0.62	MRS-GARCH-t
MS-GARCH-t	6	-1045.39	0.87	0.93	0.57	
MS-GARCH-GED	6	-1036.91	0.95	0.98	0.59	

The above Table 4.18 presents model selection and error metrics. The study conducted the in-sampling evaluation to identify the best fitted models for exchange rates of BRICS countries. The competing models are compared based on model selection criterion and error metrics. The results revealed that the MRS-GARCH model with the t-distribution performed better in modelling the exchange rates of Brazil, Russia, India, and South Africa as opposed to other distributions. The MRS-GARCH models with t-distribution are regarded as the best fitted models. The next Section 4.7 presents the model comparison results.

4.7.MODEL COMPARISON RESULTS OF MRS-GARCH MODELS

The Table 4.19 below presents the MRS-GARCH model in-sampling results for model comparison.

Table 4.19: MRS-GARCH: In-sample Results for Model Comparison

Variable	Maximum Likelihood Estimation				Bayesian MCMC Estimation				Best **Model**
	Model	MSE	RMSE	MAE	Model	MSE	RMSE	MAE	
Brazil	MRS-GARCH-t	0.66	0.81	0.55	MRS-GARCH-t	0.46	0.68	0.47	MRS-GARCH-t
Russia	MRS-GARCH-N	1.03	1.02	0.76	MRS-GARCH- t	0.97	0.98	0.74	MRS-GARCH-t
India	MRS-GARCH-t	0.73	0.85	0.64	MRS-GARCH-t	0.82	0.91	0.66	MRS-GARCH-t
China	MRS-GARCH-GED	0.87	0.93	0.65	MRS-GARCH-t	0.77	0.88	0.57	MRS-GARCH-t
South Africa	MRS-GARCH-t	0.78	0.89	0.55	MRS-GARCH-t	0.87	0.93	0.57	MRS-GARCH-t

This section compares the forecast ability of the estimated models under MLE and Bayesian MCMC estimation methods as presented in Table 4.19 above. The aim of the section is to determine the best estimated method that can be used to estimate MRS-GARCH models. The results from the data deduced 30 models (15 from MLE and 15 from Bayesian MCMC). From the 30 models, 10 models were selected as the best models. These models are deemed accurate and can be used for further forecasting. The three criteria of MSE, RMSE, and MAE as discussed in chapter 3 were used to measure forecast accuracy of the models.

The best model for each of the exchange rate of BRICS countries is presented as follows; Brazil, Russia, and China were better produced from the Bayesian MCMC estimation method, while India and South Africa were better produced from the ML estimation method. To test the accuracy of the 5 selected models, the DM test for predictive accuracy was employed. The subsequent Section 4.8 presents the results of the DM test for in-sampling on testing the accuracy of the models.

4.8.PREDICTIVE ACCURACY OF MODELS (IN-SAMPLING)

This section predicts the accuracy of the models. According to Binner *et al* (2006), Chai & Draxler (2014), Wei *et al* (2016) the model with the lowest MSE, RMSE, and MAE does not necessarily suggest a superior forecasting ability of a model, since the difference between error metrics for two different models may not be significantly different from zero. Therefore, in order to assess the predictive accuracy of the two estimated models, DM test for pairs of models for exchange rate were performed using the absolute error loss differential and the results are summarized in Table 4.20.

Table 4.20: Comparison of Model Accuracy: Diebold-Mariano Absolute Error Test

HAC standard errors and covariance (Bartlett kernel, Newey-West fixed bandwidth = 6.0000)			
Absolute Error Loss Differential		t-Statistic	Prob.
Variable	**Hypothesis**		
Brazil	MRS-GARCH-t vs. MRS-GARCH-t	3.0357	0.0013
Russia	MRS-GARCH-N vs. MRS-GARCH-t	0.7721	0.2204
India	MRS-GARCH-t vs. MRS-GARCH-t	2.0776	0.0194
China	MRS-GARCH-GED vs. MRS-GARCH-t	3.1743	0.0008
South Africa	MRS-GARCH-t vs. MRS-GARCH-t	3.1770	0.0000

According to the results in Table 4.20, the null hypothesis of the paired models MRS-GARCH-t (ML) vs. MRS-GARCH-t (MCMC), MRS-GARCH-t (ML) vs. MRS-GARCH-GED (MCMC), MRS-GARCH-GED (ML) vs. MRS-GARCH-t (MCMC), and MRS-GARCH-t (ML) vs. MRS-GARCH-t (MCMC) for exchange rates having equal predictive accuracy is rejected since their respective p-values (0.0013, 0.0194, 0.0008 and 0.0000) of the test statistics are less than 1%, 5%, and 10% levels of significance. However, the null hypothesis that the paired model MRS-GARCH-N (ML) vs. MRS-GARCH-t (MCMC) for Russia having the same predictive accuracy is not rejected since the p-value (0.2204) of the test statistic is great than 1%, 5%, and 10% levels of significance. This indicates that the accuracy of the paired models is equal. Based on the error metrics and DM Absolute error tests, the appropriate conditional distribution to be used is student-t distribution and the best estimation method is Bayesian (MCMC) for in-sample volatility forecasts. The next Section 4.9 presents the results of the out-of-sample volatility forecasts.

4.9. OUT-OF-SAMPLING EVALUATION USING MAXIMUM LIKELIHOOD

In this section, the study evaluates the models through their out-of-sample performances using Maximum Likelihood. The study presents the results of the out-of-sample evaluation of 1-month, 5-month and 10-month-ahead forecast of volatility. The next subsection presents the out-of-sample 1-month-ahead forecast performance.

4.9.1. THE 1-MONTH-AHEAD VOLATILITY FORECAST

The Table 4.21 below presents the results of the out- of-sample evaluation of 1-month-ahead volatility.

Table 4.21: Out-of-sample evaluation of 1-month-ahead volatility Results

Maximum Likelihood Estimation Method					
Variable	**Model**	**MSE**	**RMSE**	**MAE**	
Brazil	MRS-GARCH-N	701.46	26.49	26.49	MRS-GARCH-N
	MRS-GARCH-t	721.04	26.85	26.85	
	MRS-GARCH-GED	2348.24	48.47	48.46	
Russia	MRS-GARCH-N	704.61	26.54	26.54	MRS-GARCH-GED
	MRS-GARCH-t	706.30	26.58	26.58	
	MRS-GARCH-GED	701.46	26.49	26.49	
India	MRS-GARCH-N	725.36	26.93	26.93	MRS-GARCH-t
	MRS-GARCH-t	696.23	26.38	26.38	
	MRS-GARCH-GED	721.04	26.85	26.84	
China	MRS-GARCH-N	706.87	26.59	26.59	MRS-GARCH-N
	MRS-GARCH-t	711.73	26.68	26.28	
	MRS-GARCH-GED	731.57	27.04	27.04	
South Africa	MRS-GARCH-N	704.09	26.54	26.54	MRS-GARCH-t
	MRS-GARCH-t	700.93	26.48	26.48	
	MRS-GARCH-GED	705.77	26.57	26.57	

The above Table 4.21 reported the out-of-sample evaluation of 1-month ahead volatility forecasts. The three error metrics as discussed in Chapter 3 were used to evaluate the model forecasts. According to the MSE, RMSE and MAE, the MRS-GARCH model with t-distribution performed better in modelling the exchange rates of India and South Africa for 1-month ahead volatility forecasts. On the other hand, MRS-GARCH model with normal distribution performed better in modelling the exchange rates of Brazil and China. MRS-GARCH model with GED performed better in modelling the exchange rate of Russia for 1-month ahead volatility forecasts. The subsequent Subsection 4.9.2 presents the results of the out-of-sample evaluation of 5-month-ahead volatility forecast.

4.9.2. THE 5-MONTH-AHEAD VOLATILITY FORECAST

The section presents the summary of the results of the out-of-sample evaluation of 5-month ahead volatility forecasts.

Table 4.22: Out-of-sample evaluation of 5-month-ahead volatility Results

Maximum Likelihood Estimation Method					**Selected Model**
Variable	**Model**	**MSE**	**RMSE**	**MAE**	
Brazil	MRS-GARCH-N	703.12	26.52	26.51	MRS-GARCH-GED
	MRS-GARCH-t	715.90	26.76	26.76	
	MRS-GARCH-GED	701.91	26.49	26.49	
Russia	MRS-GARCH-N	684.35	26.16	26.13	MRS-GARCH-t
	MRS-GARCH-t	683.16	26.14	26.13	
	MRS-GARCH-GED	714.52	26.73	26.72	
India	MRS-GARCH-N	705.38	26.56	26.56	MRS-GARCH-t
	MRS-GARCH-t	698.03	26.42	26.42	
	MRS-GARCH-GED	699.57	26.44	26.44	
China	MRS-GARCH-N	719.11	26.81	26.81	MRS-GARCH-t
	MRS-GARCH-t	705.85	26.57	26.57	
	MRS-GARCH-GED	711.14	26.67	26.67	
South Africa	MRS-GARCH-N	712.69	26.69	26.69	MRS-GARCH-GED
	MRS-GARCH-t	706.77	26.58	26.58	
	MRS-GARCH-GED	697.42	26.41	26.41	

The Table 4.22 above results of error metrics revealed that MRS-GARCH model with GED-distribution performed better in modelling Brazil and South Africa. However, the MSE, RMSE and MAE revealed that MRS-GARCH model with t-distribution performed better in modelling the exchange rates of Russia, India and China for 5-month ahead volatility forecasts. The next Subsection 4.9.3 presents the results of the out-of-sample evaluation of 10-month-ahead volatility forecasting.

4.9.3. THE 10-MONTHS-AHEAD VOLATILITY FORECAST

This section presents the results of out-of-sample evaluation forecasts of 10-month-ahead volatility. Table 4.23 above presents the out-of-sample results of 10-month ahead volatility forecasts. The models were evaluated based on MSE, RMSE and MEA. According to the error metrics, MRS-GARCH models with student-t distribution performed better in modelling the exchange rates of Brazil and India for 10-month ahead volatility forecasts.

Table 4.23: Out-of-sample evaluation of 10-month-ahead volatility Results

Maximum Likelihood Estimation Method					Best Model
Variable	Model	MSE	RMSE	MAE	
Brazil	MS-GARCH-N	736.43	27.13	27.13	MRS-GARCH-t
	MS-GARCH-t	696.90	26.40	26.40	
	MS-GARCH-GED	714.64	26.73	26.72	
Russia	MS-GARCH-N	715.66	26.75	26.74	MRS-GARCH-GED
	MS-GARCH-t	736.40	27.14	27.13	
	MS-GARCH-GED	701.87	26.49	26.48	
India	MS-GARCH-N	714.63	26.73	26.73	MRS-GARCH-t
	MS-GARCH-t	702.45	26.49	26.49	
	MS-GARCH-GED	715.04	26.74	26.74	
China	MS-GARCH-N	673.51	25.98	25.94	MRS-GARCH-N
	MS-GARCH-t	697.21	26.39	26.39	
	MS-GARCH-GED	702.45	26.49	26.49	
South Africa	MS-GARCH-N	687.55	26.14	26.14	MRS-GARCH-GED
	MS-GARCH-t	697.21	26.39	26.39	
	MS-GARCH-GED	472.06	21.18	218.18	

On the other hand, MRS-GARCH model with GED-distribution performed better in modelling the exchange rates of Russia and South Africa for 10-month ahead volatility forecasts. Finally, MRS-GARCH model with normal distribution performed better in modelling the exchange rate of China for 10-month ahead volatility forecasts. The subsequent Section 4.10 presents the results of MRS-GARCH models using the Bayesian MCMC estimation method.

4.10. OUT-OF-SAMPLING VOLATILITY FORECASTS USING BAYESIAN MCMC

In this section, the study evaluates the models through their out-of-sample performances using Bayesian MCMC. The forecast horizons of 1, 5, and 10 months were considered in this study. The next Subsection 4.9.1 presents the results of 1-month-ahead volatility forecast.

4.10.1. THE 1-MONTH-AHEAD VOLATILITY FORECAST

Table 4.24 below presents the results of the out-of-sample evaluation of 1-month ahead volatility forecasts. The Table 4.24 below revealed the results of out-of-sample evaluation of 1-month ahead volatility forecast using Bayesian MCMC. According to error metrics, the MRS-GARCH model with the GED-distribution performed better in modelling the exchange rates of Brazil, Russia, India, China, and South Africa for 1-month ahead volatility forecasts. The next Subsection 4.10.2 presents the results of the out-of-sample evaluation of 5-month-ahead volatility forecasts.

Table 4.24: Out of sample evaluation of 1-month-ahead volatility Results

Bayesian MCMC Estimation Method					Selected Model
Variables	Model	MSE	RMSE	MAE	
Brazil	MRS-GARCH-N	581.37	24.11	24.11	MRS-GARCH-GED
	MRS-GARCH-t	633.08	25.16	25.16	
	MRS-GARCH-GED	319.81	17.88	17.88	
Russia	MRS-GARCH-N	662.16	25.73	25.73	MRS-GARCH-GED
	MRS-GARCH-t	673.99	25.96	25.96	
	MRS-GARCH-GED	642.70	25.35	25.35	
India	MRS-GARCH-N	712.48	26.69	26.69	MRS-GARCH-GED
	MRS-GARCH-t	691.73	26.30	26.30	
	MRS-GARCH-GED	315.42	17.76	17.76	
China	MRS-GARCH-N	700.64	26.47	26.47	MRS-GARCH-GED
	MRS-GARCH-t	721.45	26.86	26.86	
	MRS-GARCH-GED	309.67	17.60	17.60	
South Africa	MRS-GARCH-N	507.49	22.53	22.53	MRS-GARCH-GED
	MRS-GARCH-t	655.62	25.61	25.60	
	MRS-GARCH-GED	274.34	16.56	16.56	

4.10.2. THE 5-MONTH-AHEAD VOLATILITY FORECAST

The section presents the 5-month-ahead volatility forecast of BRICS countries' results.

Table 4.25: Out of sample evaluation of 5-month-ahead volatility Results

Bayesian MCMC Estimation Method					Selected Model
Variable	Model	MSE	RMSE	MAE	
Brazil	MRS-GARCH-N	509.71	21.71	21.71	MRS-GARCH-GED
	MRS-GARCH-t	708.25	26.41	26.41	
	MRS-GARCH-GED	176.82	12.74	12.74	
Russia	MRS-GARCH-N	616.15	24.60	24.60	MRS-GARCH-GED
	MRS-GARCH-t	722.25	26.76	26.76	
	MRS-GARCH-GED	156.57	11.46	11.46	
India	MRS-GARCH-N	500.08	22.36	22.11	MRS-GARCH-GED
	MRS-GARCH-t	679.53	26.07	25.53	
	MRS-GARCH-GED	177.92	13.34	12.88	
China	MRS-GARCH-N	547.95	23.41	23.25	MRS-GARCH-GED
	MRS-GARCH-t	575.91	24.00	23.17	
	MRS-GARCH-GED	193.74	13.92	13.37	
South Africa	MRS-GARCH-N	480.27	21.92	21.75	MRS-GARCH-GED
	MRS-GARCH-t	430.17	20.74	20.36	
	MRS-GARCH-GED	159.80	12.64	12.22	

The Table 4.25 above revealed the results of out-of-sample evaluation of 5-month ahead volatility forecast. The results revealed that MRS-GARCH model with the GED-distribution performed better in modelling the exchange rates of BRICS countries for 5-month ahead volatility forecasts. The next Subsection 4.10.3 presents the results of the out-of-sample evaluation of 10-month-ahead volatility forecasts.

4.10.3. THE 10-MONTH-AHEAD VOLATILITY FORECAST

The section presents the 10-month-ahead volatility forecast of BRICS countries' results.

Table 4.26: Out of sample evaluation of 10-month-ahead volatility Results

Bayesian MCMC Estimation Method					Selected Model
Variable	Model	MSE	RMSE	MAE	
Brazil	MRS-GARCH-N	1097	33.13	32.44	MRS-GARCH-GED
	MRS-GARCH-t	650.26	25.50	25.34	
	MRS-GARCH-GED	151.78	12.32	11.81	
Russia	MRS-GARCH-N	559.76	23.66	23.56	MRS-GARCH-GED
	MRS-GARCH-t	617.65	24.85	24.60	
	MRS-GARCH-GED	101.52	10.08	8.51	
India	MRS-GARCH-N	573.08	23.94	23.88	MRS-GARCH-GED
	MRS-GARCH-t	532.76	23.08	22.86	
	MRS-GARCH-GED	149.86	12.24	11.83	
China	MRS-GARCH-N	398.12	19.95	19.51	MRS-GARCH-GED
	MRS-GARCH-t	396.13	19.90	19.41	
	MRS-GARCH-GED	91.89	9.59	8.28	
South Africa	MRS-GARCH-N	527.08	22.96	22.16	MRS-GARCH-GED
	MRS-GARCH-t	691.61	26.30	25.83	
	MRS-GARCH-GED	89.17	9.44	8.44	

Table 4.26 above presents the results of the out-of-sample evaluation 10-months ahead volatility forecast. The results showed that MRS-GARCH with GED-distribution outperformed other competing distributions in forecasting the out-of-sample evaluation for the 10-month period. The subsequent Section 4.11 presents the results of out-of-sample volatility forecasts using Maximum Likelihood Estimation (MLE) and Bayesian MCMC estimation methods.

4.11. COMPARISON RESULTS OF MRS-GARCH MODELS

This section presents the model comparison results for 1, 5, 10-month ahead volatility forecasts. The results are presented in Tables 4.27 to 4.29 below.

Table 4.27: Out-of-sample Results for 1-month ahead

Variable	Maximum Likelihood Estimation				Bayesian MCMC Estimation			
	Model	MSE	RMSE	MAE	Model	MSE	RMSE	MAE
Brazil	MRS-GARCH-N	701.46	26.49	26.49	MRS-GARCH-GED	319.81	17.88	17.88
Russia	MRS-GARCH-GED	701.46	26.49	26.49	MRS-GARCH-GED	642.70	25.35	25.35
India	MRS-GARCH-t	696.23	26.38	26.38	MRS-GARCH-GED	315.42	17.76	17.76
China	MRS-GARCH-N	706.87	26.59	26.59	MRS-GARCH-GED	309.67	17.60	17.60
South Africa	MRS-GARCH-t	700.93	26.48	26.48	MRS-GARCH-GED	274.34	16.56	16.56

Table 4.28: Out-of-sample Results for 5-month ahead

Variable	Maximum Likelihood Estimation				Bayesian MCMC Estimation			
	Model	MSE	RMSE	MAE	Model	MSE	RMSE	MAE
Brazil	MRS-GARCH-GED	701.91	26.49	26.49	MRS-GARCH-GED	176.82	12.74	12.74
Russia	MRS-GARCH-t	683.16	26.14	26.13	MRS-GARCH-GED	156.57	11.46	11.46
India	MRS-GARCH-t	698.03	26.42	26.42	MRS-GARCH-GED	177.92	13.34	12.88
China	MRS-GARCH-t	705.85	26.57	26.57	MRS-GARCH-GED	193.74	13.92	13.37
South Africa	MRS-GARCH-GED	697.42	26.41	26.41	MRS-GARCH-GED	159.80	12.64	12.22

Table 4.29: Out-of-sample Results for 10-month ahead

Variable	Maximum Likelihood Estimation				Bayesian MCMC Estimation			
	Model	MSE	RMSE	MAE	Model	MSE	RMSE	MAE
Brazil	MRS-GARCH-t	696.90	26.40	26.40	MRS-GARCH-GED	151.78	12.32	11.81
Russia	MRS-GARCH-GED	701.87	26.49	26.48	MRS-GARCH-GED	101.52	10.08	8.51
India	MRS-GARCH-t	673.51	25.98	25.94	MRS-GARCH-GED	149.86	12.24	11.83
China	MRS-GARCH-N	673.51	25.98	25.94	MRS-GARCH-GED	91.89	9.59	8.28
South Africa	MRS-GARCH-GED	472.06	21.18	218.18	MRS-GARCH-GED	89.17	9.44	8.44

The Tables 4.27 to 4.29 above present the model comparison for 1, 5, 10-month ahead volatility forecasts. The models were evaluated based on error metrics to select better the MRS-GARCH models using MLE and Bayesian MCMC estimation methods. The results revealed that MRS-GARCH models with GED using Bayesian MCMC estimation outperformed other distributions in modelling 1, 5, 10-month ahead volatility forecasts. To compare the predictive accuracy, the study conducted a DM test. The subsequent section presents the results of the DM test to check the equal predictive accuracy of the models.

4.12. PREDICTIVE ACCURACY OF MODELS (OUT-OF-SAMPLING)

This section predicts the significant accuracy of the paired models. As indicated earlier, a lower MSE, RMSE, and MAE does not necessarily suggest a superior forecasting ability of a model, since the difference between error metrics, say MSEs, for two different models may not be significantly different from zero (Chen, *et al*., 2014) and Harris & Sollis (2003). Therefore, in order to assess the predictive accuracy of the two estimated models, the DM test for pairs of models for exchange rate was performed using the absolute error loss differential and the results summarized in Table 4.30.

Table 4.30: Comparison of Model Accuracy: Diebold-Mariano Absolute Error Test (1-month-ahead volatility Results)

HAC standard errors and covariance (Bartlett kernel, Newey-West fixed bandwidth = 6.0000)			
Absolute Error Loss Differential		t-Statistic	Prob.
Variable	Hypothesis		
Brazil	MRS-GARCH-N vs. MRS-GARCH-GED	6.5732	0.0000
Russia	MRS-GARCH-GED vs. MRS-GARCH-GED	8.2895	0.0000
India	MRS-GARCH-t vs. MRS-GARCH-GED	7.9653	0.0000
China	MRS-GARCH-N vs. MRS-GARCH-GED	6.7424	0.0003
South Africa	MRS-GARCH-t vs. MRS-GARCH-GED	9.5041	0.0000

The Table 4.30 above presents the model accuracy comparison results using DM Test for the 1-month-ahead volatility. The results of the study revealed that all the p-values of the respective models are less than 1%, 5%, and 10% levels of significance. There is a statistically significant difference between the paired models. Therefore, it can be concluded that the estimated MRS-GARCH-GED models (Bayesian MCMC) for exchange rate are more accurate than their counterpart where MLE is used to estimate the models. The next Table 4.31 presents Comparison of Model Accuracy: Diebold-Mariano Absolute Error Test (5-month-ahead volatility Results).

Table 4.31: Comparison of Model Accuracy: Diebold-Mariano Absolute Error Test (5-month-ahead volatility Results)

HAC standard errors and covariance (Bartlett kernel, Newey-West fixed bandwidth = 6.0000)			
Absolute Error Loss Differential			
Variable	Hypothesis	t-Statistic	Prob.
Brazil	MS-GARCH-GED vs. MS-GARCH-GED	7.2514	0.0000
Russia	MS-GARCH-t vs. MS-GARCH-GED	10.4630	0.0000
India	MS-GARCH-t vs. MS-GARCH-GED	7.9653	0.0000
China	MS-GARCH-t vs. MS-GARCH-GED	0.7424	0.0000
South Africa	MS-GARCH-GED vs. MS-GARCH-GED	9.5041	0.0000

Table 4.31 above presents the model accuracy comparison results using the DM Test for the 5-month-ahead volatility. The results of the study revealed that all the p-values of the respective models are less than 1%, 5%, and 10% levels of significance. There is a statistically significant difference between the paired models. Therefore, it can be concluded that the estimated MRS-GARCH-GED models using Bayesian MCMC for exchange rate are more accurate than their counterpart where MLE is used to estimate the models. The next Table 4.32 presents comparison of model accuracy: Diebold-Mariano (DM) absolute error test (10-month-ahead volatility results).

Table 4.32: Comparison of Model Accuracy: Diebold-Mariano Absolute Error Test (10-month-ahead volatility Results)

HAC standard errors and covariance (Bartlett kernel, Newey-West fixed bandwidth = 6.0000)			
Absolute Error Loss Differential			
Variable	Hypothesis	t-Statistic	Prob.
Brazil	MS-GARCH-t vs. MS-GARCH-GED	2.3564	0.0000
Russia	MS-GARCH-GED vs. MS-GARCH-GED	4.5625	0.0000
India	MS-GARCH-t vs. MS-GARCH-GED	5.6662	0.0000
China	MS-GARCH-N vs. MS-GARCH-GED	2.1123	0.0000
South Africa	MS-GARCH-GED vs. MS-GARCH-GED	5.2130	0.0000

The results of the model accuracy comparison using DM test for the 10-month-ahead volatility are presented in Table 4.32 above. The results of the study revealed that all the p-values of the respective models are less than 1%, 5%, and 10% levels of significance. There is a statistically significant difference between the paired models. Therefore, it can be concluded that the estimated MRS-GARCH-GED models using Bayesian MCMC for exchange rate are more accurate than their counterpart where MLE is used to estimate the models. Therefore, the MRS-

GARCH-GED using Bayesian MCMC is the best model to forecast the out-of-sample volatility. The last Section 4.13 presents the chapter summary.

4.13. CHAPTER SUMMARY

In this chapter, a preliminary analysis to detect nonlinear and nonlinear unit root in exchange rates of the BRICS countries was conducted. The nonlinear and nonlinear unit root tests were administered under MRS-GARCH models. The study used MLE Bayesian MCMC estimation methods to model and forecast the exchange rates. Furthermore, the study conducted comparison analysis using error metrics and Diebold-Mariano test for prediction accuracy. The comparison was meant to determine whether a forecast is significantly better than another one (for in-sample and out-of-sample forecasting evaluation).

The next chapter presents the summary of the findings, conclusion and recommendations of the study.

CHAPTER 5
CONCLUSION AND RECOMMENDATIONS

5.1.INTRODUCTION

The current chapter follows on the results presented in Chapter 4. It provides a summary of the study, the envisaged contribution to the body of knowledge, limitations and gives conclusions based on the study objectives. The chapter also provides recommendations and areas for further research. Challenges encountered during the course of the study are also highlighted.

The rest of the chapter is presented as follows: In Section 5.2, the discussion of the results according to the objectives is provided, Section 5.3 highlights the contribution of the study and Section 5.4 highlights the limitations of the study. In Section 5.5, conclusions are drawn based on the discussions emanating from individual objectives, Section 5.6 provides the recommendations and Section 5.7 gives the summary of the thesis.

5.2.DISCUSSION OF THE RESULTS

The study investigated the performance of MRS-GARCH Models using MLE and Bayesian MCMC estimation to model and forecast volatility of exchange rate of BRICS countries. The MLE and Bayesian MCMC estimation methods were modelled using three distributions (Normal, Student-t and the Generalized Error Distribution (GED)). Furthermore, the study conducted comparative analysis using model selection criterions (BIC, AIC, LOGL & DIC) and error metrics (MSE, RMSE and MAE) to determine best models to forecast the exchange rates of BRICS countries. Diebold-Mariano test was used to determine the accuracy of the selected model. The base of the forecasts was on in-sample and out-of-sample forecasting evaluation.

Preliminary analyses were conducted to examine the nature of the data in question and also to ensure that the data is fit for primary analyses. The study found that the data was not normally distributed and nonstationary. This was confirmed by conducting various tests such as the RESET, ARCH, BDS, McLeod & Li, Keenan, Tsay, and CUSUM. The use of these tests was motivated by Ismail and Isa (2006), Giannerini (2013) and Cavalheiro *et al* (2015) who discouraged the use of a single test as nonlinear behaviour can be detected in different ways. To further examine and confirm the existence or otherwise of any structural change in the data, BDS tests were conducted and the results revealed no structural change in the data, while Lagrange Multiplier (LM) tests also suggested that the exchange rates were nonlinearly dependent. The study by Ashley & Patterson (2006), Ismail & Isa (2006), Chen (2011), and Chen (2014) found similar results supporting the use of RESET, ARCH, BDS, McLeod and Li, Keenan, Tsay, and CUSUM tests. This satisfied the requirements for the use of nonlinear models. The findings of the study are discussed below as per objectives.

- **To estimate MLE and Bayesian MCMC MRS-GARCH models of the exchange rates of BRICS countries**

Maximum Likelihood and Bayesian MCMC were used to estimate the parameters of each of the exchange rates of BRICS countries. The study estimated 30 MRS-GARCH models of which 15 were estimated using Maximum Likelihood and the remaining 15 using Bayesian MCMC estimation methods. The methods were estimated using the three conditional distributions namely; the Normal, Student-t and the GED.

ML estimation findings from the two-regime GARCH models

The study found that all the parameter estimates for BRICS exchange rates are statistically significant at all levels. The estimates confirmed the existence of two regimes; the first regime is regarded as low volatility and second as high volatility. The MRS-GARCH models of Brazil and Russia showed a strong persistence in volatility under all distributions for the two regimes. The MRS-GARCH model with Student-t and GED for India showed strong persistence in volatility of the two regimes. This implies that the volatility is likely to remain high over several exchange rate periods. The MRS-GARCH models of China has both weak and strong persistence level in volatility for all distributions. This implies that the volatility is likely to remain low for a shorter exchange rate periods and longer on the high regime. The MRS-GARCH models for South Africa showed a high persistence in volatility under Student-t distribution as compared to other distributions. This implies that the volatility is likely to remain high over several exchange rate periods when MRS-GARCH model is estimated using Student-t distribution.

The unconditional probability of being in the first regime for Brazil ranges from Student-t to GED for different MRS-GARCH models. Similarly, the unconditional probability of being in the second regime ranges from Normal to Student-t & GED for different MRS-GARCH models. The unconditional probability of being in the first regime for Russia ranges from GED to Normal, whereas the unconditional probability of being in the second regime ranges from Normal to GED. The study revealed that the unconditional probabilities of being in the first regime for India ranges from Student-t to Normal, whereas the unconditional probability of being in the second regime ranges from GED to Normal. The unconditional probability of being in the first regime for China ranges from Normal & Student-t to GED for different MRS-GARCH models whereas the unconditional probability of being in the second regime ranges from GED to Normal & Student-t for different MS-GARCH models. The unconditional probability of being in the first regime for South Africa ranges from GED to Normal, whereas the unconditional probability of being in the second regime ranges from Normal to GED. The transition probabilities are all significant, implying that all regimes are particularly persistent in MRS-GARCH models. The findings of the study are supported by Naeini & Fatahi (2012), Sopipan *et al*. (2012), Günay (2015) and Iqbal (2016).

Bayesian MCMC estimation findings from the two-regime GARCH models

The parameters estimate of the BRICS exchange rate were found to be statistically significant at all levels. The MRS-GARCH models of Russia showed a strong persistence in volatility under all distributions for the two regimes. The findings imply that the volatility is likely to remain high over several exchange rate periods. The MRS-GARCH models of Brazil, China, India and South Africa have both weak and strong persistence level in volatility for all distributions. This implies that the volatility is likely to remain low for a shorter exchange rate periods and longer on the high regime. The study by Henneke *et al*. (2011) and Shiferaw (2018) also found similar results.

The unconditional probability of being in the first regime for Brazil ranges from Normal to Student-t for different MRS-GARCH models. Similarly, the unconditional probability of the second regime ranges from Normal to Student-t for different MRS-GARCH models. The unconditional probability of being in the first regime for Russia ranges from GED to Normal for different MRS-GARCH models, whereas the unconditional probability of being in the second regime ranges from Normal to GED for different MRS-GARCH models. The unconditional probability of being in the first regime for India ranges from Normal to Student-t for different MRS-GARCH models, whereas the unconditional probability of being in the second regime ranges from Student-t to Normal for different MRS-GARCH models. The unconditional probability of being in the first regime for China ranges from Normal to GED for different MRS-GARCH models, whereas the unconditional probability of being in the second regime ranges from GED to Normal for different MRS-GARCH models. The unconditional probability of being in the first regime for South Africa ranges from GED to Student-t for different MRS-GARCH models, while second regime ranges from Student-t to GED for different MRS-GARCH models. The transition probabilities were found significant in all cases showing the persistence of both regimes in MRS-GARCH models.

- **To compare the efficiency and performance of MLE and Bayesian MCMC MRS-GARCH models for in-sample and out-of-sample volatility forecasts**

The various model selection criterions and error metrics were used to compare the efficiency and the performance of MRS-GARCH models using MLE and Bayesian MCMC.

The findings on MLE and Bayesian MCMC MRS-GARCH models for in-sample volatility forecasts

The findings of the study showed that 5 models were selected out of the 15 estimated models for MRS-GARCH using MLE method. The study found that the exchange rates of Brazil, India and South Africa can be better estimated using MRS-GARCH-t models. The exchange rates of Russia can be better estimated using MRS-GARCH-N model. Lastly, MRS-GARCH-GED model can be used to estimate the exchange rate of China. This means MRS-GARCH using MLE is inconsistent in selecting the distribution. In contrary, the study by Sajjad *et al*. (2008) and Sattayatham *et al*. (2012) found MRS-GARCH models with Student-t distribution outperformed other distributions.

Similarly, the 5 models were also selected out of the 15 estimated models for MRS-GARCH using Bayesian MCMC. The findings of the study found that the MRS-GARCH model with the Student-t distribution performed better in modelling the exchange rates of BRICS countries. This implies that MRS-GARCH using Bayesian MCMC has consistently selected Student-t distribution as the best distribution amongst the three. The study by Ardia (2009) and Shiferaw (2018) also found similar results.

Furthermore, the study compared the in-sample evaluation of the estimated models for MRS-GARCH using MLE and Bayesian MCMC. The findings of the study revealed that Brazil, Russia, and China were better modelled using the Bayesian MCMC estimation method, while India and South Africa were better modelled from the ML estimation method. According to the results, MRS-GARCH with Student-t distribution appeared to be better in modelling the exchange of BRICS countries under both estimation methods. This implies that the t-distribution gives the better MRS-GARCH models forecast volatility. The Bayesian MCMC estimation method outperformed the MLE method in modelling the MRS-GARCH volatility forecasts. The results are in line with the study by Sattayathama *et al*. (2012), Günay (2015), Abounoori *et al*. (2016) and Shiferaw (2018).

The findings on MLE and Bayesian MCMC MRS-GARCH models for out-of-sample volatility forecasts

The findings of the out-of-sample volatility forecasts of 1-month, 5-month and 10-month ahead were computed using MLE and Bayesian MCMC methods. The findings of the study revealed that MLE method under the MRS-GARCH models with Student-t and Normal distributions performed better in modelling exchange rates of India & South Africa; and Brazil & China respectively for 1-month ahead volatility forecasts. The study found that MLE method under the MRS-GARCH model with GED only performed better in modelling the exchange rate of Russia for 1-month ahead volatility forecasts. The study by Marcucci (2005) and Iqbal (2016) also found similar results.

The findings on the 5-month ahead volatility forecasts revealed that MLE under the MRS-GARCH model with Student-t distribution performed better in modelling the exchange rates of Russia, India and China. The findings further revealed that Brazil & South Africa exchange rates were better modelled using MLE under MRS-GRACH model with GED. The study by Marcucci (2005) and Iqbal (2016).

The findings for 10-month ahead volatility forecasts revealed that MLE under the MRS-GARCH models with Student-t distribution and GED performed better in modelling exchange rate of the two countries (Brazil & India) and (Russia & South Africa) respectively. Furthermore, the MLE under the MRS-GARCH models with Normal distribution only performed better in modelling exchange rate of China. The findings of the study revealed that the MRS-GARCH models using the MLE method showed that the t-distribution dominated with seven out-of-sample volatility forecasts (2 in 1-month ahead, 3 in 5-month ahead and 2 in 10-month ahead). The findings clearly revealed that MRS-GARCH models with Student-t distribution performed better than other distributions. Similar results were also found by studies by Sattayathama *et al*. (2012) and Abounoori *et al*. (2016).

In a case where Bayesian MCMC was used, the results revealed that the MRS-GARCH model with the GED performed better in modelling the exchange rate of all the BRICS countries for 1-month, 5-month and 10-month ahead volatility forecasts. This clearly means that the MRS-GARCH using Bayesian MCMC has consistently selected GED as the best distribution amongst the three. GED is consistent in selecting the distribution of modelling exchange rate of the BRICS countries. The findings are supported by Sattayathama *et al.* (2012) and Iqbal (2016).

The study further compared the out-of-sample volatility forecast for MRS-GARCH using MLE and Bayesian MCMC. The findings of the study revealed that MRS-GARCH models with GED using Bayesian MCMC estimation outperformed other distributions in modelling 1-month, 5-month and 10-month ahead volatility forecasts. Similar results were also found by studies Shiferaw (2018) and Ardia *et al.* (2018).

- **To compare the predictive accuracy of MRS-GARCH models using MLE and Bayesian MCMC**

The study computed the DM test in order to assess the predictive accuracy of the two competing models for in-sample and out-of-sample volatility forecasts. The DM test for pairs of models for exchange rate was performed using the absolute error loss differential.

Predictive accuracy for in-sample volatility forecasts

The findings of the study showed that the MRS-GARCH with student-t distribution performed better in most cases in modelling exchange rates of BRICS countries as compared to the computing models. The findings further revealed that the Bayesian MCMC estimation method is the better estimation method as compared to MLE method. The results are in line with study by Ahmed and Suliman (2011) and Alemohammad *et al*. (2017).

Predictive accuracy for out-of-sample volatility forecasts

The findings of the predictive accuracy for out-of-sample volatility forecasts of 1-month, 5-month and 10-month ahead were computed using the DM test. The findings of the study revealed that MRS-GARCH models with GED using Bayesian MCMC estimation outperformed other distributions in modelling 1-month, 5-month and 10-month ahead volatility forecasts.

5.3.CONTRIBUTIONS TO THE BODY OF KNOWLEDGE

This study is novel (to the best knowledge of the researcher) in a sense that there is no similar study in the context of BRICS that used MRS-GARCH model with the proposed estimation methods (MLE and Bayesian MCMC) to model and forecast the exchange rates. Most studies focused on the stock indices, and agricultural prices. In particular, the study is unique in terms of modelling the exchange rates of BRICS countries. The study contributes literature pertaining to the said models on exchange rates of BRICS countries.

Most of the study such as Brunetti *et al*. (2008); Bala & Asemota (2013); May & Farrell (2017) and Augustsson (2014) revealed that there is no appropriate distribution to model and estimate the exchange rate of BRICS countries. The study contributes to the current literature by providing the appropriate distribution that can be used to estimate the exchange rate of BRICS countries. The current study as opposed to the above listed studies, contributes to the use of student-t distribution on in-sample and GED on out-of-sample volatility forecasts.

Most of the studies completed on the exchange rate mainly focused on the in-sample and out-of-sample volatility forecasts but not for BRICS countries (Babikir *et al*., 2012). The study contributes to the current literature by modelling exchange rate of BRICS using MRS-GARCH model, particularly for in-sample and out-of-sample volatility forecasts.

Most of the completed studies used error metrics to identify the better performing models (Yin, 2010 and Chuffart, 2015). The study contributed to the literature by using DM test to assess predictive accuracy of the paired models for both in-sample and out-of-sample volatility forecasts. This study may be beneficial to scholars, researchers, practitioners and risk managers to adopt alternative models for forecasting volatility.

5.4. LIMITATIONS

The study used monthly secondary time series data of the exchange rates consisting of 241 observations and it was obtained from Quantec. The models' requirements of the number of observations necessitated the extension period outside the inception of BRICS which was signed in April 2010. The researcher also did not cover other financial variables other than those deemed relevant for the study and proposed methods. The findings of the study may not be inferred or generalised as they only apply to the data used in the study. Due to limited available literature around the topic, the study consulted some sources older than ten years.

The application of MRS-GARCH models in the BRICS countries context is very limited to the best knowledge of the researcher. This study was an extension of other international studies such as those conducted by Sattayathama *et al*. (2012), Iqbal (2016), Abounoori *et al*. (2016), and Alemohammad *et al*. (2017). The studies by Sattayathama *et al*. (2012), Iqbal (2016), Abounoori *et al*. (2016), and Alemohammad *et al*. (2017) were only limited to MLE whilst the current study extends to Bayesian MCMC estimation method. It is important to note that searches from the databases such as the Google Scholar, Research Gate, J-store and Science Direct revealed that there are no studies that compared the estimation methods (MLE and Bayesian MCMC) for out-of-sample evaluation volatility forecasts.

5.5. CONCLUSIONS

The study investigated the performance of the hybrid models of MRS-GARCH using Maximum Likelihood & Bayesian MCMC estimation methods for three distributions (Normal, Student-t and GED) to determine the better technique to model the exchange rates of BRICS countries. The study focused on the two volatility forecasts namely: in-sample and out-of-sample. The in-sample volatility forecasts revealed that MRS-GARCH with Student-t distribution appeared to be better in modelling the exchange rates of BRICS countries using Bayesian MCMC estimation method. In the case of out-of-sample volatility forecasts, the MRS-GARCH with GED performed better in all cases compared to the competing models. As compared to Bayesian MCMC estimation method, the MLE method gave a lot of inconsistencies on the in-sample and out-of-sample volatility forecasts. The Bayesian MCMC estimation method gives better MRS-GARCH models estimates. The study by Shiferaw (2018) also found similar results. The DM accuracy test also confirmed the Bayesian MCMC estimation as the better method for modelling the exchange rates of BRICS countries.

5.6.RECOMMENDATIONS FOR FUTURE RESEARCH

The preceding section presented a summary of the results obtained in the study. Although the nonlinear GARCH modelling and forecasting technique proved to be good, there is room for further improvement.

- One can reproduce the same study by incorporating Monte Carlo on ML estimation method. This could improve estimating ability of the standard MLE since study found inconsistent results.
- A similar study can be extended by estimating MRS-GARCH models under MLE and MCMC with Skewed Student-t distribution and be compared with Normal, Student-t and GED. The reason to conduct a comparative study would assist in determining the better estimation method and distribution.
- The study recommends the use of hybrid models MRS-GARCH using Bayesian MCMC estimation method. This could help in improving the accuracy of the models.
- Further studies could use MRS-GARCH model with GED under Bayesian MCMC estimation utilizing different data sets. This could help establish if similar results will be found using a different data set.
- The Department of Higher Education and Training should look at sponsoring researchers to make attempts to develop the research areas on policies relating to the study. This will help with the implementations of the above-mentioned recommendations.

5.7.SUMMARY OF THE THESIS

The thesis was organized into five chapters in order to address the objectives stated in chapter one. Chapter 1 provided the introduction and background of the study, problem statement, and rationale of the study aim, and objectives of the study. Chapter 1 also outlined the significance of the study and scope limitations or delimitations of the study. Chapter 2 reviewed the relevant literature. The chapter initially provided the definition of MRS-GARCH model. Different modification versions of MRS-GARCH models were discussed. Chapter 2 also discussed the empirical literature of MRS-GARCH and its extensions. Chapter 3 of the study presented the research methodology applied in this study to reach the set objectives. Chapter 4 presented the data analysis and interpretation of results. Chapter 5 presented the summary of the findings, drew conclusions in respect of the objectives of the study and suggested recommendations as well as areas for future study.

REFERENCE

Abdullah, S.M., Siddiqua, S., & Siddiquee, M.S.H. (2017). Modeling and forecasting exchange rate volatility in Bangladesh using GARCH models: a comparison based on normal and Student's t-error distribution. Finance Innovations 3, 18. https://doi.org/10.1186/s40854-017-0071-z

Abounoori, E., Elmi, Z. M., & Nademi, Y. (2016). Forecasting Tehran stock exchange volatility; Markov switching GARCH approach. Physica A: Statistical Mechanics and its Applications, 445, 264-282.

Ahmed, A. E. M., & Suliman, S. Z. (2011). Modeling stock market volatility using GARCH models evidence from Sudan. International journal of business and social science, 2(23).

Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. In B. N. Petrov, & F. Csaki (Eds.), Proceedings of the 2nd International Symposium on Information Theory (pp. 267-281). Budapest: Akademiai Kiado.

Alemohammad, N, Rezakhah S, & Alizadeh, S.H. (2017). Markov switching component GARCH model: stability and forecasting. Communications in Statistics - Theory and Methods, 45, 4332-4348.

Andersen & Bollerslev (1998). DM-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer-Run Dependencies, Journal of Finance 53, 219-265.

Andersen, T.G., Bollerslev, T. Christoffersen, P.F., & Diebold, F.X. (2007). Practical Volatility and Correlation Modeling for Financial Market Risk Management. University of Chicago Press.

Anoruo, E., & Elike, U. (2008). Asymmetric dynamics in current account-interest rate nexus: evidence from Asian countries. Investment management and financial innovations, (5, Iss. 4), 103-110.

Ardia, D. (2008). Financial Risk Management with Bayesian Estimation of GARCH Models: Theory and Applications, Volume 612: Lecture Notes in Economics and Mathematical Systems. Heidelberg, Germany: Springer-Verlag, Berlin. ISBN 978-3-540-78656-6.

Ardia, D., Bluteau, K., Boudt, K. & Catania, L., (2018). Forecasting risk with Markov-switching GARCH models. Elsevier, Issue 34, 733-747.

Ardia, D., Bluteau, K., Boudt, K. & Trottier, D.A. (2016). Markov-Switching GARCH Models in R: The MSGARCH Package.

Arellano, M.A., Collantes, E. & RodrÌguez, G. (2017). Empirical Modeling of Latin American Stock and Forex Markets Returns and Volatility using Markov-Switching GARCH Models, Documento de Trabajo. Documentos de Trabajo / Working Papers 2017-436, Departamento de Economía - Pontificia Universidad Católica del Perú.

Ashley, R. A., & Patterson, D. M. (2006). Evaluating the Effectiveness of State-Switching Time Series Models for U.S. Real Output. Journal of Business and Economic Statistics, 24(3), 266-277. **Augustsson, V. (2014).** Evaluating Switching GARCH Volatility Forecasts during the Recent Financial Crisis, Master's Program in Economics, Lund University, 1-39.

Augustyniak, M. (2014). Maximum likelihood estimation of the Markov-switching GARCH model, Computational Statistics and Data Analysis 76, 61–75.

Babikir, A., Gupta, R., Mwabutwa, C. & Owusu-Sekyere, E. (2012). Structural breaks and GARCH models of stock return volatility: The case of South Africa. Economic Modelling, 29, 2435–2443.

Bala, D.A & Asemota, J.O. (2013). Exchange–Rates Volatility in Nigeria: Application of GARCH Models with Exogenous Break, CBN Journal of Applied Statistics 4, 89-116.

Bauwens, L., De Backer, B. & Dufays, A. (2014). A Bayesian method of change-point estimation with recurrent regimes: Application to GARCH models. Journal of Empirical Finance 29, 207-229.

Bauwens, L., Dufays, A. & Rombouts, J.V.K. (2011). The marginal likelihood for Markov-switching and change-point GARCH models. CORE Discussion Papers 2011013. Universite Catholique de Louvain, Center for Operations Research and Econometrics (CORE).

Bauwens, L., Dufays, A. & Rombouts, J.V.K. (2014). Marginal likelihood for Markov-switching and change-point GARCH models, Journal of Econometrics 178, 508–522.

Bauwens, L., Hafner, C.M. & Rombouts, J.V.K. (2007). Multivariate mixed normal conditional heteroskedasticity. Computational Statistics and Data Analysis 51, 3551–3566.

Bauwens, L., Preminger, A. & Rombouts, J.V.K. (2010). Theory and inference for a Markov switching GARCH model. The Econometrics Journal 13, 218–244.

Beiske, B. (2007). Research Methods: Uses and limitations of questionnaires, interviews and case studies, Munich: GRIN Verlag, 1-24.

Bierens, H.J. (1997). Testing the Unit Root with Drift Hypothesis against Nonlinear Trend Stationarity, with an Application to the U.S. Price Level and Interest Rate, Journal of Econometrics 81, 29 - 64.

Bildirici, M. & Ersin, O. (2014). Modeling Markov Switching ARMA-GARCH Neural Networks Models and an Application to Forecasting Stock Returns. Science World Journal

Billio, M., Getmansky, M. & Pelizzon, L. (2013). Dynamic Risk Exposures in Hedge Funds, Computational Statistics and Data Analysis, 56, 3517-3532.

Binner, J.M., Elger, C.T., Nilsson, B. & Tepper, J.A (2006). Predictable non-linearities in US inflation, Economics Letters 93, 323–328.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity, Journal of Econometrics, 31, 309–328.

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

Brock, W.A., W.D. Dechert, J.A. Scheinkman, & LeBaron, B. (1996). A Test for Independence Based on the Correlation Dimension, Econometric Reviews, 15, 197-235.

Brooks, C. (2008). Introductory Econometrics for Finance, (Second Ed.), Cambridge: Cambridge University Press.

Brooks, C., & Burke, S. P. (2003). Information criteria for GARCH model selection. The European journal of finance, 9(6), 557-580.

Brown, R. L., Durbin, J., & Evans, J. M. (1975). Techniques for Testing the Constancy of Regression Relationships over Time, Journal of the Royal Statistical Society. Series B (Methodological), 37, 2, 149-192.

Brunetti, C., Scotti, C., Mariano, R. S., & Tan, A. H. (2008). Markov switching GARCH models of currency turmoil in Southeast Asia. Emerging Markets Review, 9(2), 104-128.

Bryman, A. & Bell, E. (2011). Business Research Methods (3rd Ed.) Oxford: Oxford University Press.

Bryman, A. (2012). Social research methods (5th Ed.). Oxford: Oxford University Press.

Cai, J. (1994). A Markov model of switching-regime ARCH. Journal of Business and Economics, Applications 37, 482–493.

Cavalheiro, E.A., Vieira, K.M., Ceretta, P.S & Maehler, A.E. (2015). The influence of financial crisis on inefficiency and nonlinearity on Brazilian soybean prices, African Journal of Agricultural Research 10, 3554-3561.

Chai, T. & Draxler, R.R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. Manuscript prepared for Geoscientific Model Development 7, 1247–1250.

Chen, C. W., So, M. K., & Lin, E. M. (2009). Volatility forecasting with double Markov switching GARCH models. Journal of Forecasting, 28(8), 681-697.

Chen, S.W. (2014). Smooth transition, non-linearity and current account sustainability: Evidence from the European countries, Economic Modelling 38, 541-554

Cheteni, P. (2016). Stock market volatility using GARCH models: Evidence from South Africa and China stock markets.

Chuffart, T. (2015). Selection criteria in regime switching conditional volatility models. Econometrics, 3(2), 289-316.

Creswell, J. W. (2013). Steps in conducting a scholarly mixed methods study.

Creswell, J. W., & Tashakkori, A. (2007). Differing perspectives on mixed methods research.

Diebold, F. (1986). Modeling the persistence of conditional variances: a comment, Econometric Reviews, 5, 51–56.

Diebold, F. X. & Mariano, R. S. (1995). Comparing predictive accuracy, Journal of Business and Economic Statistics 13, 253-263.

Dueker, M.J. (1997). Markov switching in GARCH processes and mean-reverting stock-market volatility. Journal of Business and Economic Statistics 15, 26–34.

Engle, R. (1982). Autoregressive conditional Heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987-1007.

Epaphra, M. (2017). Modeling Exchange Rate Volatility: Application of the GARCH and EGARCH Models. Journal of Mathematical Finance, 7, 121-143.

Farhani, S. (2012). Tests of parameters instability: Theoretical study and empirical analysis on two types of models (ARMA model and market model). International Journal of Economics and Financial Issues (IJEFI), 2(3), 246-266.

Fisher, R. A. (1938). The statistical utilization of multiple measurements. Annals of eugenics, 8(4), 376-386.

Flick, U. (2011). Introducing research methodology: A beginner's guide to doing a research project. London: Sage.

Francq, C., Roussignol, M. & Zakoıan, J.M. (2001). Conditional Heteroskedasticity driven by hidden Markov chains. Journal of Time Series Analysis 22, 197–220.

Franses, P. H., & Van Dijk, D. (2000). Non-linear time series models in empirical finance. Cambridge university press.

Fruhwirth-Schnatter, S. (2006). Mixture and Markov-switching Models. Springer, New York, Statistics 12, 309–316.

Geman, S. & Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6, 721-741.

Gephart, R. (1999). Paradigms and research methods. Research Methods Forum 4,1-12.

Giannerini, S. (2012). The quest for nonlinearity in time series. Handbook of Statistics: Time Series 30, 43–63.

Gong, X., & Lin, B. (2017). Forecasting the good and bad uncertainties of crude oil prices using a HAR framework. Energy Economics, 67, 315-327.

Gooijer, J.G.D. & Yuan, A. (2016). Non parametric portmanteau tests for detecting nonlinearities in high dimensions, Communications in Statistics - Theory and Methods 45, 385-399.

Granger, C.W.J. & Newbold, P. (1976). The use of R2 to determine the appropriate transformation of regression variables, Journal of Econometrics, Elsevier 4, 205-210.

Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. Journal of Financial Economics, 42:27–62, 1996.

Gray, S.F. (1996). Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process. Journal of Financial Economics 42, 27–62.

Günay S. (2015). Markov regime switching generalized autoregressive conditional heteroskedastic model and volatility modelling for oil returns Int. J. Energy Econ. Policy 5, 979-985.

Haas, M., Mittnik, S. & Paolella, M. (2004). A new approach to Markov switching GARCH models. Journal of Financial Econometrics 2, 493–530.

Hall, A.D. (1994). Testing for a Unit Root in Time Series with Pretest Data-Based Model

Hamilton, J.D. & Susmel, R. (1994). Autoregressive Conditional Heteroscedasticity and changes in regime. Journal of Econometrics 64, 307–333.

Hamilton, J.D. (1988). Rational expectations econometric analysis of change in regime: an investigation of the term structure of interest rates. Journal of Economic Dynamics and Control 12, 385-423.

Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57, 357-384.

Hamilton, J.D. (1990). Analysis of Time Series Subject to Change in Regime," Journal of Econometrics 45, 39-70.

Hannan, E. J. & Quinn, B. G. (1979). The determination of the order of an Autoregression. Journal of the Royal Statistical Society 41, 190-195.

Hannan, E. J. (1980). The estimation of the order of an ARMA process. American Statistician 8, 1071-1081.

Harris, J. W. & Stocker, H. (1998). Maximum Likelihood Method. 21.10.4 in Handbook of Mathematics and Computational Science. New York: Springer-Verlag, 824.

Harris, R., & Sollis, R. (2003). Applied time series modelling and forecasting. Wiley.

Harvey, A. C., Ruiz, E. & Shephard, N. (1994). Multivariate stochastic variance models, Review of Economic Studies 61, 247-64.

Harvey, A.C. & Collier, P. (1977). Testing for functional misspecification in regression analysis, Journal of Econometrics 6, 103–119.

Hastings, W.K. (1970). Monte Carlo Sampling Methods Using Markov Chains and their Applications. Biometrika 57, 97-109.

Hendry, D. F. (1986). Econometric modelling with cointegrated variables: an overview. Oxford bulletin of economics and statistics, 48(3), 201-212.

Henneke, J.S, Rachev, S., Fabzozzi, F.J. & Nikolov, M. (2011). MCMC-based estimation of Markov switching ARMA-GARCH models, Applied Economics 43, 259-271.

Hillebrand, E. (2005). Neglecting parameter changes in GARCH models. Journal of Econometrics, 129(1-2), 121-138.

Humala, A., & Rodríguez, G. (2013). Some stylized facts of return in the foreign exchange and stock markets in Peru. Studies in Economics and Finance.

Hurvich, C.M. & Tsai, C.L. (1989). Regression and time series model selection in small samples. Biometrika, 76,297-307.

Hwang, A.S. (1996). Positivist and Constructivist Persuasions in Instructional Development. Instructional Science: An International Journal of the Learning Sciences 24, 343.

Iqbal, F. (2016). Forecasting Volatility and Value-at-Risk of Pakistan Stock Market with Markov Regime-Switching GARCH Models. European Online Journal of Natural and Social Sciences 5, 172–189.

Ismail, M.T. & Isa, Z. (2006). Modelling Exchange Rates Using Regime Switching Models. Sains Malaysiana 35, 55- 62.

Jacquier, E., Polson, N. G. & Rossi, P. E. (1994). Bayesian Analysis of Stochastic Volatility Models, Journal of Business and Economic Statistics 12, 371-389.

Jandoo, N., & Gonpot, P. N. (2018). Evaluating Exchange Rate Value at Risks Models for Fourteen African Currencies. Acta Universitatis Danubius. Œconomica, 14(6).

Jarque, C. & Bera, A. (1980). Efficient tests for normality homoscedasticity and serial independence of regression residuals, Econometric Letters 6, 255–259.

Kapetanios, G., Shin, Y., & Snell, A. (2003). Testing for a unit root in the nonlinear STAR framework. Journal of econometrics, 112(2), 359-379.

Keenan, D. (1985). A Tukey nonadditivity-type test for time series nonlinearity. Biometrika 72, 39-44.

Kim C.J. & Nelson, C.R. (1999). State-Space Models with Regime Switching: Classical and Gibbs Sampling Approaches with Applications, MIT Press, Cambridge, MA.

Kirch, C., & Kamgaing, T.J. (2011). An online approach to detecting changes in nonlinear autoregressive models.

Klaassen, F. (2002). Improving GARCH volatility forecasts with regime switching GARCH. Empirical Econometrics. 27, 363–394.

Korner-Nievergelt, F., Roth, T., Von Felten, S., Guélat, J., Almasi, B., & Korner-Nievergelt, P. (2015). Bayesian data analysis in ecology using linear models with R, BUGS, and Stan. Academic Press.

Kothari, C. R. (2004). Research methodology: methods and techniques. New Delhi: New Age International.

Kuan, C.M. (2002). Lecture on the Markov Switching Model, Institute Of Economics Academia Sinica.

Kutu, A. A., & Ngalawa, H. (2017). Monetary policy and industrial output in the BRICS countries: A markov-switching model. Folia Oeconomica Stetinensia, 17(2), 35-55.

Lama, A., Jha, G.K., Paul, R.K. & Gurung, B. (2015). Modeling and Forecasting of Price Volatility: An Application of GARCH and EGARCH Models, Agricultural Economics Research Review 28, 73-82.

Lamoureux, C. G., & Lastrapes, W. D. (1990). Heteroskedasticity in stock return data: Volume versus GARCH effects. The journal of finance, 45(1), 221-229.

Liew, V.K.S., Baharumshah, A.Z. & Chong, T.T.L. (2003). Are Asian Real Exchange Rates Stationary? Economics Letters, 83,313-316.

Luukkonen, R., Saikkonen, P. & Teräsvirta T. (1988). Testing linearity against Smooth Transition Autoregressive Models. Biometrika, 75, 491–499.

Marcucci, J. (2005). Forecasting Stock Market Volatility with Regime-Switching GARCH models. Studies in Nonlinear Dynamics and Econometrics 9, 1-53

Matei, M. (2009). Assessing Volatility Forecasting Models: Why GARCH Models take the Lead. Romanian Journal of Economic Forecasting, 4.

May, C. & Farrell, G. (2017). Modelling exchange rate volatility dynamics: Empirical evidence from South Africa, ERSA working paper 705.

May, T. (2011). Social research: Issues, methods and research. London: McGraw-Hill International.

Mazviona, B.W. (2012). Volatility Forecasting using Double-Markov Switching GARCH Models under Skewed Student-t Distribution, Department of Statistical Sciences A mini-thesis submitted at the University of Cape Town in partial fulfilment of the requirements for the degree of Masters of Philosophy in Mathematical Finance.

Melino, A. & Turnbull S.M. (1990). Pricing foreign currency options with stochastic volatility, Journal of Econometrics 45, 239-265.

Mikosch, T. & C. Starica, C. (2004). Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects, Review of Economics and Statistics 86, 378–390.

Miron, D. & Tudor, C. (2010). Asymmetric conditional volatility models: Empirical estimation and comparison of forecasting accuracy. Romanian Journal of Economic Forecasting 3, 74–92.

Montgomery, D.C., Gardiner, J.S. & Johnson, L.A. (1990). Forecasting and Time Series Analysis, Journal of the Operational Research Society 29, 371-375.

Naeini, M.N. & Fatahi, S. (2012). Comparing Regime-Switching GARCH Models and GARCH Models in Developing Countries (Case study of IRAN), Analysis Financial 119, 60-68

Newman, I. (1998). Qualitative-quantitative research methodology: Exploring the interactive continuum. Carbondale: Southern Illinois University Press.

Nicholls, D. F. & Quinn, B. G. (1982). Random Coefficient Autoregressive Model: An introduction, New York: Springer-Verlag.

Panagiotidis, T. (2005). Market capitalization and efficiency. Does it matter? Evidence from the Athens Stock Exchange. Applied Financial Economics, 15(10), 707-713.

Pattichis, C. (2012). Exchange rate effects on trade in services, Journal of Economic Studies 39, 697-708.

Ramsey, J.B. (1969). Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis, Journal of the Royal Statistical Society 31, 350-371.

Restoy, F. & Rockinger, G.M. (1994). On Stock Market Returns and Returns on Investment, Journal of Finance 49, 543–556.

Rossi, R. J. (2018). Mathematical statistics: an introduction to likelihood based inference. John Wiley & Sons.

Sajjad, R., Coakley, J., & Nankervis, J. C. (2008). Markov-Switching GARCH Modelling of Value-at-Risk. Studies in Nonlinear Dynamics & Econometrics 12, 1–31.

Sattayathama, P., Sopipana, N. & Premanode, B. (2012). Forecasting Volatility and Price of the SET50 Index using the Markov Regime Switching, Procedia Economics and Finance 2, 265 – 274.

Saunders, M., Lewis, P. & Thornhill, A. (2009). Research methods for business students, 5th edition. Harlow, Pearson Education.

Saunders, M., Lewis, P. & Thornhill, A. (2012). Research methods for business students, (6th Ed.) New York: Pearson Education, Inc. 19.

Saunders, M., Lewis, P., & Thornhill, A. (2007). Research Methods for Business Students, (6th Ed.) London: Pearson.

Saunders, M., Lewis, P., & Thornhill, A. (2012). Research methods for business students (6. utg.). Harlow: Pearson.

Schmidt C & Kohlmann T. (2008). When to use the odds ratio or the relative risk? International Journal of Public Health. 53:165–7.

Schumacker, R. E., & Lomax, R. G. (2004). A beginner's guide to structural equation modelling. psychology press.

Schwartz, G. (1978). Estimating the dimension of a model. Annals of Statistics 6, 461-464.

Schwert, G. W. (1990). Stock returns and real activity: A century of evidence. The Journal of Finance, 45(4), 1237-1257.

Schwert, W. (1989). Test for Unit Roots: A Monte Carlo Investigation. Journal of Business and Economic Statistics, 7, 147-159.

Selection, Journal of Business and Economic Statistics, 12,461-70.

Shiferaw, Y. (2018). The Bayesian MS-GARCH model and Value-at-Risk in South African agricultural commodity price markets (No. 2058-2018-5299).

Silverman, D. (2013). Doing Qualitative Research: A practical handbook. London: Sage.

Snieder R. & Larner, K. (2009). The Art of Being a Scientist: A Guide for Graduate Students and their Mentors, Cambridge: Cambridge University Press.

Sopipan, N., Sattayatham, P., & Premanode, B. (2012). Forecasting volatility of gold price using markov regime switching and trading strategy. Journal of Mathematical Finance, 2(01), 121.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. Journal of the royal statistical society: Series b (statistical methodology), 64(4), 583-639.

St-Hilaire, W.A. (2015). Qualit09]y Decision Control, Marketing Applied Strategic Agility and Energy Projects Performance, American Journal of Management, 15, 102-114.

Taylor, S. J. (1994). Modeling stochastic volatility: a review and comparative study, Mathematical Finance 4, 183-04.

Taylor, S.J. (1986). Modelling Financial Time Series. John Wiley and Sons, Ltd., Chichester.

Thadewald, T., & Büning, H. (2007). Jarque–Bera test and its competitors for testing normality–a power comparison. Journal of applied statistics, 34(1), 87-105.

Tsay, R. (1986). Non-linearity tests for time series. Biometrika 73, 461– 466.

Tsay, R. S. (1987). Conditional heteroskedastic time series models, Journal of American Statistical Association 82, 590-604.

Tsay, R. S. (2005). Analysis of financial time series (Vol. 543). John wiley & sons.

Vasanthi, S., Subha, M. V. & Nambi, S. T. (2011). An empirical study index trend prediction using Markov chain analysis, JBFSIR.1, 72-90

Wei, W., Jiang, J., Liang H., Gao, L., Liang B, Huang J., Zang, N., Liao, Y., Yu, J., Jingzhen Lai, J., Qin, J., Jinming Su, J., Ye, L & Chen, H. (2016). Application of a Combined Model with Autoregressive Integrated Moving Average (ARIMA) and Generalized Regression Neural Network (GRNN) in Forecasting Hepatitis Incidence in Heng County, China. PLoS ONE 11, 1-13.

Wiles, R., Crow, G. & Pain, H. (2011). Innovation in qualitative research methods: a narrative review. Qualitative Research 11, 587-604.

Winker, P., Gilli, M., & Jeleskovic, V. (2007). An objective function for simulation based inference on exchange rate data. Journal of Economic Interaction and Coordination, 2(2), 125-145.

Xekalaki, E., & Degiannakis, S. (2010). ARCH models for financial applications. John Wiley & Sons.

Zhang, Y. (2002). Testing for nonlinearities in time series with application to exchange rates. Doctoral dissertation, Iowa State University.

Zhang, Y.J., Yao, T. & He, L.Y. (2015). Forecasting crude oil market volatility: can the Regime-Switching GARCH model beat the single-regime GARCH models? International Review of Economics and Finance, 1- 31.

APPENDICES

Appendix A

Nonlinear Test

ARCH TEST

Under the null hypothesis of linearity, the residuals of a properly specified AR(p) model should be independent. Denote the autocorrelations of the residuals by $\rho_1, \rho_1, \ldots\ldots, \rho_m$, where $m = n/4$ (n=sample size), then the independence of the residuals, ε_t, can be tested based on the hypotheses (Engle, 1982):

$H_0 = \rho_1 = \cdots = \rho_1 = \rho_m$ (Residuals are independent)

vs.

$H_1: \rho_1 \neq 0$ for at least one j (Residuals are not independent)

The test statistic is the Q-statistic of squared residuals given by

$$Q(m) = n(n+2)\sum_{k=1}^{m} \frac{\rho_k^2}{n-k} \sim \chi_\alpha^2(m-p)$$

At the α level, the null hypothesis of linearity is rejected in favour of the alternative hypothesis if

$Q(m) > \chi_\alpha^2(m-p)$ or $\text{prob}[Q(m)] < \alpha$.

This same test is particularly useful in detecting conditional heteroskedasticity inY_t. A closely related test to the Q-statistic test is the Lagrange test of Engle (1982) for autoregressive conditional heteroskedasticity (ARCH) test based on the linear regression:

$$\hat{\varepsilon}_t^2 = \eta_0 + \sum_{i=0}^{m} \eta_i \varepsilon_{t-1}^2 + \upsilon_t$$

where $\eta_0, \eta_1, \eta_2, \ldots \ldots .., \eta_m$ are parameters and υ_t is independent and identically distributed random variable with mean 0 and variance σ_s^2. Testing for heteroskedasticity involves testing the hypotheses:

$H_0: \eta_0, \eta_1, \eta_2, \ldots \ldots .., \eta_m$ (Homoskedasticity)

vs.

$H_1: \eta_0 \neq 0$ for at least one j (Heteroskedasticity).

The test statistic is the usual F-statistic:

$$F^* = \frac{R^2/m}{(1-R^2)/(n-m-1)} \sim F_\alpha(m, n-2m-1)$$

At the α level, the null hypothesis of linearity is rejected in favour of the alternative hypothesis if

$F^* > F_\alpha(m, n-2m-1)$ or $pro(F^*) < \alpha$

Asymptotically, $F^* \sim \chi^2_\alpha(m)$.

Appendix B

Probability Densities

Peters (2001) argued that in financial time series data, the error or residual terms in GARCH models exhibit excess kurtosis and skewness. This section provides commonly used distributions.

Normal distribution

The probability density function (PDF) of the standard Normal distribution is given by:

$$f_N(\eta) \equiv \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\eta^2}, \qquad \eta \epsilon \mathbb{R}$$

This distribution is identified with the label "norm".

Student–t distribution

The PDF of the standardized Student–t distribution is given by:

$$f_N(\eta; v) \equiv \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{(v-2)\pi\Gamma\left(\frac{v}{2}\right)}} \left(1 + \frac{\eta^2}{(v-2)}\right)^{-\frac{v+1}{2}}, \qquad \eta \epsilon \mathbb{R},$$

where Γ(·) is the Gamma function. The constraint ν > 2 is imposed to ensure that the second order moment exists. The kurtosis of this distribution is higher for lower ν[5]. This distribution is identified with the label "std".

Generalized Error Distribution (GED)

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

$$f_{GED}(\eta; v) \equiv \frac{v e^{-\frac{1}{2}\left|\frac{\eta}{\lambda}\right|_v}}{\lambda 2^{\left(1+\frac{1}{v}\right)} \Gamma\left(\frac{1}{v}\right)}, \ \lambda \equiv \left(\frac{\Gamma(1/v)}{4^{1/v}\Gamma(3/v)}\right)^{1/2}, \eta \epsilon \mathbb{R}$$

where ν > 0 is the shape parameter. This distribution is identified with the label "ged".

APPENDIX C

Optimal Lag Selection for AR Specification

Variable	Lag	SBC
ABSA(t)	1	13.48991
	2	13.48883
	3	13.49927
	4	13.51109
CAPB(t)	1	14.08583
	2	14.09767
	3	14.1038
	4	14.11012
FIRB(t)	1	10.03122
	2	10.04261
	3	10.04663
	4	10.04781
NEDB(t)	1	13.65478
	2	13.64814
	3	13.65515
	4	13.66344
STDB(t)	1	12.67503
	2	12.67884
	3	12.68246
	4	12.69517

APPENDIX D

Optimal Lag Length Selection for KSS Unit Root Tests

D1: KSS Auxiliary Equations for BRAZIL

Dependent Variable d_BRAZIL

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
BRAZIL3_1	1	-0.0042	0.0022	-1.87	0.0629
d_BRAZIL_1	1	0.5447	0.0629	8.66	0.0000
d_BRAZIL_2	1	-0.3012	0.0717	-4.20	0.0000
d_BRAZIL_3	1	0.0808	0.0741	1.09	0.2768
d_BRAZIL_4	1	0.0198	0.0744	0.27	0.7904
d_BRAZIL_5	1	0.0166	0.0743	0.22	0.8236
d_BRAZIL_6	1	-0.0467	0.0743	-0.63	0.5296
d_BRAZIL_7	1	0.0335	0.0725	0.46	0.6446
d_BRAZIL_8	1	-0.0533	0.0635	-0.84	0.4018

Dependent Variable d_DETBRAZIL

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETBRAZIL3_1	1	-0.0041	0.0022	-1.87	0.0626
d_DETBRAZIL_1	1	0.5439	0.0628	8.67	0.0000
d_DETBRAZIL_2	1	-0.2997	0.0715	-4.19	0.0000
d_DETBRAZIL_3	1	0.0798	0.0739	1.08	0.2813
d_DETBRAZIL_4	1	0.0190	0.0741	0.26	0.7983
d_DETBRAZIL_5	1	0.0121	0.0740	0.16	0.8704
d_DETBRAZIL_6	1	-0.0303	0.0716	-0.42	0.6727
d_DETBRAZIL_7	1	0.0028	0.0631	0.04	0.9646

Dependent Variable d_DETBRAZIL

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETBRAZIL3_1	1	-0.0040	0.0022	-1.85	0.0655
d_DETBRAZIL_1	1	0.5444	0.0625	8.71	0.0000
d_DETBRAZIL_2	1	-0.2997	0.0712	-4.21	0.0000
d_DETBRAZIL_3	1	0.0790	0.0736	1.07	0.2842
d_DETBRAZIL_4	1	0.0202	0.0737	0.27	0.7840
d_DETBRAZIL_5	1	0.0103	0.0714	0.14	0.8853
d_DETBRAZIL_6	1	-0.0275	0.0624	-0.44	0.6601

Dependent Variable d_DETBRAZIL

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETBRAZIL3_1	1	-0.0039	0.0022	-1.80	0.0727
d_DETBRAZIL_1	1	0.5447	0.0622	8.74	0.0000
d_DETBRAZIL_2	1	-0.3016	0.0709	-4.25	0.0000
d_DETBRAZIL_3	1	0.0779	0.0732	1.06	0.2880
d_DETBRAZIL_4	1	0.0276	0.0710	0.39	0.6981
d_DETBRAZIL_5	1	-0.0036	0.062227	-0.06	0.9531

D1 continued: KSS Auxiliary Equations for BRAZIL

Dependent Variable d_DETBRAZIL

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETBRAZIL3_1	1	-0.0038	0.0021	-1.79	0.0741
d_DETBRAZIL_1	1	0.5441	0.0620	8.78	0.0000
d_DETBRAZIL_2	1	-0.3014	0.0705	-4.28	0.0000
d_DETBRAZIL_3	1	0.0785	0.0705	1.11	0.2665
d_DETBRAZIL_4	1	0.0263	0.0619	0.42	0.6717

Dependent Variable d_DETBRAZIL

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETBRAZIL3_1	1	-0.0040	0.0021	-1.89	0.0596
d_DETBRAZIL_1	1	0.5460	0.0615	8.87	0.0000
d_DETBRAZIL_2	1	-0.3087	0.0676	-4.56	0.0000
d_DETBRAZIL_3	1	0.0916	0.0614	1.49	0.1371

Dependent Variable d_DETBRAZIL

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETBRAZIL3_1	1	-0.0041	0.0021	-1.97	0.0499
d_DETBRAZIL_1	1	0.5218	0.0595	8.78	0.0000
d_DETBRAZIL_2	1	-0.2593	0.0593	-4.37	0.0000

Dependent Variable d_DETBRAZIL

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETBRAZIL3_1	1	-0.003792	0.002148	-1.765156	0.0787
d_DETBRAZIL_1	1	0.412902	0.055689	7.414440	0.0000

D2: KSS Auxiliary Equations for RUSSIA

Dependent Variable d_DETRUSSIA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETRUSSIA3_1	1	-0.0255	0.0067	-3.79	0.0002
d_DETRUSSIA_1	1	0.2560	0.0618	4.14	0.0000
d_DETRUSSIA_2	1	-0.0741	0.0638	-1.16	0.2465
d_DETRUSSIA_3	1	0.0898	0.0639	1.41	0.1611
d_DETRUSSIA_4	1	0.0418	0.0641	0.65	0.5153
d_DETRUSSIA_5	1	-0.0194	0.0641	-0.30	0.7629
d_DETRUSSIA_6	1	0.0011	0.0638	0.02	0.9863
d_DETRUSSIA_7	1	0.0269	0.0637	0.42	0.6737
d_DETRUSSIA_8	1	-0.0795	0.0612	-1.30	0.1952

Dependent Variable d_DETRUSSIA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETRUSSIA3_1	1	-0.0248	0.0065	-3.80	0.0002
d_DETRUSSIA_1	1	0.2561	0.0617	4.15	0.0000
d_DETRUSSIA_2	1	-0.0748	0.0637	-1.17	0.2413
d_DETRUSSIA_3	1	0.0916	0.0638	1.44	0.1525
d_DETRUSSIA_4	1	0.0379	0.0640	0.59	0.5549
d_DETRUSSIA_5	1	-0.0275	0.0638	-0.43	0.6669
d_DETRUSSIA_6	1	0.0063	0.0637	0.10	0.9218
d_DETRUSSIA_7	1	0.0052	0.0611	0.09	0.9319

Dependent Variable d_DETRUSSIA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETRUSSIA3_1	1	-0.0229	0.0064	-3.60	0.0004
d_DETRUSSIA_1	1	0.2610	0.0615	4.24	0.0000
d_DETRUSSIA_2	1	-0.0741	0.0636	-1.16	0.2453
d_DETRUSSIA_3	1	0.0914	0.0638	1.43	0.1529
d_DETRUSSIA_4	1	0.0391	0.0637	0.62	0.5390
d_DETRUSSIA_5	1	-0.0284	0.0636	-0.45	0.6553
d_DETRUSSIA_6	1	0.0128	0.0611	0.21	0.8345

Dependent Variable d_DETRUSSIA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETRUSSIA3_1	1	-0.0212	0.0062	-3.42	0.0007
d_DETRUSSIA_1	1	0.2633	0.0614	4.29	0.0000
d_DETRUSSIA_2	1	-0.0745	0.0635	-1.17	0.2422
d_DETRUSSIA_3	1	0.0941	0.0634	1.48	0.1388
d_DETRUSSIA_4	1	0.0380	0.0635	0.60	0.5503
d_DETRUSSIA_5	1	-0.0200	0.0611	-0.33	0.7434

D2 continued: KSS Auxiliary Equations for RUSSIA

Dependent Variable d_DETRUSSIA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETRUSSIA3_1	1	-0.0202	0.0061	-3.32	0.0010
d_DETRUSSIA_1	1	0.2625	0.0613	4.29	0.0000
d_DETRUSSIA_2	1	-0.0759	0.0631	-1.20	0.2296
d_DETRUSSIA_3	1	0.0955	0.0631	1.51	0.1314
d_DETRUSSIA_4	1	0.0355	0.0609	0.58	0.5602

Dependent Variable d_DETRUSSIA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETRUSSIA3_1	1	-0.0199	0.0059	-3.35	0.0009
d_DETRUSSIA_1	1	0.2671	0.0607	4.40	0.0000
d_DETRUSSIA_2	1	-0.0785	0.0627	-1.25	0.2121
d_DETRUSSIA_3	1	0.1070	0.0603	1.77	0.0772

Dependent Variable d_DETRUSSIA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETRUSSIA3_1	1	-0.0199	0.0058	-3.42	0.0007
d_DETRUSSIA_1	1	0.2630	0.0608	4.33	0.0000
d_DETRUSSIA_2	1	-0.0460	0.0604	-0.76	0.4473

Dependent Variable d_DETRUSSIA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETRUSSIA3_1	1	-0.0191	0.0057	-3.36	0.0009
d_DETRUSSIA_1	1	0.2526	0.0583	4.34	0.0000

D3: KSS Auxiliary Equations for INDIA

Dependent Variable d_DETINDIA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETINDIA3_1	1	-0.0019	0.0009	-2.04	0.0422
d_DETINDIA_1	1	0.2201	0.0629	3.50	0.0005
d_DETINDIA_2	1	-0.0473	0.0646	-0.73	0.4647
d_DETINDIA_3	1	-0.0192	0.0638	-0.30	0.7642
d_DETINDIA_4	1	-0.0507	0.0638	-0.79	0.4275
d_DETINDIA_5	1	-0.0562	0.0638	-0.88	0.3789
d_DETINDIA_6	1	0.1292	0.0641	2.02	0.0449
d_DETINDIA_7	1	-0.0055	0.0646	-0.09	0.9317
d_DETINDIA_8	1	0.0182	0.0628	0.29	0.7728

Dependent Variable d_DETINDIA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETINDIA3_1	1	-0.0021	0.0009	-2.26	0.0245
d_DETINDIA_1	1	0.2151	0.0628	3.43	0.0007
d_DETINDIA_2	1	-0.0526	0.0637	-0.83	0.4100
d_DETINDIA_3	1	-0.0209	0.0638	-0.33	0.7432
d_DETINDIA_4	1	-0.0493	0.0637	-0.77	0.4393
d_DETINDIA_5	1	-0.0611	0.0637	-0.96	0.3385
d_DETINDIA_6	1	0.1265	0.0640	1.97	0.0494
d_DETINDIA_7	1	-0.0121	0.0625	-0.19	0.8470

Dependent Variable d_DETINDIA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETINDIA3_1	1	-0.0020	0.0009	-2.17	0.0308
d_DETINDIA_1	1	0.2192	0.0617	3.55	0.0005
d_DETINDIA_2	1	-0.0516	0.0635	-0.81	0.4170
d_DETINDIA_3	1	-0.0218	0.0635	-0.34	0.7313
d_DETINDIA_4	1	-0.0463	0.0634	-0.73	0.4657
d_DETINDIA_5	1	-0.0595	0.0635	-0.94	0.3495
d_DETINDIA_6	1	0.1309	0.0617	2.12	0.0347

Dependent Variable d_DETINDIA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETINDIA3_1	1	-0.0021	0.0009	-2.26	0.0247
d_DETINDIA_1	1	0.2202	0.0623	3.53	0.0005
d_DETINDIA_2	1	-0.0623	0.0640	-0.97	0.3316
d_DETINDIA_3	1	-0.0202	0.0641	-0.31	0.7531
d_DETINDIA_4	1	-0.0517	0.0640	-0.81	0.4199
d_DETINDIA_5	1	-0.0168	0.0621	-0.27	0.7868

D3 continued: KSS Auxiliary Equations for INDIA

Dependent Variable d_DETINDIA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETINDIA3_1	1	-0.0020	0.0009	-2.21	0.0278
d_DETINDIA_1	1	0.2202	0.0620	3.55	0.0005
d_DETINDIA_2	1	-0.0604	0.0637	-0.95	0.3438
d_DETINDIA_3	1	-0.0189	0.0638	-0.30	0.7673
d_DETINDIA_4	1	-0.0517	0.0617	-0.84	0.4025

Dependent Variable d_DETINDIA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETINDIA3_1	1	-0.0019	0.0009	-2.20	0.0290
d_DETINDIA_1	1	0.2210	0.0618	3.57	0.0004
d_DETINDIA_2	1	-0.0581	0.0635	-0.92	0.3609
d_DETINDIA_3	1	-0.0341	0.0615	-0.55	0.5797

Dependent Variable d_DETINDIA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETINDIA3_1	1	-0.0018	0.0009	-2.10	0.0365
d_DETINDIA_1	1	0.2240	0.0615	3.64	0.0003
d_DETINDIA_2	1	-0.0611	0.0612	-0.10	0.3192

Dependent Variable d_DETINDIA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DETINDIA3_1	1	-0.0017	0.0009	-1.99	0.0475
d_DETINDIA_1	1	0.2138	0.0595	3.60	0.0004

D4: KSS Auxiliary Equations for CHINA

Dependent Variable d_CHINA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
CHINA 3_1	1	0.0007	0.0007	1.08	0.2800
d_CHINA _1	1	0.4416	0.0631	7.00	0.0000
d_CHINA _2	1	-0.1450	0.0683	-2.12	0.0347
d_CHINA _3	1	0.0756	0.0688	1.10	0.2733
d_CHINA _4	1	0.0933	0.0687	1.36	0.1757
d_CHINA _5	1	-0.0913	0.0686	-1.33	0.1840
d_CHINA _6	1	0.0512	0.0686	0.75	0.4563
d_CHINA _7	1	-0.1506	0.0680	-2.22	0.0277
d_CHINA _8	1	0.0765	0.0625	1.22	0.2227

Dependent Variable d_CHINA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
CHINA 3_1	1	0.0008	0.0007	1.21	0.2264
d_CHINA _1	1	0.4334	0.0625	6.93	0.0000
d_CHINA _2	1	-0.1410	0.0681	-2.07	0.0396
d_CHINA _3	1	0.0686	0.0685	1.00	0.3177
d_CHINA _4	1	0.0999	0.0682	1.46	0.1443
d_CHINA _5	1	-0.0857	0.0683	-1.25	0.2108
d_CHINA _6	1	0.0393	0.0678	0.58	0.5624
d_CHINA _7	1	-0.1159	0.0621	-1.87	0.0629

Dependent Variable d_CHINA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
CHINA 3_1	1	0.0007	0.0007	1.09	0.2779
d_CHINA _1	1	0.4374	0.0627	6.98	0.0000
d_CHINA _2	1	-0.1350	0.0683	-1.98	0.0491
d_CHINA _3	1	0.0533	0.0684	0.78	0.4362
d_CHINA _4	1	0.0979	0.0683	1.43	0.1526
d_CHINA _5	1	-0.0730	0.0680	-1.07	0.2837
d_CHINA _6	1	-0.0050	0.0623	-0.08	0.9362

Dependent Variable d_CHINA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
CHINA 3_1	1	0.0008	0.0007	1.13	0.2585
d_CHINA _1	1	0.4371	0.0624	7.01	0.0000
d_CHINA _2	1	-0.1408	0.0677	-2.08	0.0385
d_CHINA _3	1	0.0577	0.0680	0.85	0.3973
d_CHINA _4	1	0.0955	0.0675	1.41	0.1584
d_CHINA _5	1	-0.0690	0.0620	-1.11	0.2668

D4 continued: KSS Auxiliary Equations for BRAZIL

Dependent Variable d_CHINA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DET CHINA 3_1	1	0.0007	0.0007	1.05	0.2964
d_CHINA _1	1	0.4334	0.0619	7.00	0.0000
d_CHINA _2	1	-0.1454	0.0673	-2.16	0.0317
d_CHINA _3	1	0.0678	0.0673	1.01	0.3144
d_CHINA _4	1	0.0653	0.0617	1.06	0.2913

Dependent Variable d_CHINA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
CHINA 3_1	1	0.0007	0.0007	1.08	0.2805
d_CHINA _1	1	0.4358	0.0618	7.05	0.0000
d_CHINA _2	1	-0.1518	0.0668	-2.27	0.0238
d_CHINA _3	1	0.0878	0.0616	1.42	0.1556

Dependent Variable d_CHINA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
CHINA 3_1	1	0.0008	0.0007	1.24	0.2160
d_CHINA _1	1	0.4251	0.0614	6.92	0.0000
d_CHINA _2	1	-0.1114	0.0612	-1.82	0.0702

Dependent Variable d_CHINA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
CHINA 3_1	1	0.0007	0.0007	1.07	0.2843
d_CHINA _1	1	0.3809	0.0567	6.71	0.0000

D5: KSS Auxiliary Equations for South Africa (RSA)

Dependent Variable d_DET RSA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DET RSA 3_1	1	-0.0043	0.0020	-2.14	0.0337
d_DET RSA _1	1	0.2736	0.0626	4.37	0.0000
d_DET RSA _2	1	-0.1301	0.0649	-2.01	0.0460
d_DET RSA _3	1	0.1171	0.0648	1.81	0.0721
d_DET RSA _4	1	-0.0415	0.0653	-0.64	0.5253
d_DET RSA _5	1	0.0039	0.0652	0.06	0.9518
d_DET RSA _6	1	-0.1435	0.0648	-2.21	0.0277
d_DET RSA _7	1	0.0577	0.0650	0.89	0.3760
d_DET RSA _8	1	0.1153	0.0612	1.88	0.0607

Dependent Variable d_DET RSA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DET RSA 3_1	1	-0.0047	0.0020	-2.32	0.0210
d_DET RSA _1	1	0.2849	0.0625	4.56	0.0000
d_DET RSA _2	1	-0.1488	0.0643	-2.31	0.0215
d_DET RSA _3	1	0.1196	0.0650	1.84	0.0669
d_DET RSA _4	1	-0.0457	0.0653	-0.70	0.4850
d_DET RSA _5	1	0.0166	0.0650	0.26	0.7984
d_DET RSA _6	1	-0.1598	0.0644	-2.48	0.0137
d_DET RSA _7	1	0.0904	0.0610	1.48	0.1395

Dependent Variable d_DET RSA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DET RSA 3_1	1	-0.0048	0.0020	-2.43	0.0157
d_DET RSA _1	1	0.2735	0.0620	4.41	0.0000
d_DET RSA _2	1	-0.1482	0.0644	-2.30	0.0221
d_DET RSA _3	1	0.1153	0.0650	1.78	0.0771
d_DET RSA _4	1	-0.0342	0.0650	-0.53	0.5993
d_DET RSA _5	1	0.0030	0.0644	0.05	0.9633
d_DET RSA _6	1	-0.1271	0.0605	-2.10	0.0367

Dependent Variable d_DET RSA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DET RSA 3_1	1	-0.0044	0.0020	-2.24	0.0259
d_DET RSA _1	1	0.2775	0.0622	4.46	0.0000
d_DET RSA _2	1	-0.1480	0.0646	-2.29	0.0228
d_DET RSA _3	1	0.1038	0.0649	1.60	0.1111
d_DET RSA _4	1	-0.0163	0.0647	-0.25	0.8011
d_DET RSA _5	1	-0.0299	0.0607	-0.49	0.6232

D5 continued: KSS Auxiliary Equations for BRAZIL

Dependent Variable d_DET RSA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DET RSA 3_1	1	-0.0043	0.0020	-2.22	0.0272
d_DET RSA _1	1	0.2778	0.0620	4.48	0.0000
d_DET RSA _2	1	-0.1508	0.0640	-2.36	0.0193
d_DET RSA _3	1	0.1081	0.0641	1.69	0.0930
d_DET RSA _4	1	-0.0236	0.0604	-0.39	0.6962

Dependent Variable d_DET RSA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DET RSA 3_1	1	-0.0045	0.0019	-2.32	0.0209
d_DET RSA _1	1	0.2739	0.0615	4.45	0.0000
d_DET RSA _2	1	-0.1470	0.0633	-2.32	0.0210
d_DET RSA _3	1	0.0913	0.0601	1.52	0.1299

Dependent Variable d_DET RSA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DET RSA 3_1	1	-0.0046	0.0019	-2.40	0.0172
d_DET RSA _1	1	0.2632	0.0612	4.30	0.0000
d_DET RSA _2	1	-0.1135	0.0597	-1.90	0.0585

Dependent Variable d_DET RSA

NOTE: No intercept term is used. R-squares are redefined.

Parameter Estimates

Variable	DF	Estimate	Standard Error	t-Value	Pr > |t|
DET RSA 3_1	1	-0.0043	0.0019	-2.29	0.0235
d_DET RSA _1	1	0.2355	0.0581	4.06	0.0001

APPENDIX F

Choice of KSS Auxiliary Model Type - Zero-Mean, De-Meaned or De-Trended

Dependent Variable BRAZIL

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	-0.7338	0.0250	-11.39	<.0000

==

Dependent Variable RUSSIA

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	-0.7606	0.0050	-10.19	<.0000

==

Dependent Variable INDIA

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	-0.8377	0.0030	-10.87	<.0000

==

Dependent Variable CHINA

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	-0.6870	0.0110	-10.08	<.0000

==

Dependent Variable SOUTH AFRICA

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx. Pr > |t|
Intercept	1	-0.8460	0.0110	-11.85	<.0000

Appendix G

R program Codes

Nonlinearity Tests Codes

```
Exchange_rate <- read_excel("C:/Users/18027709/Desktop/Exchange rate.xlsx")

View(Exchange_rate)

dtsmp1 <- ts(Exchange_rate[,5], start = c(1997), end = c(2017), frequency = 12)

Nonlinearity Test (Exchange_rate [, 2], verbose = TRUE)

Keenan.test(sqrt(exchange rate))

r.cref <- diff(log(exchange rate))*100

McLeod.Li.test(y=r.cref)

Tsay.test(sqrt(exchange rate))
```

MRS-GARCH Models CODES

```
Exchange_rate <- Exchange_rate <- read_excel ("Exchange rate.xlsx")

####### Time series formating   #######

Xaba = Exchange_rate [, 3]

Dintle =ts (na.omit (Xaba), frequency=12, start=c (1997))

Kopano = log (Dintle)

############# Fitting Markov-Switching Generalised Autoregressive Conditional
Heteroscedasticity Model############

spec <- CreateSpec (variance.spec = list(model = c("sGARCH")),

                distribution.spec = list (distribution = c ("norm")),

                switch.spec = list (do.mix = FALSE, K = 2))

################### ML Maximum likeliHood #####################

fit = FitML(spec = spec,data = Kopano)

pred <- predict(fit, nahead = 5, do.return.draws = TRUE)
```

```
summary (fit)

#getting residual for ML method #ref link:
https://github.com/keblu/MSGARCH/issues/43#issuecomment-372157823

conditional_volatility_ML<- Volatility (fit)

residuals_ML <- c (Kopano)/conditional_volatility_ML

#getting MSE and MEA using residuals extracted earlier

MSE_ML = mean ((residuals_ML) ^2)

MAE_ML = mean (abs (residuals_ML))

MSE_ML

MAE_ML

########### MCMC Bayesian estimation method ##################

fit1 <- FitMCMC (spec = spec, data =kat, ctr = list(nburn = 500L, nmcmc = 500L))

summary (fit1)

#getting residual for MCMC method #ref link:
https://github.com/keblu/MSGARCH/issues/43#issuecomment-372157823

conditional_volatility_MCMC<- Volatility (fit1)

residuals_MCMC <- as.vector (Kopano)/conditional_volatility_MCMC

#getting MSE and MEA using residuals extracted earlier

MSE_MCMC = mean ((residuals_MCMC) ^2)

MAE_MCMC = mean (abs (residuals_MCMC))

#print all metrics

MSE_MCMC

MSE_ML

MAE_MCMC

MAE_ML

########################## DM test########################################

et=c(pred_ML)

et1=c (perd_MCMC)
```

```
dm.test (et, et2, alternative = c ("greater"), h=1, power = 2)

## INSAMPLE ##

# extracting actual realized volatility and utilize that to calculate MSE/MAE between
conditional volatility of MCMC and ML models

# refer study https://www.math.kth.se/matstat/seminarier/reports/M-exjobb11/110617.pdf:
step 6#initialize realized_volatility

Lehakwe= Kopano

realized_volatility=c(seq(1:length(Kopano)))

for (i in seq(1:length(Lehakwe)-1)){

  realized_volatility [i+1] = sqrt (255/30)*sd(c(Lehakwe[i],Lehakwe[i+1]))

}

#conditional volatility on prediction from MCMC model

mse_MCMC_Vol= mean ((conditional_volatility_MCMC [2:240]-
realized_volatility[2:240])^2)

mae_MCMC_vol = mean (abs (conditional_volatility_MCMC [2:240]-
realized_volatility[2:240]))

mse_MCMC_Vol

mae_MCMC_vol


######conditional volatility on prediction from ML model############################

mse_ML_Vol= mean ((conditional_volatility_ML [2:214] - realized_volatility [2:240]) ^2)

mae_ML_vol = mean (abs (conditional_volatility_ML [2:214] - realized_volatility [2:240]))

mse_ML_Vol

mae_ML_vol


########## realized vs predicted volatility, OUT of sample ###################

#for this we have to keep few records as out of sample while training

#Considering last 10 records Lehakwe[205:214] for realized volatility
```

####### Time series formating   #######

ja_oos = Lehakwe [1:240]

################ Fitting Markov-switching generalised autoregressive conditional heteroscedasticity model############

spec <- CreateSpec(variance.spec = list(model = c("sGARCH")),

distribution.spec = list (distribution = c ("std")),

switch.spec = list (do.mix = FALSE, K = 2))

################### ML Maximum likelihood #####################

fit_oos_ML = FitML (spec = spec, data = ja_oos)

pred <- predict(fit_oos_ML, nahead =5)

mse_oos_ML= mean ((c (pred$vol)-realized_volatility [230:240]) ^2)

mae_oos_ML= mean (abs(c (pred$vol)-realized_volatility [230:240]))

mse_oos_ML

mae_oos_ML

OUT-OF-SAMPLE FORECASTING VOLATILITY

########## realized vs predicted volatility, out-of-sample ###################

#for this we have to keep few records as out of sample while training

#considering last 10 records Lehakwe [:240] for realized volatility

####### Time series formating   #######

ja_oos = Lehakwe [1:240]

################ Fitting Markov-switching generalised autoregressive conditional heteroscedasticity model############

spec <- CreateSpec(variance.spec = list(model = c("sGARCH")),

distribution.spec = list (distribution = c ("ged")),

switch.spec = list (do.mix = FALSE, K = 2))

```
################### MCMC Bayesian estimation. #####################

fit_oos_ML = FitML (spec = spec, data =ja_oos)

pred <- predict(fit_oos_ML, nahead = 1)

mse_oos_ML = mean ((c (pred$vol)-realized_volatility )^2)

mae_oos_ML = mean (abs(c (pred$vol)-realized_volatility))

mse_oos_ML

mae_oos_ML
```

Diebold-Mariano test for predictive accuracy

```
dm.test(e1, e2, alternative = c("two.sided", "less", "greater"), h = 1,

 power = 2)
```

Arguments

e1	Forecast errors from method 1.
e2	Forecast errors from method 2.
alternative	a character string specifying the alternative hypothesis, must be one of "two.sided" (default), "greater" or "less". You can specify just the initial letter.
h	The forecast horizon used in calculating e1 and e2.
power	The power used in the loss function. Usually 1 or 2.