Stochastic mean-reverting volatility forecasting with Augmented ARMA-GARCH models

by

Albert Antwi (11622305)

A thesis Submitted in

fulfilment of the requirements for the Degree of **Doctor of Philosophy (PhD)** in Statistics at

School of Mathematical and Natural Sciences

Department of Statistics

University of Venda

South Africa

Promoter: Dr. Kwabena A. Kyei

Co-Promoters: Prof. Ryan Gill

Dr. Emmanuel N. Gyamfi

Dr. Ndogmo Jean-Claude

April 12, 2021.

Declaration

I **Albert Antwi** hereby declare that the thesis for the Doctor of Philosophy: Statistics degree at the University of Venda, hereby submitted by me, has not previously been submitted for a degree at this or any other university, and that it is my own work in design and execution and that all reference material contained therein has been duly acknowledged.

Signed (student)………………………………………Date: 12/04/2021………………………………

Signed (Promoter) ………………………………………Date. 12/04/2021……………………………

Signed (Co- promoter) ………………………………Date. 12/04/2021……………………………

Signed (Co- promoter) ………………………………Date. 13/04/2021……………………………

Signed (Co- promoter) ………………………………Date. 13/04/2021……………………………

Abstract

The unconditional forecasting structures of non-time varying GARCH models impose monotonic mean reversion paths on medium to long-run volatility forecasts when breaks and changes in the unconditional volatility are ignored or not appropriately accounted for. This leads to over-estimated or under-estimated volatilities and forecasts paths, which are unrepresentative of the underlying asset. In this thesis, an attempt is made to induce stochastic time-variations in the unconditional volatility forecasts of GARCH models by assuming that their unconditional volatility processes are driven by the levels of market uncertainties and since proxies used in literature may be inadequate, we seek to identify alternative proxies.

The identification requires an ARMA relationship assumption between exchange rate returns, thus we empirically test the assumption to ascertain its plausibility. Based on the plausibility of the assumption, we link exogenous returns to endogenous volatility and based on this link, we identify appropriate levels of returns as the proxies. Break variables are then constructed from the proxies and together with the proxies; they are passed to the variance equations of non-time varying GARCH models to augment the models. The augmented models are then used to forecast volatility and VaR of some selected currency pairs to assess their predictive and forecasting powers. Before the models are used to forecast volatility and VaR, a study of hypothetical mutual dependencies between the volatilities and the exogenous covariates (proxies and their break variables) is carried out to investigate the levels of shared mutual entropies among the variables. Daily prices of fourteen rand-denominated currency pairs spanning July 7, 2011, to July 3, 2016 are used for all the empirical studies.

The empirical evidence suggests that the ARMA assumption is plausible and that the exogenous returns have potentials to predict volatility. Furthermore, all the estimated parameters for the exogenous returns are positive and consistent with their directions of co-movements. Path analysis of the impacts of the returns confirmed that currency pairs are not in isolation on the forex market and that shocks of the same magnitude from the same-origin transmitted along different paths on the market may have different impacts. Based on the evidence to support the plausibility of the ARMA relationship assumption, the expectation of the square of the relationship indicates that it is analogous to an exogenous GARCH and that exogenous returns are related to volatility. This is

due to the fact that the exogenous returns are exposed to similar uncertainties in the market where the volatilities evolve; their absolute values are identified as alternative proxies for the levels of uncertainties surrounding the exchange rate market.

Evidence from the estimated hypothetical mutual entropies revealed substantial percentages of exchange entropies among the variables. Furthermore, the evidence from the modelling and forecasting of exchange rate volatility using the augmented models indicates that the forecasts revert along stochastic paths towards their long-run variances. In addition, the estimated volatilities are less persistent with significantly improved accuracies. The models also yielded relatively improved forecasts or insignificant loss of forecast accuracies with improved explanatory forecasting powers. Finally, the results from the VaR estimations and forecasting suggest that the models lead to lower failure rates and overall relative superior forecast accuracies when used to forecast 1% VaR , but not generally superior in the case of 5% VaR forecasts, although, it leads to lower asymmetric losses. The VaR models also produced lower mean daily capital requirements and the majority of them avoided the regulatory penalty zones imposed by the Basel II Accord while few of them slipped into the yellow zone, but with relatively less associated penalties.

Based on the results from the studies, we recommend the use of our proposed method to forecast volatility and VaR for exchange rates, to financial institutions, investors and other practitioners for risk management and policy decision-making. Specifically, in estimating and forecasting VaR, the proposed augmented models are recommended as complementary or supplementary models. It is believed that using forecasts from such models or by aggregating them with forecasts from internal models may lead to fewer bank failures, optimal exposure to market risks and banks may meet regulatory capital requirements without been sanctioned by regulatory bodies. Furthermore, since the augmentation leads to less persistent volatility, it may be used in addition to existing methods to model volatility and VaR of highly persistent returns.

Key words: Market uncertainty, mean-reversion, mutual entropy, volatility, GARCH, VaR.

Table of Contents

List of Tables

List of Figures

Acronyms

ADF : Augmented Dickey-Fuller

ADmean: Absolute mean

ADmax: Absolute maximum

ARCH : Auto regressive conditional heteroscedasticity

ARMA : Auto regressive moving average

ARMAX: Exogenous auto regressive moving average

BCMI: Bias corrected mutual information

CBOE: Chicago Board Options Exchange

EGARCH: Exponential GARCH

EMH : Efficiency Market Hypothesis

FIGARCH: Fractionally Integrated GARCH

GARCH : General Auto regressive Conditional Heteroscedasticity

GJRGARCH : Glosten, Jaganathan and Runkle GARCH

IGARCH: Integrated GARCH

LM : Lagrange multiplier

MAE : Mean Absolute Error

MAPE: Mean absolute prediction error

MCS : Model Confidence Set

MZR : Mincer-Zarnowitz Regression

OLS : Ordinary Least Square

QML : Quasi-maximum likelihood

RMSE : Root Mean Square Error

RMSPE: Mean square prediction error

SIXVX : The Swedish Model-Free Implied Volatility Index

TGARCH : Threshold GARCH

VaR : Value-at-Risk

VIX : The CBOE Volatility Index

Dedication

This thesis is dedicated to the Almighty God, the maker of heaven and earth, to my late parents Miss Rose Boatemaa and Mr Joseph Amponsah-Antwi, to my wife, Hlamalani Antwi and to my children, Symon-Miguel Brenden, Primrose Boatemaa, and Jennica Debrah Masindi.

Acknowledgements

I express my profound gratitude to Almighty God, who has granted me the wisdom, intellect, and strength to carry out this research. My gratitude also goes to my promoter Dr Kwabena A. Kyei and to my co-promoters, Professor Ryan Gill, Dr. Emmanuel N. Gyamfi, and Dr. Ndogmo Jean-Claude for their support, guidance, and constructive criticisms, which made the finalisation of this thesis a reality. My sincere gratitude also goes to Dr Phyllis K. Kaburise for proofreading this thesis within constraints of time. Finally, I thank the University of Venda and Sol Plaatje University for providing me with financial assistance to complete this study.

Summary

From the thesis has been produced the following original papers publication in internal accredited journals) and conference presentation. They are expected to contribute to literature and provide an alternative or supplementary tool for computing volatility and risk for effective decision-making. A summary of the specific contributions can be found at the Main Contributions section of the study, in chapter nine.

Publications

- **Antwi A**., Kyei, K.A & Gill, R. (2020), Forecasting long term exchange rate volatility with stochastic mean reverting unconditional volatility, *Journal of Statistics and Management Systems*, vol 0, Issue 0 pp. 1-23. Available at: https://doi.org/10.1080/09720510.2020.1816690.
- **Antwi A**., Kyei K. A. & Gill R.S. (2020), The Use of Mutual Information to Improve Value-at-Risk Forecasts for Exchange Rates, in *IEEE Access*, vol. 8: 179881-179900, 2020. Doi: 10.1109/ACCESS.2020.3027631. Available at: https://ieeexplore.ieee.org/document/9208673/metrics#metrics.

Conference Attendance

- **Antwi A**. (2017), Exchange rate dependencies under the joint ARMAX-TGARCH System: Empirical evidence from the South African Forex Market, Abstract proceedings of the 59th Annual Conference of the South African Statistical Association for 2017 held at Bloemfontein. Available: https://sastat.org.za/sites/default/files/inline-files/SASA%20Nov%202017%20Abstracts_1.pdf.

The general autoregressive conditional heteroskedastic models have become one of the most commonly used models to forecast volatility and VaR, hence, this thesis is a contribution to the literature on volatility and VaR forecasting. The thesis is organised under nine chapters. The first three chapters provide the general framework of the study. Chapter one introduces the broad sense of the thesis, chapter two addresses the basic theoretical literature underpinning the thesis, while chapter three delves into the methodological framework for the various tests to be used in the study. In chapter four, the empirical properties, and the results of the prerequisite data validation tests are detailed. In chapter five, we provide empirical evidence of the autoregressive relationship between

exchange rate returns. In chapter six, we identify alternative proxies for the levels of uncertainties surrounding the exchange rate market and use them to construct break variables. The proxies and the break variables are then used to augment GARCH models in chapter seven. Applications of the augmented models to data from the rand market, to study their forecasting powers are also presented in this chapter. In chapter eight, the augmented models from chapter seven are used to model and forecast VaR. Finally, the general observations from the empirical results, as well as the main themes discussed in the thesis with recommendations and direction of future studies are all outlined in chapter nine. All codes used in the data analysis are written in R© with the exceptions of some of the descriptive statistics which were carried out in EVIEWS©.

CHAPTER 1

General introduction

Chapter Summary

This chapter introduces the framework of the thesis. A general overview of volatility touching on the applications in finance and their effects on the economy as a whole are discussed. The background, justification, problem statement, objectives, expected outcomes and the scope of the study are also deliberated upon.

1.2 Background to the study

A class of models, which have been successfully used to model volatility, is the general autoregressive conditional heteroscedasticity (GARCH) models. GARCH models are theoretically appealing and most suitable for assets with thousands or large amounts of observations. When the cost component is ignored, it produces quality volatility forecasts when compared to any other alternative models (Matei, 2009). Due to these and other appealing features, GARCH models are more popular among researched areas in finance and economics. Extensions of GARCH models began after Engle introduced his ARCH model in 1982. Each extension attempts to address a different economic problem or the same economic problem, differently.

For example, asymmetric models such as GJRGARCH (Glosten, Jaganathan & Runkle, 1993), EGARCH (Nelson, 1991) and TGARCH (Zakoian, 1994) attempt to account for *leverage* or *asymmetric effect*s of news on volatility which were first observed by Black (1976). The ARCH model of Engle (1982) and its generalized version by Bollerslev (1986) failed to capture volatility asymmetry and, although, there is evidence of the superiority of non-asymmetric GARCH models over asymmetric models; such superiority is peculiar to series with salient response to shocks (Ho-Jin, 2009 and Hansen & Lunde, 2005). When it is more evident that asset returns have asymmetric response to news, asymmetric models have been found to outperform standard ARCH and GARCH models (Khan *et al.,* 2019; Ching & Siok, 2013 and Hansen & Lunde, 2005).

Findings from competing models converge to poor out-of-sampling performances despite the modifications or extensions of GARCH models (Chen, Dolado & Gonzalo, 2014). In an attempt to address this problem, Andersen & Bollerslev (1998) suggest the use of frequently sampled ex-post square returns. Amado & Teräsvirta (2014), Brownlees & Gallo (2010), Baillie & Morana (2009) and Bailie, Bollerslev & Mikkelsen (1996) among other researchers recommend the decomposition of the volatility process into conditional and unconditional components, while allowing the unconditional component to assume a deterministic time-varying functional component. Amado & Teräsvirta (2014) suggest modelling time-variation in the unconditional volatility with levels of uncertainty in the markets. The use of certain asset characteristics and traditional statistics such as cointegration (Kosapattarapim, 2013) and ARMA (McCrae et al., 2002) has also been recommended. These suggestions have been shown to lead to improvement in the explanatory power and accuracy of long-range volatility forecasts.

Other researchers, such as Hillebrand (2005) and Lamoureux & Lastrapes (1990) argue that financial markets are prone to extreme events such as financial crashes, flash crashes, and market disruptions that lead to large disturbances, thus, affecting returns; these result in sharp breaks in the unconditional volatility of the returns. They further argue that GARCH models for such returns, which fail to account for breaks, may yield an upward bias in the degree of volatility persistence in the estimated model. Such models suffer from systematic bias in the estimated forecast, which is either systematically higher or lower than the realized volatility of the underlying asset.

Other studies have shown that levels of uncertainty in the market affect long-run volatility (Amado & Teräsvirta, 2014), thus, neglecting the levels of market uncertainty in GARCH models may lead to structural misspecification and subsequent poor long-range forecasts. Long-range forecasts from such models are monotonic and may converge to a constant unconditional volatility forecast, which is contrary to the stochastic time-varying nature of the volatility of assets. In the wake of these findings, the objectives of the thesis are geared towards the improvements of long-range volatility forecasts by augmenting non-time-varying GARCH models with alternative proxies and breaks' variables to account for their presence in unconditional volatilities.

1.3 Problem statement

The structure of GARCH models imposes a restriction on the mean reversion property of long-range volatility forecasts, thus, the unconditional volatility forecasts evolve monotonically over long horizons and since the conditional forecasts converge towards the long-run variance, they evolve along monotonic trajectories, which are not representative of the underlying stochastic trajectories. Attempts have been made to address this problem by either accounting for breaks in the unconditional variance using dummy variables or modelling time-variations in the unconditional variance via proxies for levels of uncertainties surrounding the asset market. With these attempts the problem however persists, hence, we attempt to address it by using alternative proxies and their break variables to simultaneously account for breaks and changes in the unconditional volatilities.

1.1 Introduction

Volatility is an important concept in finance and it has several applications in financial risk management. Volatility estimates are used as inputs in option pricing, portfolio optimization, VaR computations and hedging of assets. In addition to the several applications of volatility, exchange rate volatility, particularly, affects government's monetary policies (Osei-Asibey, 2010) and cost of imports and exports, leading to relative changes in the prices of goods, as well as domestic and foreign investments (Faure, 2013).

The appreciation of a local currency makes the operational cost of firms to increase (Joseph, 2002), which results in an increase in prices of goods and services of firms. Goods and services of firms' exports abroad become expensive on the international market, consequently, export volumes of the firms on the international market decline. The firms lose their competitiveness on the domestic stock market and their attractiveness on the domestic stock market decline, resulting in a decline in their stock prices (Mlambo, Maredza & Sibanda, 2010).

The converse of this argument holds; since volatility affects exchange rate prices, it directly or indirectly affects the behaviour of investors, decisions by policy makers and the economy as a whole, thus modelling and forecasting of exchange rate volatility equips investors, portfolio managers, traders, and policy-makers to make better and informed decisions.

Widespread adoption of the independent floating exchange rate regime among countries, after the collapse of the Brent Wood agreement in 1976 led to the influence of country-specific market performances on exchange rates. Exchange rates of such countries are characterized by large unexpected volatile variations from their fundamentals and this is responsible for the difficulties in predicting exchange rates (Killian & Taylor, 2003 and Flood, 1981). Owing to the numerous applications of exchange rate volatility and the difficulties of its estimation, studies geared towards the improvements of volatility forecasting continue unabated.

1.4 Justification of the study

There is a need for accurate volatility estimates and forecasts since volatility affects the behaviour of investors, decisions by policy-makers and the economy as a whole. They equip investors, portfolio managers, traders, and policy-makers to make better and informed decisions. Attempts in literature to address the seemingly poor out-of-sampling performance of competing GARCH models have improved volatility forecasts, however, long-range forecasts do not mimic the stochastic underlying structure of the assets, which may affect the forecasts' accuracies. It is within this confine that the study has become necessary.

1.5 Objectives

The study aims at forecasting exchange rate volatility by augmenting GARCH models with alternative market uncertainty proxies and break variables in an attempt to relax the restrictive mean reverting property of out-of-sample forecast, so that the forecasts assume the stochastic path of the underlying volatility without significantly compromising the accuracy of the forecasts. This will be achieved through the following objectives:

- Re-examining the relationship between exchange rate returns via autoregressive framework.

- Identifying alternative proxies for exchange rate market uncertainties by theoretically linking exogenous returns to endogenous volatility.

- Estimating the levels of hypothetical mutual information between volatilities and the identified proxies as well as break variables, to assess their predictive potentials in modelling volatility.

- Augment univariate GARCH models with the proxies and the constructed break variables, in an attempt to relax the restrictive mean reverting property of out-of-sample forecasts.
- Use the augmented GARCH models to forecast volatility and VaR for single asset portfolios.

1.6 Expected contributions from the study

The study is expected to contribute directly to literature on exchange rate relationships, volatility forecasting, and VaR forecasting, for emerging markets. Furthermore, it intends to provide further evidence, in support or against, theoretical and empirical studies, which advocate that structural breaks have potentially important implications for estimated GARCH models and VaR forecasts. Finally, the study intends to provide an alternative or complementary approach for modelling and forecasting volatility and VaR to assist financial institutions and practitioners in making better-informed risk-management decisions.

1.7 Scope

The proposed augmentations of GARCH models in this study do not involve any functional modifications of internal structures, but requires the incorporation of exogenous processes via the external input options of the models. GARCH models used in this study are not exhaustive but are limited to the most commonly-used univariate models; subsequent VaR estimations are based on single-asset portfolios. The data used in the study is limited to 14 selected daily exchange rates from the South African inter-bank forex market. This is due to data access restrictions during the time of sourcing.

1.8 Conclusion

The chapter introduced the applications of volatility in financial risk management and the impacts of exchange rate volatility on investment, imports, exports and the economy as a whole. A comprehensive background to the study was provided. Not forgetting the popular saying of Plato, "*necessity is the mother of all inventions*', due diligence was also given to the motivation behind the study. The problem statement, objectives, expected contributions, as well as the scope of the study received due attention. Literature exposition on GARCH modelling and forecasting revealed that, although, several attempts have been made to improve volatility forecast, there remains the problem of unrepresentative volatility forecast paths, thus the study attempts to solve this problem.

CHAPTER 2

Review of properties of returns and GARCH models

Chapter Summary

Following the introduction of the ARCH model by Robert Engel in 1982, several variations have cropped-up in literature over the past decades. While some of the variations address the same economic problem in different ways, others address different problems altogether. In this chapter, we review some of the GARCH models commonly used in literature. Definition of volatility and its measurement used in this thesis are presented. We delve into the general statistical and economic regularities that sparked the interest in the development of the ARCH/GARCH type of models before elaborating on their general specifications and some of their basic properties. The maximum likelihood approach used in estimating ARCH/GARCH parameters is also discussed.

2.1 Introduction

Speculative asset returns exhibit several empirical characteristics, which are of much interest to researchers and practitioners alike. These interests have incresed over the past years, and have culminated into the developments of several models in an attempt to explain and forecast volatility of asset returns. Returns of speculative assets evolve in a random manner. This behaviour suggests the absence of autocorrelation in returns and makes it very difficult to predict the direction of asset returns (Martin, Hurn & Harris, 2012). Unlike returns, squared returns are autocorrelated, thus making it possible to predict volatility using heteroskedastic models. It is worth noting that, the sample autocorrelation of absolute returns is usually larger than that of the squared returns (Taylor, 1986). Serial correlation of volatility leads to the tendency of small movements in returns being followed by another small return in the next period (periods of relative calmness), while large movements in returns are followed by another large returns in the next period (periods of turbulence). This property of volatility of returns, first reported by Mandelbrot (1963), is termed *clustering*. Later studies, including those of Ding, Granger & Engle (1993) and Cont (2001) have confirmed this observation. Clustering of volatility implies that the autocorrelation of squared returns is positive.

The conditional distribution of returns is sometimes approximately normal. This is because there is a high probability of drawing subsequent small value of current return in the next period if the previous return is small and another large value of current return in the next period if the previous return is large. This is because the conditional variance of small values of previous returns is drawn from a relatively compact distribution with zero mean and approximately constant variance (Martin, Hurn & Harris, 2012), however, unlike the conditional distribution of returns, the unconditional distribution is non-normal. The unconditional distribution of relatively low-volatilities is relatively compact with high peaks, whereas it is relatively more dispersed with low peaks for high-volatilities. Averaging across the conditional distributions yields a non-normal unconditional distribution characterized by leptokurtosis (heavy-tails and sharp peaks). Empirical evidence of unconditional leptokurtosis can be found in Mandelbrot (1963), Fama (1965), Ding, Granger & Engle (1993), and Cont (2001). Figure 2.1 displays a plot of the theoretical distribution of normal and the empirical s*tudent's* t-distribution for USD/ZAR returns. The plot clearly indicates the heavy-tails and excess kurtosis of the USD/ZAR returns, which are common among financial returns.

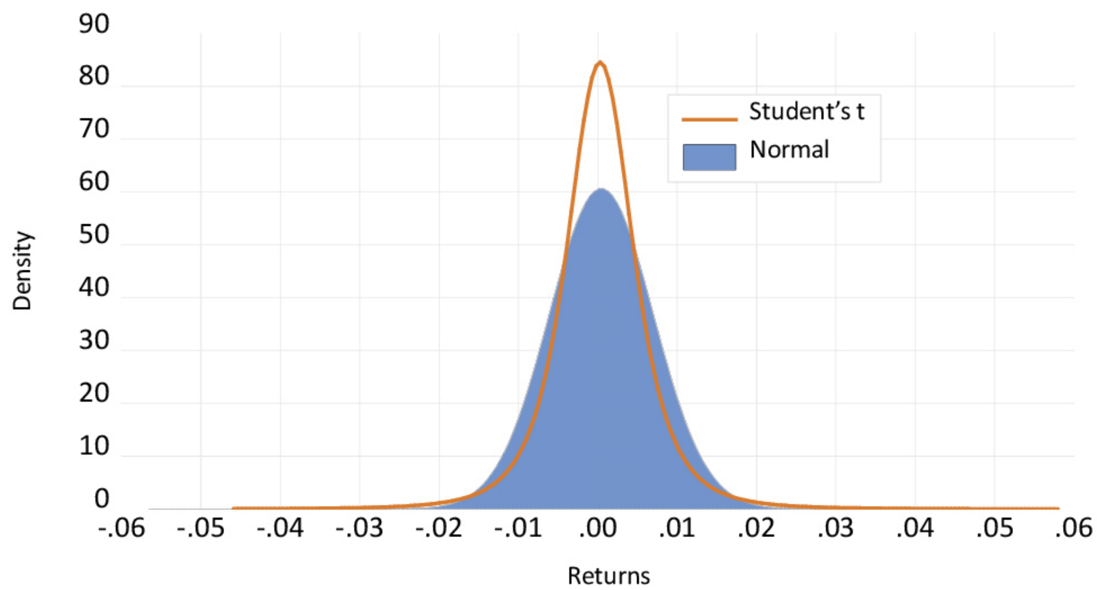

Figure 2. 1: Theoretical normal distribution versus empirical distribution of USD log-returns. The Figure was drawn using EVIEWS package.

The unconditional distribution of returns is negatively skewed, suggesting that extreme positive returns are less frequent than extreme negative returns (Nelson, 1991 and Taylor, 2005), therefore, volatility is more affected by negative returns than positive returns. This behaviour of returns is termed 'asymmetry' or 'leverage effects'. The empirical characteristics of asset returns discussed above, among others, motivated the development of the autoregressive conditional heteroscedasticity (ARCH) models and its variants, which are reviewed in the subsequent sections.

2.2 Definition and measurement of volatility

Volatility is defined as the spread of all likely outcomes of an uncertain variable (Ser-Huang, 2005). Specific to asset returns, it is an index of unexpected variability of asset returns in a period (Bucci, 2017). The spread of asset returns (volatility) is vital in predicting price movements and as an input for risk-measure computations. Statistically, volatility is computed as the sample standard deviation of returns. Let P_t denote current closing price of a currency pair and P_{t-1} the previous closing day price. If $P_t > 0$, then the current continuously compounded log return on the asset r_t is calculated as:

$$r_t = \ln(P_t) - \ln(P_{t-1}). \tag{2.1}$$

For τ days a set of log returns $r_1; r_2; r_3; \ldots; r_\tau$ are observed. If $\hat{\mu}_r$ is the average log return over the τ days, the estimate for the spread $(\hat{\sigma})$ can be statistically computed as:

$$\hat{\sigma} = \sqrt{\tfrac{1}{\tau-1}\sum_{t=1}^{\tau}(r_t - \hat{\mu}_r)^2}\,. \tag{2.2}$$

Volatility at any given time (σ_t) is a latent variable, therefore squared log return r_t^2 is used as a proxy (Ser-Huang, 2005).

2.3 Univariate ARCH and GARCH models

Suppose that the log returns of an asset can be expressed by the regression equation below:

$$r_t = \mu_t + a_t, \tag{2.3}$$

where μ_t is an estimate of the returns and a_t is the shock or mean-corrected return at time t. In modelling the conditional variance, h_t of the shock, Engle (1982) employed the basic idea that the mean-corrected return is serially uncorrelated, but dependent and that the dependency can be described by a simple quadratic function of its lagged values. Under the assumption that $z_t \sim N(0,1)$ (the standardized Student-t distribution or generalized error distribution are also commonly used) Engle (1982) used the following representations to model the conditional variance of a_t:

$$\begin{aligned} a_t &= \sqrt{h_t}.z_t, \quad t \in \mathbb{N}, \\ h_t &= \alpha_0 + \sum_{i=1}^{q} \alpha_i a_{t-i}^2 + \sum_{j=1}^{m} \zeta_j \upsilon_{jt}, \end{aligned} \tag{2.4}$$

where a_{t-i}^2 is lagged innovation, and $\alpha_0, \alpha_i, \zeta_j$ are parameters. The ARCH order is denoted q with possible m external regressors υ_j, which are passed pre-lagged. To ensure that $h_t > 0$, the parameters are conditioned, such that $\alpha_0 > 0$, $\alpha_i \geq 0$ and $\zeta_j \geq 0$.

2.3.1 Standard GARCH model

The ARCH model is plagued with several weaknesses. The model is simple, but it often requires many parameters to adequately describe the volatility process of asset returns. Asset's volatility responds differently to positive and negative shocks, however, the model assumes the same effects of positive and negative shocks on volatility because they depend on the square of the previous shocks. To obtain a finite fourth moment, the parameters of the lagged squared shocks need to be restricted to an interval. In higher-order models, the restriction is more complicated. In addition, the model does not provide any new insight for understanding the source of variations of a financial time series, but only a mechanical way to describe the behaviour of the conditional variance (Osei-Asibey, 2010). The model gives no indication of what causes such behaviour to occur. ARCH models are likely to over-predict the volatility because they respond slowly to large isolated shocks to the return series. Bollerslev (1986) generalized form of ARCH (GARCH) was proposed to address some of these drawbacks. Bollerslev proposed the addition of lagged volatilities to the ARCH model, thus the model assumes the following specification:

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i a_{t-i}^2 + \sum_{i=j}^{p} \beta_j h_{t-j} + \sum_{j=1}^{m} \zeta_j \upsilon_{jt}\,, \tag{2.5}$$

where h_{t-j} is the lagged volatility (GARCH term) with corresponding order p while β_j is a parameter. To ensure that $h_t > 0$ the parameters are conditioned such that $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$ and $\zeta_j \geq 0$. The lagged volatility term is called the "GARCH" term with the corresponding order of p. If variance targeting is used, equation (2.5) is reparametrized by replacing α_0 with:

$$\hat{h}\left(1-\hat{P}\right) - \sum_{j=1}^{m} \zeta_j \upsilon_{jt}\,, \text{ where } \hat{P} = \sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j\,.$$

The unconditional variance denoted by $\hat{h}$ is defined as $\hat{h} = \alpha_0\left(1-\hat{P}\right)^{-1}$.

2.3.2 Exponential GARCH model

The GARCH model is theoretically appealing, but its simple structure imposes an important limitation on the model. The model ignores the sign effect of unanticipated returns on the conditional volatility. It assumes that only the magnitude of the unanticipated returns affects the conditional volatility. To address this drawback, Nelson (1991) proposed the exponential GARCH. Ensuring that the $h_t > 0$, instead of modelling h_t as a linear combination of positive weighted random variables as was the case of the GARCH model, Nelson modelled h_t by making $\ln(h_t)$ linear in some functions of time and lagged random variables, such that, for an appropriate function $g(z_{t-1})$:

$$\ln(h_t) = \alpha_0 + \sum_{i=1}^{\infty} \alpha_i g(z_{t-1}) + \sum_{j=1}^{m} \zeta_j \upsilon_{jt}, \qquad \alpha_1 \equiv 1, \tag{2.6}$$

where α_0 and $\{\alpha_i\}_{t\in[1,\infty)}$ are real, non-stochastic and scalar sequences. To ensure an asymmetric response to good and bad news, Nelson sought to represent $g(z_t)$ as a function of both magnitude and direction of z_t. The choice of $g(z_t)$ that gives σ_t^2 a well-behaved moment is the linear combination of z_t and $|z_t|$, thus, the weighted innovation below was chosen:

$$g(z_t) \equiv \theta + \gamma\left(|z_t| - E|z_t|\right) = \begin{cases} (\theta+\gamma)z_t - \gamma E|z_t| & \text{if } 0 < z_t < \infty \\ (\theta-\gamma)z_t - \gamma E|z_t| & \text{if } -\infty < z_t \leq 0 \end{cases}, \tag{2.7}$$

where θ and γ are real constants and $z_t = \dfrac{a_{t-i}}{\sqrt{h_{t-i}}}$. Both $z_t \sim N(0,1)$. The infinite-moving average representation in (2.6) is complex and the parameters are not parsimonious, thus, to obtain a more simple and parsimonious model, it is re-parameterized in an ARMA form as shown below:

$$\ln(h_t) = \alpha_0 + \sum_{j=1}^{m} \zeta_j \upsilon_{jt} + \frac{\left(1+\alpha_1+\alpha_2 L^2 + \ldots + \alpha_s L^s\right)}{1-\left(\beta_1 L + \beta_2 L^2 + \ldots + \beta_m L^m\right)} g(z_{t-1}), \tag{2.8}$$

where L is a lag operator, such that $Lg(z_t) = g(z_{t-1})$. Simplification of equation (2.8) yields[1]

$$\ln(h_t) = \alpha_0 + \sum_{i=1}^{p} \gamma_i \left(\frac{a_{t-i}}{\sqrt{h_{t-i}}}\right) + \sum_{i=1}^{p} \alpha_i \left(\left|\frac{a_{t-i}}{\sqrt{h_{t-i}}}\right| - E\left|\frac{a_{t-i}}{\sqrt{h_{t-i}}}\right|\right) + \sum_{j=1}^{q} \beta_j \ln(h_{t-j}) + \sum_{j=1}^{m} \zeta_j \upsilon_{jt}. \tag{2.9}$$

If variance targeting is used, α_0 is replaced with:

$$\ln(\hat{h}_t)\left(1-\sum_{j=1}^{q}\beta_j\right) - \sum_{j=1}^{m}\zeta_j \upsilon_{jt}.$$

The necessary and sufficient condition for the existence of the second moment is $\sum_{j=1}^{q}\beta_j < 1$ and

[1] For detailed simplification, see Appendix A.

the unconditional volatility is defined as:

$$\ln\left(\hat{h}\right) = \alpha_0 \left(1 - \sum_{j=1}^{q} \beta_j\right)^{-1}.$$

2.3.3 GJR and the threshold GARCH the models

The GJR-GARCH model of Glosten, Jaganathan & Runkle (1993) is a different approach to modelling asymmetry in volatility. Unlike the exponential GARCH, it models positive and negative shocks to the conditional variance asymmetrically via the use of an indicator function. The model takes the form:

$$h_t = \alpha_0 + \sum_{i=1}^{q} (\alpha_i - \gamma_i I_{t-i}) a_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} + \sum_{j=1}^{m} \zeta_j \upsilon_{jt}, \tag{2.10}$$

where I_{t-i} is an indicator function, which takes the value of 1 if $a \leq 0$ and 0 otherwise. To ensure that the variance $h_t > 0$, the parameters are constrained such that $\alpha_0 > 0,\ \alpha_i \geq 0,\ \beta_j \geq 0,\ \zeta_j \geq 0$ and $\alpha_i + \gamma_i \geq 0$. The γ_i parameter provides information about asymmetric effects. If $\gamma_i = 0$, there is no volatility asymmetry, if $\gamma_i > 0$ negative shocks will increase volatility more than positive shocks of the same magnitude and if $\gamma_i < 0$, positive shocks increase volatility more than negative shock. The persistence parameter is given by:

$$\hat{P} = \sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j + \sum_{i=1}^{q} \gamma_i \kappa,$$

where κ is the expected value of the standardized residuals z_t below zero (effectively the probability of being below zero) defined by:

$$\kappa = \mathrm{E}\left[I_{t-j} z_{t-j}^2\right] = \int_{-\infty}^{0} f(z, 0, 1, \ldots) dz.$$

In the case of symmetric distributions, the value of κ equal to 0.5. The constant parameter α_0 is replaced with $\hat{h}_t\left(1-\hat{P}\right)-\sum_{j=1}^{m}\zeta_j \upsilon_{jt}$ if variance targeting is used. The unconditional variance is:

$$\hat{h}=\frac{\alpha_0}{1-\sum_{i=1}^{q}\alpha_i-\sum_{j=1}^{p}\beta_j-\sum_{i=1}^{q}\gamma_i\kappa}.$$

The TGARCH model of Zakoian (1994) is essentially the same as the GJR-GARCH model, however, in the case of the TGARCH model, the conditional standard deviation is modelled instead of the conditional variance. The TGARCH model is given by:

$$\sqrt{h_t}=\alpha_0+\sum_{i=1}^{q}\left(\alpha_i-\gamma_i I_{t-i}\right)a_{t-i}^2+\sum_{j=1}^{p}\beta_j\sqrt{h_{t-j}}+\sum_{j=1}^{m}\zeta_j\upsilon_{jt}, \tag{2.11}$$

2.3.4 Nonlinear Asymmetric GARCH model

The nonlinear asymmetric GARCH (NAGARCH) models can be specified based on the family GARCH model of Hentschel (1992) which is a basket of models, consisting of some of the most popular GARCH models. It allows the decomposition of the residuals in the conditional variance equation to be driven by different powers for z_t and $\sqrt{h_t}$. It also allows shifts and rotations in the news-impact curve. The main source of asymmetry for small shocks is the shift while large shocks are driven by rotation. The model is specified as:

$$h_t^{\lambda}=\alpha_0+\sum_{j=1}^{q}\alpha_j h_{t-j}^{\lambda}\left(\left|z_{t-j}-\eta_{2j}\right|-\eta_{1j}\left(z_{t-j}-\eta_{2j}\right)\right)^{\delta}+\sum_{j=1}^{p}\beta_j h_{t-j}^{\lambda}+\sum_{j=1}^{m}\zeta_j\upsilon_{jt}. \tag{2.12}$$

This specification is a Box-Cox transformation for the conditional standard deviation. The shape of the model is controlled by λ. The δ parameter transforms the absolute value function while η_{1i} and η_{2i} are parameters, which control rotations (asymmetry for large shocks) and shift (asymmetry for small shocks) in the news-impact curve respectively. We obtain the threshold GARCH model of Zakoian (1994) when $\lambda=\delta=1$, $\eta_{2j}=0$ and $\left|\eta_{1j}\right|\le 1$ and the nonlinear asymmetric

GARCH model of Engle & Ng (1993) when $\lambda = \delta = 2$ and $\eta_{1j} = 0$. The persistence of the model is given by:

$$\hat{P} = \sum_{j=1}^{q} \alpha_j \kappa_j + \sum_{j=1}^{p} \beta_j \,,$$

where κ_j is the expected value of the standardized residuals z_t under the Box-Cox transformation of the absolute value of the asymmetry term, defined by:

$$\kappa_j = \mathrm{E}\left(\left|z_{t-j} - \eta_{2j}\right| - \eta_{1j}\left(z_{t-j} - \eta_{2j}\right)\right)^{\delta} = \int_{-\infty}^{\infty} \left(\left|z_{t-j} - \eta_{2j}\right| - \eta_{1j}\left(z_{t-j} - \eta_{2j}\right)\right)^{\delta} f\left(z, 0, 1, \ldots\right) dz \,.$$

The unconditional variance is also given by:

$$\hat{h} = \left(\frac{\hat{\alpha}_0}{1 - \hat{P}}\right)^{2/\lambda}.$$

If variance targeting is used, the estimated intercept α_0 is replaced with $\hat{\sigma}^{\lambda}\left(1 - \hat{P}\right) - \sum_{j=1}^{m} \zeta_j \upsilon_{jt}$.

2.4 Parameter estimation of GARCH model via Quasi-Maximum Likelihood

Maximum likelihood (ML) estimation and quasi-maximum likelihood (QML) estimation are among the most widely used parameter estimators for GARCH models. They provide a consistent approach to parameter-estimation problems and have desirable mathematical and optimality properties. The ML seeks to maximize the actual log likelihood function while the QML seeks to maximize a related function to the log- likelihood function (often a simplified version of the actual log likelihood function).

The ML and QML estimators produce unbiased estimates in larger samples. The ML estimator is most efficient under correct model specification, but unlike the ML, QML is consistent under very mild conditions and robust to the distribution of the underlying *i.i.d* sequence (z_t) (Berkes *et al.*, 2003; Jensen & Rahbek, 2004 and Straumann, 2005). Furthermore, the observations do not required an imposed moment condition to be consistent and asymptotically normal (Francq, Horváth and Zakoïan, 2011). Owing to these merits of QML over ML, we employ the QML estimator to estimate the parameters of all the GARCH models used in this thesis. In this section, the QML estimation approach for GARCH model is briefly discussed.

Consider the GARCH (p, q) model in equation (2.5) without the external regressors. If the orders p and q are known, the vector of the parameters:

$$\theta = \left(\theta_1, ..., \theta_{p+q}\right)' := \left(\alpha_0, \alpha_1, ..., \alpha_p, \beta_1, ..., \beta_p\right)' \subset \Theta, \tag{2.12}$$

where Θ is a parameter space such that $\Theta \subset (0, \pm\infty) \times [0, \infty)^{P+q}$. The true value of the parameter is unknown, and we denote it by $\theta_0 = (\alpha_0, \alpha_{01}, \ldots, \alpha_{0q}, \beta_{01}, \ldots, \beta_{0p})'$. When writing the likelihood function of the model, instead of specifying the conditional distribution, for the purposes of illustration, we use the Gaussian quasi-likelihood function. This function coincides with the likelihood of the $z_t \sim N(0,1)$ given some initial values. If $\alpha_0, ..., \alpha_{1-q}, \tilde{\sigma}_0^2, ..., \tilde{\sigma}_{1-p}^2$ are considered to be the initial values, the conditional Gaussian quasi-likelihood of the model is defined as:

$$L_n(\theta) = L_n(\theta; a_1, ..., a_n) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\tilde{\sigma}_t^2}} \exp\left(-\frac{a_t^2}{2\left\{\alpha_0 + \sum_{i=1}^{q} \alpha_i a_{t-i}^2 + \sum_{i=j}^{p} \beta_j \tilde{\sigma}_t^2\right\}} \right), \tag{2.13}$$

Under the assumption of second-order stationarity, a reasonable choice for the unknown initial values is the unconditional variance corresponding to a given value of θ:

$$a_0^2 = ... = a_{1-q}^2 = \sigma_0^2 = ... = \sigma_{1-p}^2 = \frac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{i=j}^{p} \beta_j}. \tag{2.14}$$

Any measurable solution of $\hat{\theta}_n$ defines a QML estimator for θ if:

$$\hat{\theta}_n = \underset{\theta \subset \Theta}{\arg\max}\, L_n(\theta), \tag{2.15}$$

Maximising the likelihood function is equivalent to minimising the function, if one takes the logarithm of the likelihood function with respect to θ:

$$\tilde{I}_n(\theta) - n^{-1}t = \sum_{t=1}^{n}\left(\frac{a_t^2}{\tilde{\sigma}_t^2} + \log \tilde{\sigma}_t^2\right). \tag{2.16}$$

Similar to the ML estimator, the QML likelihood function is computationally expensive and, sometimes, encounters computational difficulties. To simplify the computational process and to alleviate the degree of the encountered numerical or computational difficulties (Francq, Horváth & Zakoïan, 2011), variance targeting is used in conjunction with the QML estimator.

Variance targeting is a technique, which relies on a re-parameterization of the model, such that the unconditional variance is estimated first before estimating the remaining parameters. The variance target estimation can be superior to the standard QML estimates for long-term prediction or VaR calculation when the model is mis-specified (Francq, Horváth & Zakoïan, 2011). The simplicity of the variance targeting procedure guarantees that the estimated unconditional variance of the GARCH model will be equal to the sample variance.

2.5 Conclusions

In this chapter, some of the empirical characteristics of speculative asset returns and the general economic reasons behind ARCH models and its generalized version were detailed. Some of the most commonly used GARCH models[2] and the maximum likelihood estimation of GARCH (1, 1) model were reviewed.

[2] For other extensions of univariate GARCH models see Bucci (2017), Andersen et al. (2009), Teräsvirta (2006), Ling & McAleer (2003), Francq & Zakoıan (2010).

CHAPTER 3

Statistical tests for GARCH models

Chapter Summary

Data validation and model diagnostics are standard procedures in statistical model-building processes. There are several specific and general statistical tests designed for these procedures for GARCH type of models. In this chapter, the procedures involved in performing some of these statistical tests as well as some forecast and model comparison tools are elaborated upon. The data-validation tests discussed, include the Dicky Fuller test and the ARCH Lagrange multiplier test. The diagnostic tests include the weighted portmanteau test, goodness-of-fit test, the likelihood ratio test, the unconditional coverage, and the Christoffersen's interval forecast. Forecast evaluation tools such as root mean square error, mean absolute error, Mincer-Zarnowitz regression, and the model set confidence test for ranking competing models are examined.

3.1 Unit root test

Stationarity is a common assumption in time-series forecasting techniques. If a series is stationary, the mean, the variance, and the autocorrelation structure do not change over time. A common observable feature of a stationary process is that it fluctuates around its mean; this property is called 'mean reversion'. In plain terminology, the mean of a stationary series has the tendency of returning or getting closer to its long-term mean after initially drifting away from it (Arefin & Ahkam, 2017 and Ribeiro, Cermeño & Curto, 2017). There are several tests used in testing stationarity and the presence of unit root in series, but more attention is given to an examination of the Dickey-Fuller (DF) test in this study. This is a common test used in literature and it is valid with large sample size. The general implicit assumption behind the unit root test requires that the time series to be tested should be decomposed as:

$$y_t = f_t + z_t + \varepsilon_t, \tag{3.1}$$

where f_t is the deterministic component, z_t is the stochastic component and ε_t is a stationary error process. In particular instances, the series to be tested can be modelled by a simple AR (1) process:

$$y_t = \phi y_{t-1} + x_t'\delta + \varepsilon_t, \tag{3.2}$$

where $\{x_t's\}$ are optional exogenous regressors, which may consist of a constant, or a constant and trend, while ϕ and δ are parameters to be estimated. The error components $\{\varepsilon_t\}$ are assumed to be white noise. From the AR (1) process the following can be deduced:

- If $|\phi| \geq 1$, y_t is a nonstationary series and the variance of y_t increases with time and approaches infinity.
- If $|\phi| < 1$, y_t is a stationary or trend-stationary series.

The hypothesis of stationarity or trend-stationarity can be evaluated by testing whether the absolute value is strictly less than one, that is, $H_0 : |\phi| = 1$ versus $H_1 : |\phi| < 1$. In the standard Dicky Fuller test, y_{t-1} is subtracted from both sides of (3. 2), hence:

$$\Delta y_t = \alpha y_{t-1} + x_t'\delta + \varepsilon_t, \tag{3.3}$$

where $\Delta y_t = y_t - y_{t-1}$ and $\alpha = \phi - 1$. The null and alternative hypotheses may be written as:

$$\begin{aligned} H_0 &: \alpha = 0 \\ H_1 &: \alpha < 0 \end{aligned}, \tag{3.4}$$

Given the estimated $\hat{\alpha}$ with standard error $se(\hat{\alpha})$, the test statistic is given by:

$$t_\alpha = \frac{\hat{\alpha}}{se(\hat{\alpha})}. \tag{3.5}$$

If the series is correlated at higher order lags, the white noise assumption of the errors is violated and the standard DF test above is invalid. Under such circumstance, the Augmented DF (ADF) test is used. The test constructs a parametric correction for higher-order correlation by assuming that the series y_t follows a $AR(P)$ process, hence, lagged difference terms (Δy_{t-i}) are added to

(3.3) to obtain:

$$\Delta y_t = \alpha y_{t-1} + x_t'\delta + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + v_t \,. \tag{3.6}$$

where β_i is a parameter, p is the order of the lagged difference term while v_t is an error component. Equation (3.6) is the ADF specification. The ADF specification is then used to test the hypothesis (3.4) subject to the test statistic in (3.5). When performing ADF test, one is faced with the option of including either a constant in the test regression, or a constant and linear time trend or none in the test regression. There is also a choice of the number of lagged difference terms to be included in the test regression. A recommended approach is to run the test with both a constant and a linear trend. Including irrelevant regressors in the regression will increase type II error. Hamilton (1994) recommends choosing a specification that is a plausible description of the data for both the null and the alternative hypotheses. With regard to the number of lagged difference terms to be included in the test regression, the usual recommendation is to include a number of lags sufficient to remove serial correlation in the residuals.

Graham, Rothenberg & Stock (1996) propose a modification of the ADF test using generalized least squares (GLS) de-trending so that the option of including explanatory variables is catered for prior to running the test regression. They define a quasi-difference of y_t that depends on the value of a which represents the specific point of the alternative against which we wish to test the null hypothesis:

$$d(y_t \mid a) = \begin{cases} y_t & \text{if } t = 1 \\ y_t - a y_{t-1} & \text{if } t > 1 \end{cases} . \tag{3.7}$$

The quasi-differenced data $d(y_t \mid a)$ is then regressed on the quasi-differenced $d(x_t \mid a)$ using the ordinary least-square (OLS) specification:

$$d(y_t \mid a) = d(x_t \mid a)' \delta(a) + \eta_t , \tag{3.8}$$

where x_t contains either a constant or both a constant and trend. In equation (3.8), one is interested in the value of $\bar{a}$ as defined in (3.9). In equation (3.8), one is interested in the value of $\bar{a}$. Given $\bar{a}$ (see equation 3.9), if T is a finite constant the OLS estimate for a from (3.8) (as recommended by Graham, Rothenberg & Stock, 1996) is defined as:

$$\bar{a} = \begin{cases} 1\text{-}7/T & \text{if } x_t = \{1\} \\ 1\text{-}27/2T & \text{if } x_t = \{1, t\} \end{cases}. \tag{3.9}$$

The GLS de-trended data y_t^d using the estimate $\bar{a}$ is defined as:

$$y_t^d \equiv y_t - x_t' \hat{\delta}(\bar{a}), \tag{3.10}$$

where $\hat{\delta}(\bar{a})$ is an estimate from the OLS regression (3.10) which is a replaced by $\bar{a}$ as defined in (3.9). The Dickey-Fuller test with GLS De-trending (DFGLS) involves estimating the augmented ADF test equation after the substitution of the GLS de-trended data:

$$\Delta y_t^d = \alpha y_{t-1}^d + \sum_{i=1}^{p} \beta_i \Delta y_{t-i}^d + v_t. \tag{3.11}$$

3.2 Weighted Portmanteau Goodness-of-Fit Tests

As a standard procedure in statistical modelling, diagnostics checks are performed to assess the fitness of the estimated model. In ARMA (p, q)-GARCH (p, q) models, one of such tests involves the checking for the absence of serial correlation and heteroscedasticity in the innovations. There are a couple of tests available in literature for these tasks, but in this study, we shall use the weighted Ljung-Box, the weighted McLeod-Li and the weighted Li & Mark Portmanteau tests of Fisher & Gallagher (2012). These tests account for the distribution of statistics of the values from the estimated models better (Ghalanos, 2019).

3.2.1 Weighted Ljung-Box Test

The weighted Ljung-Box test is used to detect the absence of serial correlation in the standardised residuals from a fitted ARMA process. The test approach is easy to implement, and the test statistic is computationally stable even if the number of lags is relatively close to the sample size (Fisher & Gallagher, 2012). Generally, the test is comparable to Mahdi and McLeod (2012), but tends to outperform the commonly used test of Ljung & Box (1978) and Monti (1994). Considering the ARMA process of the form:

$$r_t = \sum_{i=1}^{p} \phi_i r_{t-i} - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t, \tag{3.12}$$

where p is the AR order, q is the MA order, and $\varepsilon_t \sim N(0,1)$. In testing for the fitness of a model estimated with the above ARMA process using the weighted Ljung-Box testing approach, we define the sample autocorrelation function for ε_t given a sample of size m as:

$$\hat{\rho}_k = \frac{\sum_{k+1}^{m} \hat{\varepsilon}_t \hat{\varepsilon}_{t-k}}{\sum_{t=1}^{m} \hat{\varepsilon}_t^2} \quad k = 1, 2, \ldots m. \tag{3.13}$$

If the orders p and q are correctly identified, each of the above correlation coefficients should be approximately equal to zero. The weighted Ljung-Box statistic is given by:

$$\tilde{Q}_w = n(n+2) \sum_{k=1}^{m} \left(\frac{m-k+1}{m+1} \right) \left(\frac{\hat{\rho}_k^2}{n-k} \right) \geq 0. \tag{3.14}$$

If the partial-autocorrelation matrix is used instead of the sample matrix, $\hat{\rho}_k^2$ is replaced by the squared partial autocorrelation $\hat{\pi}_k^2$, thus, the weighted Monti Statistic is written as:

$$\tilde{M}_w = n(n+2) \sum_{k=1}^{m} \left(\frac{m-k+1}{m+1} \right) \left(\frac{\hat{\pi}_k^2}{n-k} \right) \geq 0. \tag{3.15}$$

The statistics $\tilde{Q}_w$ and $\tilde{M}_w$ follow the same asymptotic distribution of Peña & Rodríguez (2002, 2006). Under the null hypothesis, $\tilde{Q}_w$ and $\tilde{M}_w$ are asymptotically distributed as $\sum_{k=1}^{m} \lambda_k \chi_k^2$ where

$\{\chi_k^2\}$ are independent Chi-squared random variables with one degree of freedom and $\lambda_k\,(k=1,2,3,\ldots,m)$ are the eigenvalues[3].

3.2.2 Weighted McLeod-Li Test

The weighted McLeod-Li test is used to detect nonlinearity in the standardised squared residuals from a fitted GARCH model. The test tends to be generally more powerful in comparison to other tests. If the null hypothesis is rejected, the squared residuals are deemed approximately nonlinear and the model is deemed fit for the data. The test considers a nonlinear model of the form:

$$\varepsilon_t = g(h_t)\eta_t, \tag{3.16}$$

where $\eta_t \sim N(0,1)$ and $\{h_t\}$ follow ARMA type recursion. We can define the autocorrelation function based on the transformation $g(.)$ of the residuals as:

$$\hat{\rho}_{k(*)}^2 = \frac{\sum_{t=k+1}^{n}\left(g(\hat{\varepsilon}_t)-\bar{g}(\hat{\varepsilon}_t)\right)\left(g(\hat{\varepsilon}_{t-k})-\bar{g}(\hat{\varepsilon}_t)\right)}{\sum_{t=1}^{n}\left(g(\hat{\varepsilon}_t)-\bar{g}(\hat{\varepsilon}_t)\right)^2}\ , \tag{3.17}$$

where $\bar{g}(\hat{\varepsilon}_t)=\sum g(\hat{\varepsilon}_t)/n$. The transformation $g(.)$ of the residuals can take the form of either a squared, absolute or log-squared. Under this test, the test statistic using the autocorrelation is defined as:

$$\tilde{Q}_w^* = n(n+2)\sum_{k=1}^{m}\left(\frac{m-k+1}{m+1}\right)\left(\frac{\hat{\rho}_{k(*)}^2}{n-k}\right). \tag{3.18}$$

If the partial-autocorrelation matrix is used instead of the sample matrix, the weighted Monti Statistic is obtained by:

[3] For detailed derivation, see Fisher & Gallagher, 2012.

$$\tilde{M}_w^* = n(n+2)\sum_{k=1}^{m}\left(\frac{m-k+1}{m+1}\right)\left(\frac{\hat{\pi}_{k(*)}^2}{n-k}\right). \tag{3.19}$$

The autocorrelations $\hat{\rho}_{k(*)}^2$ and the partial autocorrelations $\hat{\pi}_{k(*)}^2$ are based on the transformations $\hat{\varepsilon}_t^2$, $|\hat{\varepsilon}_t|$ or $\log(\hat{\varepsilon}_t)$. If the series follows a stationary ARMA process, $\tilde{Q}_w^*$ and $\tilde{M}_w^*$ computed from squared residuals are asymptotically distributed as $\sum_{k=1}^{m} w_k \chi_k^2$, where $\{\chi_k^2\}$ are independent Chi-squared random variables with one degree of freedom and $w_k\,(k=1,2,3,\ldots,m)$ are the weights given by $w_k = (m-i+1)/m$.

3.2.3 Weighted Li and Mark Test

The weighted Li and Mark test is generally used to test for goodness-of-fit for the GARCH model. The test tends to have higher power than other tests in literature, particularly, in detecting long-memory of nonlinear models (Fisher & Gallagher, 2012). If the null hypothesis of the test is not rejected, then it can be concluded that there is no significant remaining heteroscedasticity in the GARCH errors, thus, implying that the model is a good fit for the data. The test relies on GARCH type of models. Considering the GARCH model defined in equation (2.3), we can define the autocorrelation function based on the standardised sample squared residuals as:

$$\hat{\rho}_k^2\left(\hat{a}_t^2/h_t\right) = \frac{\sum_{t=k+1}^{n}\left(\hat{a}_t^2/h_t - \bar{a}_t\right)\left(\hat{a}_{t-k}^2/h_{t-k} - \bar{a}_t\right)}{\sum_{t=1}^{n}\left(\hat{a}_t^2/h_t - \bar{a}_t\right)^2}, \tag{3.20}$$

where $\bar{\varepsilon} = (1/n)\sum \hat{a}_t^2/\hat{h}_t$ and $\{\hat{h}_t\}$ is the sample conditional variance. The test statistic is:

$$L_W(b,m) = n\sum_{k=b+1}^{m}\frac{(m-k+(b+1))}{m+1}\hat{\rho}_k^2\left(\hat{\varepsilon}_t^2/h_t\right). \tag{3.21}$$

where $L_W(b,m)$ is asymptotically distributed as $\sum_{k=1}^{m} w_k \chi^2{}_k$ under the null hypothesis of an adequately fitted ARCH (b) model. The $\{\chi_k^2\}$ are independent Chi-squared random variables with one degree of freedom and $w_k\left(k=1,\ldots,m\right)$, are the weights defined by $w_k=\left(m-k+\left(b+1\right)\right)/m$.

3.3 The ARCH Lagrange Multiplier test

The ARCH Lagrange Multiplier is used to test for the presence of ARCH effects in return data. The test requires an auxiliary test regression of squared residuals obtained from an ordinary least square regression of the conditional mean equation on a constant and q lags. Consider the ARMA representation below:

$$r_t=\phi_0+\phi_1 r_{t-1}-\theta_1 e_{t-1}+e_t\,. \tag{3.22}$$

For q lags, Engle (1982) defined the auxiliary test regression equation as:

$$e_t^2=\alpha_0+\sum_{i=1}^{q}\alpha_i e_{t-i}^2+\nu_t\,. \tag{3.23}$$

The null hypothesis of no ARCH effect presence is tested against the alternative of ARCH effect presence using the test statistic T^*R^2, where T is the number of observations and R^2 is the coefficient of determination. The test statistic follows a chi-squared distribution with p degrees of freedom.

3.4 Likelihood ratio test

The likelihood ratio test is used to compare the likelihood of a restricted version of a model (model without additional regressor(s)) to the unrestricted version (model with additional regressor(s)), thus any significant gains in the likelihood is because of the additional regressor or regressors. Likelihood ratio test may therefore be used to assess the effect or the contribution of an additional regressor (s) in a statistical model. Following the definition by Kalbfleisch (1985), the test is formally formulated as follows. Let Ω denote the complete parameter space of θ (where θ is the set of all possible parameters in the restricted and unrestricted models). Given that $\omega\in\theta$ and $\omega'=\{\theta\}-\{\omega\}$ is the complement of ω with respect to the parameter space Ω, we are interested in testing the null hypothesis.

$$H_o : \theta \in \omega \text{ versus } H_1 : \theta \in \omega' . \tag{3.24}$$

The test statistic is given by:

$$LRT = -2\ln\left(\frac{L_{res}\left(\hat{\theta}\right)}{L_{unres}\left(\hat{\theta}\right)}\right), \tag{3.25}$$

where $L_{res}\left(\hat{\theta}\right)$ is the likelihood of the restricted model and $L_{unres}\left(\hat{\theta}\right)$ is the likelihood of the unrestricted model. Under the null hypothesis, the test follows an asymptotic Chi-square distribution with k (number of restricted coefficients) degree of freedom. The null hypothesis is rejected when the p-value of the test statistic is less than the level of significance. In such a situation, we can conclude that the unrestricted model performs significantly better than the restricted model.

3.5 Backtesting methods for evaluating VaR Estimates

After estimating the VaR model, it is required to assess the efficiency or the accuracy of the estimates. Techniques used in these assessments are called "backtesting". Backtesting provides diagnostic checks on the quality of estimates from a risk model (Kerkhof & Melenberg, 2004). Backtesting techniques require the simulation of VaR models on past returns and then compare the predicted losses from VaR calculations to the actual realized losses at a given time horizon. The comparison identifies periods where the portfolio losses are greater than the expected VaR. If the expected returns are less than the estimated VaR, a violation or exception occurs and thus backtesting techniques are used to count the number of these violations or exceptions systematically and compared them to acceptable rates at preselected confidence intervals.

Backtesting methods can be categorized into - point forecasts, probability range forecasts or interval forecasts and forecasts of the complete probability distribution (Emmer, Kratz &Tasche, 2013). The probability range or interval forecast defines a confidence interval for the VaR forecast with predetermined probability. In this thesis, this approach which is based on the so-called violation process or hit sequence of Christoffersen (1998) is employed. The realization of asset returns over a fixed time interval $R_{t,t+1}$ with VaR is estimated at time t and the probability of α defined by

$VaR_t(\alpha)$. The hit function is then defined as:

$$I_{t+1}(\alpha)=\begin{cases}1 \text{ if } R_{t,t+1}\leq VaR_t(\alpha) & \textit{when violation occurs}\\ 0 \text{ if } R_{t,t+1}>VaR_t(\alpha) & \textit{when no violation occurs}\end{cases} \quad . \tag{3.26}$$

If day's $(t+1)$ loss is larger than the predicted VaR estimate, the hit sequence returns 1 or 0, otherwise, Christoffersen (1998) explains that the hit sequences of VaR estimates need to conform to the unconditional coverage and independence properties before they can be deemed accurate. Under the unconditional coverage property, the probability of loss in day $(t+1)$ being larger than the predicted, VaR estimate should be exactly $(1-\alpha)$ or equivalently, the probability of loss in day $(t+1)$ being smaller than the predicted VaR estimate should be exactly α. Mathematically a VaR model has correct unconditional coverage if:

$$P\left(I_{t+1}(\alpha)=1\right)=\mathrm{E}\left[I_{t+1}\right]=1-\alpha \quad . \tag{3.27}$$

It also has correct conditional coverage if:

$$P_t\left(I_{t+1}(\alpha)=1\right)=\mathrm{E}_t\left[I_{t+1}\right]=1-\alpha \, . \tag{3.28}$$

It should be noted that correct unconditional coverage is implied by correct conditional coverage, but not vice versa. If violations are more frequent than α, the VaR model systematically underestimates the actual VaR and if violations are less frequent than α, the model overestimates the expected VaR (Kosapattarapim, 2013). The independence property requires that for any $i<j$ hits, $I_{t+i}(\alpha)$ and $I_{t+j}(\alpha)$ are independent if $i\neq j$. In other words, when the exceedances are not clustered over time, VaR estimates are said to be independent. Accurate VaR estimates, therefore, have current violations at time $(t+i)$, which is independent of violations at previous time $(t+j)$ or have non-clustered exceedances over time. The hit sequences from VaR estimates are independent and identically distributed Bernoulli random variables, with success probability $(1-\alpha)$ i.e., $I_{t+1}(\alpha)\sim i.i.d.\ Bernoulli\ (1-\alpha)$, thus, in testing for VaR violation, we are interested in the null

hypotheses:

$$H_0 : E\left[I_{t+1}\right] = \pi = 1-\alpha$$
$$H_0 : E_t\left[I_{t+1}\right] = \pi = 1-\alpha \, , \tag{3.29}$$

where π is the sample average, the first null hypothesis corresponds to the unconditional coverage test and the second null hypothesis corresponds to the conditional coverage test.

3.5.1 Kupiec Likelihood Ratio Test

The Kupiec Likelihood Ratio Test (unconditional coverage) or proportion of failures test of Kupiec (1995) allows us to test if the unconditional probability (π) of violations in the risk model is consistent with the expected exceedances, at a given quantile and level of significance. Under this test the number of violations, x follow the binomial distribution:

$$f(x) = \binom{T}{x} p^x (1-p)^{T-x}, \tag{3.30}$$

where T is the total number of observations. Given $\hat{p}$ as the observed failure rate defined by $\hat{p} = x/T$ and p the theoretical failure rate, the following hypothesis is tested:

$$H_0 : p = x/T$$
$$H_1 : p > x/T \, . \tag{3.31}$$

The idea behind the test is to find out whether there is a large discrepancy between the observed failure rate, $\hat{p}$ and the theoretical failure rate p (Roccioletti, 2015). The test statistic is a likelihood ratio (LR_{uc}) defined as:

$$LR_{uc} = -2\ln\left(\frac{(1-p)^{T-x}\, p^x}{\left(1-(x/T)\right)^{T-x} (x/T)^x}\right) \to \chi_1^2 \, . \tag{3.32}$$

Under the null hypothesis of a correctly specified model, the test statistic LR_{uc} is asymptotically chi-squared distributed with one degree of freedom. If the statistic is higher than the critical value of the χ_1^2 distribution, the null hypothesis is not rejected and the observed failure rate is not significantly different from the expected at the chosen level of significance.

3.5.2 Christoffersen's interval forecast test

Time-varying volatility is one of the common regularities of asset returns (Andersen *et al.,* 2005). A risk model that fails to recognize this regularity leads to VaR estimates, which respond late to changing market conditions and cluster, over time (Roccioletti, 2015 and Pritsker, 2001). The Kupiec test does not account for this regularity, in the sense that it ignores the time losses, which occur, thus, the test may fail to reject a model that produces clustered VaR violations. The interval forecast test of Christoffersen (1998) is able to deal with this problem by jointly testing unconditional coverage and independence of consecutive violations in a composite test called - the conditional coverage or interval forecast test. The test statistic for the composite likelihood ratio LR_{cc} is given by:

$$LR_{cc} = LR_{uc} + LR_{ind} \to \chi_2^2, \tag{3.33}$$

where LR_{uc} is the test statistic for the unconditional coverage test and LR_{ind} represents the test statistic for the independence test. The LR_{uc} is asymptotically distributed as chi-square with two degrees of freedom. Under the hit function defined in (3.26), the conditional coverage test is interested in checking if the probability of violation at time t on condition that violation occurred at time $(t-1)$ is not different from the probability of violation at time t on condition that no violation occurred at time $(t-1)$. The test has four possible outcomes, which are reported in Table 3.1:

Table 3. 1: Possible outcomes for the independence of violations test

	$I_{t-1}=0$	$I_{t-1}=1$	$Total$
$I_{t=0}$	n_{00}	n_{10}	$n_{00}+n_{10}$
$I_{t=1}$	n_{01}	n_{11}	$n_{01}+n_{11}$
$Total$	$n_{00}+n_{01}$	$n_{10}+n_{11}$	N

The probability of no violation occurring at time t and $(t-1)$ is denoted by n_{00}, the probability of no violation occurring at time t given that violation which occurred at $(t-1)$ is denoted by n_{10}, the probability of violation occurring at time t given that no violation occurred at $(t-1)$ is denoted by n_{01} and the probability of violation occurring at time t and $(t-1)$ is denoted by n_{11}. Denoting the number of observations in state j given that it has already been in state i by n_{ij} and the probability of n_{ij} by p_{ij}, then:

$$p_{01}=\frac{n_{01}}{n_{00}+n_{01}},\, p_{11}=\frac{n_{11}}{n_{10}+n_{11}} \text{ and } \hat{p}_{01}=\frac{n_{01}+n_{11}}{n_{00}+n_{01}+n_{01}+n_{11}}. \tag{3.34}$$

Under the null hypothesis of independent violations, $p_{01}=p_{11}=\hat{p}$, the test statistic for the independence of violation is defined as:

$$LR_{ind}=-2\ln\left(\frac{(1-\hat{p})^{n_{00}+n_{10}}\ \hat{p}^{n_{01}+n_{11}}}{(1-\hat{p}_{01})^{n_{00}}\ \hat{p}_{01}^{n_{01}}\ (1-\hat{p}_{11})^{n_{10}}\ \hat{p}_{11}^{n_{11}}}\right)\to\chi_1^2. \tag{3.35}$$

Using test statistic (3.35), hypothesis (3.31) is tested. If the null hypothesis is not rejected, the observed failure rate is not significantly different from the expected failure rates at the chosen level of significance and the VaR estimates can respond early to changing market conditions, without clustering over time. In such a situation, the VaR estimates are deemed significantly reasonable or accurate.

3.6 Statistical Loss functions

Statistical loss functions are, generally, used in evaluating prediction and forecasting accuracies of Statistical models. When using loss functions, the predicted values or forecasts are compared to the actual or realized values. The use of loss functions to evaluate volatility forecasts have been criticised by researchers. Bollerslev, Engle & Nelson (1994) and Poon & Granger (2005) argue that the use of statistical loss functions in evaluating volatility forecasts is inappropriate. Bollerslev, Engle & Nelson (1994) observe that mean squared error (MSE) does not sufficiently penalize non-positive variance forecasts due to its symmetric nature, while Poon & Granger (2005) argue that the squaring of errors in the computation of heteroscedasticity-adjusted MSE (HMSE) values give greater weight to large errors.

In addition, the use of squared returns as proxies for actual volatility is inappropriate because square returns tend to over-exaggerate true volatility. In this regard, Andersen & Bollerslev (1998) and Christodoulakis & Satchell (1998) argue that the use of squared returns approximation leads to an inflated MSE with distorted forecasts. Awartani & Corradi (2004) suggest the use of squared mean-corrected returns to proxy true volatility when using loss functions to assess forecast accuracy of volatility models. This is believed to alleviate the problem of MSE inflation. The various statistical loss functions have their own competitive advantages, however, contemporary literature suggests that there is no standard to decide whether any of the loss functions are superior to another. In this thesis, hence, the most commonly used statistical loss functions, mean absolute Error (MAE) and root mean square error (RMSE) are used for assessing the predictive and forecasting accuracies in all the studies. Given the innovations and estimated volatility over a period, MAE and RMSE for in-sample predictions and forecasts are defined by:

$$MAE = \frac{1}{\tau}\sum_{i=1}^{\tau}\left|\sigma_{t+i}^{2} - \hat{\sigma}_{t+i}^{2}\right| \text{ and } RMSE = \sqrt{\frac{1}{\tau}\sum_{i=1}^{\tau}\left(\sigma_{t+i}^{2} - \hat{\sigma}_{t+i}^{2}\right)^{2}}, \tag{3.36}$$

where σ_{t+i}^{2} is the actual volatility, $\hat{\sigma}_{t+i}$ is the estimated or forecasted volatility and τ is the length of the estimated or forecasted sample.

3.7 Mincer-Zarnowitz regression

Mincer-Zarnowitz (1969) regression criterion is used in assessing forecast accuracy of competing models. The criteria require regressing the true volatility (or its proxy) on a constant and the forecasted volatility. Given the true volatility or its proxy h_t, a corresponding forecast $\hat{h}_t$ and parameters α_0 and α_1, the *Mincer-Zarnowitz* regression (MZR) is given by:

$$h_t = \alpha_0 + \alpha_1 \hat{h}_t + \xi_t,\ t = 1, \ldots, \tau . \tag{3.37}$$

If the model for the conditional variance is correctly specified and the volatility is unbiased for the true unobservable volatility, $\alpha_0 = 0$ and $\alpha_1 = 1$, thus, in as much as MZR is used as an accuracy tool, it is also a misspecification test. In the MZR criterion, we are interested in the Wald joint hypothesis test:

$$\begin{aligned} H_0 &: \alpha_0 = 0 \cap \alpha_1 = 1 \\ H_a &: \alpha_0 \neq 0 \cup \alpha_1 \neq 1 \end{aligned}. \quad (3.38)$$

For any two competing GARCH models, if we fail to reject the joint null hypothesis, then the model with the highest R-square has the best predictive power. GARCH errors are heteroskedastic, thus, the use of ordinary least squares to estimate the parameters of equation (3.37) may not yield optimal estimates because of autocorrelation of the errors. Ordinary least square with robust standard errors proposed by Koller & Stahel (2011) and Yohai (1987) is recommended. In this study, we employ *Mincer-Zarnowitz* regression *(MZR)* in addition to the loss functions (MAE and RMSE) to evaluate both prediction and forecast accuracies of competing models. If $\alpha_0 \approx 0$ and $\alpha_1 \approx 1$, then $\xi_t \approx 0$ and $h_t \approx \hat{h}_t$, hence, a model with a smaller mean square error from MZR (**MSEMZR**) and root mean squared MZ (**RMSEMZR**) is preferred. This measure is reported in addition to the MZR-R-square, because the null hypothesis of the joint of Wald test for both models may not be rejected simultaneously and their comparison based on MZR-R-square may not be appropriate.

3.8 Model set confidence test

Due to a large set of models available for use by financial institutions, model comparison has become an integral part of the model-building process. The comparison allows a practitioner to select one model or a set of models to use in forecasting a particular series for making an informed decision. From the onset of the year 2000, efforts have not been spared in developing new testing procedures to compare, select, and rank a set of competing models, in order of superiority relative to some specified loss metric. The procedures include the model set confidence test of Hansen, Lunde & Nason (2011), the conditional predictive ability test of Giacomini & White (2006), the superior predictive ability test of Hansen and Lunde (2005), the stepwise multiple testing procedure of Romano & Wolf (2005) and the reality check of White (2000).

There has been a long-standing problem of multiple model comparisons (Gupta & Panchapakesan, 1979 and Hsu, 1996). Multiple comparison problems are encountered when objects are compared to the best sample performance (Horrace & Schmidt, 2000 and Hansen, Lunde & Nason, 2011). When all objects being compared are selected independently of the data used for the comparison, multiple comparisons with the control are encountered (White, 2000). The model set confidence (MCS) procedure is convenient for comparisons where there is no natural benchmark. This advantage of MCS procedure makes it more suitable for use in this study because the reference models in this study are not natural benchmarks, as there have been no comprehensive studies to substantiate their superiority over existing models. Even though the MCS procedure does not require a benchmark model as in the case of multiple comparison procedure with controls, the procedure can still rank the models in order of superiority after the initial selecting of the superior set of models. One main advantage of this procedure is that after a set of superior models have been selected the models can then be used to forecast future volatility levels; predict future levels of observations, conditioned on past information (Bernardi & Catania, 2014) or forecast VaR levels (Bernardi, Catania & Petrella, 2014). Alternatively, the models can be aggregated to obtain a single better forecast measure.

The MCS procedure consists of a sequence of statistical tests, which allow for the construction of a set of superior set of models (SSM) under the null hypothesis of equal predictive ability (EPA) at pre-determined or default confidence level. The EPA statistics are calculated for arbitrary loss function satisfying the general weak stationarity conditions. Let Y_t be the observation at time t and $\hat{Y}_{i,t}$ be the output of model i at time t and define the loss function associated with the i^{th} model which measures the difference between $\hat{Y}_{i,t}$ and Y_t as:

$$\ell_{i,t} = \ell\left(Y_t, \hat{Y}_{i,t}\right). \tag{3.39}$$

The models to be compared in the study are risk models, therefore, the asymmetric VaR loss function of González-Rivera, Lee & Mishra (2004) is used. Given the historical information set $\mathcal{F}_{t-1}$ and the $\alpha - level$ predicted VaR estimate at time t denoted by VaR_t^{α}, the asymmetric VaR loss function is defined as:

$$\ell\left(y_t, VaR_t^{\alpha}\right)=\left(\tau-d_t^{\alpha}\right)\left(y_t-VaR_t^{\alpha}\right), \tag{3.40}$$

where $d_t^{\alpha}=\mathbf{1}\left(y_t<VaR_t^{\alpha}\right)$ is the $\alpha-quantile$ loss function representing the natural candidate for quantile-based risk measures. The choice of this quantile loss function is due to its ability to heavily penalise observations below the $\alpha-quantile$ level (Bernardi & Catania, 2014). The MCS procedure as defined by Bernardi & Catania,(2014) is explained below. Let $d_{ij,t}$ denote the loss differentials between models i and j defined as:

$$d_{ij,t}=\ell_{i,t}-\ell_{j,t},\ j=1,2,\ldots,m,\ \ t=1,2,\ldots,n. \tag{3.41}$$

Let also the simple loss of models i relative to j at time t be defined as:

$$d_{i.,t}=\left(m-1\right)^{-1}\sum_{j\in M} d_{ij,t}\ \ i=1,2,\ldots,m. \tag{3.42}$$

The EPA hypothesis for a set of M models can either be formulated as:

$$\begin{aligned} &H_{0,M}: c_{ij}=0,\ \text{for all } i, j=1,2,3,\ldots m \\ &H_{A,M}: c_{ij}\neq 0,\ \text{for some } i, j=1,2,3,\ldots m \end{aligned}. \tag{3.43}$$

Alternatively:

$$\begin{aligned} &H_{0,M}: c_{i}=0,\ \text{for all } i, j=1,2,3,\ldots m \\ &H_{A,M}: c_{i}\neq 0,\ \text{for some } i, j=1,2,3,\ldots m \end{aligned}, \tag{3.44}$$

where $c_{ij}=\mathbb{E}\left(d_{ij}\right)$ and $c_i=\mathbb{E}\left(d_{i.}\right)$ are assumed to be finite and independent of time. Testing the two hypotheses requires the computations of the test statistics:

$$t_{ij}=\frac{\bar{d}_{ij}}{\sqrt{\hat{\sigma}\left(\bar{d}_{ij}\right)}}\ \text{and}\ t_{i.}=\frac{\bar{d}_{i,.}}{\sqrt{\hat{\sigma}\left(\bar{d}_{i,.}\right)}}\ \text{for}\ i, j\in M, \tag{3.45}$$

where $\bar{d}_{ij}=m^{-1}\sum_{t=1}^{m} d_{ij,t}$ measures the relative sample loss between i^{th} model and j^{th} models,

$\bar{d}_{i,\cdot} = (m-1)^{-1} \sum_{j\in M} \bar{d}_{ij}$ is the simple loss of the i^{th} model relative to the averages losses across models in the M set of models while $\hat{\sigma}(\bar{d}_{ij})$ and $\hat{\sigma}(\bar{d}_{i,\cdot})$ are the bootstrapped estimates for $\text{var}(\bar{d}_{ij})$ and $\text{var}(\bar{d}_{i,\cdot})$ respectively. The two hypotheses in equations (3.43) and (3.44) can naturally map onto the two respective test statistics:

$$T_{R,M} = \max_{i,j\in M} |t_{ij}| \text{ and } T_{R,M} = \max_{i\in M} (t_{i,\cdot}). \tag{3.46}$$

The relevant distributions under the null hypothesis for the test statistics are estimated via the bootstrap procedures used in Hansen, Lunde & Nason (2003) and Hansen (2005). The MCS procedure consists of sequential procedures where the worst model is eliminated at each stage until the hypothesis of EPA is not rejected, for the remaining models belong to the superior set of models. The elimination rules used are, respectively:

$$e_{R,M} = \arg\max_i \left\{ \sup_{j\in M} \frac{\bar{d}_{ij}}{\sqrt{\hat{\sigma}(\bar{d}_{ij})}} \right\} \text{ and } e_{\max,M} = \arg\max_{i\in M} \frac{\bar{d}_{i,\cdot}}{\hat{\sigma}(\bar{d}_{i,\cdot})} \text{ for } i, j \in M . \tag{3.47}$$

3.9 Adjusted Pearson goodness-of-fit

The Pearson goodness-of-fit test is used to test whether residuals from an estimated model comes from an assumed theoretical distribution. The test compares the observed standardised residuals to the expected residual and if the observed standardised residuals come from the assumed theoretical distribution, the null hypothesis is rejected. The null hypothesis of the test is

$$\begin{aligned} &H_0: \textit{ the data follows a given distribution.} \\ &H_a: \textit{ the data does not follows a given distribution.} \end{aligned} \tag{3.48}$$

If the standardized residual is not an *i.i.d* sequence, the test tends not to perform well, thus, to correct this, Palm (1996) suggests an adjusted version of the test by categorization of the standardised residuals by their magnitudes. The test statistics version of the test as defined by Palm (1996) is computed as:

$$P(g) = \sum_{i=1}^{g} \frac{\left[\eta_i - \mathrm{E}(\eta_i)\right]^2}{\mathrm{E}(\eta_i)}, \tag{3.49}$$

where η_i is the number of observations in cell i and $\mathrm{E}(\eta_i)$ is the predicted number of observations. If the p-value is less than the conventional alpha level, the null hypothesis is not rejected and the model can be deemed a good fit for the model under consideration.

3.10 Jarque -Bera goodness-of-fit test

The Jarque –Bera (JB) test (Jarque & Bera, 1980) is used to test for normality in a sample data. If the sample data was drawn from the normal distribution, its expected skewness and excess kurtosis are not supposed to be significantly different from 0, thus the JB test aims to identify non-significant deviations of the sample skewness and excess kurtosis from the respective values of the normal distribution. The null hypothesis of the test is a joint hypothesis of zero skewness and excess kurtosis. The test statistic is defined by Jarque & Bera, (1980) as:

$$JB = \frac{n}{6}\left(S + \frac{(K\text{-}3)^2}{4} \right), \tag{3.50}$$

where S and K are the estimates of the sample skewness and kurtosis, respectively. The number of observations is denoted by n. The sample skewness and kurtosis are defined by

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}^3}, \quad \hat{\mu}_3 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3 \text{ and } \hat{\sigma}^3 = \sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^3} \tag{3.51}$$

$$K = \frac{\hat{\mu}_4}{\hat{\sigma}^4}, \quad \hat{\mu}_4 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4 \text{ and } \hat{\sigma}^4 = \left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^2 \tag{3.52}$$

where $\hat{\mu}_3$ and $\hat{\mu}_4$ are the estimates of the theoretical third and fourth central moments, respectively, $\bar{x}$ is the sample mean, and $\hat{\sigma}^2$ is the estimate of the variance. The test statistic is asymptotically chi-squared distributed with two degrees of freedom (Bowman & Shenton, 1975). This suggests that the null hypothesis is rejected at α level if $JB \geq \chi^2_{(1-\alpha,2)}$.

3.11 Conclusion

In this chapter, data validation tests for heteroscedasticity and stationarity, model diagnostics tests, forecast evaluation and model comparison tools were discussed. The Portmanteau tests of Fisher & Gallagher (2012) were selected for the post-diagnostic test for serial and autocorrelation because they are known to account better for the distribution of the statistics from the estimated GARCH models (Ghalanos, 2019). Specifically, the weighted Ljung-Box test was chosen to detect the absence of serial correlation in the standardized residuals from a fitted ARMA process because the test approach is easy to implement and the test statistic is computationally stable, even if the number of lags is relatively close to the sample size (Fisher & Gallagher, 2012). The weighted McLeod-Li test is selected to detect nonlinearity in the standardized squared residuals from the fitted GARCH models because the test tends to be generally more powerful in comparison to other tests. The weighted Li & Mark test was chosen to test for the absence of ARCH effects in the standardized squared residuals from the fitted models because the test tends to have higher power than other tests in literature, particularly, in detecting long-memory of nonlinear models (Fisher & Gallagher, 2012).

The ARCH Lagrange Multiplier was used to test for the presence of ARCH effects in the filtered returns. The ARCH LM is simple to carry out, although, the statistical size of the test is generally less than its nominal size in finite samples (Bollerslev & Wooldrige, 1992), hence, the weighted Ljung-Box is also used to make sure that the correct test conclusions are arrived at. In testing for unit root in the returns, the detrended GLS ADF test was used because it is valid with large sample sizes. The likelihood ratio test was used to evaluate the statistical significance of the unrestricted models while MAE, RMSE, the Mincer-Zarnowitz regression, and its associated loss metrics were used to assess the predictive and forecasting accuracies and superiorities of the competing models. The unconditional coverage test of Kupeic and the interval forecast test of Christoffersen were selected to backtest the VaR estimates while the MCS procedure was used to select and rank the superior set of competing VaR models because they do not require a natural benchmark.

CHAPTER 4

Data exploration and validation tests

Chapter Summary

Modelling volatility of exchange rate returns with GARCH models requires that the returns are heteroscedastic, stationary and exhibit volatility clustering, thus, we test the presence of these characteristics in the returns of our currency pairs to validate them. Returns are known to be leptokurtic with excess peaks, hence, we also assess these characteristics. The data-validations tests showed that all the return series are heteroscedastic, stationary and exhibit volatility clustering. Volatility asymmetries are found in the returns of BWP/ZAR, MWK/ZAR, and MZN/ZAR. The returns are non-normal with heavy-tails and high peaks, suggesting that heavy-tailed distributions are appropriate for modelling the innovations of the returns. Furthermore, it is observed that currencies move closely together at price-level and this translates into significant positive return-level co-movements. A substantial number of the currency-pairs exhibit strong return-level co-movements, thus, may have some predictive or forecasting powers.

4.1 Introduction

Verifications of conditions and distributional assumptions governing the estimation of GARCH models are usually the first exercise towards the process of modelling and forecasting with GARCH models. It allows one to validate the suitability of the data for use in the models. In this way, certain anticipated estimation problems can be appropriately mitigated before estimating the parameters of the model. For instance, a heavy-tailed returns modelled with normal errors may lead to estimated models which fail several diagnostic tests. It is against this background that we present this chapter. We discuss the source of the data, the reasons for choosing the data and some basic descriptions of the data. Detailed analysis and recommendations based on the empirical characteristics as well as the data validation tests are presented. All returns used in the study are continuously compounded.

4.2 Data

According to Bank of International Settlement (2019), the South African rand is the 18th world's most traded currency with trade share at par with other emerging economies like Russia, India, Brazil and Turkey. The rand recorded about $726 billion of the average daily turnover in April 2019, but, despite this huge trade share, the rand market has not received a fair share of attention in terms of volatility forecasting in comparison to developed and other emerging markets. As a contribution to literature on rand market, this study used the daily closing prices from the South African inter-bank spot forex market for all the empirical exercises.

The currencies considered include, the Swedish Kroner (SEK), United States Dollars (USD), Norwegian Kroner (NOK), Australian Dollars (AUD), Botswana Pula (BWP), Canadian Dollars (CAD), North Korean Won (KPW), Euro (EUR), British Pound Sterling (GBP), Brazilian Real (BRL), Israeli Shekel (ILS), Indian Rupees (INR), Malawian Kwacha (MWK) and the new Mozambican Metical (MZN), from July 6, 2011, to June 28, 2016. The data was sourced from OANDA (*https://www1.oanda.com*). All currencies are quoted in the South African Rand (ZAR). Due to data download restrictions at the time of sourcing, the currency pairs were selected, such that SADAC (Southern African Development Community) countries and countries from each continent, alongside with some of the most traded pairs are represented. The sample period was also chosen to coincide with 2011 geographical financial crisis. The year 2011 was particularly known for its geographical financial crisis where international credit markets showed signs of strain and economic growth around the world stalled, thus, this year was used as the starting point for the data collection. The idea was to include periods of structural breaks in as many of the currency pairs as possible because one of the reasons behind our proposed approach is the modelling of structural breaks.

All graphs, figures, data explorations and statistical analysis in this thesis were conducted using R packages except where specifically mentioned otherwise. Details of the packages and the codes used can be found in Appendix E.

4.3 Data exploration

In this section, we report and discuss the results of the descriptive analysis and correlation analysis.

Brooks (2008) argues that before modelling any data with GARCH type models, it is worthwhile to test for the presence of ARCH effects and stationarity of the GARCH innovations. In this regard, we report and discuss the ARCH Lagrange Multiplier (LM) test of Engle (1982) and the detrended-Dicky Fuller unit test. An ARMA (1, 1) filtration and 24 lags for the auxiliary regression equation were used for the ARCH LM test. The terms 'log-returns' and 'returns' are used interchangeably throughout this thesis.

4.3.1 Data exploration

Descriptive statistics are reported in Table 4.1. Within the period July 6, 2011, to June 28, 2016, it seems, on the surface, that MWK/ZAR is the most volatile currency among all the currency pairs with about 2.033 % daily volatility and average loss on investment of about 0.039 %. The most stable currency is AUD with daily volatility of 0.518% and average returns on investment of about 0.024%. KPW has the highest returns on investment (about 0.0449 %) and a daily volatility of 0. 658 %.

Figure 4.1 displays the log-returns of the selected currencies for the sampled period. Relatively, extreme negative returns appear to be more pronounced than extreme positive returns for BWP/ZAR, MWK/ZAR and MZN/ZAR currency pairs. This suggests volatility asymmetry in the log-returns, thus asymmetric models may be appropriate in modelling the log-returns. The results are confirmed by the maximum and the minimum and the negative skewness values reported in Table 4.1. These observations confirm the results the study by May & Farrell (2017), Taylor (2005) and Nelson (1991), but the observations from the other returns are in direct contrast to the results of these studies because they are positively skewed. All the returns are leptokurtic and have high Jacque-Bera test statistics which suggest that the returns are non-normal, thus agreeing with earlier results from May & Farrell (2017), Suess *et al.,* (2008), Cont (2001), Ding, Granger & Engle (1993), Fama (1965) and Mandelbrot (1963). Heavy-tailed or semi-heavy-tailed distributions are appropriate in modelling the volatilities of these returns, thus, the skewed general error distribution, the normal inverse Gaussian distribution and the Johnson SU parameterized distribution are used in modelling the GARCH innovations[4].

The skewed General Error Distribution is competitively good in modelling both long and short-tail properties of asset returns. When the true distribution is fat-tailed, estimates based on GED models are believed to be consistent (Bollerslev & Wooldridge, 1992) and yield more efficient results (Bollerslev, 1987). The normal inverse Gaussian distribution is able to simultaneously model skewness and asymmetry due to its semi-heavy tail properties. It is also closed under marginalization, conditionalization, and affine transformations; it can also capture dissimilar tail behaviours (Aas & Haff, 2006). The Johnson distribution is able to introduce time-varying mean and variance parameters (Cayton & Mapa, 2012), thus, it is able to deal with structural breaks in the returns or the innovations. The distributions can also accommodate an observed volatility smile.

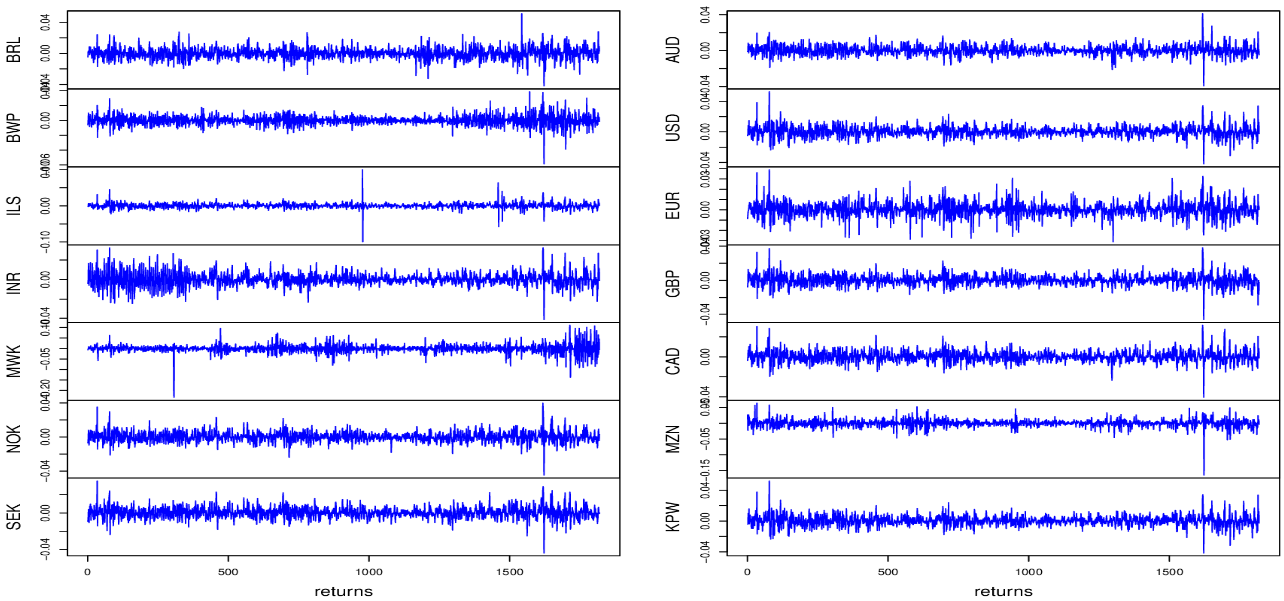

Figure 4. 1: Time plots of daily returns of rand-denominated currencies

[4]The ARMA-GARCH innovations are leptokurtic and non-normal. The innovations can be thought of as the differences between two *i.i.d* random variables (i.e. returns and mean corrected returns). The implications of the central limit theorem is that sum or differences between *i.i.d.* random variables to be approximately normal distribution as the number of random variables summed-up increases, the variables should not exhibit moment dependencies, such as serial correlation, conditional heteroscedasticity, asymmetric volatility among others. However, we know that the returns do not satisfy this assumption, hence they are non-normal. The innovations also inherit the leptokurtic properties of the returns as could be seen in Mandelbrot (1963), Fama (1965), Ding, Granger & Engle (1993) and Cont (2001) among other.

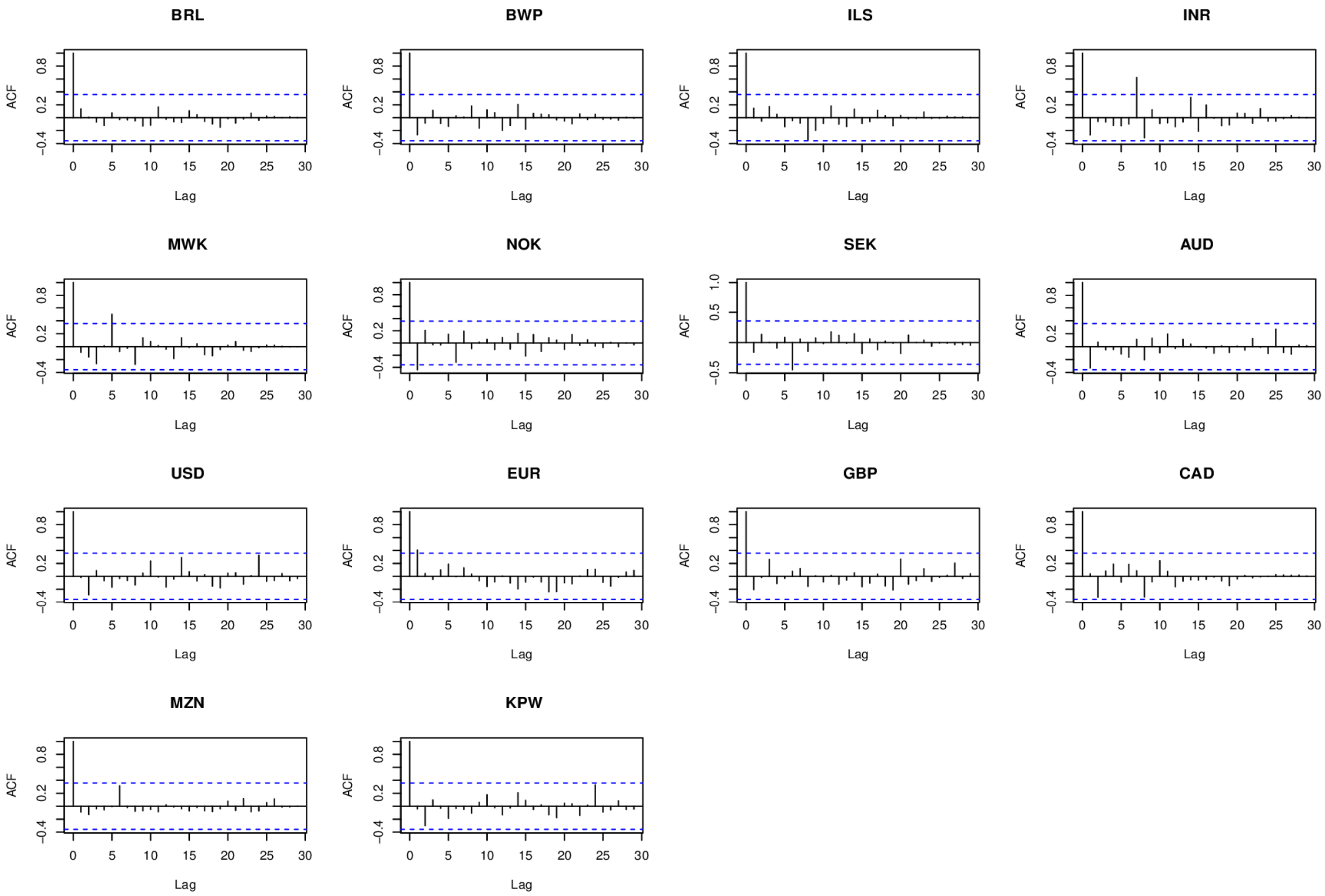

Figure 4. 2: Autocorrelation of squared returns

Table 4. 1: Descriptive statistics of the daily exchange rate returns

Currency	Mean	Std dev	Skewness	Kurtosis	JB Stat	Minimum	Maximum
AUD	0.000244	0.005182	0.260814	8.427690	2253.431	-0.03938	0.04097
BRL	0.000027	0.007122	0.146275	6.549638	96.45660	-0.04198	0.05104
BWP	0.000156	0.006438	-0.161508	10.18745	3923.260	-0.05863	0.03871
CAD	0.000283	0.005414	0.300243	7.803633	1776.214	-0.04050	0.03172
EUR	0.000304	0.007018	0.154027	6.134347	751.7798	-0.03122	0.03901
GBP	0.000346	0.006017	0.143827	8.165722	2028.748	-0.04613	0.03805
ILS	0.000378	0.007578	0.225563	42.71569	119564.3	-0.09997	0.10004
INR	0.000219	0.006543	6.730533	6.730533	1074.068	-0.04123	0.03258
KPW	0.000449	0.006580	0.387167	8.405653	2260.159	-0.04156	0.05207
MWK	-0.00039	0.020331	-1.27195	22.63265	29703.69	-0.23268	0.11123
MZN	0.000014	0.011597	-1.45600	28.65907	50543.02	-0.16503	0.06430
NOK	0.000198	0.005930	0.374495	7.887452	1853.034	-0.04480	0.03991
SEK	0.000283	0.006085	0.179220	6.448340	910.9799	-0.04383	0.03496
USD	0.000448	0.006578	0.399580	8.530526	2366.623	-0.04197	0.05247

4.3.2 Correlation analysis

A look at the time plots of the ZAR exchange rates as displayed in Figure 4.3 shows that most of the exchange rate pairs move in tandem. For instance, prices of GBP/ZAR, EUR/ZAR, AUD/ZAR, and USD/ZAR move closely together. The price-level co-movements translate into return-level co-movements as observed from Table 4.2. All the correlation coefficients are significant and positive. This suggests that during the period under consideration, on average, shocks on the South African inter-bank forex market yielded the same directional effects on the currency pairs, but with different magnitudes.

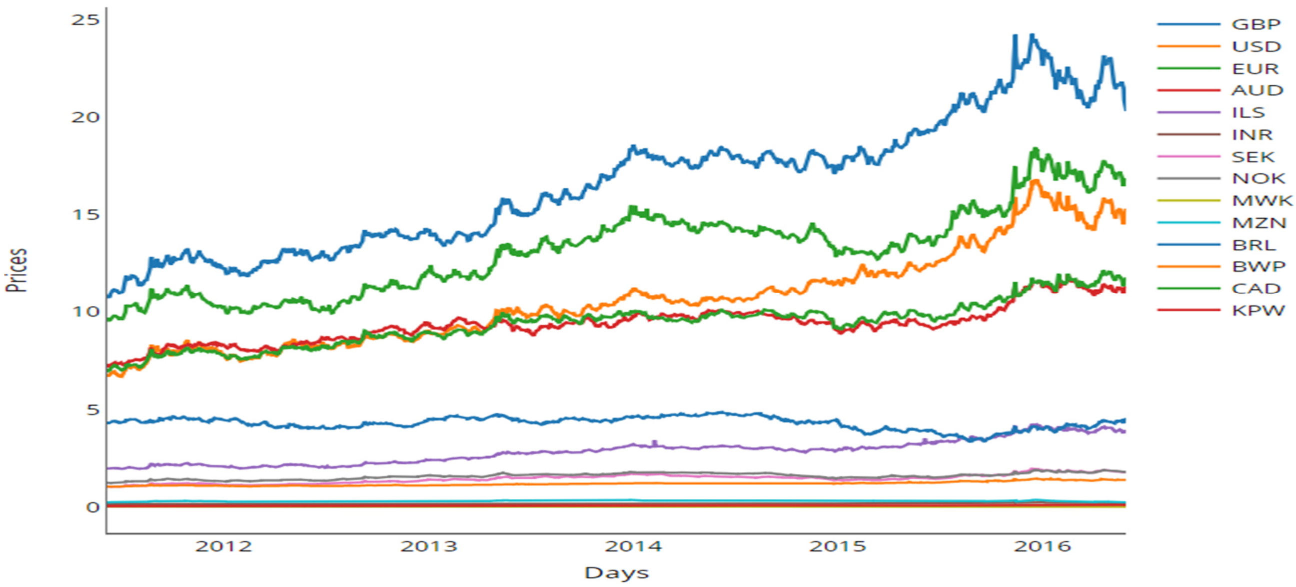

Figure 4. 3: Time plots of daily returns of rand-denominated currencies

As observed in Table 4.2, a substantial number of the currency pairs exhibit strong return-level co-movements which is in conformity with results from literature (Park & An, 2020; Marconi, 2018 and Kühl, 2008). The underlying factors of these strong co-movements may be attributed to stronger bilateral trade and financial linkages to the South African economy and greater flexibility in exchange rate regimes as observed by Park & An (2020) for renminbi-denominated currencies. Park & An (2020) conclude that the growing co-movements of renminbi-denominated currencies are the consequence of autonomous decisions at the market rather than that of management by governments or central banks, and this may be the case for the rand-denominated currencies. The implications of the exchange rate returns' co-movements suggest that exchange rate returns may share substantial mutual entropies. Furthermore, it is an indication that substantial levels of total variabilities in the returns of a particular rate may be explained by a set of returns of other rates, thus, may have some forecasting abilities.

Table 4. 2: Correlation coefficient of daily returns

	GBP	USD	EUR	AUD	ILS	INR	SEK	NOK	MWK	MZN	BRL	BWP	CAD
GBP	1												
USD	0.84	1											
EUR	0.64	0.64	1										
AUD	0.70	0.70	0.50	1									
ILS	0.67	0.72	0.53	0.53	1								
INR	0.70	0.77	0.53	0.68	0.61	1							
SEK	0.79	0.75	0.69	0.70	0.64	0.66	1						
NOK	0.76	0.71	0.64	0.70	0.62	0.66	0.85	1					
MWK	0.28	0.35	0.24	0.26	0.28	0.27	0.27	0.27	1				
MZN	0.48	0.59	0.36	0.42	0.44	0.48	0.43	0.43	0.21	1			
BRL	0.45	0.49	0.33	0.47	0.39	0.48	0.45	0.46	0.18	0.30	1		
BWP	0.60	0.62	0.46	0.51	0.51	0.53	0.59	0.588	0.27	0.34	0.34	1	
CAD	0.81	0.86	0.60	0.78	0.67	0.72	0.75	0.76	0.30	0.52	0.50	0.58	1
KPW	0.84	1.00	0.64	0.70	0.72	0.76	0.75	0.71	0.35	0.58	0.49	0.61	0.85

All correlation coefficients are significant at 5%.

4.3.3 ARCH effects

The ARCH LM test is simple to use, however, the actual statistical size of the test is generally less than its nominal size in finite samples (Bollerslev & Wooldrige, 1992), therefore, to ensure that the right conclusion of heteroscedasticity is arrived, the weighted Ljung-Box test is used in conjunction with the ARCH LM test. The weighted Ljung-Box test is also easy to implement but tends to outperform the ARCH LM test (Fisher & Gallagher, 2012). The results of the ARCH Lagrange Multiplier test and the weighted Ljung-Box test are reported in Table 4.3. Both tests reject the null hypotheses of no heteroscedasticity for all the return series. This is an indication of the presence of ARCH effects in all the returns. The test conclusion is confirmed by the time-varying volatility clustering of the log-returns displayed in Figure 4.1 and the gradual decline and persistence of autocorrelation of the square returns, displayed in Figures 4.2. The results confirm those from May & Farrell (2017), Ding, Granger & Engle (1993), and Mandelbrot (1963).

Table 4. 3: ARCH Lagrange Multiplier test

Returns	Weighted Ljung-Box Test		Lagrange Multiplier Test	
	statistic	Prob.	LM-Statistic	
AUD	183.59	<0.00001	210.0	<0.00001
BRL	72.717	<0.00001	245.0	<0.00001
BWP	314.02	<0.00001	210.0	<0.00001
CAD	135.44	<0.00001	263.0	<0.00001
EUR	177.52	<0.00001	194.0	<0.00001
GBP	193.44	<0.00001	255.0	<0.00001
ILS	354.05	<0.00001	231.0	<0.00001
INR	251.85	<0.00001	160.0	<0.00001
KPW	102.46	<0.00001	318.0	<0.00001
MWK	279.75	<0.00001	285.0	<0.00001
MZN	21.362	0.009519	1499	<0.00001
NOK	220.72	<0.00001	211.0	<0.00001
SEK	162.83	<0.00001	206.0	<0.00001
USD	106.56	<0.00001	323.0	<0.00001

Note: *Test is significant at 1% and 5%. In all the test, ARMA (1, 1) filter was used.*

4.3.4 Unit root or stationarity tests

The de-trended Dicky Fuller unit root test results are reported in Table 4.4.

Table 4. 4: Dickey-Fuller DLS test for unit root

Test	**Constant**		**Constant and Trend**		**None**	
Variable	Statistic	*p*-value	Statistic	*p*-value	Statistic	*p*-value
GBP	-7.6624	<0.01	-7.7052	<0.01	-7.3778	<0.01
USD	-8.3747	<0.01	-8.3747	<0.01	-7.8707	<0.01
EUR	-7.5739	<0.01	-7.5735	<0.01	-7.3565	<0.01
AUD	-8.0646	<0.01	-8.0797	<0.01	-7.7992	<0.01
ILS	-8.3552	<0.01	-8.3517	<0.01	-7.8888	<0.01
INR	-8.3108	<0.01	-8.3504	<0.01	-8.1618	<0.01
SEK	-7.8092	<0.01	-7.8281	<0.01	-7.5862	<0.01
NOK	-7.8758	<0.01	-7.9042	<0.01	-7.7577	<0.01
MWK	-7.3907	<0.01	-7.4118	<0.01	-7.2929	<0.01
MZN[1]	-8.8184	<0.01	-9.2272	<0.01	-8.8229	<0.01
BRL	-8.0640	<0.01	-8.0724	<0.01	-8.0659	<0.01
BWP	-8.7782	<0.01	-8.7864	<0.01	-8.4544	<0.01
CAD	-8.4234	<0.01	-8.4335	<0.01	-8.0645	<0.01
KPW	-8.3925	<0.01	-8.3924	<0.01	-7.8880	<0.01

Unit root is present if test statistic is less than 5% (i.e. reject null hypothesis).

All the p-values of the different tests for each log-returns are less than the 5% level of significance, hence, the hypotheses of no unit roots in the various log-returns are rejected. We, therefore, conclude that significant unit roots are present in all the returns, which is an indication of stationarity in the returns. The test results are in agreement with the results from May & Farrell (2017).

4.4 Conclusion

In this chapter, we presented descriptive statistics and performed data validation tests on the returns. The test results and the descriptive statistics were detailed. It was observed that currencies move closely together at price-level and this translates into significant positive co-movements at return-level. In addition, a substantial number of the currency pairs were found to exhibit strong co-movements, thus, may have some forecasting abilities. The data validations tests showed that all the return series are heteroscedastic, stationary and exhibit volatility clustering which suggests the suitability of using GARCH models in modelling the volatility of all the returns. In particular, the returns from BWP/ZAR, MWK/ZAR and MZN/ZAR exhibit volatility asymmetry, thus, asymmetric models may be appropriate for these models. Since the descriptive statistics revealed non-normality in the returns, distributions capable of capturing heavy-tails such as the skewed student-t distribution, the skewed general error distribution, the normal inverse Gaussian distribution and the Johnson SU parameterized distribution are recommended for modelling the innovations of the returns.

CHAPTER 5

Autoregressive relationships between returns of exchange rates

Chapter Summary

In this chapter, we re-examined the relationships between exchange rate returns by assuming that the relationships evolve via the ARMA process. In addition, we assume that the innovations from the ARMA follow GARCH processes. The statistical significance of the estimated empirical models is tested using the likelihood ratio test. The likelihood ratio tests confirm the non-triviality of the estimated models and the plausibility of the ARMA assumption. The estimated models were found to be more accurate in approximating the relationships between the returns for the period under consideration, thus, confirming the well-established exchange rate relationships. All the estimated parameters for the exogenous returns are positive and consistent with the respective direction of their co-movements with the endogenous returns. Higher levels of linear co-movements, however, do not necessarily translate into larger magnitudes of impacts. Path analysis of the impacts of the returns confirmed that currency pairs are not isolated on the forex market. Furthermore, shocks of the same magnitude from the same origin, transmitted along different paths on the market have different impacts.

5.1 Introduction

In finance and economics, empirical or phenomenological studies of relationships require the support of confirmatory data without necessarily a theoretical basis; such studies play a major role in economics and finance. They guide investors to identify possible profitable trading strategies (Çiçek, 2004) and provide direction to policy-makers (Kühl, 2007). Furthermore, these studies provide an opportunity to understand the nature of relationships among financial and economic variables (Holmes, 2004) and the identification of important factors affecting specific institutions (Antwi & Kyei, 2017) or some key areas of the various markets or the economy as a whole (Divya & Devi, 2014). They have also been used to test theories (Fama, 1965) and to identify factors to improve predictive or forecasting abilities of models (Kyei & Antwi, 2018 and Kosapattarapim, 2013). There are several empirical studies relating to exchange rate relationships (Firoj & Khanom, 2018; Kuhl, 2007; Wickremasinghe, 2004; Diebold, Gardeazabal, & Yilmaz, 1994 and Baillie &

Bollerslev, 1989), but much of these are indirect and concentrated on testing the efficient market hypothesis of Fama (1965) and the investigation of cointegrating relationships. Predominant methodologies used in the studies aimed at testing the efficient market hypothesis or investigating cointegrating relationships, have not considered the autoregressive framework, although, such frameworks may have applications in volatility modelling, a gap in the literature still exists. In this chapter, we re-examine the relationships between exchange rate returns via an ARMA framework and make recommendations based on their potentials to predict returns and volatilities. In re-examining the relationship, we assume that the innovations from the ARMA process follow GARCH processes.

The rest of the chapter is structured as follows. We provide a brief literature review in section 5.2, while theoretical and empirical backings for the use of ARMA models are presented in section 5.3. Data and methodology for the empirical study are introduced in section 5.4. In section 5.5, the empirical results are presented and discussed, while the procedures and the main findings of the study are summarised in section 5.6.

5.2 Empirical Exchange rate relationships

Weak-form testing of efficient market hypothesis on exchange rate markets is limited to market price informational set. The information set comprises of current and past prices. Testing approaches that rely on Granger's approach (Engle & Granger, 1983) use the information set of, at least, two exchange rates. Granger's approach relies on Granger's representation theorem. According to the theorem, two or more time series are cointegrated if they fluctuate conjointly in a long-run relationship. In other words, two or more time series are cointegrated if they have the same order of integration and a stationary linear combination of these time series exists (Kühl, 2007). In estimating cointegration relationships, Engle & Granger (1987) proposed a two-step procedure. In the first step, the first series y_t is regressed on a second series x_t and the resulting residuals z_t are tested for stationarity in the second step. Statistically, the first procedure can be represented as:

$$y_t = \alpha + \beta x_t + z_t . \tag{5.1}$$

If z_t is stationary, then y_t and x_t are cointegrated (that is, a long-run relationship exists between

y_t and x_t), therefore, any study on exchange rate market efficiency that employs the cointegration approach of Engle & Granger (1987) or its modifications thereof, is an attempt to indirectly study exchange rate relationships. Extant literature supports exchange rate relationships at price level via procedures that integrate the framework of regression, but not at returns level. In the subsequent paragraphs, we review some of these studies.

Baillie & Bollerslev (1989) studied cointegration of seven United States dollar-denominated currency pairs: Canadian Dollar, French Franc, Deutsche Mark, Italian Lira, Japanese Yen, Swiss Franc, and the British Pound via Johansen test of cointegration using nominal daily spot prices. Their study unravelled evidence of cointegration among the United States dollar-denominated currency pairs. They inferred from the observations that spot exchange rate movements must be at least partly predictable and deviations from their long-run relationship can be used in predicting future exchange rates. Using the same dataset by Baillie & Bollerslev (1989), Diebold, Gardeazabal & Yilmaz (1994), however, disputed the findings. They argue that cointegration model is sensitive to the presence of a drift assumption in the data, hence, tests with or without intercepts will result in different inferences. They concluded that there is a lack of predictive power for an out-of-sample experiment with the cointegration model. Further studies by Baillie & Bollerslev (1994), however, re-confirmed the existence of cointegration relationships among the same dataset used in their 1989 studies, but noted that the cointegrating relationships have a long memory (fractional cointegration).

Wickremasinghe (2004) examined the market efficiency of Sri Lankan foreign exchange market. The study tested weak and semi-strong form of efficiency using six bilateral exchange rates. Weak-form efficiency was examined via unit root tests while cointegration, Granger causality tests, and variance decomposition analysis were used for the semi-strong forms of efficiency. Wickremasinghe observed that 2% of the currency pairs considered were cointegrated, thus, there is evidence that the Sri Lankan foreign exchange market is weak-form efficient, although, there was evidence against the semi-strong efficiency of the Sri Lankan foreign exchange market. This result also suggests the predictability of some of the exchange rates.

Kuhl (2007) investigated the efficiency of foreign exchange rates since the introduction of the Euro. The study concentrated on bivariate cointegration analysis of daily exchange rates by drawing on the cointegration approaches of Johansen and Gregory-Hansen. The study covered the period when the Euro was introduced and the most recent period of floating exchange rates thereafter and, although, the hypothesis of no cointegration was not rejected for most bivariate exchange rates, it was rejected for seven bivariate rates. The estimation of the vector error correction model provided evidence, however, this was only consistent with market efficiency in two cases. On the application of the Gregory-Hansen cointegration test to the bivariate pairs, the hypotheses of no cointegration were rejected for the pairs EUR-GBP, EUR-SEK, GBP-SEK, and EUR-AUD. The result from the Johansen's approach in relation to the EUR-GBP pair was both convincing, therefore, breakpoint estimation by the Gregory-Hansen approach was used to generate two sub-samples and the Johansen test was then applied to the sub-samples. The result was more robust as it confirmed the rejection of cointegration in the EUR-GBP pair. The study concluded that long-run relationship exists between EUR-USD and GBP-USD bivariate pairs and that the introduction of the Euro has not resulted in an inefficient market due to the few numbers of cointegrating bivariate exchange rates.

Firoj & Khanom (2018) investigated the weak-form and semi-strong form of the efficient market hypothesis (EMH) for Bangladesh using the daily average rate of seven bilateral exchange rates: AUD, CAD, EUR, GBP, JPY, SEK, and USD from 2010/01/01 to 2017/11/30. They used the Augmented Dickey-Fuller, Phillips-Perron, Kwiatkowski–Phillips–Schmidt–Shin tests of stationary to examine the weak-form of the EMH and the Johansen cointegration test for the semi-strong form of the EMH. The results revealed that the seven exchange rates are random walks, however, the cointegration test indicates the existence of long-run relationship between the currency pairs. The above studies and many others (Alexander & Johnson, 1992; Cheung, 1993; Beran, 1994; Crowder, 1994; Caporale & Gil-Alana, 2004 and Belkacem, Meddeb & Boubaker, 2005) undoubtedly provide evidences of exchange rate relationships, but none of them directly used continuously compounded returns alongside with the ARMA-GARCH framework.

5.3 Exogenous ARMA (ARMAX) model

Empirical studies from Suess *et al.,* (2008) and Cont (2001) among others, suggest that exchange rate returns are second-order stationary process. A natural consequence of a fundamental result from Wold (1938) implies that any centred, second-order stationary process $\{R_t\}$ and '*purely non-deterministic*' process admit an infinite moving-average (MA) representation of the form:

$$R_t = \varepsilon_t + \sum_{i=1}^{\infty} \theta_i \varepsilon_{t-i} \,, \tag{5.2}$$

where $\{\varepsilon_t\}$ is a linear innovation process of $\{R_t\}$ defined by:

$$\varepsilon_t = R_t - \mathrm{E}\left(R_t \mid \mathrm{H}_R\left(t-1\right)\right) \text{ and } \mathrm{H}_R\left(t-1\right).$$

The linear innovation process denotes the Hilbert space generated by the random variables $R_{t-1}, R_{t-2}, \ldots$ and $\mathrm{E}\left(R_t \mid \mathrm{H}_R\left(t-1\right)\right)$ which denotes the orthogonal projection of R_t onto $\mathrm{H}_R\left(t-1\right)^2$. The coefficients $\{\theta_i\}$ are such that $\sum_i \theta_i^2 < \infty$. The MA representation (5.2) can be truncated at p orders, although, the truncated MA model is not usually preferred because it is not parsimonious. A parsimonious representation of the p truncated MA model is the autoregressive moving average (ARMA) model.

Definition 5.1: A second-order stationary process $\{R_t\}$ is called *ARMA (p, q)*, where *p* and *q* are integers, if there exist real coefficients $\varphi_0, \varphi_1, \varphi_2 \ldots \varphi_p, \theta_1, \theta_2 \ldots, \theta_q$ such that:

$$R_t = \varphi_0 + \sum_{i=1}^{p} \varphi_i R_{t-i} - \sum_{i=1}^{q} \theta_r a_{t-r} + a_t \,. \tag{5.3}$$

Definition 5.2: Consider the continuously compounded exchange rate returns $\{R_t\}$ and r_{jt} where $j = 1, 2, \ldots, m$ and m is the total number of currency pairs. If $\{R_t\}$ and r_{jt} are *second-order stationary* cointegrated *processes*, then there exist integers *p*, *q* and real coefficients φ_0, $\lambda_1, \lambda_2, \ldots, \lambda_m$, $\varphi_1, \varphi_2, \ldots, \varphi_p$ and $\theta_1, \theta_2, \ldots, \theta_q$ such that:

$$R_t = \varphi_0 + \sum_{j=1}^{m} \lambda_j r_{jt} + \sum_{i=1}^{p} \varphi_i R_{t-i} - \sum_{i=1}^{q} \theta_r a_{t-r} + a_t . \tag{5.4}$$

Equation (5.4) is called an *ARMAX (p, q, m)*. The a_t term represents the innovation. Definition 5.2 is a conglomeration of definition 5.1 as adapted from Tsay (2010) and the cointegration definition adapted from Engle & Granger (1987).

5.4 Data and Methodology

Daily exchange rate returns discussed in chapter four are used in the study and since the focus of this chapter is not on forecasting, the entire sample length is used in the estimation. The ARMA model in equation (5.4) is used in the empirical study. By definition, the innovation of the model is a weak-stationary process. This means that the:

i) first moment is time or shift-invariant;

ii) second moment and the variance are finite at all times, and

iii) cross moment (auto-covariance) at every lag has an associated constant covariance value associated with it.

There seem to be no time-invariant second moment imposed by the weak stationarity assumption, although, there is a further imposition of a *Gaussian white noise process* assumption on the ε_t by definition, which renders the variance of the innovations time-invariant. Empirical studies from Suess *et al.,* (2008) and Cont (2001) among others, however, have uncovered several characteristics associated with the variance of mean corrected returns of speculative assets, such as serial dependency, conditional heteroscedasticity (volatility clustering), asymmetry and heavy-tail, which are contrary to the Gaussian white noise process assumption. GARCH models are able to capture these peculiar characteristics, hence, the volatility of the innovations in (5.4) are modelled with the standard GARCH, EGARCH and GJRGARCH models. The choice of these models is based on their abilities to capture the empirical regularities specific to the data in this study.

It was observed from the descriptive analysis in chapter four that the returns for all currencies are non-normally distributed, thus, the skewed generalized error distribution is used to model the innovations. This distribution is able to capture the heavy-tail and leptokurtic characteristics exhibited by the return series. The QMLE approach is used to estimate the parameters and for computational difficulties associated with this methodology, variance targeting is used. The likelihood ratio test is used to investigate the composite impact of the covariate or exogenous returns in the models.

5.5 Empirical results and discussions

In this section, we report and discuss the empirical results. Details of the specifications of the estimated models are shown in Table 5.1. In the likelihood test, models without the exogenous returns (restricted) are used to benchmark the models with the exogenous returns (unrestricted) in assessing the composite significance of the exogenous returns.

Table 5. 1: Specifications of estimated ARMA-GARCH models [5]

Model	Returns	Exogenous regressors	GARCH	ARMA	Target	Distribution
EGARCH	SEK	AUD,NOK, BWP	(2,2)	(5,5)	YES	SGED
GJRGARCH	USD	AUD,NOK,SEK,MZN	(1,1)	(1,1)	YES	SGED
SGARCH	NOK	CAD,SEK,KPW	(2,2)	(6,2)	YES	SGED
EGARCH	SEK	None	(3,2)	(2,1)	YES	SGED
GJRGARCH	USD	None	(2,1)	(2,2)	YES	SGED
SGARCH	NOK	None	(1,1)	(5,3)	YES	SGED

Likelihood Ratio test

The results of the likelihood ratio tests are reported in Table 5.2. The test statistics are very large with probabilities far less than either 1% or 5%. The null hypotheses are therefore rejected, and we conclude that the unrestricted models are significantly better than the restricted models. The conclusions suggest that the ARMA relationships between the returns of the exchange rates are empirically non-trivial, hence, the relationship could be further used to investigate the impacts of

[5] Although exchange rates share substantial mutual entropies among themselves, including highly correlated pairs in the ARMA specification introduces serial correlation and autocorrelations, which may require higher order ARMA and GARCH terms to remove which violates the principle of parsimony. To avoid this problem, a pre-modelling via the backward feature selection approach of regression was used to select the appropriate regressors to be included in the ARMA-GARCH model.

the exogenous returns on the endogenous returns, provided the models pass the diagnostics tests.

Table 5. 2: Likelihood ratio test

Model	EGARCH	GJRGARCH	SGARCH
Test statistic	1986.02	2503.437	2544.086
p-value	<0.0001	<0.0001	<0.0001

Goodness-of-fit assessment

It is imperative to conduct diagnostic checks to assess the goodness-of-fit for any estimated statistical model before it used to make any valid deductions and conclusions. In estimating the parameters of a stationary GARCH process, the persistence parameter is restricted to be less than unity, hence, it is imperative to check if this condition is met. In addition to this restriction, if a GARCH model is a good fit to a data, there should be no more remaining serial correlation in the standardized residuals and no more ARCH effects in both the standardized and squared residuals. The empirical distributions of the residuals should be also a good approximation of the assumed theoretical distribution. The results of the diagnostic tests are reported in Table 5.3.

It is observed that all the estimated models satisfy the restriction conditions on their respective persistence parameters. The adjusted-Pearson goodness-of-fit tests have p-values, which are greater than 0.01, thus at 1% level of significance, the empirical distributions of the standardized residuals from the estimated models are deemed good approximations of the respective assumed theoretical distributions. The weighted Ljung-Box, weighted McLeod-Li, and weighted Li & Mark tests are all significant at 1% because the associated p-values are less than 0.01, therefore, no significant serial correlation and autocorrelation (ARCH effects) remain in the residuals and the squared residual. The reported adjusted R-squares reveal that the respective regressors explained at least 70% of the total variations in the respective endogenous returns. Out of these variations, the exogenous returns accounted for 86% of the variations in the NOK/ZAR returns, 90% of the variations in the SEK/ZAR returns and 84% of the variations in the USD/ZAR returns. The high levels of explained total variations in the endogenous returns indicate that the strength of the relationship among the exchange rate returns is in agreement with the observed high co-movements observed in chapter four; it is also in conformity with the result from Park & An (2020) and Kühl

(2008). Based on the results of the individual diagnostic tests discussed above, it can, reasonably, be asserted that the models are adequate fits to the data.

Table 5. 3: Diagnostic checks

Tests	ARMAX-GARCH		ARMAX-EGARCH		ARMAX-GJRGARCH	
	t-value	*p*-value	t-value	*p*-value	t-value	*p*-value
Weighted Ljung-Box	21.780	0.28221	24.0864	0.57950	4.802	0.49938
weighted McLeod-Li	11.491	0.3061	18.1436	0.02842	6.900	0.20770
Weighted Li & Mark	2.0152	0.7555	8.65600	0.05144	5.4879	0.17960
Adjusted Pearson Gof	49.91	0.4369	40.1800	0.81120	39.69	0.82630
Metric	ARMAX-GARCH		ARMAX-EGARCH		ARMAX-GJRGARCH	
Persistence	0.8540191		0.9921722		0.9969615	
Adj Rsq	0.781[0.669]		0.701[0.630]		0.766[0.647]	

Note: *Gof means Goodness-of-fit. Test is significant at 1%. Values in square brackets indicates differences between the unrestricted and the restricted models.*

Accuracy assessments for the estimated models

Models are not complete representations of realities or situations, but mere approximations, thus it is necessary to assess the error of approximation to determine how accurate an estimated model is able to approximate the intended reality. The mean absolute prediction error (MAPE) and the root mean square prediction error (RMSPE) metrics are used to assess the accuracy of the estimated mean relationships. The MAPE and the RMSPE metrics were computed using prediction errors from the ARMA/ARMAX portions of the estimated models because the study is interested in the first moment relationship. Since it is very difficult to predict the directions of returns (Francq & Zakoıan, 2010), we also computed the percentages of correctly-predicted directions (CPD) to assess how well the models are able to predict the directions of the returns. The results are reported in Table 5.4.

The MAPE and the RMSPE values from the estimated ARMAX (ARMA model with exogenous regressors) models are relatively lower than the ARMA models. Comparatively, there is a substantial average reduction of approximately 48% in these values. The CPD percentages of the ARMAX model are quite higher than those from the ARMA models. The tight tracking of the paths of the realized returns by the fitted returns from the ARMAX models, as displayed in Figure 5.1 and the less tight tracking of the paths of the realized returns by the fitted returns from the ARMA models also displayed in Figure 5.2, confirm these results. Based on the assessments above, it can be concluded that the ARMAX models are relatively accurate in approximating the relationships between the returns for the sample period under consideration.

Table 5. 4: Statistical metrics for restricted and unrestricted models

Parameter	% CPD	MAPE	RMSPE
ARMAX-SGARCH	75.9208	0.00156	0.00196
ARMA-SGARCH	59.3183	0.00321	0.00404
ARMAX-EGARCH	79.7691	0.00167	0.00210
ARMA-EGARCH	60.0330	0.00326	0.00399
ARMAX-GJRGARCH	73.3370	0.00214	0.00266
ARMA-GJRGARCH	56.5700	0.00373	0.00477

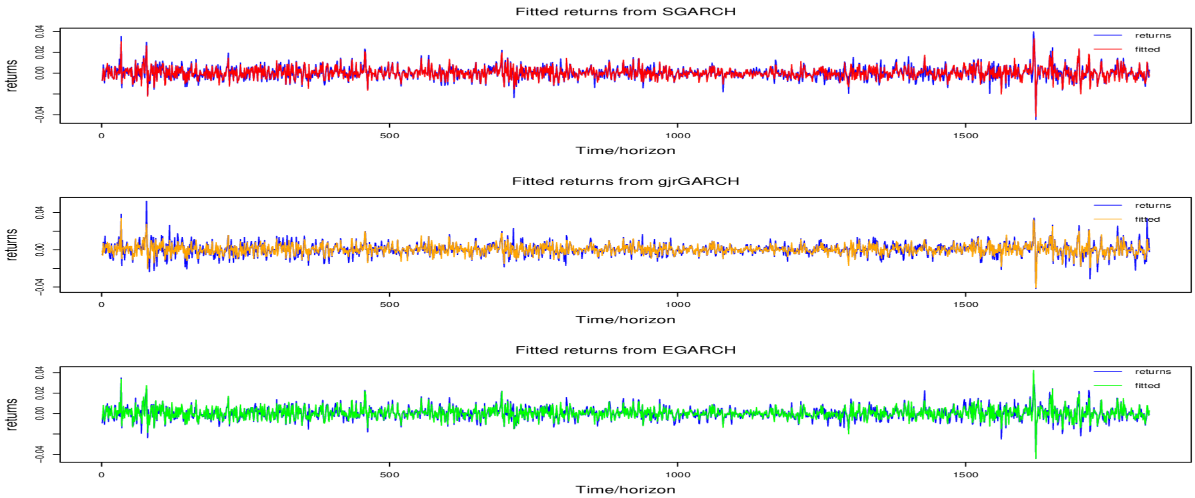

Figure 5. 1: Time plots of fitted returns from ARMAX models versus realized returns

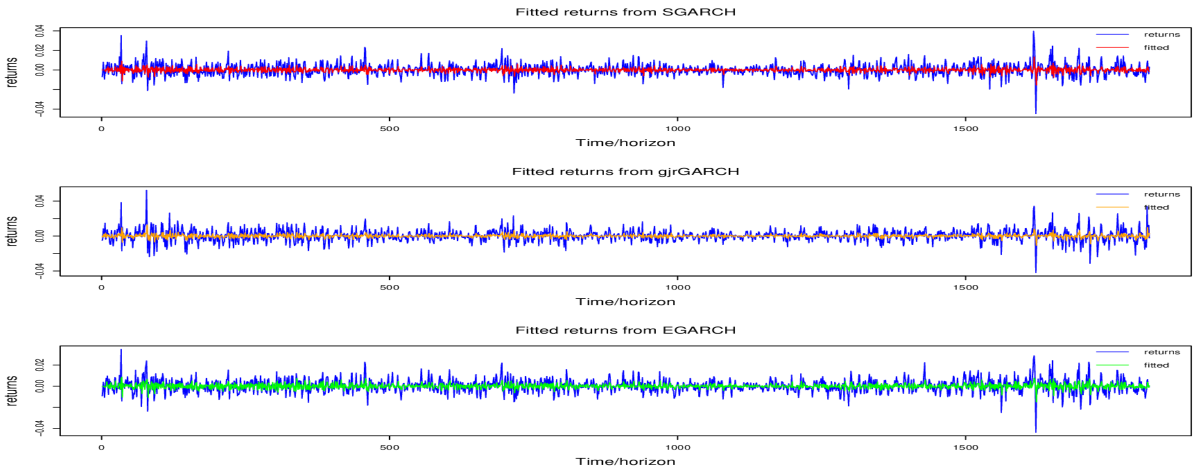

Figure 5. 2: Time plots of fitted returns from ARMA models versus realized returns

Discussions

In Table 5.5, the estimated parameters for the exogenous returns are reported, whereas the parameters for the ARMA and GARCH terms are reported in Appendix B. The standardized errors of estimations for all the parameters are reasonably small. All the estimated parameters for the exogenous returns are positive, significant, and consistent with the directions of their co-movements. The estimated direction of impacts agree with the impact of GBP/USD and Yen/USD on INR/USD from the studies conducted by Gadwala & Mathur (2014), but is in contrast to the effects of EUR/USD on INR/USD from the same study. In addition, the results indirectly support exchange rate co-movements unearthed by Park & An (2020), Marconi (2018) and Kühl (2008).

In conclusion, the results suggest that currency pairs can be expressed as a linear combination of other currency pairs. They further imply that the appreciations of the prices of the exogenous currency pairs lead to the appreciations in the endogenous pairs and vice versa. Higher levels of linear co-movements, however, do not necessarily translate into larger magnitudes of impacts as observed by a comparison of the standardized estimates with the correlation coefficients in Table 5.5. This observation may be explained by the fact that exchange rate co-movements is nonlinear, thus, its relationship with the estimated impacts may be also nonlinear. Among the currency pairs considered, one standard deviation shock from the Swedish market is likely to have more impact on the Norwegian and the USA markets than it has on the other respective markets. Similarly, one standard deviation shock from the Norwegian market has more impact on the Swedish market than it has on the Australian and the Botswana markets.

Table 5. 5: Estimated parameters for ARMAX-GARCH models and correlation coefficients

Model	Parameter	Estimate	Std error	Stdd est	Coref	*p*-value
ARMAX-SGARCH (NOK)	CAD	0.285640	0.001802	0.3129	0.76	<0.001
	SEK	0.643449	0.004097	0.6271	0.71	<0.001
	KPW	0.018256	0.000672	0.0165	0.85	<0.001
ARMAX-EGARCH (SEK)	AUD	0.161471	0.001809	0.16578	0.70	<0.001
	NOK	0.725405	0.000774	0.74436	0.66	<0.001
	BWP	0.084076	0.002531	0.07946	0.51	<0.001
ARMAX-GJRGARCH (USD)	SEK	0.382044	0.013765	0.48494	0.75	<0.001
	AUD	0.303781	0.012976	0.32841	0.70	<0.001
	NOK	0.129583	0.010849	0.14375	0.71	<0.001
	MZN	0.156945	0.006888	0.08902	0.59	<0.001

Stdd denotes standardised estimate and Coref denotes correlation coefficient. Test is significant if p-value<0.05.

A directed acyclic graph for the impacts among the returns is displayed in Figure 5.3. The graph depicts the interaction of impacts on the exchange rate market, thus confirming the statement "*no currency pair exists in isolation on the market*." The transmission of significant shocks originating from intra-market or inter-market events are initiated by the forces of demand and supply, which are then absorbed by market microstructures such as order flow (Kleinbrod & Li 2017). The market microstructures then transmit the shock across the currencies in the market, possibly through one of the many network paths, such as the one indicated in Figure 5.3.

The effect of a shock of 1% magnitude emanating from the BWP[6] market transmitted along the path $BWP \rightarrow SEK \rightarrow USD$ has 0.03212 impact on the returns of USD (i.e., 0.382044×0.084076), while the same magnitude of shock has 0.00701 impact on the USD market (i.e. $0.084076 \times 0.643449 \times 0.129583$) if it is transmitted along the alternative path $BWP \rightarrow SEK \rightarrow NOK \rightarrow USD$. Similarly, a shock of the same magnitude from the AUD market has 0.06398 impact on the USD market if it is transmitted along the path $AUD \rightarrow SEK \rightarrow USD$, however, if the same magnitude of shock is transmitted along the path of $AUD \rightarrow SEK \rightarrow NOK \rightarrow USD$, it has 0.01346 impact on the USD market. A shock of 1% magnitude from the CAD market transmitted along the $CAD \rightarrow NOK \rightarrow USD$ path has 0.03701 impact on the USD market but a 0.07916 impact on the same market, if transmitted via the $AUD \rightarrow SEK \rightarrow NOK \rightarrow USD$ path. A shock of 1% magnitude originated from the KPW market has 0.00237 impact on the USD market via the $KPW \rightarrow NOK \rightarrow USD$ path but a 0.00506 impact on the same USD market if transmitted along the path $KPW \rightarrow NOK \rightarrow SEK \rightarrow USD$.

It can be deduced from the foregoing analysis that shock transmission along different market paths has different effects on a particular market. The conclusion may explain why events from developing markets have less effect on developed markets than the effects developed markets

[6] The of BWP/ZAR is simply denoted BWP. This is done for all the other currency pairs hereafter,

have on developing markets.

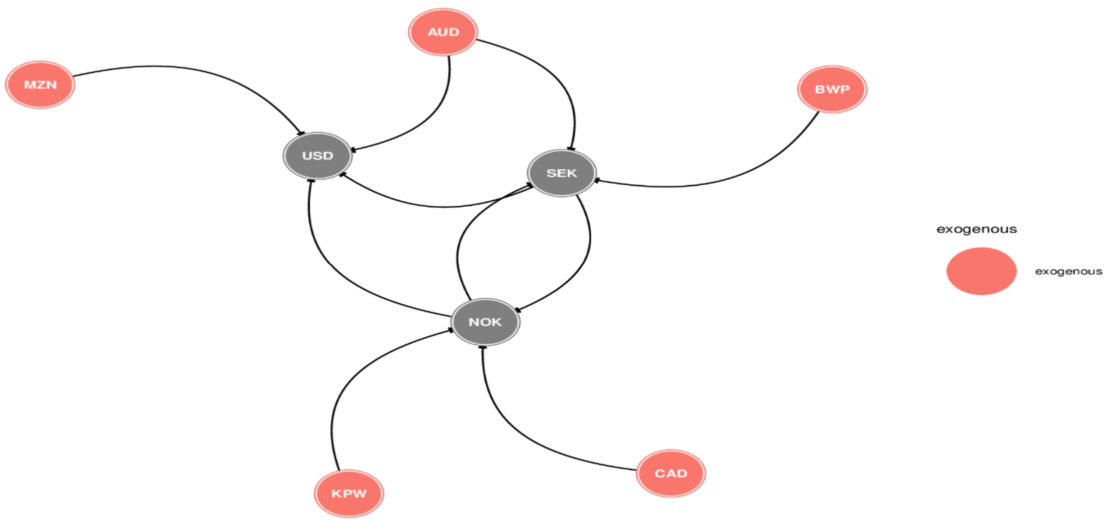

Figure 5. 3: A directed acyclic graph for inter-exchange rate dependencies (endogenous variables are shaded in grey)

5.6 Conclusions

In this chapter, we re-examined the relationships between exchange rate returns using ARMAX-GARCH framework. The statistical significance of the estimated empirical models was tested using the likelihood ratio test. Results from the likelihood ratio tests confirm the significance of the estimated models. The respective exogenous exchange rate returns accounted for at least 86% of the 70% total average variations explained by the collective regressors in each of the estimated models. The estimated models were found to be more accurate in approximating the relationships between the returns for the period under consideration. These results indirectly confirm the well-established price-level exchange rate relationship. All the estimated parameters for the exogenous returns are positive and consistent with the respective direction of their co-movements, however, higher levels of linear co-movements do not necessarily translate into larger magnitudes of estimated impacts. Path analysis of the impacts of the returns confirmed that currency pairs are not isolated on the market. Furthermore, shocks of the same magnitude from the same origin transmitted along different paths on the market have different impacts. This observation may explain why events from developing markets have less effect on developed markets than the effect the developed markets have on the developing markets.

CHAPTER 6

Alternative market uncertainty proxies for modelling time-variations in the unconditional volatility of exchange rates

Chapter Summary

In modelling time-variations and breaks in the unconditional volatilities of GARCH models, VIX and SIXVX have been used prominently, however, these proxies may be inadequate because they do not always take into account certain market factors. They are also single measures that attempt to represent a basket of market uncertainties, thus in this chapter, we attempt to employ a theoretical framework to identify alternative proxies that are likely to overcome these shortfalls. We then use the identified proxies to construct break variables. A study of hypothetical mutual dependencies between pre-estimated volatilities, the proxies, and the break variables is conducted to investigate the levels of shared mutual entropies to assess their perceived potential in modelling time-variations and breaks in the unconditional volatility process.

The theoretical exercise indicates that, the expectation of the ARMAX process for squared returns is analogous to an exogenous GARCH process, hence, the returns of covariate or exogenous currency pairs are linked theoretically to the endogenous volatilities. The exogenous returns are exposed to similar uncertainties in the market where the volatilities evolve and the fact that levels of uncertainties in the market is a positive measure, absolute exogenous returns are identified as alternative proxies for the levels of uncertainties surrounding the exchange rate market. On investigation, we found evidence of substantial mutual dependencies between volatilities and the proxies as well as constructed break variables, thus, they are recommended for modelling changes and breaks in the unconditional volatilities of exchange rates.

6.1 Introduction

The risk taxonomy proposed by Gomery (1995) classifies uncertainty into the ‘known’, the ‘unknown’ and the ‘unknowable’. The known part is termed “volatility”, the unknown part is termed “market uncertainty”, while the unknowable part may be termed “uncertainty of market uncertainty." The efficient market hypothesis implies that known factors are already reflected in efficient market prices, but not market uncertainty and uncertainty about market uncertainty. Uncertainty about market uncertainty cannot be measured because the events defining the probability space cannot be identified in advance (Diebold, Doherty& Herring, 2010), although to some extent, market uncertainty can be measured. The degree of uncertainties in the markets affects the unconditional volatility process (Amado & Laakkonen, 2014).

In measuring the degree of market uncertainty, proxies are used due to its latent nature. Common proxies used in literature to model time-variations in the unconditional volatility include the VIX (Ezpeleta, 2015 and Amado & Laakkonen, 2014) and the SIXVX (Ezpeleta, 2015 and Kambouroudis & McMillan, 2016). Including VIX and SIXVX in the variance equations of GARCH models improves volatility estimates (Kambouroudis & McMillan, 2016), although, the proxies may be inadequate in modelling changes in the unconditional variance of exchange rate volatility; this is because they do not always take into account certain market factors such as inflation and money supply, among others. In addition, they reveal little that is not already reflected in the current and past prices, hence, they may have very limited predictive abilities. The proxies are also single measures, which attempt to represent a basket of market uncertainties and may not be adequate, thus, the need for alternative proxies has become necessary.

In choosing an alternative form of proxy, its direct theoretical or empirical links to volatility, multiplicity availability and mutual dependency with volatility should be of utmost importance. Risk premium theory links the returns of an asset to its volatility (Black, 1976) and assets are known to move together (Baybogan, 2013). Exchange rate co-movements and the relationship between an asset's volatility and its returns may imply a possibility of a relationship between absolute exogenous returns and endogenous volatility. Due to this perceived relationship and the fact that there are multiple returns available, levels of exogenous returns could be better candidates to alternatively proxy the uncertainties surrounding the exchange rate markets for the sole purpose of modelling time-variations in the unconditional volatility of GARCH models.

In this chapter, we show that indeed, there is a direct theoretical link between covariate exchange rate returns and exogenous volatility. We further show that these proxies are not only linked to volatility, but they also share substantial mutual entropies with volatility to buttress their potentials in modelling time-variations in the unconditional volatility of exchange rates. Furthermore, we construct break variables from these proxies and similarly show that they have potentials to model breaks in the unconditional volatility of exchange rates. In computing the mutual entropies, we employ the Jackknife-bias corrected Kernel density estimator. We also computed the Pearson's correlation coefficients to measure the strength of the linear dependency among the variables. Simple t-tests were then used to test the significance of the linear dependencies.

The rest of the chapter is structured as follows. In section 6.2, we present the theoretical framework and the identification of the alternative proxies. The concept of mutual information is then introduced in section 6.3; we provide data and the methodology for the empirical study in section 6.4. In section 6.5, the empirical results are presented and discussed, finally, the salient points and the main findings of the study are summarised in section 6.6.

6.2 Alternative proxy for market uncertainty and its break variable

The identification of the alternative proxies for market uncertainty requires that we first link the dynamics of the exogenous returns to endogenous volatility, thus, we make plausible assumptions substantiated by either theory or empirical evidence to help us in the linking exercise. The assumptions are

1. Exchange rate returns are second-order stationary processes.
2. Exchange rate returns are co-integrated.
3. The relationship between two or more exchange rate returns evolve via an ARMA process.
4. The expectation of the square of the ARMA process in assumption 3 is analogous to an exogenous GARCH process.

The plausibility of assumption 1 is derived from empirical studies (Suess, *et al.,* 2008) among others. Assumption 2 can be indirectly inferred from the empirical evidence of exchange rate cointegration (Engle & Granger, 1987; Kuhl, 2007 and Firoj & Khanom, 2018); assumption 3 can be deduced from Whittle (1951) and inferred from the study in chapter five. Finally, since assumption 4 has not been considered in literature, we intend to derived it statistically based on assumptions 1-3. In the previous chapter, the plausibility of assumption 3 was empirically verified, thus based on this evidence, we link the dynamics of the exogenous exchange rate returns to the volatility of endogenous returns. Recalling equation (5.4), if we square both sides, we have:

$$R_t^2 = \left(\sum_{j=1}^{m} \lambda_j r_{j,t} + \sum_{i=1}^{p} \varphi_i R_{t-i} \right)^2 + 2\left(\sum_{j=1}^{m} \lambda_j r_{j,t} + \sum_{i=1}^{p} \varphi_i R_{t-i} \right)\left(\varphi + a_t - \sum_{r=1}^{q} \theta_r a_{t-r} \right) + \left(\varphi + a_t - \sum_{j=1}^{q} \theta_r a_{t-r} \right)^2. \quad (6.1)$$

Definition 6.1: Consider two variables $x, y \in \mathbb{R}$ then:

$$\sum_{i=1}^{m} x_i \sum_{j=1}^{n} y_j = \left(x_1 y_1 + x_2 y_1 + \ldots + x_m y_1 \right) + \left(x_1 y_2 + x_2 y_2 + \ldots + x_m y_2 \right) + \ldots + \left(x_1 y_n + x_2 y_n + \ldots + x_m y_n \right) = \sum_{i=1}^{m} \sum_{j=1}^{n} x_i y_j \text{ and}$$

$$\sum_{i=1}^{m} x_i \sum_{i=1}^{m} x_i = \left(x_1 x_1 + x_2 x_1 + \ldots + x_m x_1 \right) + \left(x_1 x_2 + x_2 x_2 + \ldots + x_m x_2 \right) + \ldots + \left(x_1 x_m + x_2 x_m + \ldots + x_m x_m \right) = \sum_{i=1}^{m} \sum_{j=1}^{m} x_i x_j .$$

Using the results from definitions 6.1, the expanded forms of the individual terms in (6.1) are:

$$\left(\sum_{j=1}^{m} \lambda_j r_{j,t} + \sum_{i=1}^{p} \varphi_i R_{t-i} \right)^2 = \sum_{i=1}^{m} \sum_{j=1}^{m} \lambda_{ij} r_{i,t} r_{j,t} + \sum_{i=1}^{m} \sum_{j=1}^{p} 2\varphi_i \lambda_j R_{t-i} r_{j,t} + \sum_{i=1}^{p} \sum_{j=1}^{p} \varphi_{ij} R_{t-i} R_{t-j}, \quad (6.2)$$

$$\begin{aligned} 2\left(\sum_{j=1}^{m} \lambda_j r_{j,t} + \sum_{i=1}^{p} \varphi_i R_{t-i} \right)\left(\varphi + a_t - \sum_{r=1}^{q} \theta_r a_{t-r} \right) &= \sum_{j=1}^{m} \left(2\varphi \lambda_j \right) r_{j,t} + a_t \sum_{j=1}^{m} 2\lambda_j r_{j,t} + \sum_{i=1}^{p} \left(2\varphi \varphi_i \right) R_{t-i} + \\ & a_t \sum_{i=1}^{p} \left(2\varphi_i \right) R_{t-i} + \sum_{j=1}^{m} \sum_{r=1}^{q} \left(-2\lambda_j \theta_r \right) a_{t-r} r_{j,t} + \sum_{i=1}^{p} \sum_{r=1}^{q} \left(-2\varphi_i \theta_r \right) a_{t-r} R_{t-i} \end{aligned}, \quad (6.3)$$

$$\left(\varphi + a_t - \sum_{r=1}^{q} \theta_r a_{t-r} \right)^2 = \varphi^2 + 2\varphi a_t + a_t^2 + \sum_{r=1}^{q} -2\varphi \theta_r a_{t-r} + a_t \sum_{r=1}^{q} -2\theta_r a_{t-r} + \sum_{r=1}^{q} \sum_{j=1}^{q} \left(\theta_r \theta_j \right) a_{t-r} a_{t-j}. \quad (6.4)$$

Substituting (6.2), (6.3), and (6.4) into (6.1) yields:

$$\begin{aligned} R_t^2 = \varphi^2 &+ \sum_{r=1}^{q} \sum_{j=1}^{q} \left(\theta_r \theta_j \right) a_{t-r} a_{t-j} + \sum_{i=1}^{m} \sum_{j=1}^{m} \lambda_{ij} r_{i,t} r_{j,t} + \sum_{i=1}^{p} \sum_{j=1}^{p} \varphi_{ij} R_{t-i} R_{t-j} + \sum_{i=1}^{p} \left(2\varphi \varphi_i \right) R_{t-i} + \sum_{j=1}^{m} \left(2\varphi \lambda_j \right) r_{j,t} + \sum_{i=1}^{p} \sum_{r=1}^{q} \left(-2\varphi_i \theta_r \right) a_{t-r} R_{t-i} \\ &+ \sum_{j=1}^{m} \sum_{r=1}^{q} \left(-2\lambda_j \theta_r \right) a_{t-r} r_{j,t} + \sum_{r=1}^{q} -2\varphi \theta_r a_{t-r} + \sum_{i=1}^{m} \sum_{j=1}^{p} 2\varphi_i \lambda_j R_{t-i} r_{j,t} + \left(\sum_{j=1}^{m} 2\lambda_j r_{j,t} + \sum_{i=1}^{p} \left(2\varphi_i \right) R_{t-i} + \sum_{r=1}^{q} -2\theta_r a_{t-r} + 2\varphi + a_t \right) a_t \end{aligned}. \quad (6.5)$$

The last term of equation (6.5) is a function of current error, thus, without losing track of notations, the term is simply denoted by a_t. Practically for distinct i, j and r, the $a_{t-r} a_{t-j}$, $R_{t-i} R_{t-j}$,

$R_{t-i}R_{t-j}$ and $r_{i,t}r_{j,t}$ terms may have insignificant effects on square returns, thus it is assumed that their effects are negligible. Similarly, the effects of the interaction terms: $a_{t-r}R_{t-i}$, $a_{t-r}r_{j,t}$ and $R_{t-i}r_{j,t}$ as well as the term $\{2\varphi\theta_r a_{t-r}\}$ are assumed negligible. If we let $\varphi^2 = \delta$, $\theta_{rr} = \psi_r$, $\varphi_{ii} = \phi_i$, $\lambda_{jj} = \theta_j$, $2\varphi\lambda_j = \lambda_k$ and $2\varphi\varphi_i = \varpi_i$, equation (6.5) reduces to:

$$R_t^2 = \delta + \sum_{r=1}^{q}\psi_r a_{t-r}^2 + \sum_{i=1}^{p}\phi_i R_{t-i}^2 + \left(\sum_{i=1}^{p}\varpi_i R_{t-i} + \sum_{j=1}^{m}\theta_j r_{j,t}^2 + \sum_{k=1}^{m}\lambda_k r_{k,t}\right) + a_t \,. \tag{6.6}$$

The random behaviour of speculative asset returns suggests the absence of autocorrelation in the levels of returns and this indicates the unpredictability of the direction of asset returns, thus, for a stationary asset returns R_t, the autocorrelation function up to lag 1 is given by $\rho = 0$. Applying the law of total variance and taking expectation with respect to the information set $\mathcal{F}_{t-1}$, the volatility h_t can be expressed as

$$h_t = \delta + \sum_{r=1}^{p}\psi_r a_{t-r}^2 + \sum_{i=1}^{q}\phi_i h_{t-i} + \left(\sum_{i=1}^{p}\varpi_i E_{t-1}\left[R_{t-i}\right] + \sum_{j=1}^{m}\theta_j E_{t-1}\left[r_{j,t}^2\right] + \sum_{j=1}^{m}\lambda_j E_{t-1}\left[r_{j,t}\right]\right). \tag{6.7}$$

It can be seen that equation (6.7) is analogous to the standard GARCH process and that the last three terms are external variables, which are theoretically linked to volatility. Market uncertainty is a positive measure, the absolute values of the last exogenous term (covariate returns) are used as potential proxies for the levels of uncertainties surrounding the exchange rate market. We do not consider the square returns as suitable candidates because of their noisy nature.

The chosen proxies overcome the shortfalls of the VIX and SIXVX proxies. Firstly, the dynamics of exchange rate are driven by macroeconomic fundamentals, market microstructure and several other market factors (Kleinbrod & Li, 2017). Since returns are direct computations or realizations from prices, they inherit these myriads of factors, thus, the proxies are much more appropriate for volatility modelling. Secondly, because there are several pairs of currencies traded on any forex market, the proxies are multiply abundant, thus, a representative choice could be used to model time-variations in the unconditional volatility of exchange rates comprehensively. To construct the break variables from the proxies, consider the positive returns from m covariate returns with N sample length, then, the lag-one break variables are constructed as follows:

$$B_{j,t-1} = \begin{cases} 1, \text{ if } \left|r_{j,t-i-1}\right| < \left|r_{t-i}\right| \\ 0, \text{ if } \left|r_{j,t-i-1}\right| > \left|r_{t-i}\right| \end{cases} \text{ for } i = 1,2,\ldots,N \text{ and } j = 1,2,\ldots,m\,. \tag{6.8}$$

The break variables and the proxies will be simultaneously used to account for breaks and time-variations in the unconditional volatilities of the exchange rates, comprehensively. Before this is done, we investigated the predictive abilities of the break variables and the proxies in the subsequent sections by estimating the percentages of information they exchange with the volatilities of the exchange rates whose volatilities are to be modelled in this thesis.

6.3 Mutual information

Information theory, a fabricated attempt to study communication systems, originated from the study by Shannon (1948). One of the major applications of the theory is its use in the study of statistical dependency between random variables, using mutual information. Mutual information (MI) measures the reduction in the uncertainty about a random variable conditioned on the knowledge of another variable, thus, it is more closely related to the concept of entropy introduced by Shannon (1948). There are several ways of measuring dependency (Zenga, Xiaa & Tong, 2018), however, MI is by far a perfect statistic in several ways (Ross, 2014). Unlike linear correlation measures, MI is more general in the sense that it contains all information about the dependency of variables including linear and nonlinear and it is very effective in measuring any kind of relationship (Cover & Thomas, 2012). MI also has a straightforward interpretation, grounded in information theory and insensitive to the size of data sets. Mutual information between bivariate variables can either be zero (corresponding to independence) or positive (corresponding to dependency). Mutual information can be defined as follows:

Definition 6.2: Let (X,Y) be a pair of continuous random variables with values over the space $\chi \times \mathcal{y}$. If the joint probability density function is $P_{(X,Y)}$ and the marginal distributions are P_X and P_Y, the mutual information is defined as:

$$MI(X,Y) = \int_{\mathcal{y}}\int_{\chi} p_{(X,Y)}(x,y)\log\left(\frac{p_{(X,Y)}(x,y)}{p_{(X)}(x)\,p_{(Y)}(y)}\right)dxdy\,. \tag{6.9}$$

MI is a positive unbounded measure $(MI \in [0,\infty))$ with unit in bits (base 2) or nats (base *e*). Mutual information is a function of the expected amount of information contained in a variable, called entropy (Scheuerell, 2017). For continuous variables (X,Y), if the probability density function for X is $f(x_i)$ and the marginal probability density function of (X,Y) is $f(x_i,y_i)$, then the entropy of X and the joint entropy of (X,Y) are respectively defined as:

$$H(X) = E\left[-\log\left(f(X)\right)\right] = -\int_{\chi} f(x_i)\log_b f(x_i)\,dx, \tag{6.10}$$

$$H(X,Y) = -\int_{\chi,y} f(x_i,y_i)\log_b f(x_i,y_i)\,dxdy. \tag{6.11}$$

The mutual information between X and Y is therefore given by:

$$MI(X,Y) = H(X) + H(Y) - H(X,Y). \tag{6.12}$$

6.4 Data and Methodology

6.4.1 Data

The variables of interest in this chapter are absolute returns, break variables, and volatility. Lag-one of the entire sample lengths for all the return discussed in chapter three are used. The lag-one break variables are constructed based on equation (6.8) and since volatility is latent, they are pre-estimated using the return data. The lagging is required because we are interested in using the variables for forecasting in the subsequent chapters, thus, we were able to assess the volatility predictive abilities of the variables.

The latent nature of volatility implies that actual shared entropies cannot be computed, thus the mutual information computed in this chapter are hypothetical in a broader sense. We intend to forecast the volatilities and VaR for the currency pairs: MWK/ZAR, BWP/ZAR, BRL/ZAR, ILS/ZAR, and SEK/ZAR[7], therefore, it is more appropriate to use pre-estimated volatilities for these pairs using similar specifications in Table 7.1. Individual estimation of the volatilities of the above currency pairs is time-consuming, thus, the simultaneous estimation approach implemented in rugarch package (Ghalanos, 2019) is utilized. The GARCH estimation procedures used for computing the volatilities and VaR in this thesis are based on a moving window, thus simple moving averages of the variables are used. MI is a positive unbounded measure, hence, to be able to compare two MI estimates we normalized the estimated estimates using equation (6.13)-adapted from Cover & Thomas (2012).

$$\text{Normalized BCMI} = \frac{BCMI(X,Y)}{\sqrt{H(X)H(Y)}}. \tag{6.13}$$

6.4.2 Methodology

Among the commonly-used estimation approaches for mutual information are the probability density-based methods, such as the Burg's maximum entropy method (Burg's MEM), the kernel density estimation (KDE), and the nearest-neighbour approach. For instance, Benedetto *et al.,* (2020) and Benedetto, Mastroeni & Vellucci (2019) used the Burg's MEM to model the flow of information between financial variables. Pardy (2013) also developed a new KDE for application to large high-dimensional datasets, frequently used in genomic experiments. In as much as the Burg's MEM of Benedetto *et al.,* (2020) and Benedetto, Mastroeni & Vellucci (2019) and the KDE of Pardy (2013) are non-parametric estimators and are suitable for high-dimensional datasets, they differ in four major aspects.

In the first instance, the new Kernel density estimation is based on Shannon's definition of entropy and joint entropy. The probability density function is estimated by filtering the data with a kernel, which is then normalized with an integral of one, which is usually symmetric and localized (Benedetto *et al.,* 2020). The Burg's MEM on the other hand is based on the extrapolation of the autocorrelation function of variables by using the entropy rate definition, after which the estimation of

[7] The pairs were chosen based on the condition that after the inclusion of appropriate proxies and dummy variables through the backward elimination approach of variable selection, the models is still adequate fit to the data and also parsimonious.

the power spectral density is estimated by the Fourier transformation of the extended autocorrelation function (Pardy, 2013).

Secondly, the new KDE is purposely designed to model static relationship between variables while the Burg's MEM is designed to model dynamic relationship and to predict future entropy of a time series by exploiting the unknown, but predicted autocovariance function of the future time interval (Benedetto, Mastroeni & Vellucci, 2019 and Pardy, 2013).

Thirdly, the Burg's MEM relies on the assumption of second-order stationarity. If this hypothesis is not satisfied, the series can be partitioned into smaller epochs, which are approximately stationary, or the series can be represented by alternative functions instead of the usual sine and cosine functions. Simulation studies, however, have shown that the Burg's approach to spectral analysis is robust in the presence of non-stationarity. Unlike the Burg's MEM, the KDE does not rely on stationarity assumption, although, it relies heavily on the choice of the tuning parameters, thus the corresponding estimators may be very unstable or extremely biased. This problem is addressed by using the Jackknife-bias corrected algorithm. The Jackknife-bias corrected Kernel density estimator automates the bandwidth selection, such that the optimal bandwidth is estimated. This helps to reduce the bias at the boundary region, thus, improve the efficiency of estimation (Zenga, Xiaa & Tong, 2018). Unlike the randomized resampling approaches for correcting bias in Kernel estimation, the Jackknife approach is deterministic, in the sense that it gives the same result when reapplied to any given data. In addition, by restricting the resampling to a specific group of n subsamples, substantial computational costs could be avoided. The approach puts an upper limit on the number of subsamples and the relationships between Jackknife repetitions can be exploited to avoid redundant computations (Pardy, 2013).

Lastly, unlike the Burg's MEM, the Kernel density estimation models the distribution of a continuous variable as a mixture of conditional distributions for each level of a categorical variable, thus it is suitable for the estimation of mutual information between a mixture of continuous and discrete variables. Due to this, KDE can be used to model the relationship between a mixture of discrete or categorical and continuous variables without the need for variable transformation (Pardy, 2013).

In comparison to other density-based estimation approaches, such as the mirrored KDE, ensemble

KDE, copula-based generalized nearest-neighbour graphs and the mixed generalized nearest-neighbour graphs, the Jackknife-biased corrected KDE is more computationally efficient. In addition, the procedure is completely data-driven and there is no need for a predetermined tuning parameter. It does not necessitate boundary correction and yet it retains the same estimation efficiency because the boundary biases are eliminated automatically. Furthermore, the estimates are numerically stable. Due to these advantages over the existing methods, the Jackknife-biased corrected KDE is employed in this study (Zenga, Xiaa & Tong, 2018).

6.5 Empirical results

The empirical results of the mutual dependency estimates are reported in Table 6.1 and 6.2, as well as Figure 6.2. From Table 6.1, it is observed that the returns of SEK/ZAR have the highest percentage of exchanged information with the volatilities of both MWK/ZAR and BRL/ZAR pairs. Between the two volatilities, the information exchange is highest with MWK/ZAR, suggesting that returns of SEK/ZAR may have a better chance of predicting the volatility of MWK/ZAR than the volatility of BRL/ZAR. Similarly, returns of INR/ZAR may have a better chance of predicting the volatility of ILS/ZAR than the volatility of SEK/ZAR. Concerning the volatility of BWP/ZAR, the returns of MWK/ZAR stand a better chance of predicting the volatility of BWP/ZAR than the other returns because they have the highest BCMI, although there is only 6% normalised shared information.

In Table 6.2, the endogenous break variables of MWK/ZAR and SEK/ZAR exchange more percentages of mutual information with their respective volatilities than the corresponding exogenous break variables. In relation to the volatility of MWK/ZAR, the non-normalized shared information with the exogenous SEK/ZAR break variable is, however, slightly higher than that of the exogenous break variable. In respect of the volatility of BWP/ZAR, the break variables for SEK/ZAR and INR/ZAR have the highest percentage of exchanged information with the volatility. A similar statement can be made for the break variables for MWK/ZAR and SEK/ZAR with respect to the volatility of ILS/ZAR, although, the non-normalized exchanged information is relatively higher for the break variable for SEK/ZAR. Finally, the break variable for MWK/ZAR exchanges the highest percentage of information with the volatility of BRL/ZAR. In general, among the variable pairings considered in the study, the exogenous break variables tend to exchange relatively higher mutual information with volatility in comparison to the endogenous break variables. This suggests that break variables constructed from exogenous returns have higher likelihoods of volatility predictive abilities, hence, are more likely to be adequate than the endogenous break variables in accounting for breaks in the unconditional volatilities of exchange rates.

There is quite a lot of variable pairing with insignificant linear dependencies, although their respective BCMI (biased-corrected MI) are substantial. For example, the strength of the linear dependency between the returns of BWP/ZAR and the volatility of MWK/ZAR is abysmally low and insignificant, although the normalized BCMI is about 32%. This suggests that it is very unlikely that the two variables will move together linearly, thus, the nature of the relationship between the two variables is more likely to be nonlinear. This observation is consistent with the rationale behind the concept of mutual information in measuring dependency (Cover, 2012). Based on the foregoing deduction, the nature of the relationships between the variable pairings in Tables 6.1 and 6.2 (indicated by cells with asterisks) are more likely to be nonlinear, hence nonlinear volatility models may be more appropriate when estimating volatility of exchange rate returns.

When the joint distribution of paired variables is a bivariate normal, there is an exact logarithmic relation between mutual information and linear correlation coefficient (ρ), which is defined by Gel'fand & Yaglom (1957) as:

$$MI = -0.5\log\left(1-\rho^2\right) . \tag{6.14}$$

It is observed from Figure 6.1 that the relationship between MI and correlation coefficient do not assume the form of equation (6.14), but rather it evolves randomly. This implies that the joint distributions of the paired variables are not bivariate normal. The individual variables are characterized by heavy-tailed distributions (see the data description section), thus, the bivariate joint distributions may be a mixture of heavy-tailed distributions. In addition, the relationship suggests that increased levels of mutual information are not associated with an increased strength of linear dependencies, thus, there may be low levels of shared information between variables, although, the strength of the linear dependency structure may be comparatively higher. This result confirms the earlier results obtained in chapter five.

In conclusion, the levels of exchange mutual information between volatilities and their respective lagged covariates provide substantial evidence of the predictive abilities of the covariates and these may be adequate in modelling variations and breaks in the unconditional volatilities of GARCH models for exchange rates. The results indirectly support the findings of returns dependencies from Owusu, Adam & Tweneboah (2017) and exchange rate co-movements unravel by several studies including Park & An (2020), Marconi (2018) and Kühl, (2008).

Table 6. 1: Estimated BCMI and NBCMI between volatilities and returns.

Volatility	MWK/ZAR	BWP/ZAR	BRL/ZAR	ILS/ZAR	SEK/ZAR	INR/ZAR	NOK/ZAR
MWK/ZAR		0.743 (0.323) -0.009*	0.600 (0.28) 0.114	0.820 (0.35) -0.059	**0.917** (**0.39**) -0.077	0.868 (0.37) **-0.228**	0.869 (0.37) -0.124
BWP/ZAR	**0.155** (**0.06**) 0.127		0.086 (0.04) 0.200 L	0.142 (**0.06**) 0.056	0.127 (0.05) 0.084	0.142 (**0.06**) -0.160	0.128 (0.05) 0.022*
BRL/ZAR	0.581 (0.23) 0.138	0.506 (0.21) NL -0.051		0.637 (0.26) NL -0.031*	**0.660** (**0.27**) -0.064	0.634 (0.26) **-0.193**	0.644 (0.26) -0.096
ILS/ZAR	0.321 (0.15) NL 0.055	0.318 (0.16) 0.075	0.310 (0.16) **0.118**		0.415 (0.20) 0.109	**0.486** (**0.23**) -0.006*	0.420 (0.20) 0.087
SEK/ZAR	0.313 (0.14) -0.101	0.281 (0.13) 0.247	0.241 (0.12) **0.260**	0.362 (0.16) 0.141		**0.386** (**0.17**) 0.114	0.326 (0.15) 0.233

Normalized BCMI (NBCMI) estimates are reported in brackets while the strength of linear dependencies are reported in the third line of each row. .All estimated information measures are measured in bits. Asterisks indicate that the correlation coefficient is not significant at 5%.

Table 6. 2: Estimated BCMI and NBCMI between volatilities and break variables.

Metric	MWK/ZAR	BWP/ZAR	BRL/ZAR	ILS/ZAR	SEK/ZAR	INR/ZAR	NOK/ZAR
MWK/ZAR	0.666 (**0.29**) -0.036*	0.434 (0.19) -0.070	0.627 (0.28) 0.081	0.490 (0.21) -0.067	**0.686** (0.26) **0.082**	0.627 (0.24) -0.061	0.548 (0.23) -0.036*
BWP/ZAR	0.122 (0.05) -0.025*	0.163 (0.07) **0.141**	0.133 (0.06) 0.029*	0.170 (0.07) 0.058	**0.244** (**0.08**) 0.063	0.236 (**0.08**) 0.062	0.132 (0.05) 0.060
BRL/ZAR	**0.550** (**0.23**) **-0.069**	0.391 (0.17) 0.035*	0.475 (0.21) 0.036*	0.375 (0.15) -0.030*	0.534 (0.19) **0.069**	0.489 (0.18) -0.044*	0.396 (0.16) -0.024*
ILS/ZAR	0.295 (**0.15**) 0.045*	0.202 (0.10) **0.114**	0.282 (0.14) 0.011*	0.235 (0.11) 0.055	**0.348** (**0.15**) 0.042*	0.251 (0.11) 0.069	0.259 (0.12) 0.073
SEK/ZAR	0.285 (0.13) 0.066	0.214 (0.10) 0.180	0.249 (0.12) -0.043*	0.196 (0.09) **0.186**	**0.355** (**0.14**) -0.011*	0.254 (0.11) 0.147	0.265 (0.12) 0.035*

The first line of each row reports the estimated BCMI while the second lines (values in brackets) indicate the normalized BCMI estimates. The strength of linear dependencies are reported in the third lines of each row. .All estimated information measures are measured in bits. MWK/ZAR, BWP/ZAR, BRL/ZAR, ILS/ZAR, and SEK/ZAR are endogenous break variables while INR/ZAR and NOK/ZAR are exogenous. Asterisks indicate that the correlation *coefficient is not significant at 5%. Note that normalized BCMI can be interpreted in terms of percentages by multiplying each value by 100.*

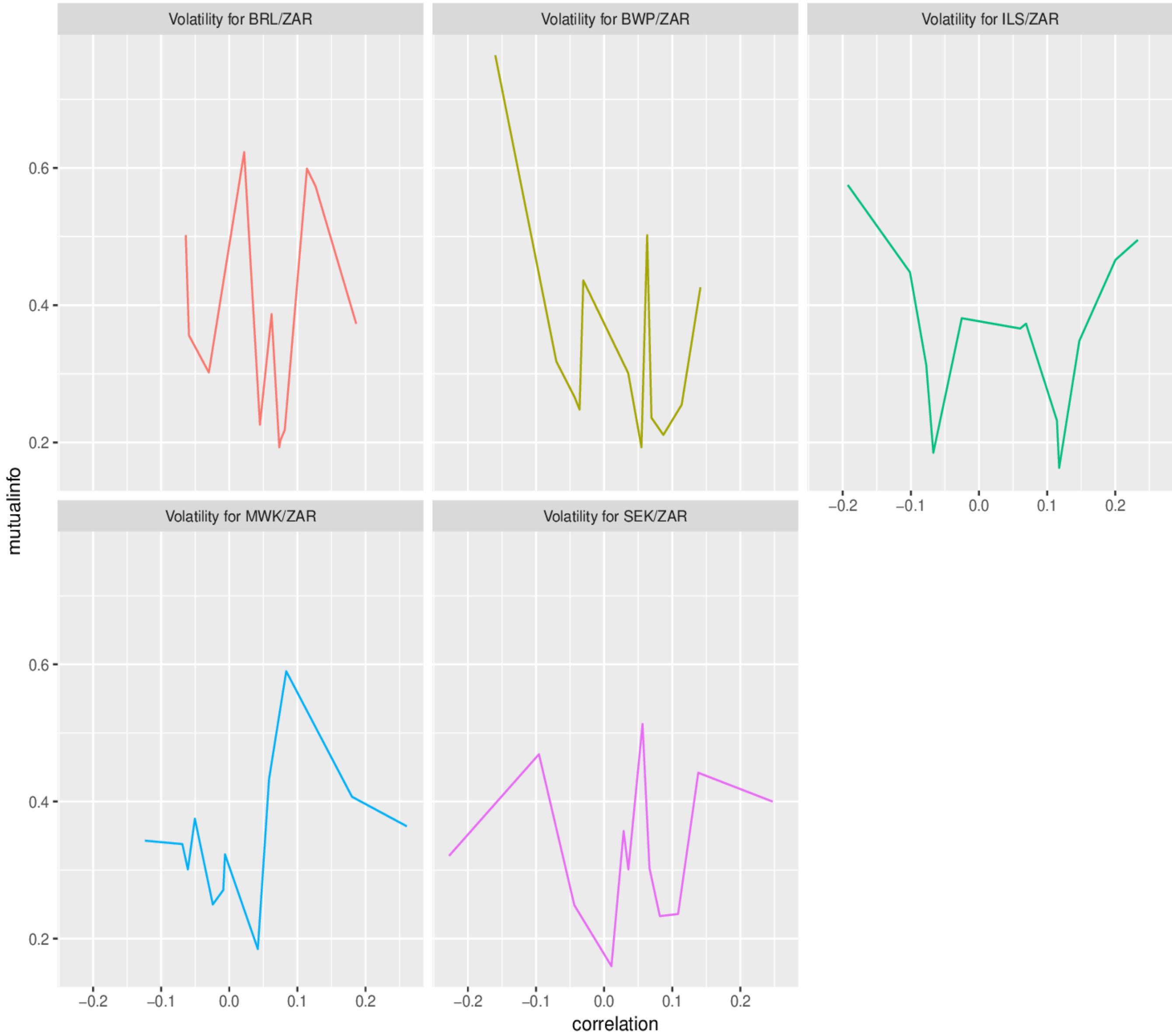

Figure 6. 1: Dependency plot for the five volatilities and their respective lagged covariates

6.6 Conclusion

In this chapter, we sought to identify alternative proxies for the levels of uncertainties surrounding the exchange rate market, which could be used to model changes in the unconditional variance of exchange rate volatilities. The proxies were identified from a theoretically derived model, which is analogous to the exogenous GARCH model. We also constructed break variables from these proxies, which could be used to model breaks in the unconditional volatilities of exchange rates. A study of the hypothetical mutual dependencies between pre-estimated volatilities and lagged break variables (those constructed from the proxies and those from the endogenous returns) was conducted to investigate the levels of the dependencies and their predictive abilities. The Jack-knife-biased corrected KDE was used in computing the mutual entropies while Pearson's method was employed in computing the strength of the linear dependency. Simple t-tests were then used to test the significance of the estimated strengths of the linear dependencies.

The results indicate that the expectation of a derived ARMAX process for squared returns is analogous to an exogenous GARCH process, thus returns of covariate currency pairs are theoretically linked to the volatility of exchange rates. From this derived process, absolute returns were identified as alternative proxies for the levels of uncertainties surrounding the exchange rate market. Subsequently, the proxies were used to construct break variables. Findings from the hypothetical studies indicate that the joint distributions of all the paired variables are not bivariate normal and since the individual variables are characterized by heavy-tailed distributions, the bivariate joint distributions may be a mixture of heavy-tailed distributions. In addition, there is evidence of substantial percentages of exchanged information between the lagged exogenous variables, suggesting that there is a better chance of predicting exchange rate volatility with these variables. In general, the exogenous break variables tend to exchange relatively higher mutual information with volatility in comparison to the endogenous break variables, suggesting that break variables constructed from exogenous returns have higher likelihood of volatility predictive abilities. They are, hence, more likely to be adequate in accounting for breaks in the unconditional volatilities of exchange rates.

CHAPTER 7

Forecasting exchange rate volatility with alternative proxies for market uncertainty and exogenous break variables

Chapter Summary

Monotonic paths of long-range unconditional volatility forecasts from GARCH models are inconsistent with the underlying stochastic path, thus, in this study we attempt to induce stochastic mean reversion in the unconditional volatilities of non-time varying GARCH models by augmenting the models with the proxies for market uncertainty and break variables identified from the previous chapter. Using data from the rand market, the augmented models are used to forecast long-range volatilities of selected exchange rates.

Evidence from the empirical studies indicates that the unconditional forecasts from the augmented models are stochastic, thus, long-range forecasts revert along stochastic path towards the long-run variance. In addition, the estimated volatilities from the models are less persistent with significantly improved accuracies. Furthermore, the models yielded relatively improved forecasts or insignificant loss of forecast accuracies with improved explanatory forecast power. We recommend the approach for forecasting medium to long term volatility of exchange rate and for modelling and forecasting volatility of series exhibiting IGARCH effects.

7.1 Introduction

The general autoregressive conditional heteroscedasticity (GARCH) is economical and theoretically appealing. It provides a more real-world context for modelling, predicting, and forecasting volatility because of its ability to describe the dynamic structures of volatility clustering and other empirical regularities associated with volatility (Amado &Teräsvirta, 2014). The models are primarily built to forecast volatility, and since these forecasts are used as inputs in other financial tasks, their accuracies are non-negotiable. Several GARCH models are able to produce good in-sample estimates, although, findings from competing GARCH models converge towards seemingly poor volatility forecasts in all GARCH models. The models are characterized by large mean squared prediction errors; explain very little of the variability in asset returns (typically below

10%) and appear to be biased (Chen, Dolado & Gonzalo, 2014).

The seemingly poor GARCH forecasts have been attributed to the approximation of unobserved true volatility with square returns (Andersen & Bollerslev, 1998 and Christodoulakis & Satchell, 1998). The aforementioned scholars argue that the approximation leads to an inflated mean square error with distorted forecasts. The presence of structural breaks in the returns of assets has serious implications on the GARCH estimates. Models, which fail to account for breaks in the unconditional variance lead to sizable upward biases in the degree of persistence in the estimated models, thus the fitted models may fail to track changes in the unconditional variance and produce forecasts that systematically underestimate or overestimate volatility on average, over long horizons (Rapach, Strauss & Wohar, 2007). Amado & Teräsvirta (2014) have also identified its failure to account for time-variations in the unconditional volatility as another cause of the seemingly poor GARCH forecasts.

A vast amount of theoretical and empirical studies have cropped up in an attempt to improve GARCH forecasts, however, the out-of-sample forecasts from the models do not mimic the stochastic behaviour of the underlying volatility. The unconditional forecasts from such models evolve monotonically, hence, the conditional forecasts revert toward the long-run variance along a non-stochastic path. Inducing stochastic mean-reversion on the conditional volatility or time-variations on the unconditional volatility can be done by structurally modifying the GARCH model or by using external variables. Amado & Laakkonen (2014) showed that levels of uncertainties in the markets affect the unconditional volatility process, hence, they are plausible candidates for externally inducing stochastic mean-reversion in the exchange rate volatility process. Levels of market uncertainty are latent, thus, proxies are used in literature. Several attempts in literature have addressed the problem of structural breaks and time-varying unconditional volatility, although, the alternative proxies for the exchange rate market uncertainty identified in chapter six and their break variables have not been given much attention, thus, a gap in literature exists. This chapter is an attempt to narrow this gap.

In as much as the proxies are exogenous, the break variables are also exogenous, unlike the endogenous breaks used in studies like Karlsson (2016). An advantage of the proxies is that they are theoretically linked to volatility and they share substantial levels of mutual entropies with volatility, thus, it is anticipated that they will yield improved volatility forecasts. The approach does not require functional modifications of existing models, therefore, its implementation is simple and may be easily integrated into various volatility modelling frameworks.

The rest of the chapter is organised as follows. In section 7.2, a literature review on some of the attempts to improve GARCH forecasts are presented. In section 7.3, we revise the mean-reversion concept as it applies to volatility forecasts from GARCH models. The proposed augmented GARCH models are formulated in section 7.4, which also explains the forecasting structure of the augmented GARCH models. Empirical applications of the proposed method to the rand forex market are presented in section 7.5 and the study is summarized in section 7.6.

7.2 Literature review

In an era of increased volatility for speculative assets and its subsequent effects on the various sectors of the economies around the globe, the need for improved volatility estimates and forecasts has necessitated several studies. While some of these studies consider structural modifications of existing GARCH models, others use non-structural modification approaches, such as the external incorporation of the several statistical properties of returns such as long-range dependency, structural breaks, co-integration (Kosapattarapim, 2013) and ARMA (McCrae *et al.,* 2002), which are not captured by standard GARCH models. In some of these studies, exogenous covariates, dummy variables, or varying estimation windows have been proposed to capture these statistical properties. The simultaneous use of absolute returns and their break variables to account for variations and breaks in the unconditional variance of GARCH models, however, are not conspicuous in literature, although these have potentials to induce stochastic mean-reversion on the volatility forecasts. In the next few paragraphs, we look at some of the recent and prominent studies on structural and non-structural modifications of existing GARCH models.

Shi & Yang (2018) recently proposed an adaptive Hyperbolic EGARCH (A-HYEGARCH) model to estimate the long memory of high-frequency time series with potential structural breaks. Based on the original HYGARCH model of Shi & Kin-Yip (2016), they used logarithm transformation to ensure the positivity of the conditional variance. The structural change was further allowed via a flexible time-dependent intercept in the conditional variance equation. To demonstrate the effectiveness of their proposed method, they tested the model empirically using high-frequency returns from standard and poor 500 (S&P), financial times stock exchange 100 (FTSE), Australian security exchange (ASX), and Nikkei 225 indexes. The result shows that the proposed model outperforms various competing specifications and can effectively account for structural breaks.

Cantamessa, Gautam & Xiang (2016) forecasted volatility of ten stocks from the Eurostoxx50, using daily stock prices from July 11, 2012 to June 10, 2016. The vanilla GARCH model with exogenous market data, news-sentiment impact scores, and news volume were used. The estimated model was then compared to the realized volatility. The result revealed that the inclusion of scheduled news and news volume characterized by negative sentiment improve volatility forecasts. The results of the study confirm previous results obtained by Li & Engle (1998).

Kambouroudis & McMillan (2016) modelled volatility of stocks from the US, UK and France using the symmetric GARCH, threshold-GARCH, EGARCH, asymmetric power-ARCH, component GARCH and the IGARCH models with VIX and trading volumes as variance regressors. They also considered regressing VIX and trading volumes directly on volatility as well as the estimation of appropriate models for VIX and used the forecasts thereof to represent the volatility of stock return. The data used consisted of daily closing prices of the main indexes, S&P 500, FTSE 100, and CAC40 (Cotation Assistée en Continu). The result suggests that VIX and volume have capabilities to improve the forecasting performance of GARCH models and the other forecasting models considered. The improvements were much better when the two variables were simultaneously incorporated into the volatility models.

Karlsson (2016) suggested the identification of the existence of breaks and the subsequent use of dummy variables in the mean return equation to account for structural breaks. In defining the break variables, he considered two states of volatility regimes. The first state a denoted 0 was used as a benchmark state for normal volatility and the second state denoted by 1 was used to represent volatile regime with rapid movements in the market. Karlsson then applied the methodology to returns of the main Nordic indices: OMXS30, OMXC20 and OMXH25 spanning from January 2, 2008, to December 30, 2015 to a set of commonly-used GARCH models. Evidence from the study indicates that the hypothesis, that the models incorporating structural breaks have better predictive accuracies was rejected. One of the drawbacks of the approach is that states are only categorised as either normal or volatile with no consideration of the actual economic state, hence, even though financial events may differ in some characteristic, crises are modelled in the same way as periods of increased volatility. This problem may have contributed to the inability of the breaks to improve the predictive and forecasting abilities of the models as other breaks variables did in other studies.

Lee (2015) modelled nonlinearities in stock market volatility using FIGARCH (1, d, 1) with different estimation windows to account for long-range dependence and level shifts in the unconditional variance process. Daily data from S&P 500 and the Korea Composite Stock Price Index were used. The results indicate that the out-of-sample forecasting performances of models with structural breaks are superior to GARCH models without breaks. It was further revealed that the out-of-sample forecasting performances of GARCH models with structural breaks were superior to GARCH models with expanding window. It was noted that, although, the incorporation of structural breaks in GARCH models improves forecasts, the performances of the prediction models could be tainted with uncertainties related to statistical tests and estimation methodologies.

Amado & Teräsvirta (2014) proposed the decomposition of the variance structure of GARCH process into conditional and unconditional components and allowed the unconditional component to evolve smoothly over time, via a linear combination of logistic transition functions with time transitional variable. The approach was then used to model time variations and structural breaks in the non-stationary variance component of GJR-GARCH and GARCH models using a sample data of daily returns from the Dow Jones Industrial Average (DJIA) index from January 2, 1920, to May 31, 2011. Evidence from the study shows that the constancy assumption of unconditional variance is rejected and the dependency structures of the return series are well explained by the smooth logistic transitional changes in the unconditional variance. The path of the unconditional volatility evolution is however monotonic, although, formal statistical tests adjudged the accuracy of the volatility forecast from the new models as superior to the GJR-GARCH model at all horizons, for a subset of the long return series.

Rapach, Strauss & Wohar (2007) investigated the empirical relevance of structural breaks for GARCH models by adjusting the estimation window for GARCH (1, 1) forecasting models to accommodate potential structural breaks. The adjustments considered included the rolling windows of various sizes, as well as a method where the estimation window for the GARCH (1, 1) forecasting model is determined by applying the modified iterated cumulative sum of squares algorithm to the observations available at the time of the forecast formation. They used eight daily U.S. dollar-denominated exchange rate returns from 1980 to 2005. The results indicated that, accounting for structural breaks in the unconditional variance of exchange rate returns by adjusting the estimation window helps to improve the out-of-sample forecasts of exchange rate volatility. Similar to this study, Babikir *et al.,* (2012) used a rolling window estimation method to account for structural breaks in the unconditional volatility of GARCH (1, 1) model using daily returns from the Johannesburg Stock Exchange All Share Index from July 2, 1995 to August 25, 2010. Their findings however were not in direct agreement with Rapach, Strauss & Woha (2007) as there were no statistical gains from using competing models that explicitly account for structural breaks relative to GARCH(1,1) model with an expanding window.

Other studies have also shown that increasing the sampling frequency and the use of other external factors may help improve the volatility estimates. For instance, using GARCH (1, 1), GJR-GARCH (1, 1) models, and monthly data from April 1982 to November 2011, Jabeen & Khan (2014) modelled the volatility of Pak Rupee-denominated currencies (US Dollar, British Pound, Canadian Dollar, and Japanese Yen). Macroeconomic variables: relative real income, foreign reserves, inflation rate differential, terms of trade, and productivity and trade restrictions were passed to the variance equation. Their findings indicated that these variables have significant influence on the volatility process. Similar to this study, Chipili (2012) used AR-GARCH (1, 1), AR-TGARCH (1, 1), and AR-EGARCH to model volatility of eight real Kwacha-denominated currencies over the period 1968-2008. The study used macroeconomic variables in the variance equation. Evidence from the study indicated that the volatility process is influenced by openness, terms of trade, money supply, real interest rate, real productivity or shocks, foreign reserves and exchange rate regime.

Andersen & Bollerslev (1998) have shown that increasing the sampling frequency of ex-post square returns in GARCH (1, 1) from one day to five-minutes, dramatically improves the explanation power of ex-post variability of square returns from 4.7% to 47.9% for Deutsche Mark to US dollar returns and from 2.5% to 39.2% for Japanese Yen to the US dollar returns. They further found out that uncertainties in the market and economic variables affect volatility. Particularly, they observed that the time of the announcement of the USA macroeconomic indicators like employment report, the producer price index or the quarterly GDP affects the volatility of the Deutsche Mark-USD exchange rates. Glosten, Jaganathan & Runkle (1993) on the other hand found that October and January indicator variables, as well as private information assist in explaining some of the dynamics of the conditional volatility of equity returns. Other factors, such as market microstructures (Kleinbrod & Li, 2017), holidays, weekends and deterministic time-of-day (Engle & Patton, 2002) also affect exchange rate volatility and may therefore be useful in volatility forecasting.

7.3 Mean reversion in assets returns and GARCH models

Mean-reversion is defined as the change of market return in the direction of a reversion level in response to a reaction to a prior change in the market return (Hillebrand, 2003). In simple terms, it suggests that after the returns have reached certain extreme point, they have the tendency to revert to their long-run average (Arefin & Ahkam, 2017 and Ribeiro, Cermeño & Curto, 2017). Hillebrand (2003) argues that mean reversion varies from one equity returns to another and they follow paths determined by a past reaction to changes, hence, there is a general belief that past history of asset prices can be used to forecast prices.

Volatility inherits the mean reversion property of the underlying asset returns, hence, although volatility has the tendency to return to the long-run variance, the probability of it reverting to the exact variance is negligible. Volatility, therefore, could only get closer and closer to the long-run variance almost all the time, thus, the concept of mean-reversion is not treated in the strict sense in empirical studies. To illustrate this, plots of the GBP/ZAR volatility for 1378 days, divided into three samples are shown in Figure 7.1. There seems to be more than 1 day where the volatilities reverted to the exact long-run variance (represented by the straight blue lines), however, this is not the case. In fact, over the whole sample, the probability that a return reverted exactly to the long-run variance is only 1 out of the 1378 days (i.e., 0.07%).

GARCH models are able to capture this property of volatility for the in-sample estimation, but the same cannot be said about the out-of-sample estimates. There is an imposed exponential decay of shocks on the GARCH model, thus, the unconditional forecasts inherit this structure. Conditional volatilities converge or revert to the unconditional volatility in the long-run, however, the mean reversion is restricted along the exponential path, hence, the forecasts revert monotonically toward a constant long-term average in the long-run. To examine the monotonic behaviour of GARCH models, consider the GARCH (1, 1) model with a forecasts structure adapted from Tsay (2010) as shown below:

$$\hat{h}_{t+k} = \alpha_0 (1-P)^{-1} (1-P^{k-1}) + P^{k-1}\hat{h}_{t+1}, \tag{7.1}$$

where $P = \alpha_1 + \beta_1 < 1$, $\alpha_1 \geq 0$ and $\beta_1 \geq 0$. Note that P is constant and $k = 1, 2, 3, \ldots$ is a deterministic instantaneous forecast horizon. The variable $\{P^{k-1}\}$ is monotone decreasing. As $k \to \infty$, $\{P^{k-1}\} \to 0$ and $(1-P^{k-1}) \to 1$, thus, the forecasts revert to the long-run variance $\alpha_0 (1-P)^{-1}$ along an increasing monotonic (exponential) path. The empirical monotonic path of the GARCH (1, 1) model is demonstrated in Figure 7.2. It is clearly observed from the Figure that, the monotonic mean-reversion trajectory of the conditional forecasts do not in any way mimic the underlying stochastic volatility path. In an attempt to address this problem, we use the exogenous proxies and their break variables introduced in chapter six to augment the GARCH models.

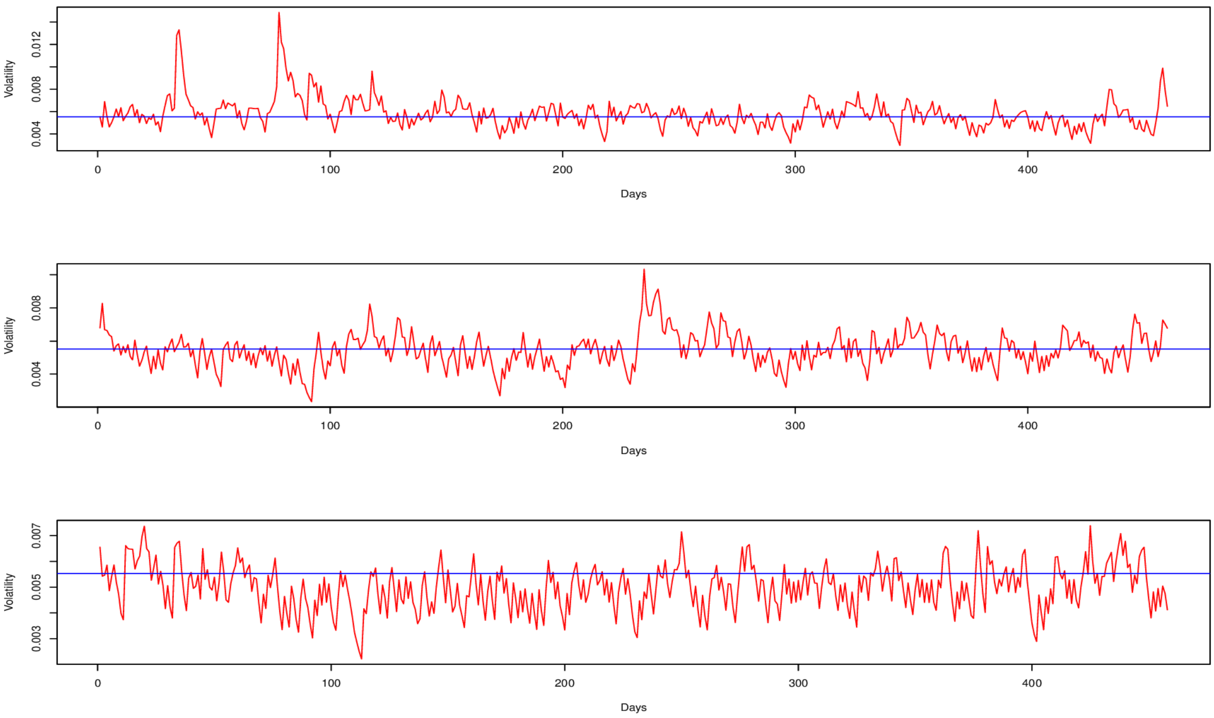

Figure 7.1: Estimated volatility of GBP/ZAR from GARCH (1, 1) model

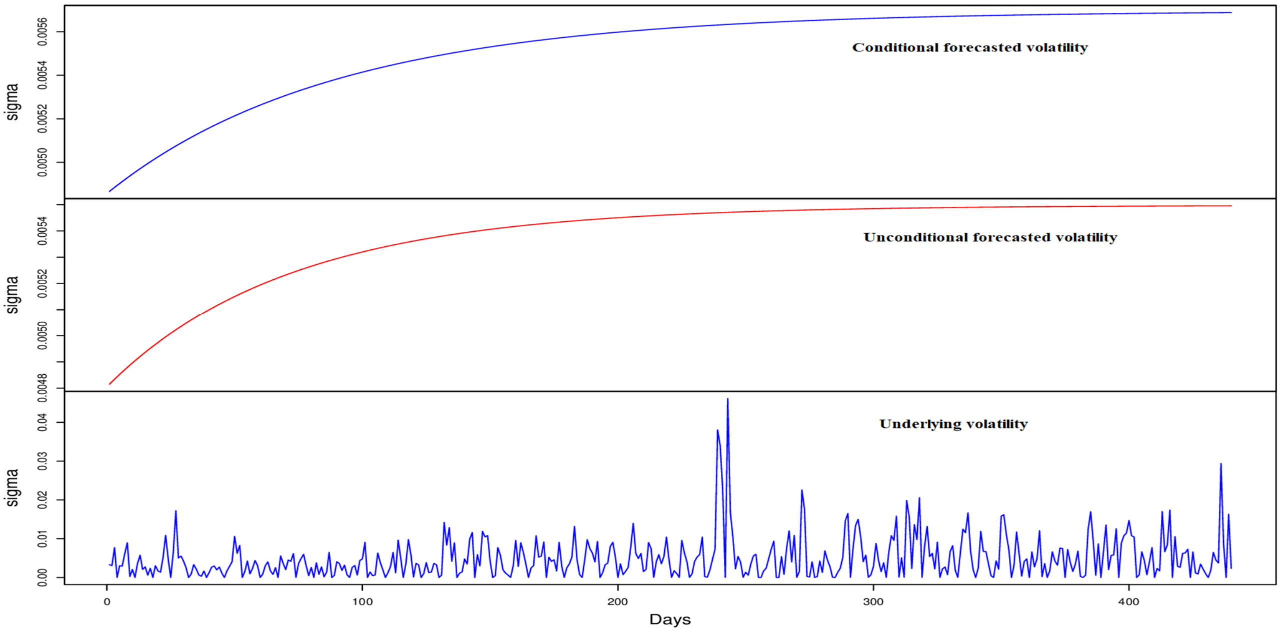

Figure 7. 2: Forecast and underlying volatility of GBP/ZAR from GARCH (1, 1) model

7.4 Augmentation of GARCH models

The proposed augmentation in this thesis requires the incorporation of the market uncertainty proxies and their break variable obtained in the previous chapter into the volatility modelling process. The GARCH models selected for this exercise include the standard GARCH, EGARCH, GJRGARCH, TGARCH, and NAGARCH. The augmentation is carried out by passing the variables to the variance equations of the above models as shown below:

$$h_t = \alpha_0 + \sum_{m=1}^{k} \theta_m B_{t-1} + \sum_{l=1}^{k} \lambda_l \left| r_{l,t-1} \right| + \sum_{i=1}^{q} \left(\alpha_i a_{t-i}^2 - \gamma_i I_{t-i} a_{t-i}^2 \right) + \sum_{i=1}^{p} \beta_i h_{t-i}, \tag{7.2}$$

$$\ln(h_t) = \alpha_0 + \sum_{m=1}^{k} \theta_m B_{t-1} + \sum_{l=1}^{k} \lambda_l \left| r_{l,t-1} \right| + \sum_{n=1}^{n} \gamma_n \left(\frac{\varepsilon_{t-n}}{\sqrt{h_{t-n}}} \right) + \sum_{i=1}^{p} \beta_i \left(\left| \frac{\varepsilon_{t-i}}{\sqrt{h_{t-n}}} \right| - E \left| \frac{\varepsilon_{t-i}}{\sqrt{h_{t-n}}} \right| \right) + \sum_{j=1}^{m} \alpha_j \ln(h_{t-j}), \tag{7.3}$$

$$h_t^{\lambda} = \omega + \sum_{m=1}^{k} \theta_m B_{t-1} + \sum_{l=1}^{k} \lambda_l \left| r_{l,t-1} \right| + \sum_{j=1}^{q} \alpha_j h_{t-j}^{\lambda} \left(\left| z_{t-j} - \eta_{2j} \right| - \eta_{1j} \left(z_{t-j} - \eta_{2j} \right) \right)^{\delta} + \sum_{j=1}^{p} \beta_j h_{t-j}^{\lambda}, \tag{7.4}$$

$$h_t = \omega + \sum_{m=1}^{k} \theta_m B_{t-1} + \sum_{l=1}^{k} \lambda_l \left| r_{l,t-1} \right| + \sum_{j=1}^{q} \alpha_j h_{t-j} \left(\left| z_{t-j} \right| - \left(z_{t-j} - \eta_{2j} \right) \right) + \sum_{j=1}^{p} \beta_j h_{t-j}, \tag{7.5}$$

$$h_t^2 = \omega + \sum_{m=1}^{k} \theta_m B_{t-1} + \sum_{l=1}^{k} \lambda_l \left| r_{l,t-1} \right| + \sum_{j=1}^{q} \alpha_j h_{t-j}^2 \left(\left| z_{t-j} - \eta_{2j} \right| \right)^2 + \sum_{j=1}^{p} \beta_j h_{t-j}^2 . \tag{7.6}$$

Equations (7.2) to (7.6) assume the same notations in section (2.3) and equation (6.8). To understand how the augmented models may relax the restrictive mean-reversion property of GARCH models and thus induce randomness on the paths of the forecasted unconditional and the conditional volatilities, we analysed the forecast structure of the augmented GARCH (1, 1) model. For horizon with length k, subject to the information set $\mathcal{F}_{t-1}$, following similar proof from Tsay (2010), the $k-step$ ahead forecast h_{t+k} is given by:

$$\hat{h}_{t+k} = \operatorname{var}\left[a_{t+k} \mid \mathcal{F}_{t-1}\right] = \mathrm{E}\left[a_{t+k}^2 \mid \mathcal{F}_{t-1}\right] = \mathrm{E}\left[\mathrm{E}\left(h_{t+k} z_{t+k}^2 \mid F_{t+k-1}\right) \mid \mathcal{F}_{t-1}\right] = \mathrm{E}\left[h_{t+k} \mid \mathcal{F}_{t-1}\right]. \tag{7.7}$$

Substituting (7.3) into (7.7) yields:

$$\hat{h}_{t+k} = \alpha_0 + \sum_{i=1}^{\max(p,q)} \left(\alpha_i + \beta_i\right) \mathrm{E}\left[h_{t+k-i}\right] + \sum_{j=1}^{m} \lambda_j \mu_{r_{j,t+k-1}} + \sum_{j=1}^{m} \psi_j \mu_{B_{j,t+k-1}} . \tag{7.8}$$

Note that $\mathrm{E}\left[h_{t+k}\right] = a_{t+k}^2$ for $k < 0$. If we consider the GARCH (1, 1) with returns from only one exchange rate along with its break variable, equation (7.8) reduces to:

$$\hat{h}_{t+k} = \alpha_0 + P h_{t+k-1} + \lambda_1 \mu_{r_{1,t+k-1}} + \psi_1 \mu_{B_{t+k-1}} , \tag{7.9}$$

where $P = \left(\alpha_1 + \beta_1\right)$. Recursive substitution into (7.9) with subsequent simplification yields:

$$\hat{h}_{t+k} = \left\{\alpha_0 \left(1-P\right)^{-1}\right\} + \left(\lambda_1 \mu_{r_{1,t+k-1}} + \psi_1 \mu_{B_{1,t+k-1}}\right)\left(1 + P^{k-1}\right) + P^k \left\{\left(1-P\right)^{-1}\left(\left(1-P\right) h_t - \alpha_0\right)\right\}. \tag{7.10}$$

Note that $\left\{\mu_{r_{1,t+k-1}}\right\}$ and $\left\{\psi_1 \mu_{B_{1,t+k-1}}\right\}$ are stochastic, $\lambda_1 \geq 0$ and $\psi_1 \geq 0$, thus $\left\{\lambda_1 \mu_{r_{1,t+k-1}} + \psi_1 \mu_{D_{tk-1}}\right\} \geq 0$. For positivity and convergence of $\hat{h}_{t+k}$, $\alpha_0 > 0$, $\alpha_1 > 0$, $\beta_1 > 0$ and $P < 1$. As $k \to \infty$, $\left\{1 + P^{k-1}\right\} \to 1$ and $P^{k-1} \to 0$, thus, the multi-step ahead forecast over k horizons converges to a sum of the unconditional variance and external processes as indicated below:

$$\hat{h}_{t+k} \xrightarrow[k \to \infty]{} \left\{\alpha_0 \left(1-P\right)^{-1} + \lambda_1 \mu_{r_{1,t+k-1}} + \psi_1 \mu_{B_{1,t+k-1}}\right\}. \tag{7.11}$$

The sum $\left\{\lambda_1 \mu_{r_{1,t+k-1}} + \psi_1 \mu_{B_{1,t+k-1}}\right\}$ is stochastic, therefore, the long-range forecasts revert along stochastic path towards the long-run variance $\alpha_0 \left(1-P\right)^{-1}$. The sum serves as an external drifter that controls the reversion trajectory. When $\left(\lambda_1 \mu_{r_{1,t+k}} + \psi_1 \mu_{B_{1,t+k-1}}\right) > \left(\lambda_1 \mu_{r_{1,t+k}} + \psi_1 \mu_{B_{1,t+k}}\right)$, the conditional volatility is drifted above the long-run variance and vice versa due to the positivity constraint imposed on the drifter $\left(\lambda_1 \mu_{r_{1,t+k}} + \psi_1 \mu_{B_{1,t+k-1}}\right) \neq \left(\lambda_1 \mu_{r_{1,t+k}} + \psi_1 \mu_{B_{1,t+k}}\right)$. The stochastic drifter relaxes the restrictive mean-reversion property of GARCH forecasts. This is because the volatility forecasts revert towards the long-run variance from either directions of the unconditional variance, unlike the restrictive version (which allows unidirectional mean-reversion), thus, it induces randomness on the path of the forecasts. The associated parameters λ_1 and ψ_1 control the degree and trajectory of the mean-reversion. If $\psi_1 > 0$ and or $\lambda_1 > 0$ the volatility reverts along the stochastic path towards the unconditional variance. If $\psi_1 = \lambda_1 = 0$, the volatility reverts towards the unconditional variance along the monotonic path (as in the restrictive case), therefore, the necessary condition for the augmented model to relax the restrictive property of GARCH models is when $\phi_1 > 0$ and or $\lambda_1 > 0$.

7.5 Empirical application of the Augmented GARCH models

Data and Methodology

The practical relevance of any extension or augmentations of GARCH model can be assessed when applied to real data. The focus of this section is, therefore, the application of the augmented versions of the standard GARCH, EGARCH, GJRGARCH, TGARCH and NAGARCH models to the closing returns of daily spot-ask prices from the South African inter-bank forex market prices from July 7, 2011, to July 3, 2016. The data consist of 1819 days of closing returns from fourteen selected currency pairs. The data is divided into two sub-samples - the first sample consisting of 1379 returns from July 7, 2011 to April 15, 2015, which are used for the in-sample estimation, while the remaining 440 returns from April 16, 2015 to June 28, 2016 are used for assessing the out-of-sample forecasts. Empirical analysis and descriptions of the data have already been addressed in chapter four. The mean components of the above augmented GARCH models are assumed to follow the ARMA processes defined below.

$$r_t = \varphi + \sum_{i=1}^{p} \lambda_i r_{t-i} - \sum_{i=1}^{q} \theta_r a_{t-r} + a_t , \tag{7.12}$$

where φ, λ_i and θ_r are parameters, r_t is endogenous returns, r_{t-i} is the i^{th} autoregressive terms with corresponding order p, a_{t-r} is the lag r^{th} moving average terms with order q and a_t as the innovation term. Table 7.1 summarises the specifications of the different types of models to be fitted. For each specification, two models will be fitted - the augmented version and the non-augmented version. Due to parameter instability associated with the QMLE approach, variance targeting is used in all the models in a bid to alleviate this problem.

Based on the results from chapter four, the following distributions are considered: Normal inverse Gaussian (NIG), Johnson SU Reparametrized distribution (JSU) and skewed generalized error distribution (SGED). In addition to the diagnostic tools discussed in chapter three, the adjusted Pearson goodness-of-fit test is also used to assess the distributional assumptions made about the GARCH innovations. Evidence from the preceding chapter suggests adequacies of the proxies and the dummy or break variables, thus, we anticipate that the augmentations may yield improved estimates and forecasts.

Table 7. 1: Specifications of fitted models[8]

Model	Returns	Proxy	Dummy	GARCH	ARMA	Target	Distribution
GJRGARCH	MWK	BRL	BRL	(1,1)	(5,5)	YES	NIG
NAGARCH	BWP	INR	INR	(3,0)	(2,1)	YES	JSU
SGARCH	BRL	MWK	MWK	(2,1)	(5,4)	YES	JSU
TGARCH	ILS	None	MWK	(2,2)	(1,1)	YES	NIG
EGARCH	SEK	INR, NOK	INR	(2,1)	(1,1)	YES	SGED
GJRGARCH	MWK	None	None	(1,1)	(5,5)	YES	NIG
NAGARCH	BWP	None	None	(3,0)	(2,1)	YES	JSU
SGARCH	BRL	None	None	(2,1)	(5,4)	YES	JSU
TGARCH	ILS	None	None	(2,2)	(1,1)	YES	NIG
EGARCH	SEK	None	None	(2,1)	(1,1)	YES	SGED

[8] The proxies and the dummy variables were selected based on the principle of parsimony via the backward elimination approach of variable selection. Model adequacy was also taking into consideration.

Empirical results and discussions

a) In-sample estimation

The diagnostic tests, summary statistics for estimated models, and the in-sampling volatility forecast performances are reported in Tables 7.2, 7.3 and 7.4 respectively. It is clearly seen from Table 7.2 that the empirical distributions are good approximations to the assumed theoretical distributions because the reported *p*-values from the adjusted-Pearson goodness-of-fit tests are above the conventional 5% significance level. Furthermore, the results from the Weighted Ljung-Box, weighted McLeod-Li, and Weighted Li & Mark tests suggest that no significant serial correlation and autocorrelation (ARCH effects) are remaining in the residuals, thus, it can reasonably be concluded that the estimated augmented models are good fits to the respective return series.

Table 7. 2: Diagnostic test

Model	SGARCH	EGARCH	NAGARCH	TGARCH	GJRGARC
WLB Test	0.928700	0.090560	0.064580	0.476200	0.100357
WML Test	0.331800	0.418700	0.567300	0.982920	0.261090
WLM Test	0.501900	0.210200	0.165300	0.997900	0.997900
APG Test	0.188000	0.399400	0.692900	0.731000	0.553100

The table reports the p-values for the Weighted Ljung-Box (WLB), weighted McLeod-Li (WLM) and the Weighted Li and Mark (WLM).

Table 7.3 displays the information criteria, unconditional variance, half-life, and persistence values. A model with a higher absolute AIC value is preferred since it is able to reduce information leakage. Observations from the Table suggest that the absolute AIC values for the augmented models are consistently higher than those of the non-augmented models (reference models), hence the augmented models have the tendency to reduce information leakage. Furthermore, it is observed that shocks persist more in the reference models than they do in the corresponding augmented models. Consequently, it takes a shorter time for half of the shocks in the augmented models to decay than it takes in the reference models. The results are in agreement with the results from Rapach, Strauss & Wohar (2007). The proposed method has the tendency to reduce the persistence of shocks, thus, it may be helpful in modelling series with IGARCH effect (Lamoureux & Lastrapes, 1990 and Bailie, Bollerslev & Mikkelsen, 1996).

Table 7. 3: Summary statistics for estimated models

Estimate	SGARCH	EGARCH	NAGARCH	TGARCH	GJRGARCH
AIC	**-7.382125**	**-7.807806**	**-7.944363**	**-7.694865**	**-5.915052**
	-7.378340	-7.737600	-7.890463	-7.607815	-5.914817
$\sigma \times 10^{-5}$	4.093307	2.361060	2.530314	1.882270	27.65188
	4.096637	2.516949	2.535729	1.904645	27.65382
Half-Life	**2.030559**	**1.025993**	**0.961771**	**3.218797**	**7.565556**
	2.063441	27.80431	9.135474	7.273614	8.959483
Persistence	**0.710805**	**0.508858**	**0.486412**	**0.806264**	**0.912453**
	0.714682	0.975379	0.926933	0.909104	0.925552

The values in the first rows of each estimate correspond to the augmented models while the values in the last rows correspond to the non-augmented models. Estimates for best models are in bold.

In Table 7.4, the Mincer-Zarnowitz's R-squares for the augmented models are larger than those from the reference models, hence, relatively higher proportions of the total variations in the augmented models are better explained than in the reference models and these proportions are quite substantial in the case of EGARCH, NAGARCH and TGARCH models. By implication, the consistent low MAE and RMSE metrics observed for the augmented model, in comparison to the reference models indicate that predictions from the augmented models are more accurate than predictions from the reference models.

Table 7.4: In-sampling volatility forecast performances

Model	SGARCH	EGARCH	NAGARCH	TGARCH	GJRGARCH
$MAE \times 10^{-5}$	**372.3573**	**274.2166**	**2.643384**	**5.153822**	**919.3562**
	373.0483	291.7725	2.777578	5.669522	919.7994
$RMSE \times 10^{-5}$	**454.8136**	**337.8631**	**0.002536**	**0.024081**	**1353.972**
	456.4939	354.7085	0.002591	0.028025	1356.719
MZ- Rsq	**1.098000**	**13.57000**	**5.200000**	**13.20701**	**6.980000**
	0.908000	3.470000	1.630000	4.038363	6.330000

The values in the first rows of each estimate correspond to the augmented models while the values in the last rows correspond to the non-augmented models. Estimates for best models are in bold.

Theoretical GARCH models may be correctly specified, although, the same cannot be said for the estimated models. A mis-specified model may yield inaccurate forecasts with misleading standard errors, thus, it is worth testing for the correct specification of an estimated model before using it to generate any forecasts. The GMM type moment (orthogonality) tests of Hansen (1982) were used in this exercise. Under a correctly specified estimated GARCH model, certain population moment conditions about the standardized residuals should hold in the sample, the moment conditions can be tested using individual t-test or jointly using the Wald test. Under these tests, the following moment conditions are tested:

$$\begin{array}{l} M_1 : \mathrm{E}\left[z_t\right]=0 \\ M_2 : \mathrm{E}\left[z_t^2-1\right]=0 \\ M_3 : \mathrm{E}\left[z_t^3\right]=0 \\ M_4 : \mathrm{E}\left[z_t^4-3\right]=0 \end{array} \text{ and } \begin{array}{l} Q_1 : \mathrm{E}\left[\left(z_t^2-1\right)\left(z_{t-j}^2-1\right)\right]=0 \\ Q_2 : \mathrm{E}\left[\left(z_t^3\right)\left(z_{t-j}^3\right)\right]=0 \\ Q_3 : \mathrm{E}\left[\left(z_t^4-3\right)\left(z_{t-j}^4-3\right)\right]=0 \end{array}, \tag{7.13}$$

where $j=1,\ldots,p$ are the lags, M_1 to M_4 denote the individual moment conditions and Q_1 to Q_3 are the joint conditional moment conditions corresponding to variance, skewness, and kurtosis, which are Chi-square distributed with p degrees of freedom. When the moment conditions are tested jointly using a Wald test, the joint distribution is also Chi-square distributed, but with $3p+4$ degrees of freedom. In this study, individual t-test was employed and the results are reported in Table 7.5. All the t-values are less than their corresponding critical values, thus, we can conclude with 95% confidence that the estimated augmented models are correctly specified.

Table 7.5: The GMM Orthogonality Test for Augmented models

Model	t-value	Q2/M1	Q3/ M2	Q4/ M3	Joint/ M4
	Critical(1)	3.841459	3.841459	3.841459	14.06714
	Critical(2)	5.991465	5.991465	5.991465	18.30704
SGARCH	JCMC	0.628780	0.631081	0.337932	8.717651
	IMC	0.101775	-0.005563	0.151219	-1.296606
EGARCH	JCMC	0.380442	0.538286	1.307274	9.577889
	IMC	0.577075	0.026608	-1.035767	0.671571
NAGARCH	JCMC	0.144517	5.076979	3.237231	18.06104
	IMC	0.918320	-0.151426	-0.374556	0.360243
TGARCH	JCMC	1.107339	2.419303	1.905318	8.675098
	IMC	-0.698064	0.667217	0.916918	-1.052868
GJRGARCH	JCMC	1.107339	2.415090	2.671087	9.390035
	IMC	-0.698064	0.667217	-0.922053	0.801745

JMC denotes joint conditional moment conditions and IMC denotes individual moment conditions. The test is based on the entire out-of-sample data and a lag length of 1 for SGARCH and 2 for the rest of the models. The 1 and 2 in front of the critical values indicates the lag used in computing the critical values.

b) Out-of-sampling forecasting

The augmented models are the best-fitted models, although, simulation studies from Kosapattarapim, Lin & McCrae (2011) and Shamiri & Isa (2009) suggest that the best-fitted GARCH models do not necessarily produce the best forecasts. It is, therefore, essential to check if there are trade-offs in forecast accuracies and whether such trade-offs are statistically significant or not before the augmented models can be adjudged superior to the reference models, in forecasting. In this regard, paired t-tests (or unpaired t-tests if the comparing metrics are not correlated) with the following hypotheses were carried out:

$$\begin{aligned} H_0 &: \mu_m - \mu_n = 0 \\ H_a &: \mu_m - \mu_n > 0 \end{aligned}, \tag{7.14}$$

where μ_m denotes the MAE or MSE of the augmented model and μ_n denotes the MAE or MSE of the non-augmented model. If we fail to reject the null hypothesis at a 5% level of significance, we can conclude that the forecast accuracy trade-offs are statistically insignificant and the model, which produces stochastic heteroskedastic, forecast is preferred for forecasting medium to long-term volatility. The *p*-values for the paired t-test for MAE, RMSE, MAEMZR, RMSEMZR, and the MZR R-Square values are reported in Tables 7.6 to 7.8. Considering the 1-day ahead forecasts in Table 7.6, the MAE for all the augmented models are consistently smaller than the non-augmented versions of Zakoian (1994), Engle & Ng (1993), Glosten, Jaganathan & Runkle (1993), Nelson (1991) and Bollerslev (1986), except for the GJRGARCH model.

Table 7. 6: 1-day ahead forecasts metrics based on MAE

Model	SGARCH	EGARCH	NAGARCH	TGARCH	GJRGARCH
Augmented	7.6937E-03	3.5945E-03	3.6974E-03	3.6161E-03	1.2424E-03
Reference	8.3548E-03	4.2054E-03	4.9786E-03	4.5817E-03	1.2293E-03

At horizon 180, four out of the five forecast performance metrics favour the augmented SGARCH model. The difference between the only metric (MAEMZR) where the reference model is favoured is not significant at 5%. The null hypothesis of the Wald test for the coefficients of the MZR is rejected at 5% for the SGARCH models and the augmented model has the highest R-square values. Similar results are also obtained for the GJRGARCH models, except that the null hypothesis of the Wald test for the coefficients of the MZR is not rejected at 5%. For the EGARCH, NAGARCH, and the TGARCH models, the null hypotheses of the Wald tests are not rejected, however, the augmented models have the highest R-square values. It is interesting to note that in addition to the fact that, the augmented models produced the highest R-square values, the majority of the metrics for all models, with the exception of the NAGARCH and TGARCH models do favour the superiority of the augmented models over the non-augmented versions of Glosten, Jaganathan & Runkle (1993), Nelson (1991) and Bollerslev (1986). In the case of the NAGARCH and the TGARCH models, where the non-augmented models are favoured, there is, however, only one metric for NAGARCH and two for TGARCH where the differences in metrics are significant at 5%.

Table 7.7: 180-days ahead forecast metrics based on MAE, RMSE and MZR based metrics

Model	Metric	SGARCH	EGARCH	NAGARCH	TGARCH	GJRGARCH
Augmented	MAE$\times 10^{-5}$	413.8929	344.3483	393.3816	431.0459	1111.496
Reference		417.5581	300.0124	378.5148	421.2554	1113.902
		(0.21407)	(0.00043)	(0.08329)	(0.28270)	(0.69657)
Augmented	RMSE$\times 10^{-5}$	579.6573	403.2133	479.4797	872.0268	1389.378
Reference		580.7808	367.9401	474.2117	856.6003	1389.472
		(0.64838)	(0.00589)	(0.74045)	(0.02690)	(0.9888)
Augmented	MAEMZR	391.1884	282.7771	1381.367	416.4231	781.9995
Reference		390.0674	285.0617	1381.325	417.9018	786.7399
	$\times 10^{-5}$	(0.62819)	(0.59136)	(0.98087)	(0.54068)	(0.36149)
Augmented	RMSEMZR	579.3001	367.5536	2340.027	869.0112	1337.508
Reference		579.7909	369.9846	2343.104	864.5541	1336.206
	$\times 10^{-5}$	(0.78905)	(0.56203)	(<0.0001)	(0.00150)	(0.83611)
Augmented	MZR-RSQ	0.56680*	2.142100	4.0083	0.6651	1.5495
Reference		0.12620*	0.105000	3.6120	0.2564	0.2949
	Votes	4/5	3/5	2/5	2/5	4/5

*Reported p-values t-tests for the differences in loss functions and MZR related metrics. The test is significant at 5%. MZR-RSQ values with * means the null hypothesis of Wald test is rejected at 5%.*

The majority of the metrics for 440-days ahead forecasts were in support of the superiority of all the forecasts from augmented models, except for the EGARCH model. The differences in three of the metrics for the EGARCH model are significant at 5%. With the exception of the SGARCH and the GJRGARCH models, the null hypotheses of the Wald test are rejected for all models and the augmented models recorded the highest R-square values for all the five models.

From the discussions above, it can be concluded that, in terms of the five forecast performance metrics, the augmented version of the SGARCH model of Bollerslev (1986) produced superior forecasts at all horizons in comparison to the non-augmented version. In the case of the EGARCH model of Nelson (1991), the non-augmented version produced superior forecasts at horizon 440, while the augmented version produced superior forecasts at horizons 1 and 180. In the case of the NAGARCH model of Engle & Ng (1993) and the TGARCH model of Zakoian (1994), the non-augmented versions produced superior forecasts at horizon 180, while the augmented versions produced superior forecasts at both horizon 1 and 440. Finally, in the case of the GJRGARCH model of Glosten, Jaganathan & Runkle (1993), the non-augmented version produced better one-day ahead forecast than the augmented version, but the conclusion for 180 and 440 days ahead forecasts are the opposite. In general, the augmentation of the GARCH models yielded improved volatility forecasts at all horizons, which is in agreement with the results from Cantamessa, Gautam & Xiang (2016), Kambouroudis & McMillan (2016), Rapach, Strauss & Wohar (2007) and Li & Engle (1998), but in direct contrast to the results from Karlsson (2016) and Babikir *et al.,* (2012).

Table 7. 8: 440-days ahead metrics based on MAE, RMSE and MZR based metrics

Model	Metric	SGARCH	EGARCH	NAGARCH	TGARCH	GJRGARCH
Augmented	$MAE\times10^{-5}$	456.7868	403.3447	472.9303	432.7812	1484.435
Reference		463.4463	372.5623	466.6622	411.6663	1491.331
		(0.00259)	(<0.0001	(0.20588)	(0.00019)	(0.05684)
Augmented	$RMSE\times10^{-5}$	618.1244	541.6421	680.2091	748.2984	2217.511
Reference		619.9167	522.3589	689.7097	726.2627	2219.186
		(0.36075)	(0.00172)	(0.03738)	(<0.0001)	(0.59625)
Augmented	MAEMZR	437.9258	362.0686	464.4777	410.028	1381.367
Reference	$\times10^{-5}$	440.4601	361.7976	464.3997	410.682	1381.325
		(0.31589)	(0.81083)	(0.97065)	(0.55413)	(0.98087)
Augmented	RMSEMZR	627.8748	535.6042	690.4555	728.9321	2340.027
Reference	$\times10^{-5}$	629.2035	533.7110	693.3301	729.3406	2343.104
		(0.53339)	(0.04359)	(<0.0001)	(<0.0001)	(0.04109)
Augmented	MZR-RSQ	1.421600*	0.173400	0.5717	0.3339	0.1677*
Reference		0.00700*	0.222190	0.4720	0.2234	0.0135*
	Votes	5/5	1/5	3/5	3/5	3/5

*Reported p-values t-tests for the differences in loss functions and MZR related metrics. The test is significant at 5%. MZR-RSQ values with * means the null hypothesis of Wald test is rejected at 5%.*

The main motivation for the thesis is the relaxation of the restrictive mean reversion property of GARCH models so that, the forecasts revert to the long-run volatility stochastically. In this regard, it is appropriate to check if the augmentations were able to achieve this objective. GARCH forecasts converge towards unconditional volatility it is, hence, sufficient to observe the unconditional volatility forecasts to assess the mean reversion trajectory of the forecasts. Figure 7.3 displays the forecasted unconditional standard deviations of the augmented models. It can be observed that, the unconditional variances (smooth random lines) revert to the long-run variances (black straight lines) stochastically from either direction. Unlike the non-augmented versions, the FIGARCH model of Baillie & Morana (2009) and the time-varying models of Amado & Teräsvirta (2014) that use time as a transitional variable, the mean-reversion paths of the conditional volatilities are bi-directional and stochastic, therefore, are consistent with the underlying volatility paths.

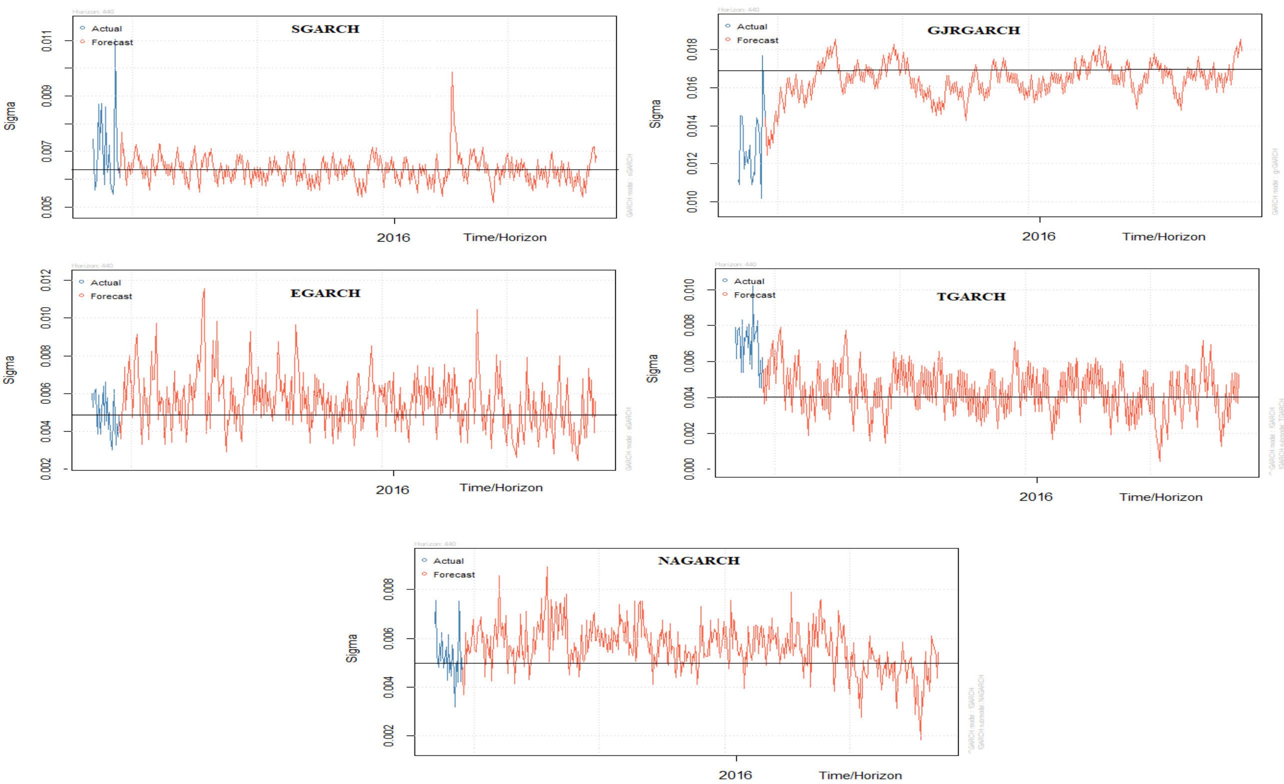

Figure 7. 3: Time plots for forecasted unconditional volatilities

7.6 Conclusion

In this chapter, proxies for market uncertainty and their constructed break variables were used to augment the univariate standard GARCH, EGARCH, GJRGARCH, TGARCH, and NAGARCH models in an attempt to account for changes and breaks in their unconditional volatilities. The augmented models were then used to model and forecast the volatilities of rand-denominated currencies and their performances were compared to the non-augmented versions using the mean absolute error and root mean squared error metrics as well as the Mincer-Zarnowitz related forecast metrics.

Evidence from the empirical study indicates that unlike the non-augmented versions of the Zakoian (1994), Engle & Ng (1993), Glosten, Jaganathan & Runkle (1993), Nelson (1991) and Bollerslev (1986), the augmented models reduce information leakages and persistence of shocks in volatility, which is in direct agreement with the results from Rapach, Strauss & Wohar (2007). Subsequently, the times taken for half of the shocks in the volatilities from the augmented models to decay were reduced. All the augmented models were adjudged the best-fitted models. The augmentation also relaxed the restrictive mean reversion property of GARCH forecasts so that the forecasts revert stochastically to the long-run volatility, which is consistent with the underlying volatilities of the currency pairs considered. The augmentation also improved the explanatory powers of the generated forecasts.

The augmented SGARCH model produced superior forecast in terms of the MAE and RMSE metrics at all horizons with improved explanatory powers and the forecasts also revert to the long-run volatility stochastically, thus, the model is recommended for volatility estimates at all horizons. The reference EGARCH model outperformed the augmented model in terms of the long-range forecasts, although, it seems more appropriate and plausible to recommend the augmented EGARCH model for long term volatility forecasting. This is because the number of forecast metrics with significant trade-off accuracies is small, and the augmented model has higher explanatory forecasts power than the reference model. In addition, the forecasts from the augmented models revert to the long-run volatility stochastically. The augmented NAGARCH and TGARCH models are the best-fitted models with plausible underlying volatility paths and since there are no significant trade-offs forecast accuracies, the models are preferred when forecasting both short and long-range volatilities. The augmented GJRGARCH model is best suited for long term forecasting, but not for a day ahead forecasts. In general, the proposed augmentations produced best-fitted models and stochastic mean-reverting forecasts at longer horizons with improved explanatory powers.

CHAPTER 8

Estimation and forecasting of VaR using augmented GARCH models.

Chapter Summary

One of the immediate applications of volatility forecasts is its use as an input in VaR computations. Accurate volatility forecast yields better VaR forecasts, which may save financial institutions from regulatory sanctions, and overexposure to market risks. In this study, the proposed augmented GARCH models introduced in the preceding chapter are used to forecast 1% and 5% VaR and their forecasting abilities are assessed and are compared to some benchmark models.

Evidence from the empirical study suggests that the augmented models lead to fewer violations, improved 1% VaR forecasts, and optimal daily capital requirements for all the models. There is, however evidence of relative superiority of the majority of the models for the 5% VaR forecasts from the augmented models, although they have relatively higher failure rates. The superiority results based on the 1% VaR estimates of the augmented VaR models confirm the superiority of the volatility models established in chapter seven. We cannot, however, make such generalization based on the 5% VaR estimates, although the majority of the augmented models outperformed the non-augmented version. Based on these results, we recommend the integration of our approach into existing risk modelling frameworks. It is believed that such models may lead to, fewer bank failures, exposure of banks to optimal market risks, and assist them in computing optimal regulatory capital requirements and minimize penalties from regulators.

8.1 Introduction

Risk measurement is one of the most essential tasks in financial risk management for banks, corporate treasuries, portfolio management firms, and other financial institutions and practitioners. In financial institutions, risk measures are used to specify capital requirements - the amount of capital that must be added to a position to make its risk acceptable to regulators (Francq & Zakoıan, 2010). In addition, risk measures are employed in decision-making regarding hedging and portfolio optimization. One of the commonly-used risk measures among financial institutions is the VaR -with the underlying asset's volatility as an input. Inaccurate volatility forecasts may lead to underestimation or overestimation of the actual VaR forecast and financial institutions may lose the opportunity cost or cannot recover losses from crisis periods (Su & Hung, 2018). Furthermore, financial institutions may be over- exposed to risk and face regulatory sanctions.

VaR computation requires volatility input, since VaR accuracy is dependent on accurate volatility forecasts. Accurate volatility forecasts yield optimal VaR forecasts and less violations, thus VaR's forecasting assessment provides an indirect assessment of the predictive abilities of competing volatility models (Bucci, 2017). A common approach for computing the volatility input is the use of GARCH models (Aridi, Cheong & Hooi, 2018), thus we use the augmented models in chapter seven to compute the volatility input. Our main aim in this chapter is to forecast VaR, but since accurate volatility forecasts yield accurate VaR forecasts and the volatility component is computed using the augmented models, we also seek to use the assessment of the estimated VaR models to confirm the superiority of the augmented forecast established in chapter seven.

Procedures used in backtesting VaR estimates are predominantly centred on the dynamic quantile test of Engle & Manganelli (2004), the conditional coverage test of Christoffersen (1998), and the unconditional coverage test of Kupiec (1995). Recent and past studies from Bayer (2018), Laporta, Luca & Petrella (2018), Su & Hung (2018), Bams, Gildas & Thorsten (2017), Su (2015) and Lee & Su (2012), among others, have used these procedures to assess the accuracies of individual VaR estimates. For competing VaR models, the model confidence set test (MCS) procedure of Hansen, Lunde & Nason (2011) and the superior predictive ability (SPA) test of Hansen (2005) are among the most commonly used tests. In this study, the conditional coverage test of Christoffersen (1998) and the unconditional coverage test of Kupiec (1995) were used to assess the accuracies of the individual VaR estimate, while the MCS procedure is used to compare competing models.

The rest of the chapter is structured as follows - a comprehensive overview of VaR concept is discussed in section 8.2, while VaR estimation procedures are presented in section 8.3. Empirical applications of the VaR estimation procedures proposed in this study are presented in section 8.4 and finally, the results of this chapter are summarised in section 8.5.

8.2 The concept of VaR

VaR is concerned with the possibilities of losses associated with a portfolio, at a given time. It is a downside risk metric, which measures the risk as to whether the actual return will fall below or above the expected returns. In other words, VaR is the uncertainty about the magnitude of the differences in returns and the expected returns (Roccioletti, 2015). It is the preferred risk metric by many experts because of its perceived superiority in backtesting the estimated losses (Roccioletti, 2015). Technically, VaR is defined as the maximum portfolio loss at a given confidence level, α in a time interval, in a scenario where there is a portfolio of risky assets held over a fixed time in the horizon Δ (Tsay, 2010). If the loss distribution associated with this portfolio has distribution function $F_L(l) = P(L \leq l)$, the maximum possible loss $inf\{l \in R : FL(l) = 1\}$ evaluates the level of risk associated with holding the portfolio over time Δ. This scenario leads to the technical definition of VaR below.

Definition 8.1 Given a confidence level $\alpha \in (0,1)$, the VaR of a portfolio is given by the smallest number l such that the probability that the loss L exceeds l is no larger than $(1-\alpha)$ (Roccioletti, 2015), that is:

$$\mathrm{VaR}_\alpha = \inf\{l \in \mathbb{R} : P(L > l) \leq 1-\alpha\} = \inf\{l \in \mathbb{R} : F_L(l) > \alpha\}. \tag{8.1}$$

In probabilistic terms, VaR is a quantile of the loss distribution. Given losses L, the generalized inverse $F\leftarrow$ is called the quantile function of L such that:

$$\mathrm{VaR}_\alpha(L) = q_\alpha(L) := F^{\leftarrow}(L) = \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\}. \tag{8.2}$$

Alternatively, VaR can be constructed from the probabilistic function of the underlying returns.

Definition 8.2 Consider the returns of an asset r_t with the change in the value of the asset over the next k periods defined by $\Delta r_t = \Delta V(k) = r(t+k) - r(k)$. The VaR over time horizon k associated with the left tail probability α of the returns' distribution is defined by (Roccioletti, 2015) as:

$$\alpha = P[\Delta V(k) \leq VaR_t] = P[\Delta r \leq VaR_t]. \tag{8.3}$$

VaR is a function of time and the left tail quantile of the distribution with probability α. Throughout this chapter, all computations and assessments relating to VaR models are based on the definition (8.2). Typically, if the returns are normally distributed, the $(1-\alpha)$ confidence level VaR estimate (adapted from Roccioletti, 2015) is computed as:

$$VaR_t = Z_\alpha \hat{\sigma}_t + \mu, \tag{8.4}$$

where VaR_t is the $(1-\alpha)\%$ estimated VaR at time t, Z_α denoted by $P[Z < Z_\alpha] = \alpha$ is the left quantile probability from the standard normal distribution and $\hat{\sigma}_t$ refers to the estimated standard deviation at time t. In as much as the VaR computations in equation (8.4) is not computationally cumbersome, asset returns exhibit heavy-tails with high peaks, thus, in conformity with literature and previous studies in the preceding chapter, the returns of assets in this study are modelled with non-normal distributions.

VaR has several attractive features, which make it the preferred risk metric among financial institutions. The metric corresponds to an amount that could be lost at some preselected probability. It also measures the nature of the risk factors and the risk factor sensitivities. The metric applies to all activities and types of risks in financial institutions and it can be compared across different markets at different exposures. In addition, it can be measured at any level, from a single trade or portfolio case up to a single enterprise-wide metric covering all the risks in the firm, as a whole. It can be used to find the total VaR of a very large portfolio in aggregated form or to isolate component risks corresponding to different types of risk factors in disaggregated form. This would be in addition to the metric accounts for the dependencies between the component of assets or portfolios (Alexander, 2008).

One of the main drawbacks of VaR metric is its non-sub-additivity property. This property of VaR contradicts the principles of diversification, hence, the foundations of modern portfolio theory (Szegö, 2004). Evidence from literature, however, suggest that this is not a serious issue in many practical applications provided the underlying risks have a finite variance or, in some cases, a finite mean (Emmer *et al.,* 2015).

VaR also estimates the upper bound on the losses that occur with a given frequency, thus, we do not know anything about the sizes of the potential losses, which is of much interest to financial risk practitioners. To address these theoretical drawbacks of VaR, an alternative risk measure, the expected shortfall (ES) is used. Given an integrable loss function L $with$ $E(|L|)<\infty$ having a continuous distribution function F_L and a confidence level $\alpha \in (0,1)$, the expected shortfall is defined by (Roccioletti, 2015) as:

$$ES_\alpha = E(L|L \geq VaR_\alpha) = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(L)\,du\,. \quad (8.5)$$

In this study, we compute both the VaR and the ES estimates, however, the ES estimates are used for comparison purposes.

8.3 Desirable Properties of VaR

A good risk measure is required to possess the following desirable properties: coherency, elicitability, conditional elicitability, and robustness. VaR possess all these properties except that in the general sense, it is not coherent due to its non-subadditivity, however, in certain instances the measure is sub-additive. The following are some of the instances when VaR is sub-additive:

i. the random variables are independent and identically distributed, as well as positively regularly varying,

ii. the random variables have an elliptical distribution,

iii. the random variables have an Archimedean survival dependence structure etc., and

iv. although the sub-additivity of VaR is conditional, it is comonotonic additive.

In this section, we briefly discuss these properties. For details on the above instances and other possibilities where VaR is sub-additive see Emmer *et al.,* 2015, Daníelson *et al.* (2013) and Embrechts *et al.,* (2013).

8.3.1 Coherency

Definition 8.3: A risk measure ρ is said to be coherent if it is homogenous, subadditive, monotonic and translationally invariant (Emmer *et al.*, 2015). Mathematically, coherency of ρ indicates that the following axioms hold:

i. Homogeneity: If for all loss variables L and $h \geq 0$, $\rho(hL) = h\rho(L)$,

ii. Subadditivity: If for all loss variables L_1 and L_2, $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$,

iii. Monotonicity: If for all loss variables L_1 and L_2, $L_1 \leq L_2 \Rightarrow \rho(L_1) \leq \rho(L_2)$,

iv. Translation invariance: If for all loss variables L and $a \in \mathbb{R}$, $\rho(L-a) = \rho(L) - a$.

A complementary property of subadditivity is comonotonic additivity.

Definition 8.4: According to Emmer *et al.,* (2015) A risk measure ρ is said to be comonotonically additive if for any given two real-valued random variables, L_1 and L_2, there is a real-valued random variable X and non-decreasing functions f_1 and f_2:

$$L_1 = f_1(X),\ L_2 = f_2(X) \text{ and } \rho(L_1 + L_2) = \rho(L_1) + \rho(L_2). \tag{8.6}$$

A subadditive and comonotonically additive risk measure rewards diversification but, do not attach any diversification benefits to comonotonic risks (Emmer et al., 2015).

8.3.2 Elicitability

Definition 8.5: Let ν be a functional on a class of probability measures $\mathcal{P}$ on $\mathbb{R}$ such that $\nu : \mathcal{P} \to 2^{\mathbb{R}}$ (power set of $\mathbb{R}$) and $P \mapsto \nu(P) \subset \mathbb{R}$ is said to be elicitable relative to the $\mathcal{P}$ if and only if there is a scoring function s that is strictly consistent for ν relative to $\mathcal{P}$.(Emmer et *al.,* 2015)

The scoring function is defined by $s : \mathbb{R} \to [0, \infty)$ and $(x, y) \to s(x, y)$ if x and y are the point forecasts and observations, respectively. The scoring function is consistent for v relative to the class $\mathcal{P}$ if and only if, for all $P \subset \mathcal{P}, t \in v(P)$ and $x \in \mathbb{R}$ $\mathcal{P}$,

$$\mathrm{E}_P\left[s(t,L)\right] \leq \mathrm{E}_P\left[s(x,L)\right], \tag{8.7}$$

where L is a real-valued random variable with distribution P. The scoring function is strictly consistent if it is consistent and:

$$\mathrm{E}_P\left[s(t,L)\right] = \mathrm{E}_P\left[s(x,L)\right] \Rightarrow x \in v(P). \tag{8.8}$$

Elicitability is useful in determining optimal point forecasts (Emmer et al., 2015). In addition, the property is useful in comparing the performance of different forecast methods (Gneiting, 2011).

8.3.3 Conditional elicitability

Every elicitable risk measure is also conditionally elicitable (i.e. second-order elicitable), hence, the elicitability of VaR implies that it is also conditionally elicitable (Emmer *et al.,* 2015).

Definition 8.6: A functional v belonging to the probability measures $X_n \sim P_n, n \geq 1$ $\mathcal{P}$ is conditionally elicitable if the functionals $\tilde{\gamma}$ and $\gamma\!:\!\mathrm{D} \to 2^{\mathbb{R}} \big| \mathrm{D} \subset \mathcal{P} \times 2^{\mathbb{R}}$ exist (Emmer *et al.,* 2015) such that:

i. $\tilde{\gamma}$ is elicitable relative to $\mathcal{P}$,

ii. $\left(P, \tilde{\gamma}(P)\right) \in \mathrm{D}$ for all $P \in \mathcal{P}$,

iii. for all $c \in \tilde{\gamma}(\mathcal{P})$, the functional $\gamma_c : \mathcal{P}_c \to 2^{\mathbb{R}}, P \mapsto \gamma(P,c) \subset \mathbb{R}$ is elicitable relative to $\mathcal{P}_c = \left\{P \in \mathcal{P} : (P,c) \in \mathrm{D}\right\}$,

iv. $\nu(P) = \gamma\left(P, \tilde{\gamma}(P)\right)$ for all $P \in \mathcal{P}$.

In backtesting and forecast comparison of VaR models, conditional elicitability allows us to decompose complex forecast method into components, which are validated separately. This approach is attractive in the sense that it allows for an opportunity to make complex forecasting methods tractable. The main drawback of this approach is that the optimal choice of forecast models may not be necessarily involved (Emmer *et al.,* 2015).

8.3.4 Robustness

Robustness is another desirable property of VaR measure. This property suggest that small measurement errors in the loss distribution have insignificant impact on the VaR estimate, thus, robustness is investigated in terms of continuity (Emmer *et al.,* 2015). Unlike other risk measures, VaR is both continuous with respect to weak topology and other stronger notion of convergence, such as the Wasserstein distance (Emmer *et al.,* 2015 and Bellini *et al.,* 2014).

The motivation behind the use of Wasserstein distance for robustness check lies in the fact that large observations which are neither outliers nor measurement errors do occur in finance and insuranc (Emmer *et al.,* 2015).

Definition 8.7: Let $P_n, n \geq 1$ and P be probability measures (Emmer *et al.,* 2015). Given that $X_n \sim P_n, n \geq 1$ and $P \sim X$, a risk measure ρ is continuous at X with respect to the Wasserstein distance if:

$$\lim_{n\to\infty} d_w\left(X_n, X\right) = 0 \Rightarrow \lim_{n\to\infty}\left|\rho\left(X_n\right) - \rho\left(X\right)\right| = 0. \tag{8.9}$$

where d_w is the Wasserstein distance defined for two probability measures P and Q by:

$$\lim_{n\to\infty} d_w\left(P, Q\right) = \inf\left\{\mathrm{E}\left(\left|X - Y\right|\right) : X \sim P, Y \sim Q\right\}. \tag{8.10}$$

8.3 Estimation of VaR

There are several methods used in computing VaR in financial literature. The most common ones are the mean-variance approach, historical simulation method, Monte Carlo simulations, and Extreme Value Theory (Kuester, Mittnik & Paolella, 2006 and Dowd, 2002). The historical simulation method uses historical returns to construct sample quantiles. One of the advantages of this approach is that, it determines the joint probability distribution of the market variables and avoids the need for cash-flow mapping, however, it is computationally slow and does not easily allow volatility updating schemes to be used (Hull & White, 1998). Notwithstanding the disadvantage of the historical simulation method, the study focuses on this approach to estimate and forecast VaR, the reason being that the models that are being assessed are built on historical simulation approaches. Consider the return series $\{R_t\}$ (adapted from Tsay, 2010), such that:

$$\begin{aligned} R_t &= \varphi_0 + \sum_{j=1}^{k} \lambda_j r_{jt-1} + \sum_{i=1}^{p} \varphi_i R_{t-i} - \sum_{i=1}^{q} \theta_r \varepsilon_{t-r} + \varepsilon_t, \\ \varepsilon_t &= \eta_t \sqrt{h_t} \end{aligned} \tag{8.11}$$

where h_t assume the forms discussed in section 7.4, with the associated non-normal distributions for the innovations. By definition, VaR estimates based on these models are computed from the formula:

$$VaR_t^D = Z_\alpha^D \sqrt{\hat{h}_t} + \mu, \tag{8.12}$$

where VaR_t^D denotes the VaR estimate based on GARCH models with an appropriate distribution, $\hat{h}_t$ is the forecasted volatility at time t and μ is the estimated mean from the GARCH model. Using the appropriate distributions associated with the augmented models, VaR estimates in this study are computed using the following specifications:

$$VaR_t^{NIG}(A) = Z_\alpha^{NIG} \sqrt{\hat{h}_t} + \mu, \tag{8.13}$$

$$VaR_t^{JSU} = Z_\alpha^{JSU} \sqrt{\hat{h}_t} + \mu, \tag{8.14}$$

$$VaR_t^{SSTD} = Z_\alpha^{SSTD} \sqrt{\hat{h}_t} + \mu, \tag{8.15}$$

where NIG is the Normal inverse Gaussian, JSU is the Johnson SU Reparametrized distribution and SSTD is the skewed student's t-distribution. The volatility component in equation (8.13) assumes the GJRGARCH and TGARCH models, that equation (8.14) assumes NAGARCH model and the SGARCH and EGARCH models for equation (8.15). This is because in as much as we want to evaluate the forecasting performance of the VaR estimates based on volatility forecasts from the augmented models, we are also interested in the assessment of the estimated VaR models to confirm the superiority of the augmented forecast as established in the previous chapter.

8.4 Empirical results for VaR estimations

In this section, we applied data from the inter-bank rand forex market to the VaR estimation procedures discussed in section 8.3. The same data, sub-samples, returns, and volatility model specifications used in chapter seven are used here; hereafter, the augmented VaR models are denoted VaR_A while the benchmarking models are denoted VaR_B. For each model, one-step-ahead VaR forecasts at 5% and 1% confidence levels corresponding to the risk metrics methodology and the requirements of the Basel II Accord were obtained. The rolling window with a re-fitting interval of 100 days was used.

8.4.1 Backtesting the estimated VaR models

Figures D1 to D5 in Appendix D display the plots of realized returns, out-of-sample VaRs predictions, and VaR violations. Inspections of the plots give impressions that there are very few VaR violations in all the models (represented by the red dots) for the 99% quantile, but relatively more violations for the 95% quantile. There is pre-indication that all the models may respond early to changing market conditions. This is due to the fact that the evolutions of the VaR estimates seem to be non-clustering, however, albeit useful, the visual inspections do not constitute proper backtest analysis, hence we proceed to discuss the Kupeic's likelihood ratio test and the Christoffersen's interval forecast tests which are reported in Table 8.1.

The critical values corresponding to 1% and 5% Chi-Squares with one degree of freedom are 6.635 and 3.841, respectively. Similarly, the corresponding two degrees of freedom critical values are 9.21 and 5.99 respectively. A good VaR estimate must pass both the unconditional coverage and the independence tests, thus, our interests lie in its failing to reject the null hypotheses of these tests. The null hypotheses are rejected when the test statistics are higher than the corresponding critical values. It is noted from Table 8.1 that the test statistics for all the models are below the corresponding critical values, except for the *VaR (GJR)* and $VaR_A(S)$ models for the 5% VaR estimates, therefore, the null hypotheses for correct exceedances and independence of failures are not rejected for the models with smaller test statistics. These signify that the respective models passed the tests at 1% and 5% confidence levels. Observations from the same Table indicate that all the models for the 99% VaR estimates, with the exception of the $VaR_B(T)$, have LR_{UC} test statistics which are less than the 6.635 critical value, thus we failed to reject the null hypotheses of correct exceedances of VaR violations from the models. Similarly, we fail to reject the null hypotheses of correct exceedances of VaR violations from the 95% VaR estimated models, with the exception of the VaR_A *(GJR)*, VaR_B *(GJR)*, and $VaR_A(S)$ models.

Table 8. 1: In-sample Backtesting of the estimated VaR models

Models	violations (VaR)	Kupeic's likelihood ratio test		Conditional coverage test	
		LR_{UC}	Reject	LR_{CC}	Reject
99% Confidence level-VaR estimation					
VaR_A(GJR)	4	0.038	No	0.111	No
VaR_B(GJR)	4	0.038	No	0.111	No
VaR_A(NA)	3	0.507	No	0.548	No
VaR_B(NA)	3	0.507	No	0.548	No
VaR_A(S)	10	5.292	No	5.758	No
VaR_B(S)	10	5.292	No	5.758	No
VaR_A(T)	10	5.292	No	5.758	No
VaR_B(T)	11	7.059	Yes	7.624	No
VaR_A(E)	6	0.528	No	0.694	No
VaR_B(E)	9	3.730	No	4.107	No
95% Confidence level-VaR estimation					
VaR_A(GJR)	34	5.949	Yes	11.67	Yes
VaR_B(GJR)	34	5.949	Yes	11.67	Yes
VaR_A(NA)	23	0.047	No	0.088	No
VaR_B(NA)	21	0.049	No	2.159	No
VaR_A(S)	32	4.222	Yes	5.315	No
VaR_B(S)	30	2.763	No	3.509	No
VaR_A(T)	27	1.119	No	1.193	No
VaR_B(T)	26	0.725	No	0.869	No
VaR_A(E)	26	0.725	No	4.001	No
VaR_B(E)	25	0.413	No	3.434	No

Correct conditional correct unconditional coverage is implied by correct conditional coverage, but not vice versa. All VaR estimates are based on 1% and 5% short positions corresponding to 99% and 95% confidence levels respectively.

The above results suggest that, all the models are significantly accurate and acceptable for making risk decisions, however, we cannot categorically make this conclusion. This is due to the fact that, the unconditional test does not account for time-varying volatility, in the sense that it ignores the time losses which occur, thus, the test may fail to reject a model that produces clustered VaR violations (Roccioletti, 2015). To address this problem, we used the conditional coverage test. Observations from Table 8.1 indicate that the null hypotheses of correct exceedances of VaR violations and independence are not rejected for all the models except the 95% *VaR (GJR)* models.

The observed failure rates of all the models, except for the 95% *VaR (GJR)* models, hence, are not significantly different from the corresponding expected rates. We can, therefore, categorically conclude that, with the exception of the 95% *VaR (GJR)* models, the VaR estimates from all the models are significantly accurate and would respond early to changing market conditions, without clustering over time.

8.4.2 Comparing VaR models

In comparing the superiority of the estimated VaR models, the MSC procedure alongside with the number of VaR violations, VaR violation failure rates and absolute mean (ADmean) and absolute maximum (ADmax) of VaR violating returns contemplated in McAleer & da Veiga (2008) were-used. In theory, the number of expected violations with 95% and 99% confidence levels for 1-step ahead forecasts are 22 and 4.4 (5% of 440 and 1% of 440) respectively. Models with violations closer to the expected violations tend to have fewer failure rates. In addition to the VaR estimates, we report the mean ES values, as well as their MSE values for comparison purposes. A model, with smaller mean VaR, mean VaR loss, mean ES and the MSE for expected shortfall, fewer failure rate, superior MSC rank, as well as a minimum ADmax and ADmean is preferred. It may not be possible for a model to be superior in terms of all the eight metrics, hence a voting criterion based on these metrics is introduced to help in selecting the model with the overall superior ability (i.e., a model with the majority of the votes is adjudged superior). The results of the comparison metrics are reported in Tables 8.2.

A look at the MCS procedures reveal that all the augmented models were selected into the superior set of models at both 1% and 5% confidence levels, and they are consistently ranked number one. The VaR_B *(NA)* and the VaR_B *(T)* models were eliminated by the procedure, thus in terms of this test, the corresponding augmented models have superior predictive abilities. On average, the augmented models yield lower failure rates in comparison to the referenced models at 1% confidence level, but not at 5%, which is in agreement with the results from Kumar (2020), hence, the 1% confidence level VaR models may lead to fewer bank failures.

In addition, the augmented models tend to have lower MSE values for the expected shortfalls estimates and lower mean VaR losses, hence the augmented models have superior predictive abilities in comparison to the reference models, however, there are exceptions to this generalization, which are seen in the *VaR(S)* and *VaR (E)* models. It is further observed that the mean absolute VaR estimates from the augmented models are consistently lower than the estimates from the reference models. Furthermore, the mean ES for all the augmented models are lower than the values for the corresponding reference models, except for the *VaR(S)* and *VaR(E)* models, therefore, on average, it is anticipated that the use of these augmented models may lead to lower bank costs. The augmented models also tend to have lower mean absolute deviations for VaR violating estimates, but produce large maximum absolute deviations in some instances.

Table 8. 2: Model comparison metrics

Model	VaR Mean Loss	Mean VaR	Mean ES	MSE (ES)	VaR Rank	Failure rate	AD of VaR violations Maximum	AD of VaR violations Mean	%Votes
99% Confidence level-VaR and ES estimations									
VaRA(GJR)	0.059416	-0.06035	-0.08227	0.00923	1	0.009	0.034237	0.017118	**100**
VaRB(GJR)	0.059729	-0.06067	-0.08244	0.00926	2	0.009	0.034296	0.017148	00
VaRA(NA)	0.019871	-0.01977	-0.02484	0.00077	1	0.007	0.001414	0.000943	**100**
VaRB(NA)	0.020608	-0.02052	-0.02609	0.00084	0	0.007	0.002098	0.001399	00
VaRA(S)	0.019770	-0.01949	-0.02613	0.00082	1	0.023	0.010186	0.003008	**57**
VaRB(S)	0.019772	-0.01951	-0.02594	0.00081	2	0.023	0.012205	0.002970	43
VaRA(T)	0.017405	-0.01682	-0.02198	0.00065	1	0.023	0.020258	0.008171	**88**
VaRB(T)	0.017642	-0.01707	-0.02214	0.00066	0	0.025	0.022355	0.007944	12
VaRA(E)	0.014298	-0.01373	-0.01677	0.00042	1	0.014	0.005688	0.002783	**88**
VaRB(E)	0.014340	-0.01379	-0.01776	0.00040	2	0.020	0.007131	0.002809	12
95% Confidence level-VaR and ES estimations									
VaRA(GJR)	0.059419	-0.06035	-0.05080	0.00387	1	0.077	0.034237	0.017118	**86**
VaRB(GJR)	0.059729	-0.06067	-0.01646	0.00389	2	0.077	0.034296	0.017148	14
VaRA(NA)	0.018306	-0.01818	-0.01727	0.00038	1	0.052	0.001414	0.000943	**88**
VaRB(NA)	0.020608	-0.02052	-0.01706	0.00041	0	0.048	0.002098	0.001399	12
VaRA(S)	0.019770	-0.01949	-0.01694	0.00040	1	0.073	0.010186	0.003008	50
VaRB(S)	0.019772	-0.01951	-0.01457	0.00039	2	0.068	0.012205	0.002970	50
VaRA(T)	0.017405	-0.01682	-0.01468	0.00033	1	0.061	0.020258	0.008171	**88**
VaRB(T)	0.017642	-0.01707	-0.01224	0.00034	0	0.059	0.022355	0.007944	12
VaRA(E)	0.014298	-0.01373	-0.01291	0.00026	1	0.059	0.005688	0.002783	**75**
VaRB(E)	0.014340	-0.01379	-0.08227	0.00024	2	0.057	0.007131	0.002809	25

GJR, NA, S, T and E denotes GJRGARCH, NAGARCH, standard GARCH, TGARCH and EGARCH respectively. Zero rank indicates that a model was eliminated by the MSC procedure at 99% confidence level based on the asymmetric mean loss. When there is a tie, the particular metric is excluded when computing the percentage votes.

The overall predictive abilities of the models based on the voting patterns indicate that the augmented models are relatively superior to the reference models for all the 1% VaR estimates. The same conclusion cannot be made for the 5% VaR estimates. This is because there is a split of votes among the *VaR(S)* models. It is worth noting that, although, the *VaR_A (GJR)* model was decisively adjudged superior to the *VaR_B (GJR)*, neither can be used in making risk decisions because they all failed the independence and the unconditional coverage tests. In general, the available evidence suggests the superiority of the proposed method in forecasting VaR, which agrees with the results from Kumar (2020), Reddy *et al.,* (2017) and Rapach, Strauss & Wohar (2007). Classifying them as superior results is based on the 1% VaR estimates of the augmented VaR models which confirm the superiority of the volatility models established in chapter seven, however, we cannot make such generalization based on the 5% VaR estimates, although the majority of the augmented models outperformed the non-augmented version.

8.4.3 Capital requirement analysis

Our approach yielded superior VaR and ES estimates for all models at 1% and majority at 5%, however, they are of no practical importance to risk practitioners, especially the banks when they are unable to use the estimates to compute acceptable regulatory capital requirements (as set out by the Basel II Accord). The capital requirements are used to control and monitor market-risk exposure of financial institutions and they act as a buffer for adverse market conditions. Overestimation of VaR forces institutions to hold significant amounts of capitals and lose opportunity costs, while underestimation overexposes institutions to market risks and losses in their balance sheets that cannot be recovered at crisis periods (Bucci, 2017), which may lead to repercussions on their positions, on the market. Basel II Accord allows banks to use their internal models to compute VaR estimates. McAleer & Veiga (2008), however, emphasize that banks have the responsibility to demonstrate the accuracy of their models sufficiently through backtest analysis based on the number of VaR violations. In addition to this, the Basel II Accord has instituted penalty zones (see Table 8.3) to penalize bad models in terms of a multiplicative factor k, based on the VaR estimates over the last 250 business days. Based on the penalty zones, the capital requirement is defined by the Bank for International Settlements, (2011) as:

$$\text{Capital requirements} = \max\left\{-VaR_{t-1}, -(3+k)\overline{VaR_{60}}\right\}, \qquad (8.11)$$

where $\overline{VaR_{60}}$ is the average VaR over the last 60 business days. In line with the standards prescribed by the Basel Committee, we focus the backtest analysis on the 99% quantile VaR estimates. Table 8.4 reports the mean daily capital requirements (MDCR) for the out-of-sample period. A result that immediately emerges from the Table is that, the augmented models consistently produced lower MDCR in comparison to non-augmented versions. The superior performance of the augmented models in terms of the MDCR coincides with our previous analysis based on the eight model comparison metrics. Furthermore, we can observe that the majority of the augmented models avoided the regulatory penalty zone while a few of them slipped into the yellow zone with the associated penalties, however, the converse of these observations is the case for the non-augmented models. These observations suggest that most of the augmented models would easily (no imposed penalty) pass the scrutiny of regulatory bodies unlike the non-augmented. The outcomes in the yellow range are plausible for both accurate and inaccurate models, however, the presumption that a model is inaccurate grows as the number of exceptions increase in the range (Bank for International Settlements, 2011). Looking at Table 8.1, the expected violations for the $VaR_A\ (T)$ model is smaller than that of $VaR_B\ (T)$, the $VaR_A\ (T)$ models, thus, would face fewer hurdles (penalties) in passing the scrutiny of regulatory bodies, unlike the $VaR_B\ (T)$. Both versions of the *VaR(S)* models, however, will face similar challenges before being certified accurate by the regulatory bodies.

Table 8. 3: Penalty thresholds based on Basel II Accord

Zones	250 business days		440 business days
	violations	Increase in *k*	violations
Green	0-4	0	0-8
Yellow	5	0.4	9
	6	0.5	10
	7	0.65	11
	8	0.75	12
	9	0.85	13
Red	≥ 10	1	14 or more

The Basel II Accord penalty zones are normally based on 250 business-trading days. To obtain an equivalent threshold for 440 days as indicated in Table 8.3, we use the binomial probabilities associated with the considered true levels of coverage (99%) for a sample size of 444 days. "The yellow zone begins at the point such that the probability of obtaining at maximum that number of exceptions equals or exceeds 95%. The starting point of the red zone is the one for which the same probability equals or exceeds 99.99%".

Table 8. 4: Mean daily capital requirement

	Models	VaR(GJR)	VaR(NA)	VaR(S)	VaR(T)	VaR(E)
MDCR	Benchmark	0.06014	0.01977	0.01949	0.01680	0.01361
	Augmented	**0.06047**	**0.02052**	**0.01951**	**0.01706**	**0.01379**
Penalty Zone	Benchmark	Green	Green	Yellow	Yellow	Green
	Augmented	Green	Green	Yellow	Yellow	Yellow

8.5 Conclusion

There are undesirable consequences associated with inaccurate VaR estimations for banks, thus accurate forecasting of VaR forms an integral part of decision-making and long-term stability of financial institutions. In this chapter, 1% and 5% VaR were estimated using univariate GARCH models augmented with exogenous variables. In assessing the individual accuracies of the VaR estimates, the conditional coverage and the unconditional coverage tests were used. Competing models, which passed both tests, were then ranked using the model set confidence test procedure. Several other model comparison tools were also employed to assist in selecting the best models. We also conducted a capital requirement analysis to assess the usefulness of the models to banking institutions in computing mean daily capital requirements. The main findings of the study include the following:

- Our approach led to a significant reduction of information leakages in the in-sample fitted models and perceived reductions of information leakages in the out-of-sample models. In addition, the approach also yielded less persistent volatilities, reduced half-life, and improved in-sample explanatory powers of the models. In addition, there were improvements in the predicted volatilities from all the models, however, the same cannot be said about the forecasted volatilities.

- Our approach yielded better forecasts for all the 1% VaR models and the majority of the 5% VaR models. Accurate volatility is implied by an accurate VaR forecast (Bucci, 2017), thus, our approach yielded similar superiorities in terms of the volatility forecasts.

- On the usefulness of the VaR estimates in computing daily capital requirements, our approach consistently produced lower MDCR for all the models. Furthermore, the majority of the models built on our approach, avoided the regulatory penalty zones, while few of them slipped into the yellow zone with relatively less associated penalties.

In conclusion, our approach led to fewer VaR violations, improved 1% VaR forecasts, lower ES forecasts, and optimal daily capital requirements, thus, the models are preferred from regulatory and institutional point of views, because they would lead to optimal bank costs and fewer bank failures. The 5% VaR forecasts for the *VaR (NA)*, *VaR (T),* and *VaR (E)* models, however, may not be preferred from regulatory point of view, although, they yielded improved VaR and ES forecasts. This is because the models have relatively higher violations, hence, they may lead to frequent bank failures or severe regulatory penalties, however, from an institutional point of view, they are recommended because of their perceived lower bank costs. It should be noted, however, that our proposed methodology *per se* might not be the cause of the relative higher violations for the 5% VaR estimates. This may be due to the use of inappropriate volatility model specifications and or inadequate market uncertainty proxies and exogenous break variables to estimate the volatility inputs of the VaR model. In a broader sense, the results are in support of studies, which advocate that failure to account for breaks in the unconditional variance leads to sizable upward biases in the degree of persistence in the estimated GARCH models. These forecast systematically underestimate or overestimate volatility and the subsequent VaR on average, over long horizons (Rapach, Strauss & Wohar, 2007 and Reddy *et al.,* 2017).

CHAPTER 9

General conclusions, contributions, recommendation, limitations and further research

9.1 General conclusions

The use of inaccurate volatility estimates leads to misinformed policy decisions. In a financial institutional context, it may lead to regulatory sanctions, overexposure to market risk and loss of revenue due to opportunity cost; these hamper the growth of financial institutions and their market positions as a whole. In terms of policymaking, the effects are felt in the economic pressures on daily living. Accurate volatility forecasts are, therefor, non-negotiable in the financial industry, and policy formulations.

The ARCH-GARCH framework is one of the most popular methods used in forecasting volatility by financial institutions, researchers, and practitioners, however, models built under these frameworks have been found to converge to poor volatility forecasts (Chen, Dolado & Gonzalo, 2014). One of the several factors attributed to this seemingly poor forecast is the unaccounted structural breaks and changes in the unconditional volatility (Chen, Dolado & Gonzalo, 2014; Amado & Teräsvirta, 2014; Rapach, Strauss & Wohar, 2007 and Andersen & Bollerslev, 1998).

As an attempt to improve volatility forecasts and to allow long-range volatility forecasts to revert along stochastic paths towards their long-run variance, the study assumed that the levels of uncertainties surrounding the exchange rate market affect the unconditional volatility of exchange rates, thus using absolute exchange rate returns as proxies, break variables were constructed. The proxies and the break variables were then incorporated into the variance equation of GARCH models to account for changes and breaks in the unconditional volatility processes. In the ensuing sections, we present chapter-by-chapter, conclusions of the thesis.

In the first chapter, a general introduction of the study focussing on an overview of exchange volatility, background, motivation, problem statement, scope, and the anticipated contributions of the study to literature were discussed. Reviews of the properties of asset returns and GARCH models were presented in chapter two.

In chapter three, we discussed data validation tests, model diagnostics tests, forecast evaluation, and model comparison tools. The Portmanteau test of Fisher & Gallagher (2012) was selected for post diagnosis of serial and autocorrelations in the standardized errors. For the data validations, the ARCH Lagrange Multiplier test was selected to test for the presence of ARCH effects in the return, and the ADF test for unit root testing. In the evaluation of the statistical significance of the unrestricted models, the likelihood ratio test was chosen, while the Mincer-Zarnowitz regression's R-square, MAE and RMSE were selected to assess the forecasting accuracies, and superiority of the competing models. In backtesting the VaR estimates, the unconditional coverage test of Kupeic, and the interval forecast tests of Christoffersen were selected and the MCS test was chosen to select and rank the superior set of competing VaR models.

In chapter four, we presented a descriptive analysis of the data as well as the validation of the data for use in GARCH models. It was observed that currencies move closely together at price-level and this translates into significant positive co-movements at the return-level. A substantial number of the currency pairs were also found to exhibit strong return-level co-movements. All the return series were found to be heteroskedastic, stationary and exhibit volatility clustering. Heavy-tailed distributions including the skewed student-t distribution, the skewed general error distribution, the normal inverse Gaussian distribution and the Johnson SU parameterized distribution, were used to model the innovations of the returns because they were found to be non-normal, leptokurtic, and heavy-tailed.

In chapter five, under ARMAX-GARCH framework, we re-examined the relationships between exchange rate returns. The statistical significance of the estimated models was investigated using the likelihood ratio test. The likelihood ratio tests confirmed that the estimated relationships are significant. The ARMAX-GARCH models were found to be more accurate in approximating the relationships between the returns for the period under consideration. These results confirm the well-established exchanged rate relationships. All the estimated parameters for the exogenous returns were positive and consistent with the direction of their respective co-movements with the endogenous returns. Path analysis of the impacts of the returns confirmed that currency pairs are not in isolation on the market. It was further observed from the path analysis, that shocks of the same magnitude, from the same origin, transmitted along different paths on the market may have different impacts.

Before modelling volatility with the augmented models, in chapter six, a pre-study of the hypothetical mutual dependencies between volatility, lagged proxies, and breaks variables were computed; this was to investigate the levels of the dependencies and their possible significance in modelling volatility. Results from the study provided evidence of substantial levels of shared mutual information between the volatilities and the exogenous variables and, although the strengths of the linear dependencies were weak, the strengths were significant for substantial number of variable pairings. In addition, it was observed that the exogenous break variables had more potential to account for breaks in the unconditional volatility of exchange rates adequately, than the endogenous break variables (those constructed from endogenous returns).

In chapter seven, in an attempt to improve volatility forecasts and to induce forecasts to revert along stochastic paths, we assumed that the unconditional volatilities of GARCH models were driven by the levels of uncertainties in the market. Since other proxies used ithe literature were reasoned to be inadequate in modelling changes in the unconditional volatility, absolute returns from exogenous currency pairs were identified as alternative proxies. The proxies and break variables, constructed from the proxies were passed to the variance equations of standard GARCH, EGARCH, GJRGARCH, TGARCH, and NAGARCH models. The augmented models were then used to estimate and forecast the volatilities of exchange rates. The results showed that shocks in the augmented models decay faster than the non-augmented models. It was further observed that forecasts from the augmented models revert along stochastic paths towards their unconditional volatilities, which is consistent with the underlying volatility paths. The explanatory powers of the augmented models were relatively improved in relation to the non-augmented version with an overall improved forecast accuracies or insignificant loss of accuracies.

One of the practical applications of volatility estimates or forecasts is their usage as inputs in VaR models. In order to assess the significance of the estimated forecasts in chapter seven and to also forecast risks of single-asset currency portfolios, volatility estimates and forecasts from the augmented models were used to estimate and forecast VaR of the currency pairs used; this was discussed in chapter eight. The conditional coverage tests and the unconditional coverage tests were used to assess the accuracies of the individual VaR estimate, while the MCS procedure was used to compare the augmented models to the non-augmented versions.

Evaluations from the backtesting methods confirmed that four out of the five augmented models, generally performed well in comparison to the non-augmented models at 1% level of confidence. For the 5% VaR estimates, only two out of the five augmented models performed better while two had similar predictive abilities as the non-augmented models. The augmented GARCH models lead to fewer violations with improved, but lower 1% VaR forecasts, thus, the models are preferred from regulatory and institutional point of views, as they lead to fewer bank failures and perceived lower bank costs. The models are, however, not preferred from the regulatory point of view for 5% VaR estimates because they have relatively higher violations and may lead to frequent bank failures, but from an institutional point of view, the models are recommended because of their perceived lower costs.

9.2 Main contributions of the study

The main contributions of the study in relations to existing studies are summarised below.

- In studying dependencies among the variables (volatility, returns, and dummy or break variables constructed from returns), to the best of our knowledge, this is the first study to employ the concepts of mutual entropy and the Jackknife-biased corrected KDE approach to data from the rand forex market. This aspect of the study, hence, serves as a contribution to literature on the relationships among the dynamics of exchange rate markets of emerging economies from the perspective of mutual information. The mutual information concepts can be used to re-examine and re-affirm the various established forms of theoretical and phenomenological relationships in literature. In addition, due to the insensitiveness of mutual information to the size of data sets, it can be used to study the relationships between financial variables with limited or small data samples.
- In speculative asset markets, asset returns are related to their volatilities via the risk premium theory (Black, 1976). This relationship is the underlying basis for computing risk premium and can be used to study the effects of an asset's volatility on its returns. The reverse of this relationship is used to study the concept of volatility asymmetry. Assets move in tandem, hence, a cross-asset returns-volatility relationship (the relationship between the volatility of one asset and the returns of another asset) and its reverse may exist. These relationships may be useful in studying the effects of cross-asset risk on asset returns and asymmetric effects of exogenous

returns on volatility, respectively. The reverse of cross-asset returns-volatility relationship has been useful in studying spillover of volatility asymmetry across different speculative asset markets (Yarovayaa, Brzeszczyński & Lauc, 2017), as well as the improvement of multivariate volatility forecasts (Santos, Nogales & Ruiz, 2013). Cross-asset returns-volatility relationship is a theoretical construct, thus empirical evidence of substantial mutual dependencies between volatility and the exogenous returns support the plausibility of this hypothesis.

- The study also contributes to the extant literature on VaR estimation. In its unique contribution, it brings on-board, a simple, but a novel approach to account for breaks and changes in the unconditional volatility of GARCH-type models. The approaches used in Rapach, Strauss & Wohar (2007) and Reddy et al., (2017) to model breaks and time-variations involve relatively complicated procedures which are sometimes not easy to be incorporated into other modelling frameworks. In comparison to these approaches, our methodology, however, is simple and easy to incorporate into other volatility frameworks, such as the stochastic volatility framework. Furthermore, since the break variables used are exogenous, unlike Karlsson (2016), the variables prevent the compounding of bias, which may be introduced by consecutive endogenous outliers in the parameter estimation. In addition, our approach takes into consideration the actual economic state of volatility in constructing the break variables unlike Karlsson (2016), hence, periods of crises are modelled differently from periods of increased volatility. Again, due to the superiority of the models built on our approach and the fact that the MCS procedure ranked all models built on our approach number one, our approach, thus, provides alternative or complementary tools, which can be used to mitigate risk in financial institutions, comprehensively. It is also useful to individual traders and investors who may not have any standard approach of computing financial risk associated with their daily decision-making.
- Finally, our approach relaxed the restrictive mean-reversion property of GARCH forecasts, thus, long-range forecasts revert to their long-run volatility along stochastic path. This observation is consistent with the underlying volatility of exchange rate.

9.3 Recommendations

The proposed augmentations of GARCH models lead to forecasts, which revert towards their long-run variance along stochastic paths with an improved explanatory power, and overall improved forecasts accuracies, or insignificant loss of accuracies, the approach, thus, is recommended for volatility forecasting, VaR forecasting, and the computations of other risk metrics that require volatility estimates as inputs. VaR forecasts from the proposed method belong to a superior set of models based on the model confidence set test of Hansen, Lunde & Nason (2011). Combined forecasts from superior set of models have been found to perform better than the individual forecasts (Bernardi & Catania, 2014 and Bernardi, Catania & Petrella, 2014). It is, therefore recommended, that institutions, traders and other financial practitioners may combine forecasts obtained from their internal models with those obtained from the proposed method to obtain a better and improved forecasts to help in making better informed decisions.

9.4 Limitations of the study

The proposed method depends on substantial levels of mutual dependency among assets, thus, it is unlikely to yield improved forecasts and or induce the forecasts to revert along stochastic paths, when there is no substantial mutual dependency. This, however, is not seen as a major challenge because due to favourable macro-driven environment and increased high frequency trading activities, the levels of mutual dependency of assets are unlikely to reduce to negligible levels. Including irrelevant proxies may lead to unstable parameters, persistent ARCH effects and significant serial correlation as well as information leakage. It is recommended, therefore, that appropriate feature selection algorithm is applied when deciding on which exogenous covariates to be used as proxies and to construct the break variables. This task is, however arduous due to the latent nature of volatility. In addition, the approach is not parsimonious because it sometimes requires more exogenous covariates and higher order ARMA-GARCH terms to guarantee optimal parameters.

9.5 Future studies

The methodology can be extended to multivariate case, where common exogenous covariates can be incorporated into the volatility processes to improve VaR forecasts for multi-asset portfolio. The methodology may also be useful in forecasting volatility and VaR for other speculative class of assets, which are known to be mutually dependent. In future, attention could be focussed on applying the methodology to a broader exchange rate markets in an attempt to generalize the findings. Alternative modelling framework and approaches, such as stochastic volatility and extreme value theory and other GARCH models or combinations, thereof may be explored. In such instances, additional model performance comparisons tools beyond those included in this thesis could be explored. Future studies should also focus on the forecasting abilities of known proxies, such as the CBOE volatility index and the Swedish model-free implied volatility index in comparison to the market uncertainty proxies used in this study. Similar studies may also be conducted to compare our approach of accommodating breaks in the unconditional volatility of GARCH models and other known approaches.

References

[1] Alam M.Z. (2012) Forecasting the BDT/USD Exchange Rate using Autoregressive Model. *Global Journal of Management and Business Research*, 12(19). Accessed on 07/08/2020.

[2] Alexander C (2008) *Market Risk Analysis Volume IV: Value-at-Risk Models*. John Wiley and Sons Ltd.

[3] Alexander C.O. Johnson A. (1992) Are Foreign Exchange Markets Really Efficient? *Economics Letters* 40: 449-453.

[4] Alexopoulos M. and Cohen J. (2015) The power of print: uncertainty shocks, markets, and the economy. *International Review of Economics Finance*, 40:8–28.

[5] Amado C. and Laakkonen H. (2014) Modelling Time-Varying Volatility in Financial Returns: Evidence from the Bond Markets. *Essays in Nonlinear Time Series Econometrics*, 139-160.

[6] Amado C. and Terasvirta T. (2011) *Modelling Volatility with Variance Decomposition*. CREATES Research Paper 2011-1, Aarhus University.

[7] Amado C. and Teräsvirta T. (2014) Modelling changes in the unconditional variance of long stock return series. *Journal of Empirical Finance*, 25(1):15-35. Doi:10.1016/j.jempfin.2013.09.003.

[8] Andersen T. and Bollerslev T (1998) Deutsche Mark–Dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. *Journal of Finance*, 53: 219-265.

[9] Andersen T., Bollerslev T., Christoffersen P. and Diebold F. (2005) *Practical volatility and correlation modelling for financial market risk management*. M. Carey and R. Stulz (eds.), Risks of Financial Institutions. University of Chicago Press for NBER.

[10] Andersen T.G., Bollerslev T., Christoffersen P.F. and Diebold F.X. (2007) *Practical Volatility and Correlation Modelling for Financial Market Risk Management*. University of Chicago Press.

[11] Andersen T.G., Davis R.A., Kreiß J.P. and Mikosch T. (2009) *Handbook of Financial Time Series*. Springer-Verlag Berlin Heidelberg.

[12] Antwi A. and Kyei K.A. (2017) Model for estimating profitability targets: a case study of Amalgamated bank of South Africa (ABSA). *Actual problems of economics,* 11(197), 66-76.

[13] Antwi A., Kyei K. A. and Gill R.S. (2020) The Use of Mutual Information to Improve Value-at-Risk Forecasts for Exchange Rates, in *IEEE Access*, 8: 179881-179900. Doi: 10.1109/ACCESS.2020.3027631.

[14] Antwi A., Kyei, K.A. and Gill, R. (2020) Forecasting long term exchange rate volatility with stochastic mean reverting unconditional volatility. *Journal of Statistics and Management Systems*, 0(0): 1-23. Doi: 10.1080/09720510.2020.1816690.

[15] Arefin M. K. and Ahkam, S. N. (2017) Return and volatility spillover between financial market participants of Dhaka stock exchange using asymmetric G.A.R.C.H. methods. International Journal of Trade, Economics, and Finance, 8(3):33–140.

[16] Aridi, N. A., Cheong, C. W., & Hooi, T. S. (2018) An Estimation of Value at Risk using GARCH Models for the Conventional and Islamic Stock Market in Malaysia. *International Journal of Academic Research in Business and Social Sciences*, 8(11), 2054–2065. Doi: 0.6007/IJARBSS/v8-i11/5568.

[17] Awartani B.M.A. and Corradi V. (2004) Predicting the volatility of the S&P 500 stock index via GARCH models: the role of asymmetries. *International Journal of Forecasting*, 21:167–183.

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

[19] Bailie R.T., Bollerslev T. and Mikkelsen H.O. (1996) Fractionally Integrated Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics*, 74:3-30.

[20] Baillie R.T. and Bollerslev T. (1989) Common stochastic trends in a system of exchange rates. *Journal of Finance*. Doi: 10.1111/j.1540-6261.1989.tb02410.x. Accessed on 02/02/2020.

[21] Baillie R.T. and Bollerslev T. (1994) Co integration, Fractional Cointegration and Exchange Rate Dynamics. *The Journal of Finance*, XLIX (2): 737-745.

[22] Baillie R.T. and Morana C. (2009) Modelling long memory and structural breaks in conditional variances: An adaptive FIGARCH approach. *Journal of Economic Dynamics and Control,* 33: (1577)-1592.

[23] Baillie R.T., Bollerslev T. and Mikkelsen H.O. (1996) Fractionally integrated generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 74:3–30.

[24] Baker S.R., Bloom N., and Davis S.J. (2016) Measuring economic policy uncertainty. *The Quarterly Journal of Economics*. Doi:10.1093/qje/qjw024. Accessed on 11/11/2019.

[25] Bams D., Gildas B. and Thorsten L. (2017) Volatility measures and Value-at-Risk. *International Journal of Forecasting,* 33:848–63.

[26] Bank for International Settlements (2011) *Basel Committee on Banking Supervision Revisions to the Basel II market risk framework*. Available: at: http://www.bis.org/publ/bcbs193.pdf. Accessed on 14/09/2020.

[27] Bank of international Settlement (2019) *Triennial Central Bank Survey Foreign exchange turnover in April 2019*. Available at: https://www.bis.org/statistics/rpfx19_fx.pdf. Accessed on 17/09/2019.

[28] Battiti R. (1994) Using Mutual Information for Selecting Features in Supervised Neural Net Learning. *IEEE Transactions on Neural Networks,* 5:537–50.

[29] Baybogan B. (2013) Empirical Investigation of MGarch Models. *Journal of Statistical and Econometric Methods, vol. 2, no.3, 2013, 75-93*. Available at: http://www.scienpress.com/Upload/JSEM/Vol%202_3_7.pdf. Accessed 08/04/2021.

[30] Bayer S. (2018) Combining Value-at-Risk forecasts using penalized quantile regressions. *Econometrics and Statistics*, 8:56-77.

[31] Belkacem L., Meddeb Z.E., and Boubaker H. (2005) Foreign Exchange Market Efficiency: Fractional Co integration Approach. *International Journal of Business*, 10 (3): 285-302. Available at: https://ssrn.com/abstract=778444. Accessed on 02/03/2019.

[32] Benedetto F. Mastroeni L. Quaresima G. and Vellucci P. (2020) Does OVX affect WTI and Brent oil spot variance? Evidence from an entropy analysis. *Energy Economics*, 89(C): 104815. Doi:10.1016/j.eneco.2020.104815.

[33] Benedetto F., Mastroeni L. and Vellucci P. (2019) Modeling the flow of information between financial time-series by an entropy-based approach. *Annals of Operations Research*, 1-18.

[34] Beran J. (1994) Statistics for Long-Memory Processes. *Monographs on Statistics and Applied Probability*. New York, Chapman and Hall.

[35] Berkes I. Horváth L. Kokoszka, P. (2003) GARCH processes: structure and estimation. *Bernuolli*, 9:201–227.

[36] Bernardi M. and Catania L. (2014) *The Model Confidence Set package for R*. Available at: http://arxiv.org/abs/1410.8504. Accessed on 01/02/2020.

[37] Bernardi M., Catania L., and Petrella L. (2014) Are News Important toPredict Large Losses ?Working Paper, Arxiv Preprint. Available at: https://arxiv.org/abs/1410.6898. Accessed on 08/06/2020.

[38] Bird R. and Yeung D. (2011) *How do Investors React Under Uncertainty*. Unpublished Manuscript. Available at: https://www.uts.edu.au/sites/default/files/PaperBirdRon.pdf. Accessed on 02/03/2020.

[39] Black, F. (1976) Studies of stock price volatility changes. Proceedings of the 1976 Meeting of Business and Economics Statistics Section of the American Statistical Association **27**(1), 399–418.

[40] Bloom N. (2009) The Impact of Uncertainty Shocks, *Econometrica*, 77: 623-685.

[41] Bollerslev T and Melvin M (1994) Bid—ask spreads and volatility in the foreign exchange market: *An empirical analysis. Journal of International Economics*, 36: 355-372.

[42] Bollerslev T. (1986) Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31:307-327.

[43] Bollerslev T. and Ghysels E. (1996) Periodic autoregressive conditional heteroscedasticity. *Journal of Business & Economic Statistics*, 14(2):139-151.

[44] Bollerslev T. and Wooldrige J.M. (1992) Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models With Time-Varying Covariances, *Econometric Reviews*, 11, 143-172.

[45] Bollerslev T., Engle R.F. and Nelson D.B (1994) *ARCH Models, in Engle, R.F. and McFadden, D., eds. The Handbook of Econometrics*, (4th ed). Amsterdam: North

[46] Boubaker H. and Belkacem L. (2010) Interdependence between Exchange Rates: Evidence from Multivariate Fractional Co integration. International *Journal of Business*, 15(3).

[47] Brooks C. (2008) *Introductory economics for finance*. U.K: Cambridge University Press.

[48] Brownlees C.T. and Gallo G.M. (2010) Comparison of volatility measures: A risk management perspective. *Journal of Financial Econometrics*, 8(1):29–56.

[49] Bucci A. (2017) Forecasting realized *volatility: a review*. MPRA Paper No. 83232. Available at: https://mpra.ub.uni-muenchen.de/83232. Accessed on 01/02/2020.

[50] Cantamessa A., Gautam M., and Xiang Y. (2016) *Volatility Forecast with GARCH Model and News Analytics*. Available at SSRN: https://ssrn.com/abstract=3406083.

[51] Caporale G.M. and Gil-Alaña L.A. (2004) Fractional Co integration and Real Exchange Rates. *Review of Financial Economics*, 13:327-340.

[52] Carriero, A., Clark, T. E., & Marcellino M. (2015) *Common drifting volatility in large Bayesian vars. Journal of Business and Economic Statistics*. http://dx.doi.org/10.1080/07350015.2015.1040116. Accessed on 21/01/2020.

[53] Cayton P.J and Mapa D. (2012) Time-varying conditional Johnson SU density in value-at-risk (VaR) methodology, MPRA Paper 36206, University Library of Munich, Germany. Available at: https://mpra.ub.uni-muenchen.de/36206/1/MPRA_paper_36206.pdf. Accessed on 25/09/2020.

[54] Chen L, Dolado, J.J. and Gonzalo J. (2014) Detecting big structural breaks in large factor models. *Journal of Econometrics*, 180: 30-48. Available at: http://hdl.handle.net/1814/32939. Accessed on 05/01/2019.

[55] Chen L., Dolado J. and Gonzalo J. (2014) Detecting big structural breaks in large factor models. *Journal of Econometrics*. Doi: 180. 10.1016/j.jeconom.2014.01.006.

[56] Cheung Y.W. (1993) Long Memory in Foreign-Exchanges Rates. Journal of Business and *Economic Statistics,* 11:93-101.

[57] Ching M.L. and Siok K.S. (2013) Comparing the performances of GARCH-type models in capturing the stock market volatility in Malaysia. *Procedia Economics and Finance*, 5 (2013): 478-487.

[58] Chipili J. (2012) Modelling Exchange Rate Volatility in Zambia. *The African Finance Journal*, 14 (2), 85-107.

[59] Christodoulakis G.A. and Satchell S.E. (1998) *Forecasting (LOG) Volatility Models*. Discussion Papers 9814, University of Exeter, Department of Economics. Available at: https://ideas.repec.org/p/exe/wpaper/9814.html. Accessed on 25/08/2019.

[60] Christoffersen P. (1998) Evaluating interval forecasts. *International Economic Review,* 39: 841–62.

[61] Chuliáa H., Guilléna M. and Uribeb J.M. (2017) Measuring uncertainty in the stock market. *International Review of Economics and Finance,* 48 (2017) 18–33.

[62] Çiçek M. (2014) A cointegration test for Turkish foreign exchange market efficiency. *Asian Economic and Financial Review*, 4(4): 451-471.

[63] Cont R. (2001) Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative finance*, 1(2001): 223-236.

[64] Cover T.M and Thomas J.A. (2012) *Elements of Information Theory*. John Wiley & Sons.

[65] Crampton J. and Loizou G. (2001) The completion of a poset in a lattice of antichains. *International Mathematical Journal,* 1(3): 223-238.

[66] Crowder W. (1994) Foreign Exchange Market Efficiency and Common Stochastic Trends. *Journal of International Money and Finance*, 13:551-564.

[67] Daníelson J., Jorgenson B., Samorodnitsky G., Sarma M., and de Vries C. (2013) Fat tails, VaR and subadditivity. *Journal of Econometrics,* 172(2), 283–291.

[68] Dick C.D., Schmeling M. and Schrimpf A. (2013) Macro Expectations, aggregate uncertainty, and expected term premia. *European Economic Review*, 58(C), 58–80.

[69] Diebold F., Doherty N. and Herring R. (2010) *The known, the unknown, and the unknowable in financial risk management: measurement and theory advancing practice.* New York: Princeton University Press.

[70] Diebold F.X., Gardeazabal J., and Yilmaz K. (1994) On the cointegration and exchange rate dynamics. *Journal of finance*, 49:727-735.

[71] Ding Z., Granger C.W.J. and Engle R.F. (1993) A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1):83-106.

[72] Divya K.H. and Devi V.R. (2014). A Study on Predictors of GDP: Early Signals. *Procedia Economics and Finance*, 11, 375-382. Doi: 10.1016/S2212-5671(14)00205-6.

[73] Dowd K. (2002) *Measuring Market Risk*. Chichester and New York: John Wiley and Sons.

[74] Duan J.C. (1997) Augmented GARCH (p; q) process and its diffusion limit. *Journal of Econometrics*, 79:7-127.

[75] Easley, D., and O'Hara, M., (2010) Liquidity and Valuation in an Uncertain World. *Journal of Financial Economics*, 97: 1-11.

[76] Embrechts P., Puccetti G., and Rüschendorf L. (2013) Model uncertainty and VaR aggregation. *Journal of Banking and Finance*, 37(8), 2750–2764.

[77] Emmer S., and Tasche D. (2004) Calculating credit risk capital charges with the one-factor model. *Journal of Risk*, 7(2), 85–103.

[78] Emmer S., Kratz M. and Tasche D. (2013) What is the best risk measure in practice? A comparison of standard measures. *Journal of Risk*, 18(2): 31-60, 2015.

[79] Engle R.F and & Patton A. J (2002) What good is a volatility model? *Journal of quantitative finance*, 1(2001): 237–245.

[80] Engle R.F. (1982) Auto-regressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4), 987-1007.

[81] Engle R.F. (2001) *An introduction to the use of ARCH models in applied econometrics*. NYU working paper. Available at: https://www.stern.nyu.edu/rengle/GARCH101.PDF. Accessed on 05/01/2019.

[82] Engle R.F. and Granger C.W. (1987) Cointegration and error correction: Representation, estimation and testing. *Econometrica*, 55:251-256.

[83] Engle R.F. and Manganelli S. (2004) CAVaR: Conditional autoregressive Value-at-Risk by regression quantiles. *Journal of Business and Economic Statistics*, 22: 367–81.

[84] Engle R.F. and Mezrich J. (1996) GARCH for Groups. *Risk*, 9:36-40.

[85] Engle R.F. and Ng V.K. (1993) Measuring and testing the impact of news on volatility. *Journal of Finance*, 48(5):1749-1778.

[86] Engle R.F. and Rangel J. G. (2008) The spline-GARCH model for low frequency volatility and its global macroeconomic causes. *Review of Financial Studies*, 21:1187–1222.

[87] Engle R.F. and Rangel J.G. (2004) The Spline-Garch Model for Low Frequency Volatility and its Global Macroeconomic Causes. *Review of Financial Studies*, 21, 2008.

[88] Engle R.F. and Sokalska M.E. (2012) Forecasting intraday volatility in the US equity market with multiplicative component GARCH. *Journal of Financial Econometrics*, 10(1):54-83. Doi: 10.1093/jj_nec/nbr005.

[89] Engle R.F. Granger C.W. (1983) Cointegrated variables and error correction models. *Econ*, 55(2):251-276. Available at: https://www.jstor.org/stable/1913236. Accessed on 19/05/2019.

[90] Epstein L., and Schneider M. (2010) Ambiguity and Asset Markets. *Annual Review of Financial Economics* 2:15-346.

[91] Ezpeleta E.V. (2015) Modeling volatility for the Swedish stock market. Unpublished thesis, Uppsala University, Department of Statistics. Available at http://uu.diva-portal.org/smash/get/diva2:898640/FULLTEXT01.pdf. Accessed on 07/07/2020.

[92] Fama E. (1965) The behavior of stock market prices. *Journal of Business*, 38:34-105.

[93] Faure A.P. (2013) *Exchange Market: An introduction* (1st ed). Quoin institute (pty) Limited and Bookboon.com.

[94] Firoj M. and Khanom S. (2018) Efficient Market Hypothesis: Foreign Exchange Market of Bangladesh. *International Journal of Economics and Financial Issues*, 8(6): 99-103. Available at: https://doi.org/10.32479/ijefi.7097. Accessed on 10/03/201.

[95] Fisher T.J. and Gallagher C.M. (2012) New Weighted Portmanteau Statistics for Time Series Goodness of Fit Testing. *Journal of the American Statistical Association*, 107:498, 777-787. Doi: 10.1080/01621459.2012.688465.

[96] Flood R.P. (1981) Explanations of Exchange rate volatility and other empirical regularities in some popular models of foreign exchange market. *Carnegie-Rochester conference series on public policy*, 15: 219-250.

[97] Francq C. and Zakoian J (2019) *GARCH Modes, Structure, Statistical Inference and Financial Applications* (2nd ed). John Wiley & Sons Ltd.

[98] Francq C. and Zakoıan J.M. (2010) GARCH Models. *Structure, Statistical Inference and Financial Applications*. John Wiley and Sons, Ltd. ISBN 978-0-470-68391-0.

[99] Francq C., Horváth L. and Zakoïan J.M (2011) Merits and drawbacks of variance targeting in GARCH models. *Journal of Financial Econometrics*, 9(4):619-656. Available at: https://doi.org/10.1093/jjfinec/nbr004. Accessed on 20/02/2020.

[100] Freedman D. and Diaconis P. (1981) On the histogram as a density estimator: L2 theory. Z. *Wahrscheinlichkeitstheorie verw Gebiete,* 57, 453-476.

[101] Gadwala S.B and Mathur S.K. (2014). Modelling & Forecasting of Re/$ Exchange rate – An empirical analysis.

[102] Galan's A. (2019) Rug arch*: Univariate GARCH models. R package version 1.4-1.*

[103] Garcia S., Luengo J., Sáez J.A., Lopez V. and Herrera F. (2013) A Survey of Discretization Techniques: Taxonomy and Empirical Analysis in Supervised Learning. *IEEE Transactions on Knowledge and Data Engineering,* 25 (4): 734-50.

[104] Geek J.Y.P. (1986) Modelling the persistence of conditional variances: a comment. *Econometric Reviews*, 5:57-61.

[105] Gel 'and I.M. and Yalow A.M. (1957) Calculation of amount of information about a random function contained in another such function. *American Mathematical Society Translations*, 2(12): 199–246.

[106]Giacomini R. and White H. (2006). Tests of Conditional Predictive Ability, *Econometrica*, 74:1545-1578.

[107]Glosten L.R., Jagannathan R. and Runkle D.E (1993) On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48(5):1779-1801.

[108]Gneiting T. (2011) Making and evaluating point forecasts. *Journal of the American Statistical Association*, 106(494), 746–762.

[109]Gomery R. (1995) The known, the unknown and the unknowable. *Scientific American*.

[110]González-Rivera G., Lee T.H. and Mishra S. (2004) Forecasting volatility: A reality check based on option pricing, utility function, Value-at-Risk, and predictive likelihood. *International Journal of Forecasting*, 20(4):629-645. Doi: 10.1016/j.ijforecast.2003.10.003.

[111]Graham E., Rothenberg T.J. and Stock J.H. (1996) Efficient Tests for an Autoregressive Unit Root. *Econometrica,* 64:813-836.

[112]Granger C. & Poon, Ser-H. (2005) Practical Issues in Forecasting Volatility. *Financial Analysts Journal,* 61(1): Doi: 10.2469/faj.v61.n1.2683.

[113]Gupta S.S. and Panchapakesan S. (1979) Multiple Decision Procedures. John Wiley & Sons, New York.

[114]Hamilton J.D. (1994) *Time Series Analysis*. Princeton University Press.

[115]Hansen L. (1982) Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 50(4), 1029-1054. doi:10.2307/1912775.

[116]Hansen P.R. (2005) A test for superior predictive ability. *Journal of Business & Economic Statistics,* 23:365–80.

[117]Hansen P.R. and Lunde A. (2005) A forecast comparison of volatility models: does anything beat a GARCH (1, 1)? *Journal of Applied Econometrics*, 20(7):873-889. Available at: Doi:10.1002/jae.800. URL http://dx.doi.org/10.1002/jae.800. Accessed on 25/01/2020.

[118]Hansen P.R., Huang Z. and Shek H.H. (2012) Realized GARCH: a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, 27(6):877–906.

[119]Hansen P.R., Lunde A. and Nason J.M. (2003) *Choosing the best volatility models*: The model confidence set approach. Oxford Bulletin of Economics and Statistics, 65(s1):839-861.

[120]Hansen P.R., Lunde A. and Nason J.M. (2011) The model confidence set. *Econometrica*, 79(2):453–497.

[121]Hentschel L. (1992) All in the family nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics*, 39(1):71-104.

[122]Higgins M.L. and Bera A. (1992) A class of nonlinear arch models. *International Economic Review*, 33(1):137–58.

[123]Hillebrand E. (2003) Mean reversion models of financial markets. Unpublished PhD thesis, University of Bremen. Available at: https://media.suub.uni-bremen.de/bit-stream/elib/1906/1/E-Diss549_diss02.pdf. Accessed on 09/12/2020.

[124]Hillebrand E. (2005) Persistence in Variance, Structural Change and the GARCH Model. *Journal of Econometrics*, 129(1-2):121-138.

[125]Hinich, M.J. (1982) Testing for Gaussianity and linearity of a stationary time series. *Journal of Time Series Analysis*, 3(3):169-176.

[126]Ho-Jin L. (2009) Forecasting Performance of Asymmetric GARCH Stock Market Volatility Models. *Journal of International Economic Studies*, 13(2).

[127] Holland B.C. (2008) *Introductory economics for finance*. U.K Cambridge University Press.

[128]Holmes M.J. (2004) Is There Nonlinear Real Exchange Rate Adjustment for the Asian Economies?. *ASEAN Economic Bulletin* 21(2), 198-212.

[129]Horrace W.C. and Schmidt P. (2000) Multiple Comparisons with the Best, with Economic Applications. *Journal of Applied Econometrics*, 15:1-26.

[130]Hsu J.C. (1996) *Multiple Comparisons*. Chapman & Hall/CRC, Boca Ranton, Florida.

[131]Hu K.T. (1962) On the Amount of Information. *Theory Probab. Appl*, **7**(4): 447-455.

[132]Hull J. and White A. (1998) Incorporating volatility updating into the historical Simulation method for value at risk. *Journal of Risk*. Doi:10.1.1.35.8358.

[133]Ince A. A. R. (2016) Measuring multivariate redundant information with pointwise common change in surprisal. *Entropy* 2017, 19:7, 318. https://doi.org/10.3390/e19070318.

[134]Jabeen M. & Khan S.A (2014) Modelling Exchange Rate Volatility by Macroeconomic Fundamentals in Pakistan. *International Econometric Review (IER)*, 6(2): 58-76. Available at: http://www.era.org.tr/makaleler/12080099.pdf. Accessed 02/09/2018.

[135]Jarque C.M. & Bera A.K (1980) Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*. 6 (3): 255–259. Doi:10.1016/0165-1765(80)90024-5.

[136]Jensen S.T. Rahbek A. (2004) Asymptotic inference for nonstationary GARCH. *Econometric Theory*, 20:1203-1226.

[137]Joseph N. (2002) Modelling the impacts of interest rate and exchange rate changes on UK Stock Returns. *Derivatives use, trading and regulation*, 794:306-323.

[138]Jurado K., Ludvigson S.C. and Ng S. (2015) Measuring uncertainty. *American Economic Review*, 105(3):1177–1216.

[139]Kalbfleisch J.G. (1985) *Probability and Statistical Inference*, (2nd ed), Springer-Verlag.

[140]Kambouroudis D.S. and McMillan D.G. (2016) Does VIX or volume improves GARCH volatility forecasts? *Applied Economics*, 48(13):1210-1228.

[141]Karlsson O. (2016) *Volatility forecasting under structural breaks*. Unpublished thesis. Available at: http://www.diva-portal.org/smash/get/diva2:957411/FULLTEXT01.pdf . Accessed on 06/01/2019.

[142]Kerkhof J. and Melenberg B. (2004) Backtesting for risk-based regulatory capital. *Journal of Banking & Finance*, 28(8): 1845-1865.

[143]Khan M.S., Khan K.I., Mahmood S. and Sheera M. (2019) Symmetric and asymmetric volatility clustering via GARCH family models: An evidence from religion-dominant countries. *Paradigms*, 13(1):20-25.

[144]Killian L. and Taylor M.P. (2003) Why is it so difficult to beat the random walk model? *Journal of international Economics*, 60: 85-107.

[145]Kleinbrod, V. and Li, X. (2017) Order Flow and Exchange Rate Co-Movement (January 3, 2017). *Asian Finance Association (AsianFA) 2017 Conference*. doi.org/10.2139/ssrn.2893295.

[146]Koller M. and Stahel W.A. (2011) Sharpening Wald-type inference in robust regression for small samples. *Computational Statistics & Data Analysis*, 55(8):2504–2515.

[147]Kosapattarapim C. (2013) *Improving volatility forecasting of GARCH models: applications to daily returns of emerging stock market*. Unpublished PhD thesis. Available at:https://ro.uow.edu.au/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=5006&context=theses. Accessed on13/12/2029.

[148]Kosapattarapim, C., Lin, Y. & McCrae, M. (2012). Evaluating the volatility forecasting performance of best fitting GARCH models in emerging Asian stock markets. International Journal of Mathematics & Statistics, 12 (2), 1-15.

[149]Kuester K., Mittnik S. and Paolella M. (2006) Value-at-Risk prediction: a comparison of alternative strategies. *Journal of Financial Econometrics*, 4 (1):5-89.

[150]Kühl M. (2008) Strong co-movements of exchange rates: *Theoretical and empirical cases when currencies become the same asset*. EcoMod2008 23800071, EcoMod. Available at: https://www.econstor.eu/bitstream/10419/31989/1/585596816.pdf. Accessed on 07/08/2020.

[151]Kühl, M (2007) *Cointegration in the foreign exchange market and market efficiency since the introduction of the Euro: Evidence based on bivariate cointegration analyses*. cege Discussion Papers, No. 68, University of Göttingen, Center for European, Governance and Economic Development Research (cege), Göttingen. Available: http://hdl.handle.net/10419/31987. Accessed on 31/10/2019.

[152]Kumar D. (2020) Value-at-Risk in the Presence of Structural Breaks Using Unbiased Extreme Value Volatility Estimator. *Journal of Quantitative Economics, Springer; The Indian Econometric Society (TIES),* vol. 18(3): 587-610. Doi: 10.1007/s40953-020-00197-w.

[153]Kupiec P.H. (1995) Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, 3:73–84.

[154]Kyei K.A. and Antwi A. (2017) Internal and External Factors of Future Returns in the Banking Business: Time Series Analysis of Interrelationship. *Journal of Economics and Behavioral Studies,* 9(1), 82-89.

[155] Kyei K.A. and Antwi A. (2018) Determination of banks profitability: modelling some factors. *Actual Problems of Economics*, 2(200):108-123.

[156]Lahiri K. and Sheng X. (2010) Measuring forecast uncertainty by disagreement: the missing link. *Journal of Applied Econometrics*, 25(4):514–538.

[157]Lamoureux C. and Lastrapes W. (1990) Persistence in Variance, Structural Change and the GARCH Model. *The Journal of Business & Economic Statistics*, 8(2):225-234.

[158]Laporta A.G., Luca M. and Petrella L. (2018) Selection of Value-at-Risk models for energy commodities. *Energy Economics*, 74: 628-43.

[159]Lee C. and Su J. (2012) Alternative statistical distributions for estimating Value-at-Risk: Theory and evidence. *Review of Quantitative Finance and Accounting*, 39:309–31.

[160]Lee G.J. and Engle R.F. (1999) *A permanent and transitory component model of stock returns volatility*. In Cointegration Causality and Forecasting a Festschrift in Honor of Clive WJ Granger. Pages 475-497, Oxford University Press.

[161]Lee H. (2015) The Effect of Level Shift in the Unconditional Variance on Predicting Conditional Volatility. *Journal of Economic Theory and Econometrics*, 26(2): 36–56.

[162]Li L. and Engle R. F. (1998) Macroeconomic Announcements and Volatility of Treasury Futures. *UCSD Economics Discussion Paper 98-27*.

[163]Li H. (2012) The impact of China's stock market reforms on its international stock market linkages. *Quarterly Review of Economics and Finance*, 52:358–68.

[164]Li W.K. and Mark T. K. (1994) On the Squared Residual Autocorrelations in Non-

[165]Linear Time Series with Conditional Heteroscedasticity. *Journal of Time Series Analysis*, 15(6):627–636.

[166]Ling S. and M. McAleer (2003) Asymptotic theory for a vector ARMA-GARCH model. *Economic Theory*, 19(2): 280-310.

[167]Ljung G.M. and Box G.E.P. (1978) On a Measure of Lack of Fit in Time Series Models. *Biometrika*, 65(2): 297–303.

[168]Lo A.W. and MacKinlay C.A. (1988) Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test. *The Review of Financial Studies*, 1:41–66.

[169]Mahdi E. and McLeod I.A. (2012) Improved Multivariate Portmanteau Test. *Journal of Time Series Analysis*, 33(2): 211–222.

[170]Mandelbrot B. (1963) The variation of certain speculative prices. *Journal of business, 36:394-419*.

[171]Marconi D. (2018) Currency co-movements in Asia-Pacific: The regional role of the renminbi. *Pacific Economic Review,* 23(2), 150-163. https://doi.org/10.1111/1468-0106.12266.

[172]Martin V.L., Hurn A.S. and Harris D. (2012) *Econometric Modelling with Time Series, Specification, Estimation and Testing*. Cambridge University Press.

[173]Matei M. (2009) Assessing volatility forecasting models: why GARCH models take the lead. *Romanian Journal of Economic Forecasting*. Available at: http://www.ipe.ro/rjef/rjef4_09/rjef4_09_3.pdf. Accessed on 06/01/2020.

[174]May C. and Farrell G. (2017) Modelling exchange rate volatility dynamics: Empirical evidence from South Africa. ERSA working paper 705. Available at: https://www.econrsa.org/system/files/publications/working_papers/working_paper_705.pdf.

[175]McAleer M and da Veiga B (2008) Single-index and portfolio models for forecasting value-at risk thresholds. *Journal of Forecasting*, 27(3):217-235. ISSN 1099-131X. Doi: 10.1002/for.1054.

[176]McCrae M., Lin Y.X., Pavlik D. and Gulati C. (2002) Can cointegration based forecasting outperform univariate models? An application to Asian exchange rates. *Journal of forecasting*, 20:273-286.

[177]McGill W.J. (1954) Multivariate information transmission. In: *Psychometrika* 19 (2): 97-116. DOI: 10.1007/BF02289159.

[178]McLeod A.I. and Li W.K. (1984) Diagnostic Checking ARMA Time Series Models Using Squared-Residual Auto-correlations. *Journal of Time Series Analysis*, 4, 269-273.

[179]Mikosch T. and Stărică C. (2004) Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects. *The Review of Economics and Statistics*, 86:378-390.

[180]Mincer J.A. and Zarnowitz V. (1969) *The Evaluation of Economic Forecasts, NBER Chapters, in: Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, pages 3-46, National Bureau of Economic Research, Inc.

[181]Mlambo C., Maredza A. and Sibanda k. (2010) Effects of Exchange Rate Volatility on the Stock Market: A Case Study of South Africa. *Mediterranean Journal of Social Sciences.* Doi:10.5901/mjss.2013.v4n14p561.

[182]Monti, A. C. (1994) A Proposal for a Residual Autocorrelation Test in Linear Models. *Biometrika*, 81(4): 776–780.

[183]Nelson. D.B. (1991) Conditional heteroscedasticity in asset returns: a new approach. *Econometrica*, 59(2):347-70.

[184]Osei-Asibey K.P. (2010) *Exchange rate volatility in LDCs, some findings from the Ghanaian, Mozambican and Tanzanian markets.* Unpublished doctoral thesis. Available at: https://discovery.dundee.ac.uk/en/studentTheses/exchange-rate-volatility-in-ldcs. Accessed on 01/07/2016.

[185]Owusu P.J, Adam A.M. and Tweneboah G. (2017) Co-movement of real exchange rates in the West African Monetary Zone. *Cogent Economics & Finance*. Doi: /10.1080/23322039.2017.1351807.

[186]Palm F.C. (1996) GARCH models of volatility, Handbook of statistics (14th ed). Elsevier.

[187]Pantula S.G. (1986) Comment: Modelling the persistence of conditional variances. *Economet-ric Re-views*, 5(1):71-74.

[188]Pardy C. (2013) Mutual information as an exploratory measure for genomic data with discrete and continuous variables. Unpublished thesis, The University of New South Wales. Available: http://handle.unsw.edu.au/1959.4/52767. Accessed on 03/09/2020.

[189]Park B. and An J. (2020) What Drives Growing Currency Co-movements with the Renminbi? *East Asian Economic Review*, 24(1), 31-59. Doi: 10.11644/KIEP.EAER.2020.24.1.371.

[190]Peňa D. and Rodríguez J. (2006) The Log of the Determinant of the Autocorrelation Matrix for Testing Goodness of Fit in Time Series. *Journal of Statistical Planning and Inference*, 136(8): 2706–2718.

[191]Peňa, D. and Rodríguez J. (2002) A Powerful Portmanteau Test of Lack of Fit for Time Series. *Journal of the American Statistical Association*, 97(458): 601–610.

[192]Peng H., Fuhui L. and Ding C (2005) Feature Selection Based on Mutual Information Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 27 (8): 1226-38.

[193]Poon S. & Granger, C. (2005) Practical Issues in Forecasting Volatility. *Financial Analysts Journal*, 61(1), 45-56. Available at: http://www.jstor.org/stable/4480636. Accessed on 08/09/2021.

[194]Pritsker M. (2001) The hidden dangers of historical simulation. *Finance and Economics. Dis-cussion Series,* 2001-27.

[195]Rapach D., Strauss J. and Wohar M. (2007) Forecasting Stock Return Volatility in the Presence of Structural Breaks. *Frontiers of Economics and Globalization.* Doi: 10.1016/S1574-8715(07)00210-2.

[196]Reddy L., Pillai D., Huang, C. & Huang, C-K. (2017) Forecasting Stock Market Volatility in the Presence of Structural Breaks: An Application to Value-at-Risk Estimation in South Africa. *59th Annual Conference of the South African Statistical Association (2017), Bloemfontein, South Africa*: Zenodo. Doi: 10.5281/zenodo.3357652.

[197]Ribeiro P. P., Cermeño R. and Curto J. D. (2017) Sovereign bond markets and financial volatility dynamics: Panel-G.A.R.C.H. evidence for six-euro area countries. Finance Research Letters, 21:107–114.

[198]Roccioletti S. (2015) *Backtesting Value-at-Risk and expected shortfall*. Springer Gabler.

[199]Romano J.P. and Wolf M. (2005) Stepwise multiple testing as formalized data snooping. *Econometrica*, 73(4):1237, 1282.

[200]Ross B.C. (2014) Mutual Information between Discrete and Continuous Data Sets. *PLoS ONE*, 9(2):e87357. Doi:10.1371/journal.pone.0087357.

[201]Santos A.A.P., Nogales F.J., and Ruiz E. (2013) Comparing univariate and multivariate models to forecast portfolio Value-at-Risk. *Journal of financial econometrics*, 11(2): 400-441.

[202]Scheuerell M. D. (2017) muti: An R package for computing mutual information. Available at: http://doi.org/doi.org/10.5281/zenodo.439391. Accessed on 20/05/2020.

[203]Schwert G.W. (1990) Stock volatility and the crash of '87. *Review of Financial Studies*, 3(1):77.

[204]Scott D.W. (2010) Scott's Rule. Wiley Interdisciplinary Reviews: *Computational Statistics*, 2 (4): 497–502.

[205]Scotti C. (2016) Surprise and uncertainty indexes: real-time aggregation on real-activity macro surprises. *Journal of Monetary Economics*, 82:1–19.

[206]Ser-Huang P. (2005) *A Practical Guide to Forecasting Financial Market Volatility*. John Wiley & Sons.

[207]Shamiri A. and Isa Z. (2009) Modelling and forecasting volatility of the Malaysian stock market. *Journal of Mathematics and Statistics*, 5:234-240.

[208]Shannon C.E. (1948) A mathematical theory of communication. *The Bell Systems Technical Journal* 27 (3): 379–423.

[209]Shi Y. and Kin-Yip H. (2016) Addressing the Confusion Between Hyperbolic Memory and Regime Switching: The Markov Regime-Switching Hyperbolic Garch Model. *Working Paper*. Available at: http://dx.doi.org/10.2139/ssrn.2469028. Accessed on 21/09/2020.

[210]Shi Y. and Yang Y. (2018) Modelling High Frequency Data with Long Memory and Structural Change: A-HYEGARCH Model. *Risks*, 2018: 6-26. Doi: 10.3390/risks6020026. Available at: http://www.mdpi.com/2227-9091/6/2/26/pdf. Accessed on 07/01/2020.

[211]Srivastava M. S. (2005) Some Tests Concerning the Covariance Matrix in High Dimensional Data. *Journal of Japan Statistical Society*, 35(2): 251–272.

[212] Starica C. and Granger C. (2005) Nonstationarities in stock returns. *The review of Economics and Statistics*, 87 (3): 503–522. Doi: 10.1162/0034653054638274.

[213]Straumann D. (2005) *Lecture Notes in Statistics: Estimation in Conditional Heteroscedastic Time Series Models*. 181. Springer, New York.

[214]Su J. (2014) The interrelation of stock markets in China, Taiwan and Hong Kong and their constructional portfolio's Value-at-Risk estimate. *Journal of Risk Model Validation*, 8:69-127.

[215]Su J. (2015) Value-at-Risk estimates of the stock indices in developed and emerging markets including the spillover effects of currency market. *Economic Modelling*, 46:204–24.

[216]Su J. and Hung J. (2018) The Value-At-Risk Estimate of Stock and Currency-Stock Portfolios' Returns. *Journal of Risks*, 6:133. Doi: 10.3390/risks6040133.

[217]Suess S., Jondeau E., Poon S. and Rockinger M. (2008) Financial modelling under non-Gaussian distributions. *Financ Mark Portfolio Manag*, 22:91–92. Doi: 10.1007/s11408-007-0071-5.

[218]Szegö G. (2004) *Risk Measures for the 21st Century.* John Wiley & Sons, Ltd, Chichester.

[219]Taylor S.J. (1986) *Modelling financial time series*. Wiley, Chichester.

[220]Taylor S.J. (2005) *Asset Price Dynamics, Volatility, and Prediction* (1st ed). Princeton

[221]Teräsvirta T. (2006) *An Introduction to Univariate GARCH Models. University of Aarhus and Stockholm School of Economics*. SSE/EFI Working Papers in Economics and Finance, No. 646. Available at: https://www.econstor.eu/bitstream/10419/56283/1/521138485.pdf. Accessed on 15/04/2019.

[222]Tsay R.S. (2010) *Analysis of Financial Time Series*, (3rd ed). John Wiley & Sons.

[223]Whaley R. (2000) The Investor Fear Gauge. *Journal of Portfolio Management*, 26(3):12-17.

[224]White H. (2000) A Reality Check for Data Snooping. *Econometrica*, 68:1097-1126.

[225]Wickremasinghe G.B. (2004) *Efficiency of foreign exchange markets: a developing country perspective*. International Finance 0406004, University Library of Munich, Germany. Available at: https://econwpa.ub.uni-muenchen.de/econ-wp/if/papers/0406/0406004.pdf. Accessed on 15/04/2019.

[226]Wold H. (1938) *A Study in the Analysis of Stationary Time Series*. Uppsala: Almqvist & Wiksell.

[227] Yarovayaa L., Brzeszczyński J. and Lauc C.K.M. (2017) Asymmetry in spillover effects: Evidence for international stock index futures markets. *International Review of Financial Analysis*, 53: 94-111. doi.org/10.1016/j.irfa.2017.07.007.

[228] Yohai V.J. (1987) High breakdown-point and high efficiency estimates for regression. *The Annals of Statistics*, 15:642–65.

[229] Zakoian J.M. (1994) Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, 18(5):931-955.

[230] Zenga X., Xiaa Y. and Tong H (2018) Jackknife approach to the estimation of mutual information. PNAS, 115(40):9956-9961. Available at: https://doi.org/10.1073/pnas.1715593115. Accessed on 05/06/2020.

Appendices

Appendix A

Simplification of the ARMA representation of EGARCH model

Consider the ARMA representation

$$\ln\left(h_t\right)=\alpha_0+\frac{\left(1+\alpha_1+\alpha_2L^2+...+\alpha_sL^s\right)}{1-\left(\beta_1L+\beta_2L^2+...+\beta_qL^q\right)}g\left(z_{t-1}\right). \tag{A1}$$

Writing (A1) in a compact form and simplifying yields:

$$\left(1-\sum_{j=1}^{q}\beta_jL^j\right)\ln\left(h_t\right)=\alpha_0\left(1-\sum_{j=1}^{q}\beta_jL^j\right)+g\left(z_{t-1}\right)+\sum_{i=1}^{s}\alpha_iL^ig\left(z_{t-1}\right). \tag{A2}$$

Expanding the left-hand side term of (A12) and solving for $\ln\left(h_t\right)$ yields

$$\ln\left(h_t\right)=\alpha_0\left(1-\sum_{j=1}^{q}\beta_jL^j\right)+\sum_{j=1}^{q}\beta_j\ln\left(h_{t-j}\right)+g\left(z_{t-1}\right)+\sum_{i=1}^{s}\alpha_iL^ig\left(z_{t-1}\right). \tag{A3}$$

Now, expanding the last term of (A3) and applying the lag operator on $g\left(z_{t-1}\right)$ while recalling that $\alpha_1\equiv1$, we have:

$$g\left(z_{t-1}\right)+\sum_{i=1}^{s}\alpha_iL^ig\left(z_{t-1}\right)=\alpha_1g\left(z_{t-1}\right)+\alpha_1g\left(z_{t-2}\right)+\alpha_2g\left(z_{t-3}\right)+...+\alpha_sg\left(z_{t-1-s}\right). \tag{A4}$$

Since it is unlikely for the parameters of $g\left(z_{t-1}\right)$ and $g\left(z_{t-2}\right)$ in (A4) to be equal at any given time, a shifting of the parameters of the $g\left(z_{t-2}\right)$ term yields:

$$g\left(z_{t-1}\right)+\sum_{i=1}^{s}\alpha_iL^ig\left(z_{t-1}\right)=\sum_{i=1}^{p}\alpha_ig\left(z_{t-i}\right)\text{ where }p=s+1. \tag{A5}$$

Substituting (A5) into (A4) yield:

$$\ln\left(h_t\right) = \alpha_0 + \sum_{i=1}^{q} \beta_j \ln\left(h_{t-1}\right) + \sum_{i=1}^{p} \alpha_i g\left(z_{t-i}\right). \tag{A6}$$

Now, recalling the definition from (2.5) in section 2.2.2, we can write

$$g\left(z_{t-i}\right) \equiv \theta\left(\frac{a_{t-i}}{\sqrt{h_{t-i}}}\right) + \gamma\left(\left|\frac{a_{t-i}}{\sqrt{h_{t-i}}}\right| - E\left|\frac{a_{t-i}}{\sqrt{h_{t-i}}}\right|\right). \tag{A7}$$

Thus, substitution of (A7) into (A6) yields:

$$\ln\left(h_t\right) = \alpha_0 + \sum_{j=1}^{q} \beta_j \ln\left(h_{t-j}\right) + \sum_{i=1}^{p} \alpha_i \left\{\theta\left(\frac{a_{t-i}}{\sqrt{h_{t-i}}}\right) + \gamma\left(\left|\frac{a_{t-i}}{\sqrt{h_{t-i}}}\right| - E\left|\frac{a_{t-i}}{\sqrt{h_{t-i}}}\right|\right)\right\}. \tag{A8}$$

Expanding (A8) while substituting $\alpha_i\theta = \gamma_i$ and $\alpha_i\gamma = \alpha_i$ yields

$$\ln\left(h_t\right) = \alpha_0 + \sum_{i=1}^{p} \gamma_i \left(\frac{a_{t-i}}{\sqrt{h_{t-i}}}\right) + \sum_{i=1}^{p} \alpha_i \left(\left|\frac{a_{t-i}}{\sqrt{h_{t-i}}}\right| - E\left|\frac{a_{t-i}}{\sqrt{h_{t-i}}}\right|\right) + \sum_{j=1}^{q} \beta_j \ln\left(h_{t-j}\right). \tag{A9}$$

Appendix B

Additional Tables for chapter 5

Table B 1: Estimates for the unrestricted ARMA-GARCH model in chapter 5

Parameter	SGARCH (NOK)	EGARCH(SEK)	GJRGARCH (USD)
mu	-0.000116(0.000002)	0.000092(0.000012)	0.000249(0.000058)
ar1	2.007179(0.002849)	-0.500072(0.000277)	-0.099735(0.012830)
ar2	-1.442425(0.001362)	-0.838518(0.000512)	
ar3	0.621684(0.000142)	0.054731(0.000079)	
ar4	-0.283773(0.000095)	-0.190297(0.000051)	
ar5	0.142013(0.000189)	-0.174758(0.000044)	
ar6	-0.045818(0.000244)		
ma1	-1.697832(0.000176)	0.834044(0.000505)	0.254623(0.016170)
ma2	0.696838(0.000074)	0.999849(0.000050)	
ma3		0.175028(0.000402)	
ma4		0.074439(0.000411)	
ma5		0.158533(0.000502)	
alpha1	0.212661(0.053823)	0.063924*(0.040530)	0.031944(0.001885)
alpha2	0.000072*(0.082407)	-0.01851*(0.041731)	
beta1	0.367595*(0.329633)	0.992678(0.000031)	0.971434(0.000015)
beta2	0.273691*(0.238927)	-0.00051*(0.000779)	
gamma1		0.448906(0.069202)	-0.012679(0.002360)
gamma2		-0.38664(0.067442)	
skew	0.945101(0.069774)	0.998628(0.027610)	1.014359(0.014946)
shape	1.299872(0.064338)	1.284112(0.063022)	1.067958(0.050420)
omega	0.000001	-0.09136	3.920459e-08

Note: *Standard erros for the estimates are reported in brackets. External regressors are excluded.*

Table B 2: Estimated mean parameters for Augmented GARCH models in chapter 7

Variable	SGARCH	EGARCH	NAGARCH	TGARCH	GJRGARCH
mu	-0.000066	-0.000037	-0.000014	0.000327*	0.000103
	0.000153	**0.000147**	**0.000062**	**0.000153**	**0.000267**
ar1	0.774224*	-0.291449*	0.528180*	0.006775	-0.460186*
	0.044286	**0.063142**	**0.061269**	**0.110981**	**0.205887**
ar2	-1.299056*		-0.122854*	0.270761*	-0.337352
	0.027614		**0.031277**	**0.113134**	**0.227680**
ar3	0.904104*				-0.064145
	0.043043				**0.243899**
ar4	-0.785211*				0.369821
	0.034719				**0.196932**
ar5	0.068343*				-0.251870
	0.029806				**0.169254**
ma1	-0.649640*	0.563105*	-0.657827*		0.438884*
	0.026208	**0.055761**	**0.057382**		**0.208237**
ma2	1.146804*				0.195655
	0.023425				**0.222677**
ma3	-0.775374*				-0.036076
	0.004444				**0.230855**
ma4	0.606622*				-0.428506*
	0.023583				**0.181420**
ma5					0.206926
					0.158590

Note: *Asterisk indicates parameter is significant at 5% and standard errors of the estimates are in bold.*

Table B 3: Variance estimates for Augmented GARCH models in chapter 7

Variable	SGARCH	EGARCH	NAGARCH	TGARCH	GJRGARCH
omega	0.000006	-6.000952*	-0.000001	-0.000334	0.000007
		1.044131			
alpha1	0.218730*	0.013471	0.208907*	0.187734*	0.322879*
	0.042644	**0.036919**	**0.000373**	**0.030340**	**0.075778**
alpha2	0.000000		0.097020*	0.024648	
	0.023527		**0.002335**	**0.013100**	
alpha3			0.001762*		
			0.000001		
beta1	0.492075*	0.508858*		0.468314*	0.706584*
	0.043159	**0.092016**		**0.092819**	**0.077732**
beta2				0.095668	
				0.177326	
gamma1		0.078687*			-0.233429*
		0.204994			**0.072345**
eta1			-0.104569	-0.443649*	
			0.079599	**0.130831**	
eta2			-0.059777	0.999994	
			0.044993	**0.755942**	
eta3			-9.998326*		
			0.713925		
vxreg1	0.000157	47.41490*	0.001768*	0.002367*	<0.000001
	0.000113	**12.86954**	**0.000002**	**0.000229**	**0.001099**
vxreg2	0.000009*	1.098537*	0.000013*		0.000033
	0.000000	**0.105444**	**0.000000**		**0.000017**
Vxreg3		4.580441			
		13.51892			
skew	0.134367	1.031916*	-0.084694	0.129281*	0.008224
	0.115712	**0.040150**	**0.112667**	**0.054974**	**0.066660**
shape	1.876654*	1.594535*	1.842916*	1.952084*	1.237084*
	0.194511	**0.092146**	**0.162921**	**0.406497**	**0.063830**

Note*: Asterik indicates parameter is significant at 5% and standard errors of the estimates are in bold.*

Appendix C

Additional Figures for chapter 7

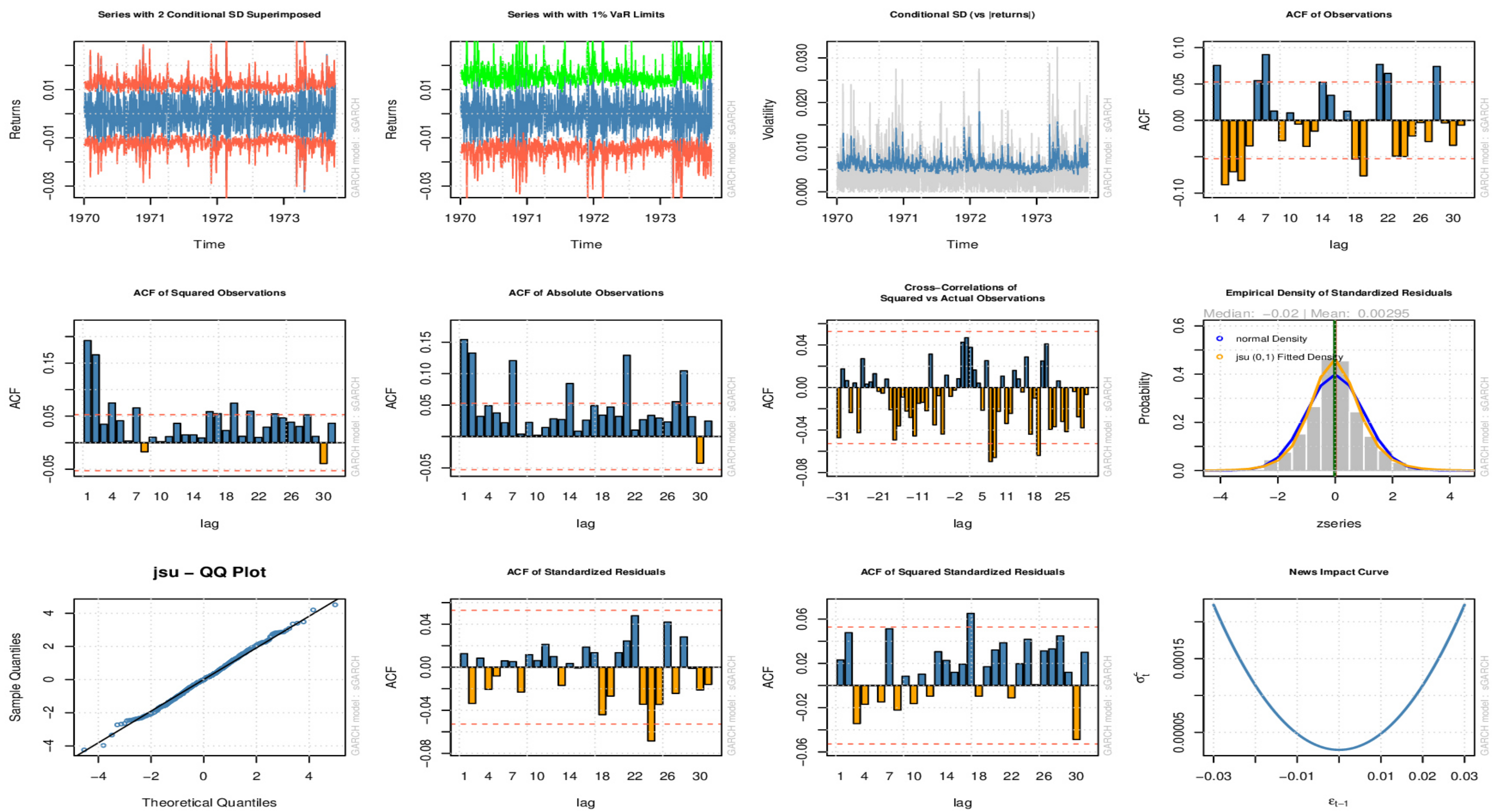

Figure C1: In-sampling plots for augmented standard GARCH model

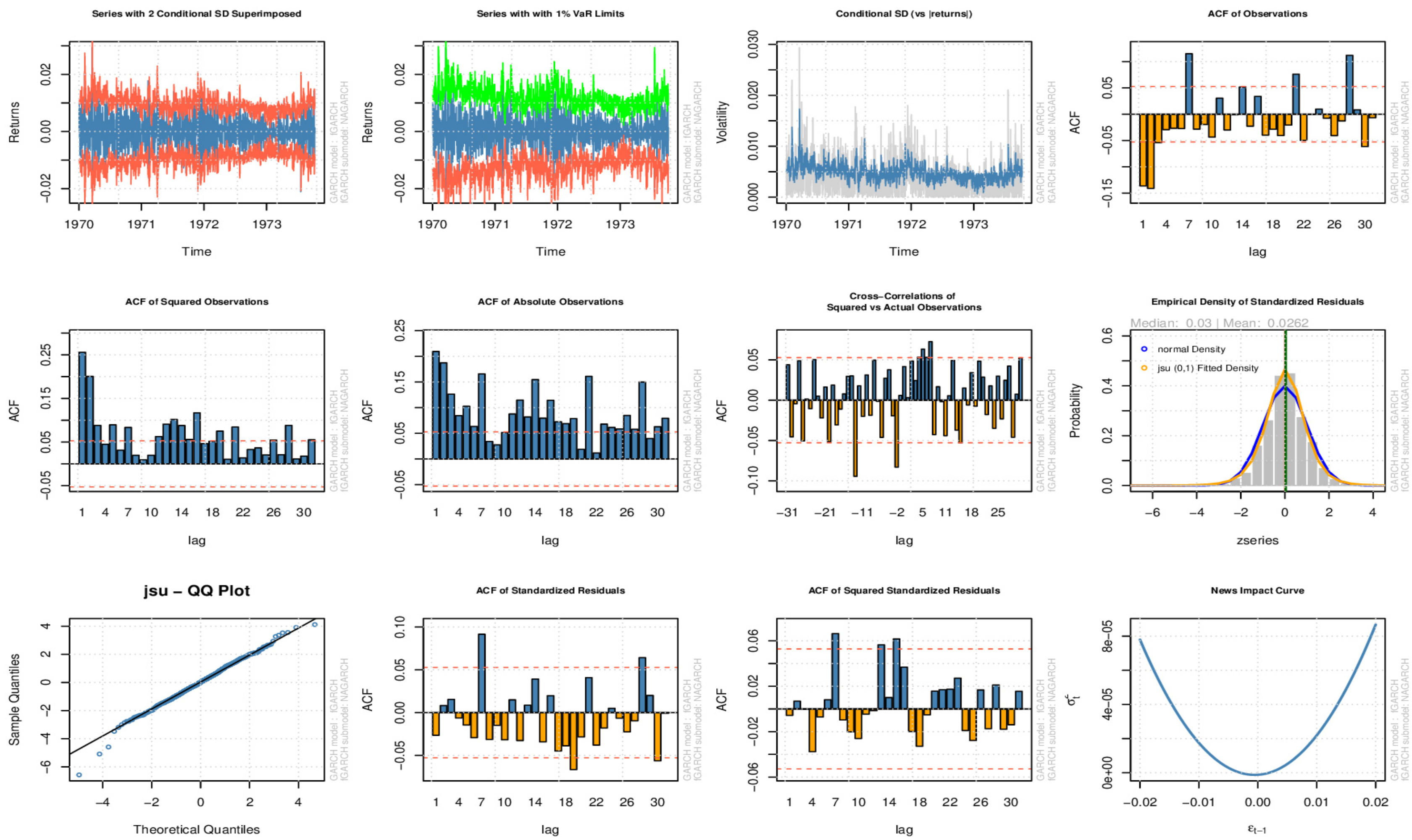

Figure C2: In-sampling plots for Augmented NAGARCH model

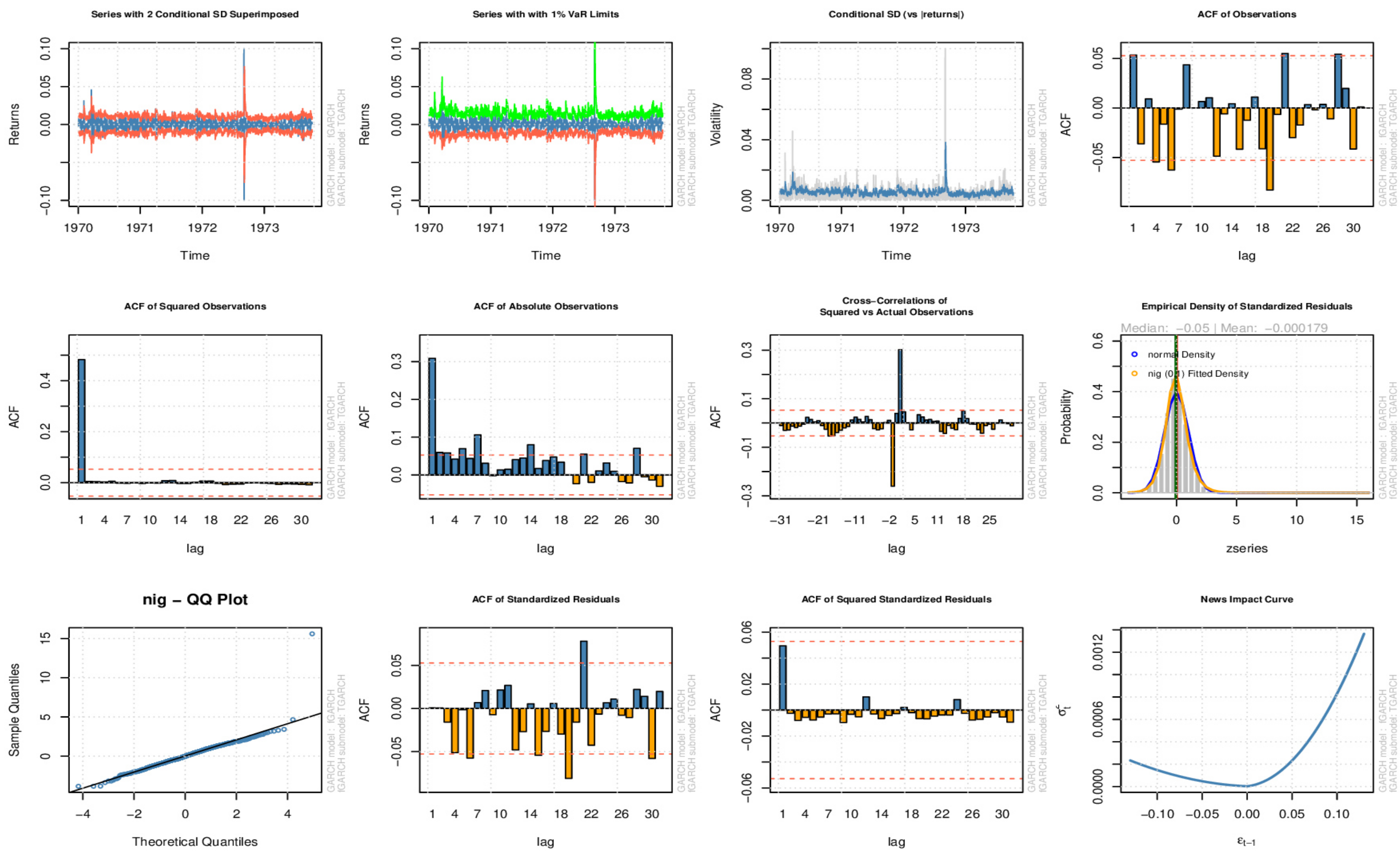

Figure C3: In-sampling plots for Augmented TGARCH models

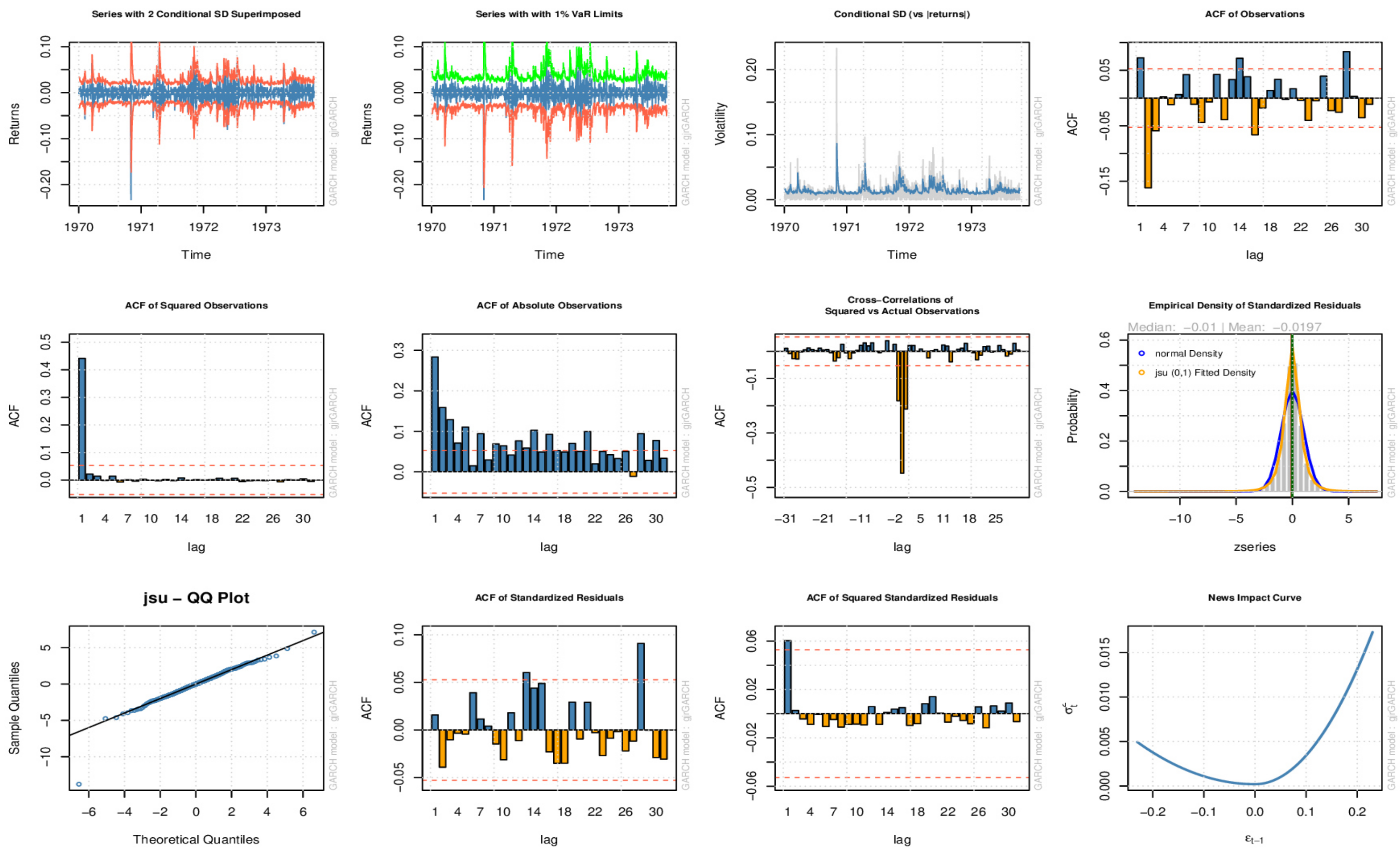

Figure C4: In-sampling plots for augmented gjrGARCH model

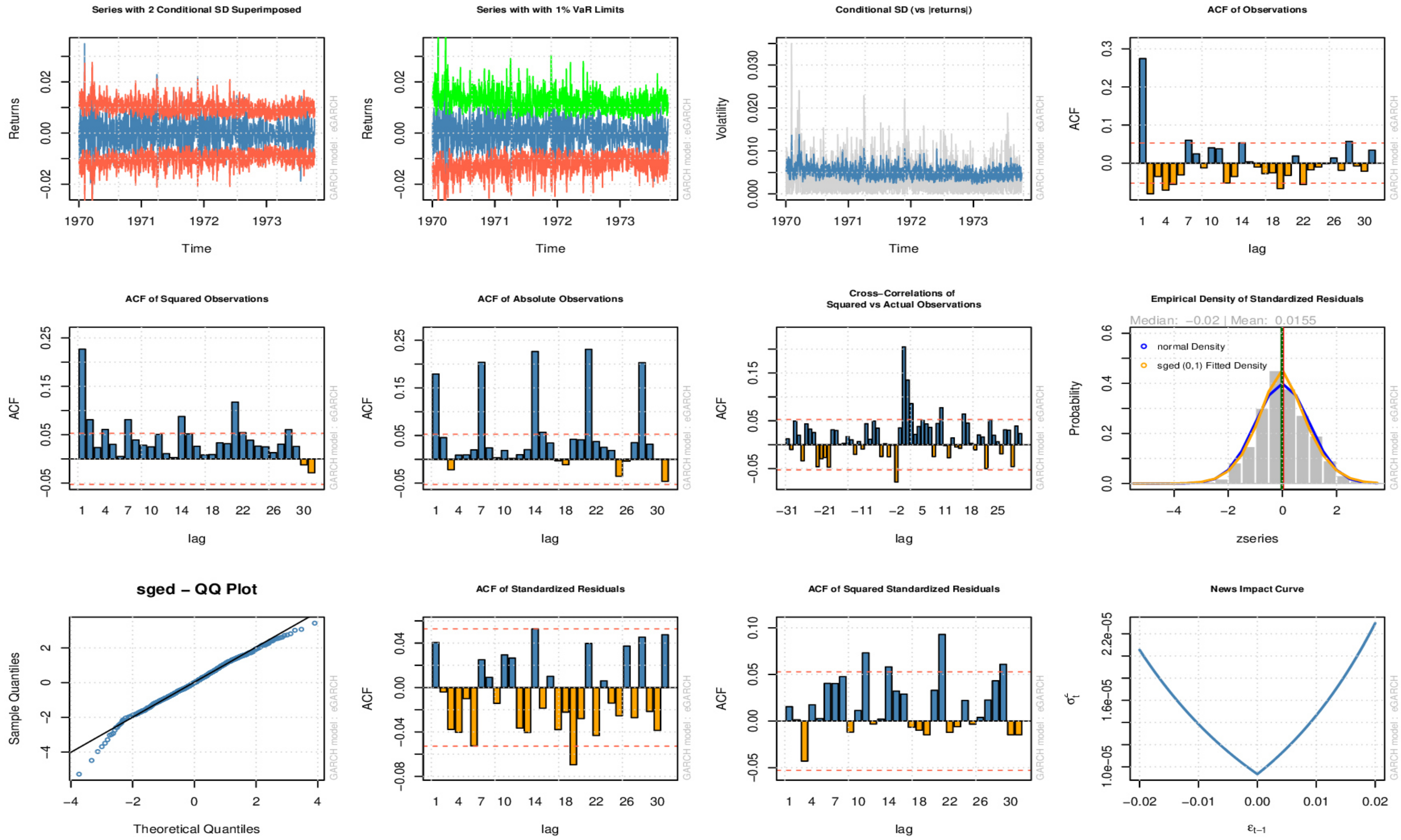

Figure C5: In-sampling plots for Augmented EGARCH model

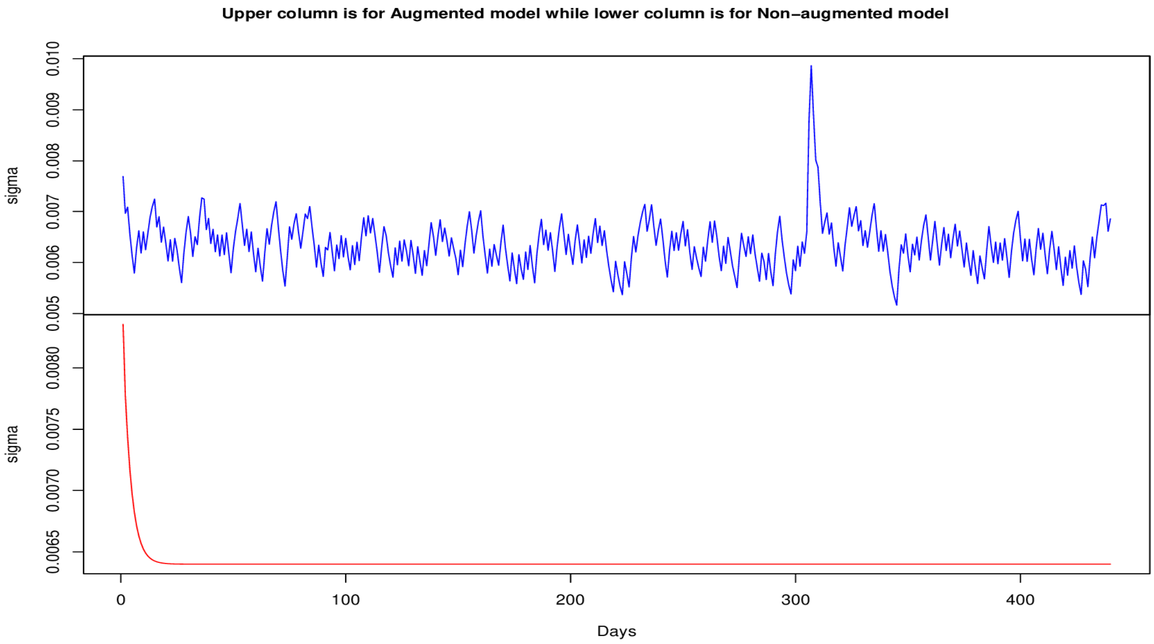

Figure C6: Conditional forecasted variances for SGARCH

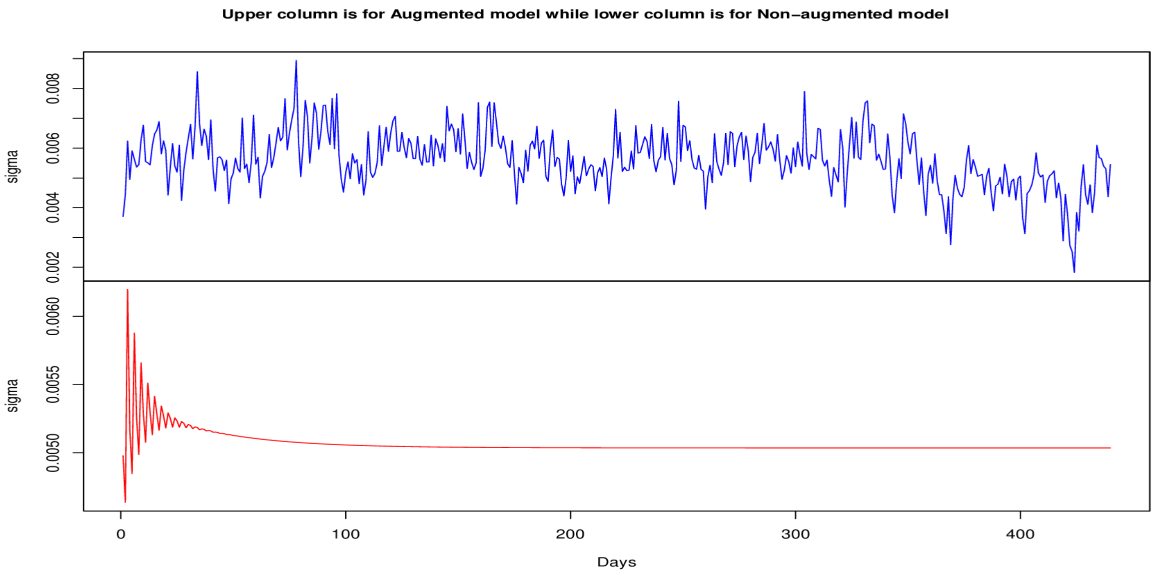

Figure C7: Conditional forecasted variances for NAGARCH

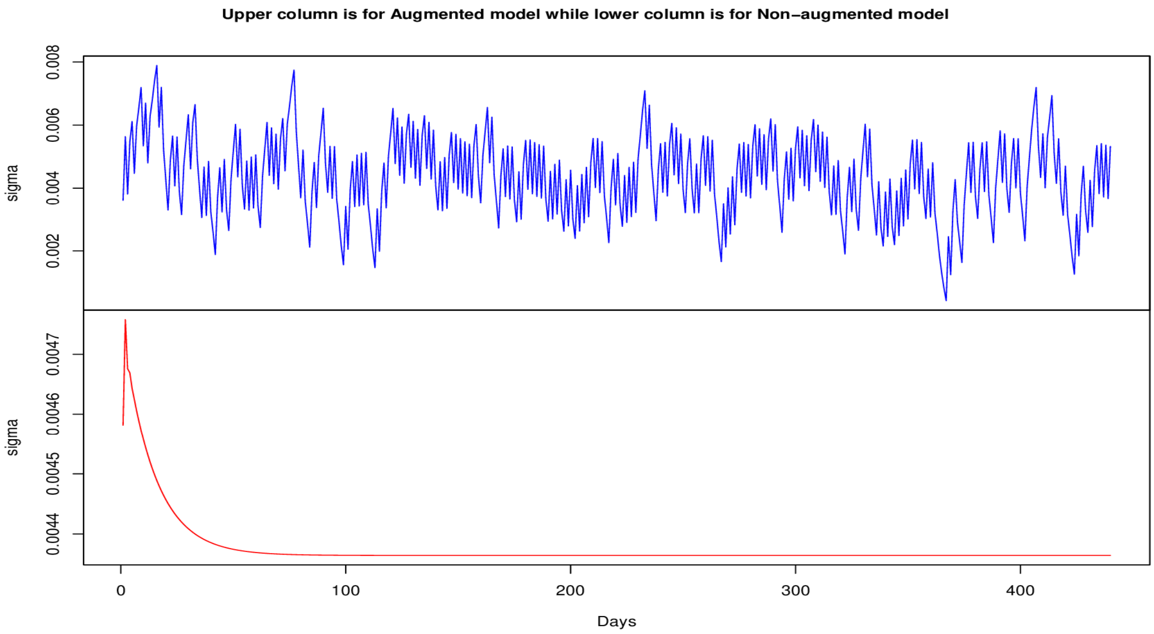

Figure C8: Conditional forecasted variances for TGARCH

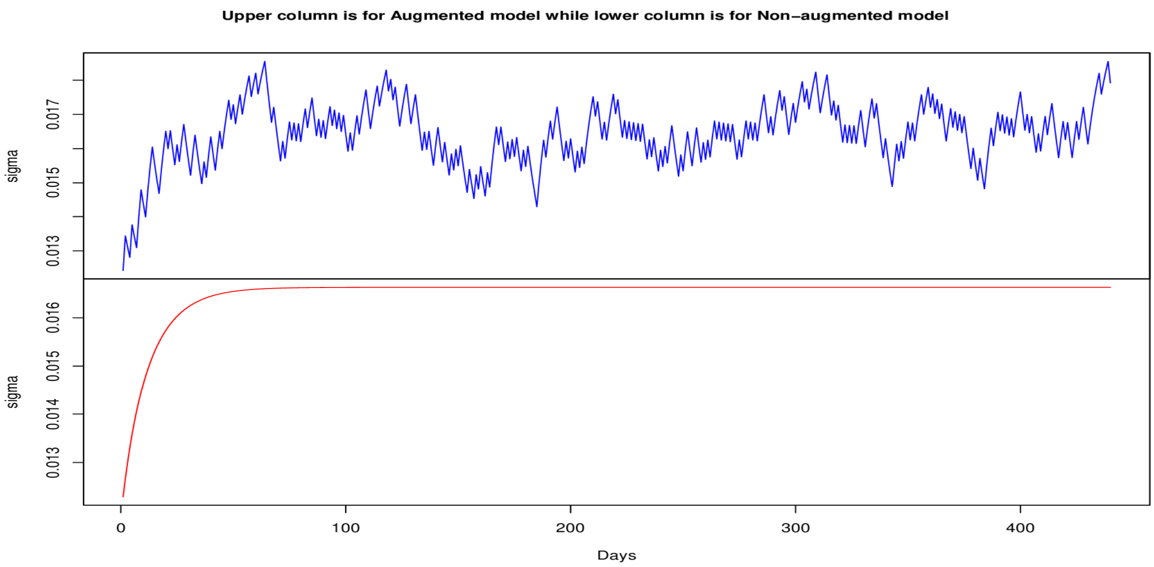

Figure C9: Conditional forecasted variances for GJRGARCH

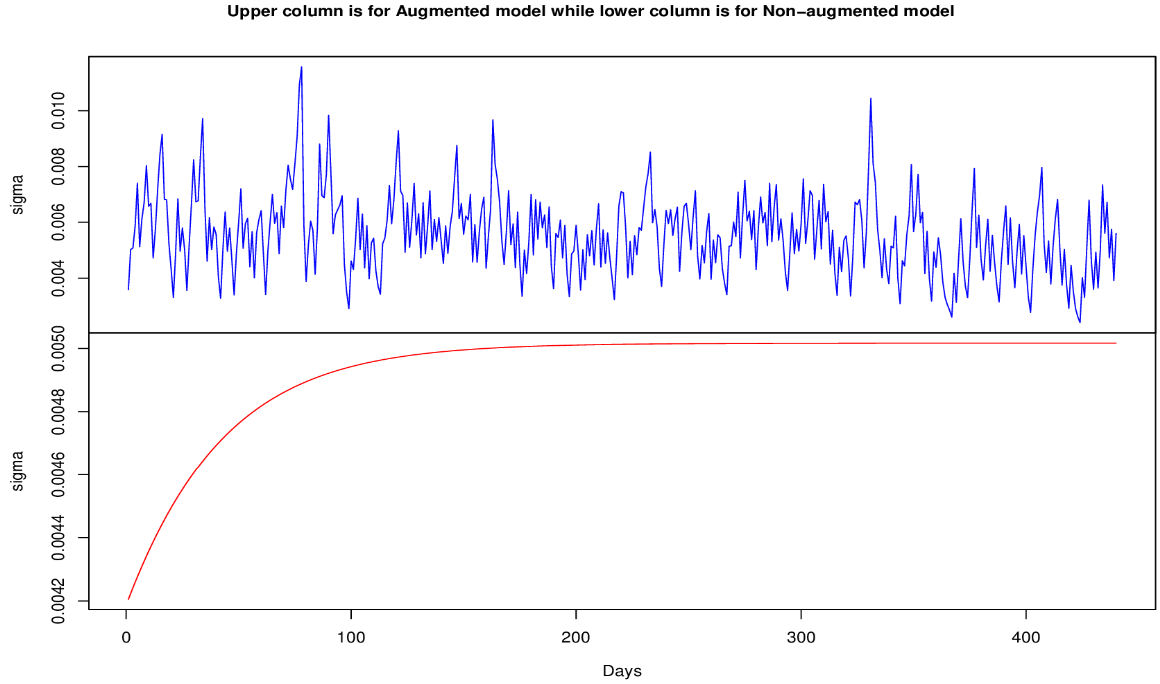

Figure C10: Conditional forecasted variances for EGARCH

Appendix D

Figures for chapter 8

The first rows of each of the graphs in Appendix A1 to A5 represent 99% VaR models while the second rows are for the 95% VaR models. In each row, the first model represents the augmented model while the last represent the non-augmented versions.

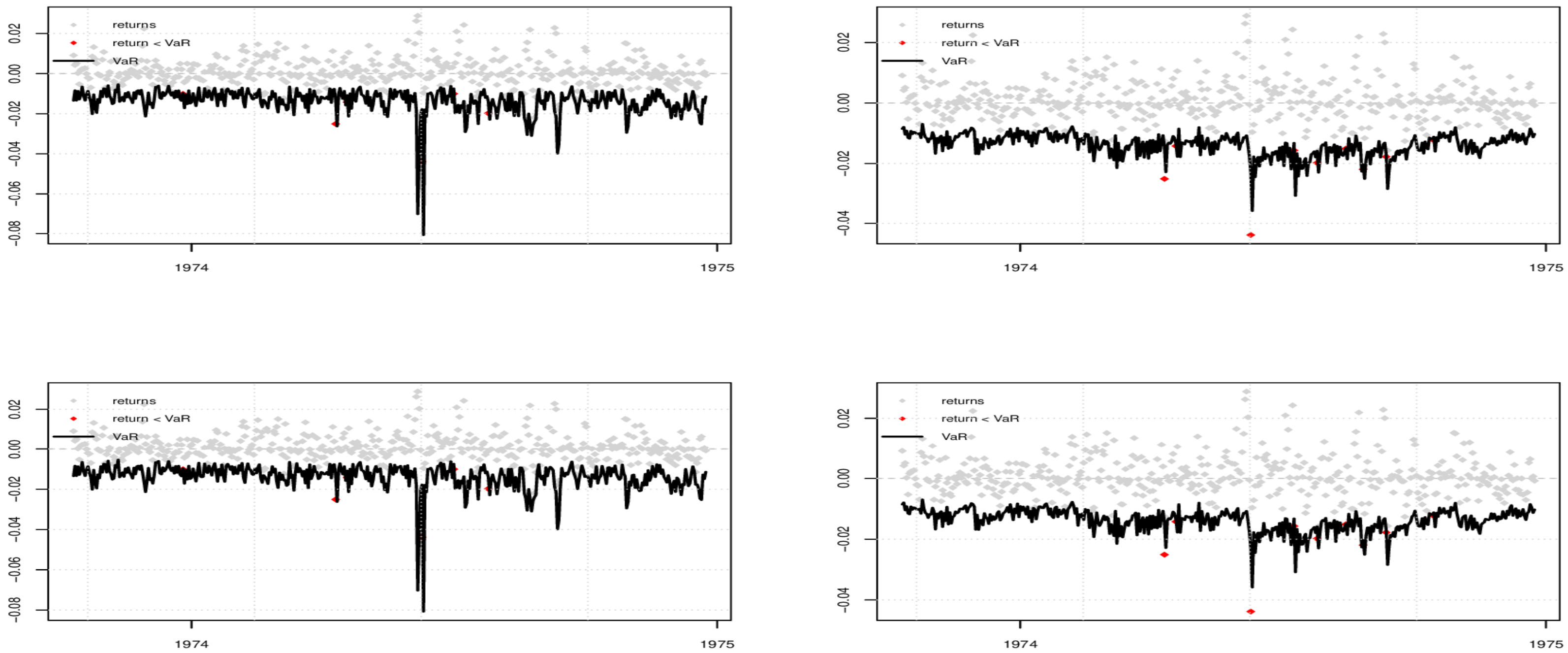

FIgure D1: **Forecasted VaR violations plot for EGARCH**

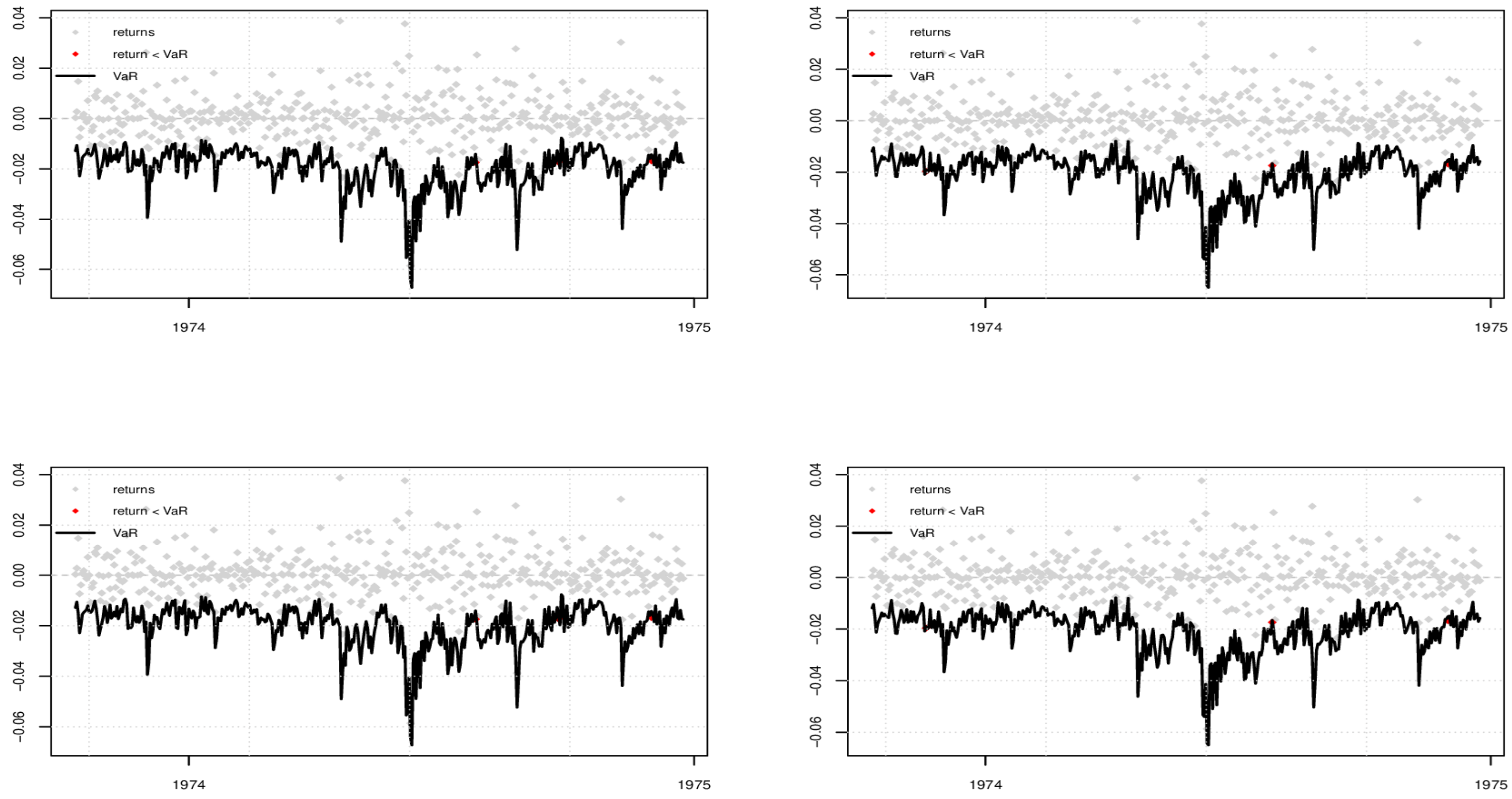

Figure D2: Forecasted VaR violations plot for NAGARCH

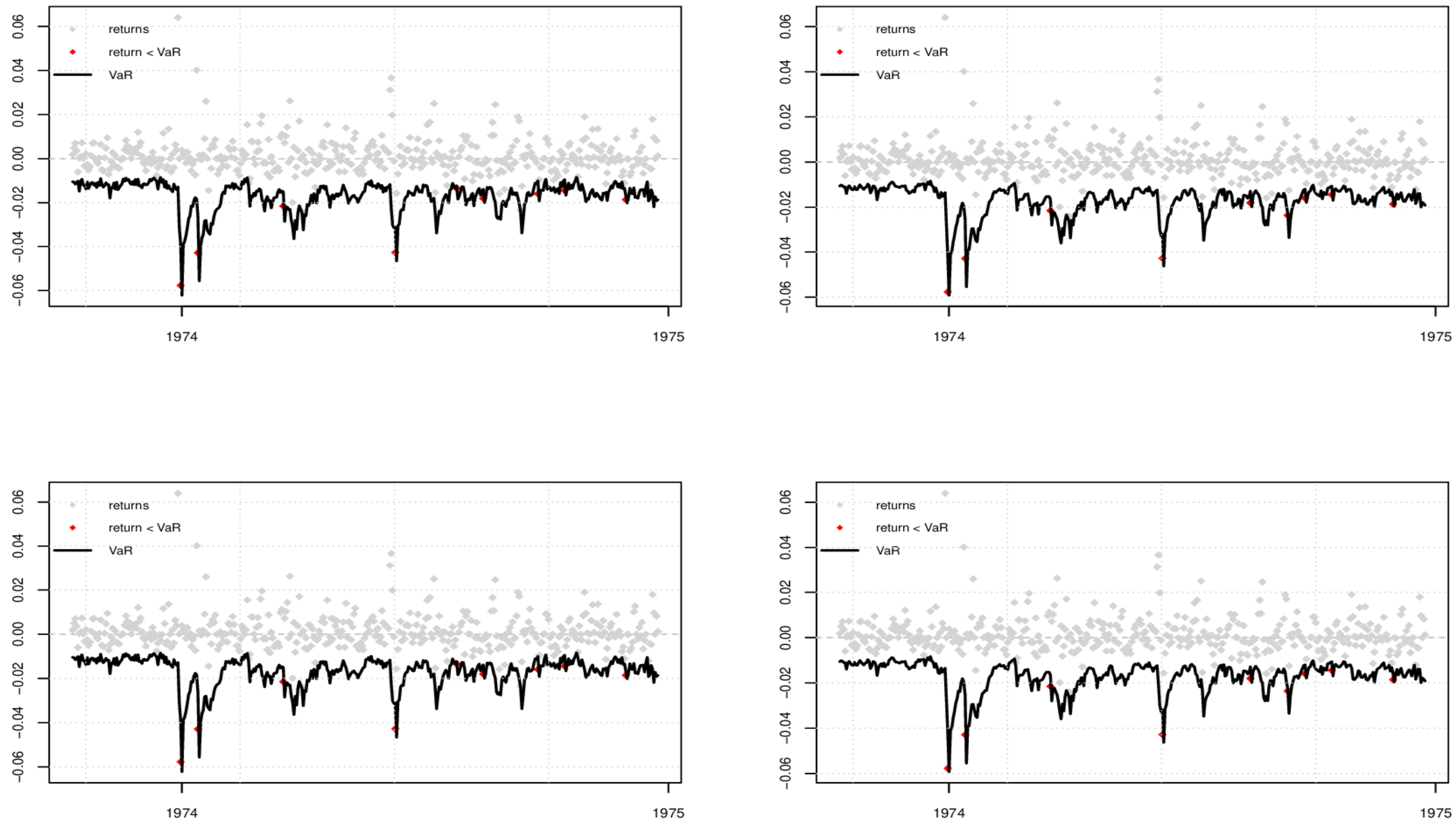

Figure D3: Forecasted VaR violations plot for TGARCH

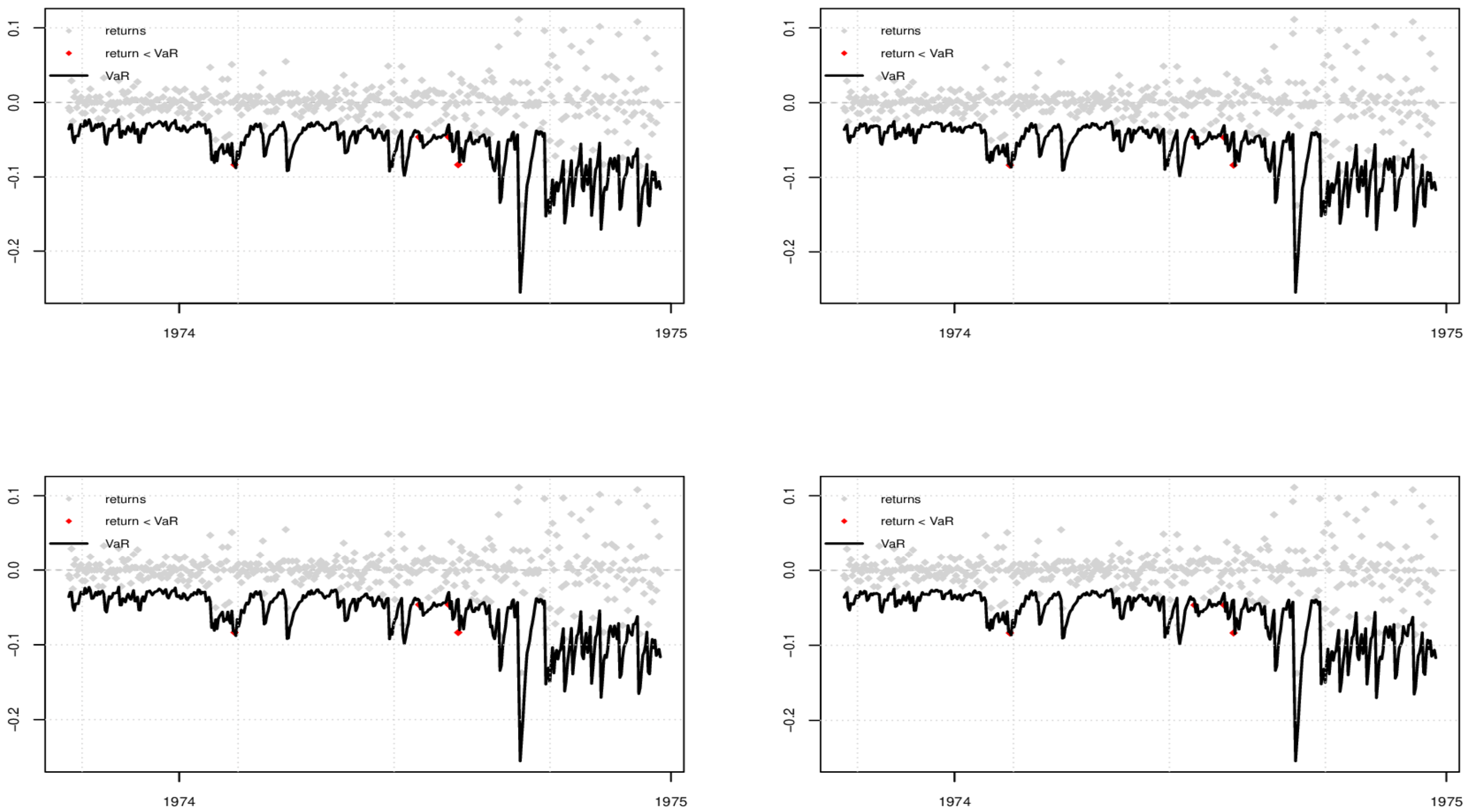

Figure D4: Forecasted VaR violations plot for GJRGARCH

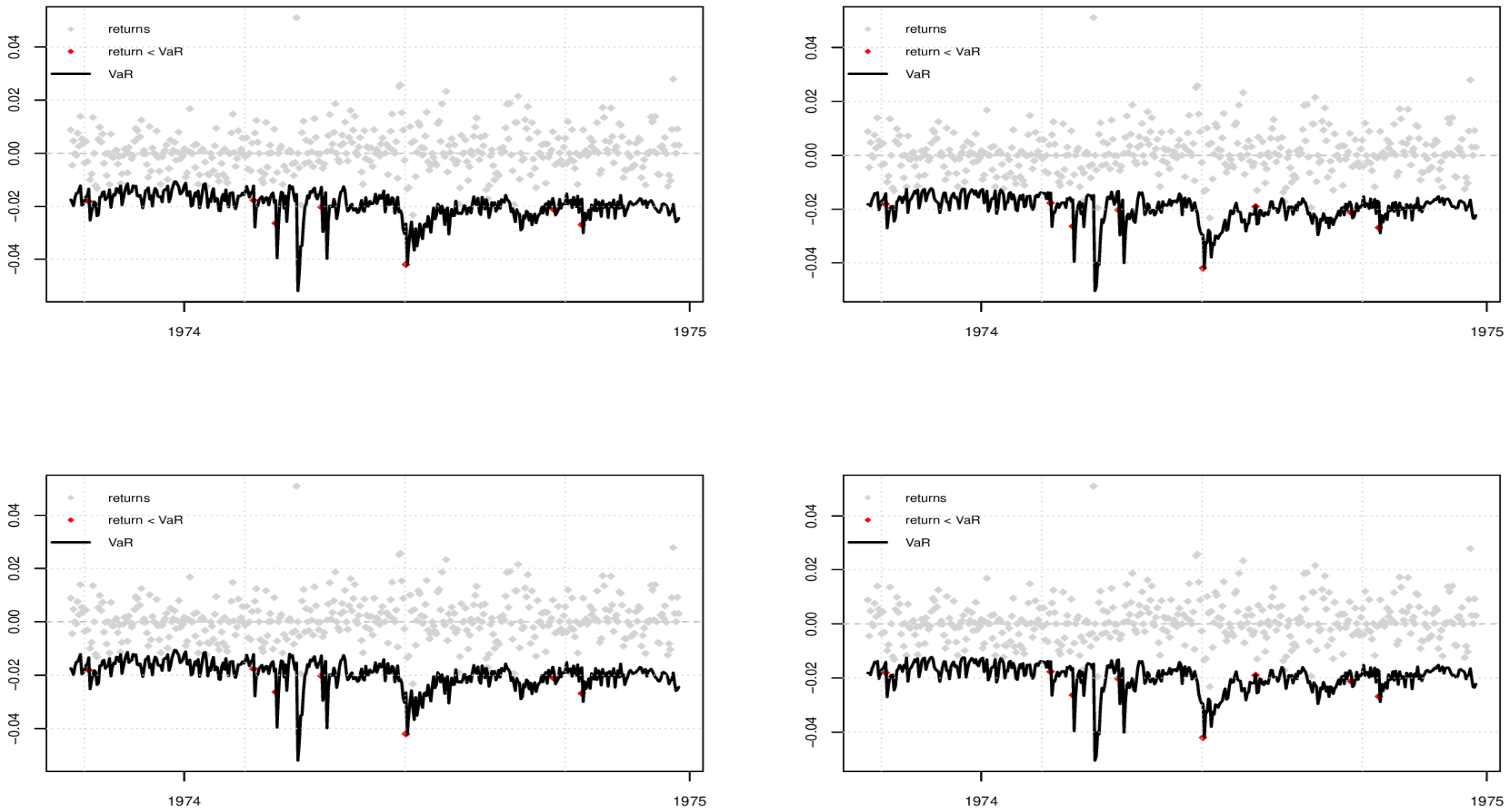

Figure D5: Forecasted VaR violations plot for SGARCH

Appendix E

Program files (Rcodes)

Appendix E1: Codes for data validation tests and miscellaneous analysis

NB: EVIEWS software was used for computing the descriptive analysis, unit root tests and for plotting Figure 2.1

```
####################### A Directed Acyclic Graphs ################
#############################################Packages
Require(aTSA);require(vrtest)
library(ggdag); require(DescTools)
attach(rates); require(zoo)
require(plotly); require(TSstudio)
require(xts)
############################################ Returns
BRL=diff(log(brl));BWP=diff(log(bwp));ILS=diff(log(ils))
INR=diff(log(inr));MWK=diff(log(mwk));NOK=diff(log(nok))
SEK=diff(log(sek));AUD=diff(log(aud));USD=diff(log(usd))
EUR=diff(log(eur));GBP=diff(log(gbp));CAD=diff(log(cad))
MZN=diff(log(mzn));KPW=diff(log(kpw))
###################### ARCH LM and Portmanteau Tests
fit.brl<-estimate(BRL, p=1,q=1); arch.test(fit.brl)
fit.bwp<-estimate(BWP, p=1,q=1); arch.test(fit.bwp)
fit.ils<-estimate(ILS, p=1,q=1);  arch.test(fit.ils)
fit.inr<-estimate(INRL, p=1,q=1); arch.test(fit.inr)
fit.mwk<-estimate(MWK, p=1,q=1); arch.test(fit.mwk)
fit.nok<-estimate(NOK, p=1,q=1); arch.test(fit.nok)
fit.SEK<-estimate(SEK, p=1,q=1); arch.test(fit.SEK)
fit.AUD<-estimate(AUD, p=1,q=1);  arch.test(fit.AUD)
fit.USD<-estimate(USD, p=1,q=1); arch.test(fit.USD)
fit.EUR<-estimate(EUR, p=1,q=1); arch.test(fit.EUR)
fit.GBP<-estimate(GBP, p=1,q=1); arch.test(fit.GBP)
fit.CAD<-estimate(CAD, p=1,q=1); arch.test(fit.CAD)
```

```
fit.MZN<-estimate(MZN, p=1,q=1); arch.test(fit.MZN)
fit.KPW<-estimate(KPW, p=1,q=1); arch.test(fit.KPW)
##################################A directed Acyclic graph
theme_set(theme_dag())
dag<-dagify(SEK~AUD+NOK+BWP,  NOK~CAD+SEK+KPW,
                                        USD~AUD+NOK+SEK+MZN)
ggdag_exogenous(dag,node_size = 23,text_size = 3.88,edge_type = "arc",
                node = TRUE,stylized = TRUE,text = TRUE,label_size = text_size)
###########################################Plot of returns
data.ret<-cbind(BRL,BWP,ILS,INR,MWK,NOK,SEK,AUD,USD,EUR,GBP,CAD,MZN,KPW)
par(mfrow=c(4,4)); par(bg="white")
names.var<-c("BRL","BWP","ILS","INR","MWK","NOK","SEK","AUD",
          "USD","EUR","GBP","CAD","MZN","KPW")
linear = matrix(data.ret [1:30,], ncol=14)
             sapply(1:ncol(linear), function(i) {
                   acf(linear[,i], main=paste("", names.var[i]), lag.max=nrow(linear))})
###############################################Correlation
cor(data.ret)
 ##############################################Plotting prices
data.prices<-cbind(aud,eur,gbp,usd,bwp,ils,mzn,nok,kpw,brl,sek,mwk,cad,inr)
Prices<-data.frame(time=seq(as.Date('2011-06-07'),by='day',length=1820), data.prices)
ts_plot(Prices,title = "Prices of rand denominated currencies", Xtitle = "Days",
      Ytitle = "Prices",Xgrid =FALSE,Ygrid = FALSE) %>%
       layout(paper_bgcolor = "white", plot_bgcolor = "white",font = list(color = "black"),
       yaxis = list(linecolor = "#6b6b6b",  zerolinecolor = "#6b6b6b", gridcolor= "#444444"),
        xaxis = list(linecolor = "#6b6b6b", zerolinecolor = "#6b6b6b",gridcolor= "#444444"))
############################ Plots for mean reversion for GARCH(1,1)
#Data attachment and calling of required packages into memory
attach(rates);library(rugarch); require(forecast);require(zoo);
############################################### Creating varaibles
extret=abs(diff(log(mwk)))[1:1818]
direct1<-ifelse(extret[-1]>extret,1, 0)
```

```
exvaret<-cbind(extret,direct1)
usdret=diff(log(gbp))[2:1819]
#################################################### Estimating Garch
usdspec<-ugarchspec(variance.model = list(model = "sGARCH",submodel="NULL",
            garchOrder = c(1,1),external.regressors=exvaret,variance.targeting=TRUE),
          mean.model = list(armaOrder = c(0,0),include.mean = TRUE),distribution.model='jsu')
usdfit<-ugarchfit(usdspec,usdret, solver = 'hybrid',out.sample=440)
bboot <-ugarchforecast(usdfit, n.ahead=440, external.forecasts = list(vregfor = exvaret))
vol.dat<-data.frame(time=seq(as.Date('2011-06-07'),by='day',length=1378),sigma(usdfit))
par(mfrow=c(3,1))
plot.zoo(vol.dat[1:460,2],col=rainbow(1), ylab="Volatility",xlab="Days")
abline (a=matrix(sqrt(uncvariance(usdfit))), b=0,col="blue")
plot.zoo(vol.dat[461:920,2],col=rainbow(14), ylab="Volatility",xlab="Days")
abline (a=sqrt(uncvariance(usdfit)), b=0,col="blue")
plot.zoo(vol.dat[921:1378,2],col=rainbow(14), ylab="Volatility",xlab="Days")
abline (a=sqrt(uncvariance(usdfit)), b=0, col="blue")
a=ifelse(sigma(usdfit)==sqrt(uncvariance(usdfit)), 1,0)
sum(a)
 (1/length(sigma(usdfit)))*100
```

Appendix E2: Codes for estimating parameters of ARMAX-GARCH

```
#####################################SGARCH####################################
#################################### Calling and attachment
attach(rates);library(rugarch)
require(forecast);require(zoo);
require(stats);require(sandwich)
require(rms);require("aod")
library(robustbase)
##########################################Creeating regressors
nokret<-diff(log(nok))
nokextret1<-(diff(log(cad)))
nokextret2<-(diff(log(sek)))
```

```
nokextret3<-(diff(log(kpw)))
nokexvaret<-cbind(nokextret1,nokextret2,nokextret3)
######################################### Estimating parameters
nokspec<-ugarchspec(variance.model = list(model = "sGARCH",submodel="NULL", garchOrder = c(2,2), variance.targeting=TRUE), mean.model = list(armaOrder = c(6,2), include.mean = TRUE,external.regressors=nokexvaret), distribution.model='sged')
nokfit<-ugarchfit(nokspec,nokret, solver = 'hybrid',out.sample=0)
show(nokfit)
plot(nokfit,which = "all")
nokspec1<-ugarchspec(variance.model = list(model = "sGARCH",submodel="NULL",
      garchOrder = c(2, 2),variance.targeting=TRUE),mean.model = list(armaOrder = c(6, 2),
 include.mean = FALSE),distribution ='sged')
nokfit1 <-ugarchfit(nokspec1,nokret, solver = 'hybrid',out.sample=0)
show(nokfit1)
############################################In-sampling forecasts
uncvariance(nokfit);uncvariance(nokfit1)
halflife(nokfit);halflife(nokfit1)
persistence(nokfit);persistence(nokfit1)
infocriteria(nokfit);infocriteria(nokfit1)
###############################Calculating root mean absolute error
nokerror<-abs(nokfit@fit$residuals)-sigma(nokfit)
nokerror1<-abs(nokfit1@fit$residuals)-sigma(nokfit1)
mean(abs(nokerror));mean(abs(nokerror1))
sqrt(mean(nokerror^2));sqrt(mean(nokerror1^2))
nokerror<-nokret-(fitted(nokfit))
nokerror1<-nokret-(fitted(nokfit1))
mean(abs(nokerror));mean(abs(nokerror1))
sqrt(mean(nokerror^2));sqrt(mean(nokerror1^2))
nokret-(fitted(nokfit)+nokfit@fit$residuals)
#################################################### T tests
cor(nokerror,nokerror1)
cor(nokerror,nokerror1)
(t.test( abs(as.vector(nokerror1)),abs(as.vector(nokerror)),alternative =  "two.sided", mu=0,
```

```
      paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
(t.test(as.vector(nokerror1^2),as.vector(nokerror^2),alternative =  "two.sided", mu=0,
     paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
### Calculating explanatory power-R-square
summary(lmrob(nokret~fitted(nokfit)))$r.square
summary(lmrob(nokret~fitted(nokfit1)))$r.square
##################################SGARCH####################################
###########################################Creeating regressors
usdret<-diff(log(usd))
USDRET<-diff(log(usd))
extret1<-(diff(log(aud)))
extret2<-(diff(log(sek)))
extret3<-(diff(log(nok)))
extret4<-(diff(log(mzn)))
exvaret<-cbind(extret1,extret2,extret4,extret3)
#################################### Estimating GJRGarch parametrs
usdspec<-ugarchspec(variance.model = list(model = "gjrGARCH",
        garchOrder = c(1,1),variance.targeting=TRUE),
   mean.model = list(armaOrder = c(1,1),include.mean = TRUE,external.regressors=exvaret),
   distribution.model='sged')
usdfit<-ugarchfit(usdspec,usdret, solver = 'hybrid',out.sample=0)
show(usdfit)
plot(usdfit,which = "all")
usdspec1<-ugarchspec(variance.model = list(model = "gjrGARCH",
            garchOrder = c(1,1),variance.targeting=TRUE),
          mean.model = list(armaOrder = c(1,1),include.mean = FALSE),distribution ='sged')
usdfit1 <-ugarchfit(usdspec1,usdret, solver = 'hybrid',out.sample=0)
show(usdfit1)
############################################## In-sampling forecast
uncvariance(usdfit);uncvariance(usdfit1)
halflife(usdfit);halflife(usdfit1)
persistence(usdfit);persistence(usdfit1)
```

```
infocriteria(usdfit);infocriteria(usdfit1)
################################### calculating root mean absolute error
error<-abs(usdfit@fit$residuals)-sigma(usdfit)
error1<-abs(usdfit1@fit$residuals)-sigma(usdfit1)
mean(abs(error));mean(abs(error1))
sqrt(mean(error^2));sqrt(mean(error1^2))
cor(error,error1)
cor(error,error1)
############################################################ T tests
(t.test( abs(as.vector(error1)),abs(as.vector(error)),alternative =  "two.sided", mu=0,
         paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
(t.test(as.vector(error1^2),as.vector(error^2),alternative =  "two.sided", mu=0,
        paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
################################# Calculating explanatory power-R-square
summary(lmrob(abs(usdfit@fit$residuals)~sigma(usdfit)))$r.square
summary(lmrob(abs(usdfit1@fit$residuals)~sigma(usdfit1)))$r.square
##################################EGARCH################################
################################################ Creating regressors
sekret<-diff(log(sek))
sekRET<-diff(log(sek))
sekextret1<-(diff(log(aud)))
sekextret2<-(diff(log(nok)))
sekextret3<-(diff(log(bwp)))
sekexvaret<-cbind(sekextret1,sekextret2,sekextret3)
#################################### Estimating the parameters of EGarch
sekspec<-ugarchspec(variance.model = list(model = "eGARCH",garchOrder = c(2,2),
  variance.targeting=TRUE),  mean.model = list(armaOrder = c(5,5),
  include.mean = TRUE,external.regressors=sekexvaret),distribution.model='sged')
sekfit<-ugarchfit(sekspec,sekret, solver = 'hybrid',out.sample=0)
show(sekfit)
plot(sekfit,which = "all")
sekspec1<-ugarchspec(variance.model = list(model = "eGARCH",
```

```
           garchOrder = c(2, 2),variance.targeting=TRUE),
   mean.model = list(armaOrder = c(5, 5),include.mean = FALSE),distribution ='sged')
sekfit1 <-ugarchfit(sekspec1,sekret, solver = 'hybrid',out.sample=0)
show(sekfit1)
####################################################In-sampling forecast
############################################### calculating other statistics
uncvariance(sekfit);uncvariance(sekfit1)
halflife(sekfit);halflife(sekfit1)
persistence(sekfit);persistence(sekfit1)
infocriteria(sekfit);infocriteria(sekfit1)
####################################### calculating root mean absolute error
sekerror<-abs(sekfit@fit$residuals)-sigma(sekfit)
sekerror1<-abs(sekfit1@fit$residuals)-sigma(sekfit1)
mean(abs(sekerror));mean(abs(sekerror1))
sqrt(mean(sekerror^2));sqrt(mean(sekerror1^2))
##################
cor(sekerror,sekerror1)
cor(sekerror,sekerror1)
############################################################### T tests
(t.test( abs(as.vector(sekerror1)),abs(as.vector(sekerror)),alternative =  "two.sided", mu=0,
         paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
(t.test(as.vector(sekerror1^2),as.vector(sekerror^2),alternative =  "two.sided", mu=0,
        paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
################################3# Calculating explanatory power-R-square
summary(lmrob(abs(sekfit@fit$residuals)~sigma(sekfit)))$r.square
summary(lmrob(abs(sekfit1@fit$residuals)~sigma(sekfit1)))$r.square
##############################################Fitted Plots for all three models
### ARMAX models
par(bg = "white")
par(mfrow=c(3,1))
plot.zoo(x = cbind(nokret, nokfit@fit$fitted.values),  xlab="Time/horizon",cex.lab=1.5,
                ylab =  "returns",  main =list("Fitted returns from SGARCH",font = 1,cex=1.5),
```

```
                 col=c("blue","red"), lty=c(1,1), screens = 1)
legend(1700, 0.05,legend=c("returns","fitted"), bty = "n",lty = c(1,1),col=c("blue","red"),cex=1)
plot.zoo(x = cbind(usdret, usdfit@fit$fitted.values),   xlab="Time/horizon",cex.lab=1.5,
                 ylab ="returns", main =list("Fitted returns from gjrGARCH",font = 1,cex=1.5),
                 col=c("blue","orange"), lty=c(1,1), screens = 1)
legend(1700, 0.063, legend=c("returns","fitted"),  bty = "n", lty = c(1,1),
                 col=c("blue","orange"),cex=1)
plot.zoo(x = cbind(sekret, sekfit@fit$fitted.values),        xlab="Time/horizon",cex.lab=1.5,
                 ylab   = "returns",   main =list("Fitted returns from EGARCH",font = 1,cex=1.5),
                 col=c("blue","green"), lty=c(1,1), screens = 1)
legend(1700, 0.055,legend=c("returns","fitted"),lty = c(1,1), bty = "n",
                 col=c("blue","green"),cex=1)
###ARMA models
plot.zoo(x = cbind(nokret, nokfit1@fit$fitted.values),  xlab="Time/horizon",cex.lab=1.5,
                ylab = "returns",  main =list("Fitted returns from SGARCH",font = 1,cex=1.5),
                col=c("blue","red"), lty=c(1,1), screens = 1)
legend(1700, 0.05,legend=c("returns","fitted"), bty = "n",lty = c(1,1),col=c("blue","red"),cex=1)
plot.zoo(x = cbind(usdret, usdfit1@fit$fitted.values),   xlab="Time/horizon",cex.lab=1.5, ylab =
               "returns", main =list("Fitted returns from gjrGARCH",font = 1,cex=1.5),
                col=c("blue","or      ange"), lty=c(1,1), screens = 1)
legend(1700, 0.063,legend=c("returns","fitted"), bty = "n", lty = c(1,1),
               col=c("blue","or ange"),cex=1)
plot.zoo(x = cbind(sekret, sekfit1@fit$fitted.values),        xlab="Time/horizon",cex.lab=1.5,
               ylab = "returns",   main =list("Fitted returns from EGARCH",font = 1,cex=1.5),
               col=c("blue","green"), lty=c(1,1), screens = 1)
legend(1700, 0.045,legend=c("returns","fitted"),lty = c(1,1), bty = "n",
               col=c("blue","green"),cex=1)
################################################ Likelihood ratio tests
LRSGARCH1=2*(likelihood(nokfit)-likelihood(nokfit1))
LRSGARCH; 1-pchisq(LRSGARCH,df=3)
LRGJRGARCH=2*(likelihood(sekfit)-likelihood(sekfit1))
LRGJRGARCH; 1-pchisq(LRGJRGARCH,df=4)
```

```
LREGARCH=2*(likelihood(usdfit)-likelihood(usdfit1))
LREGARCH; 1-pchisq(LREGARCH,df=3)
################################################### Adjusted R-squares
rsq1<-1-sum((fitted(nokfit)-nokret)^2)/sum((nokret-mean(nokret))^2)
rsq11<-1-sum((fitted(nokfit1)-nokret)^2)/sum((nokret-mean(nokret))^2)
adjnok1<-1-((1-rsq1)*(1819-1))/(1819-12-1)
adjnok2<-1-((1-rsq11)*(1819-1))/(1819-9-1)
##############################################################
rsq2<-1-sum((fitted(usdfit)-usdret)^2)/sum((usdret-mean(usdret))^2)
rsq21<-1-sum((fitted(usdfit1)-usdret)^2)/sum((usdret-mean(usdret))^2)
adjusd1<-1-((1-rsq2)*(1819-1))/(1819-6-1)
adjusd2<-1-((1-rsq21)*(1819-1))/(1819-2-1)
##############################################################
rsq3<-1-sum((fitted(sekfit)-sekret)^2)/sum((sekret-mean(sekret))^2)
rsq31<-1-sum((fitted(sekfit1)-sekret)^2)/sum((sekret-mean(sekret))^2)
adjsek1<-1-((1-rsq3)*(1819-1))/(1819-13-1)
adjsek2<-1-((1-rsq31)*(1819-1))/(1819-10-1)
##############################################################
adjnok1; adjnok2
adjusd1;adjusd2
adjsek1;adjsek2
################################# % of correctly predicted directions
summ<-(ifelse(sign(fitted(nokfit1))==sign(nokret),1, 0))
sum(as.matrix(summ))
(1381)/(1381+438)*100
(1079)/(1381+438)*100
summ<-(ifelse(sign(fitted(usdfit1))==sign(usdret),1, 0))
sum(as.matrix(summ))
(1334)/(1819)*100
(1029)/(1819)*100
summ<-(ifelse(sign(fitted(sekfit1))==sign(sekret),1, 0))
sum(as.matrix(summ))
```

```
(1451)/(1819)*100
(1092)/(1819)*100
summ<-(ifelse(sign(fitted(nokfit1))==sign(nokret)&sign(fitted(nokfit1))+sign(nokret)==-2,1, 0))
sum(as.matrix(summ))
```

Appendix E3: Codes for mutual entropy estimation

```
#########################################Packages and data attachment
attach(rates)
require(rugarch);require(zoo)
require(VGAM);require(mpmi)
require(Hmisc);require( pracma)
require(tidyverse)
################################################# Preparing Variables
MWK<-(diff(log(rates$usd)))
BWP<-(diff(log(rates$eur)))
BRL<-(diff(log(rates$gbp)))
ILS<-(diff(log(rates$aud)))
SEK<-(diff(log(rates$cad)))
INR<-(diff(log(rates$inr)))
NOK<-(diff(log(rates$nok)))
MWKR<-movavg(abs(diff(log(rates$mwk))), 1818, type="s") [1:1818]
BWPR<-movavg(abs(diff(log(rates$bwp))), 1818, type="s") [1:1818]
BRLR<-movavg(abs(diff(log(rates$brl))), 1818, type="s") [1:1818]
ILSR<-movavg(abs(diff(log(rates$ils))), 1818, type="s") [1:1818]
SEKR<-movavg(abs(diff(log(rates$sek))), 1818, type="s") [1:1818]
INRR<-movavg(abs(diff(log(rates$inr))), 1818, type="s") [1:1818]
NOKR<-movavg(abs(diff(log(rates$nok))), 1818, type="s") [1:1818]
################################### Computing Simple Moving Averages
MWKB<-movavg( ifelse(abs(MWK[1:1819][-1])>abs(MWK[1:1818]),1, 0), 1817,
        type="s")[1:1818]
BWPB<-movavg( ifelse(abs(BWP[1:1819][-1])>abs(BWP[1:1818]),1, 0), 1817,
       type="s")[1:1818]
```

```
BRLB<-movavg( ifelse(abs(BRL[1:1819][-1])>abs(BRL[1:1818]),1, 0), 1817,
          type="s") [1:1818]
ILSB<-movavg( ifelse(abs(ILS[1:1819][-1])>abs(ILS[1:1818]),1, 0), 1817, type="s") [1:1818]
SEKB<-movavg( ifelse(abs(SEK[1:1819][-1])>abs(SEK[1:1818]),1, 0), 1817, type="s")[1:1818]
INRB<-movavg( ifelse(abs(INR[1:1819][-1])>abs(INR[1:1818]),1, 0), 1817, type="s") [1:1818]
NOKB<-movavg( ifelse(abs(NOK[1:1819][-1])>abs(NOK[1:1818]),1, 0),1817,
          type="s") [1:1818]
##########################################################Computing Volatility
Returns<-cbind(MWK,BWP,BRL,ILS,SEK)
mwkspec<-ugarchspec(variance.model = list(model = "gjrGARCH", garchOrder = c(1, 1),
           variance.targeting=TRUE), mean.model = list(armaOrder = c(5, 5),
           include.mean = TRUE) , distribution ='jsu')
bwpspec<-ugarchspec(variance.model = list(model = "fGARCH",submodel="NAGARCH",
           garchOrder = c(3,0),variance.targeting=TRUE), mean.model = list(armaOrder = c(2,1),
           include.mean = TRUE),distribution.model='jsu')
brlspec<-usdspec<-ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(2,1),
           variance.targeting=TRUE), mean.model = list(armaOrder = c(5,4),
            include.mean = TRUE), distribution.model='jsu')
ilsspec<-ugarchspec(variance.model = list(model = "fGARCH",submodel="TGARCH",
            garchOrder = c(2,2),variance.targeting=TRUE),
            mean.model = list(armaOrder = c(2,1),include.mean = TRUE),distribution ='nig')
sekspec<-ugarchspec(variance.model = list(model = "eGARCH", garchOrder = c(1, 1),
            variance.targeting=FALSE), mean.model = list(armaOrder = c(1, 1),
            archm = FALSE,include.mean = TRUE),distribution ='sged')
mspec = multispec(c(mwkspec,bwpspec,brlspec,ilsspec,sekspec))
fitlist = multifit(multispec = mspec, data =Returns)
#show(fitlist)
Mwkv<-sigma(fitlist)[,1][2:1819]
Bwpv<-sigma(fitlist)[,2][2:1819]
Brlv<-sigma(fitlist)[,3][2:1819]
Ilsv<-sigma(fitlist)[,4][2:1819]
Sekv<-sigma(fitlist)[,5][2:1819]
```

```
####################################################Binding Volatility data
names.var<-
c("Mwkv","Bwpv","Brlv","Ilsv","Sekv","MWKR","BWPR","BRLR","ILSR","SEKR","INRR","NOKR","MWKB","BWPB","BRLB","ILSB","SEKB","INRB","NOKB")
data.mit<-setNames(data.frame(Mwkv,Bwpv,Brlv,Ilsv,Sekv,MWKR,BWPR,
BRLR,ILSR,SEKR,INRR,NOKR,MWKB,BWPB,BRLB,ILSB,SEKB,INRB,NOKB),names.var)
#######################################################Entropy estimation
############################computing MIM (mutual information matrix)
all.bcmi<-matrix(data=round(cmi(data.mit)$bcmi,3),
        nrow=19,ncol=19,byrow=TRUE,dimnames=list(names.var,names.var))
all.mi<-matrix(data=round(cmi(data.mit)$mi,3),
                nrow=19,ncol=19,byrow=TRUE,dimnames=list(names.var,names.var))
round((all.bcmi[1:5,6:19]/c(all.bcmi[1,1],all.bcmi[2,2],all.bcmi[3,3]
                          ,all.bcmi[4,4],all.bcmi[5,5])),4)
round(all.bcmi[1:5,6:19],3)
###########################################computing Normalized mutual
round(all.bcmi[1,6:19]/(sqrt(diag(all.bcmi)*all.bcmi[1,1]))[6:19],3)
round(all.bcmi[2,6:19]/(sqrt(diag(all.bcmi)*all.bcmi[2,2]))[6:19],3)
round(all.bcmi[3,6:19]/(sqrt(diag(all.bcmi)*all.bcmi[3,3]))[6:19],3)
round(all.bcmi[4,6:19]/(sqrt(diag(all.bcmi)*all.bcmi[4,4]))[6:19],3)
round(all.bcmi[5,6:19]/(sqrt(diag(all.bcmi)*all.bcmi[5,5]))[6:19],3
###################################computing Correlation matrix and t-test
c=rcorr(as.matrix(cbind(data.mit[,1:5],data.mit[,6:19])))
round(c$P[1:5,6:19],3)
round(c$z[1:5,6:19],3)
s=round((all.bcmi[1:5,6:19]),3)
rc=round(c$r[1:5,6:19],3)
#####################Preparing data for correlation vers mutual information plotting
a.data=cbind(correlation=rc[1,][-1][-7],mutualinfo=s[1,][-1][-7])
b.data=cbind(correlation=rc[2,][-2][-8],mutualinfo=s[2,][-2][-8])
c.data=cbind(correlation=rc[3,][-3][-9],mutualinfo=s[3,][-3][-9])
d.data=cbind(correlation=rc[4,][-4][-10],mutualinfo=s[4,][-4][-10])
e.data=cbind(correlation=rc[5,][-5][-11],mutualinfo=s[5,][-5][-11])
```

```
new.data=rbind(a.data,b.data,c.data,d.data,e.data)
group=rep(c("Volatility for MWK/ZAR","Volatility for BWP/ZAR",
          "Volatility for BRL/ZAR","Volatility for ILS/ZAR",
          "Volatility for SEK/ZAR"),12)
dat.f=data.frame(new.data,group)
dat.f=dat.f[order(dat.f$correlation),]
##################################Correlation vers mutual information plotting
 #ggplot(data = dat.f$, mapping = aes(x = correlation, y = mutualinfo,color=group)) +
 # geom_line()
ggplot(data = dat.f, mapping = aes(x =correlation, y = mutualinfo,color=group)) +
  geom_line() + facet_wrap(facets = vars(group))
plot.zoo(dat.f[,1],dat.f[,2])
round(abs(all.bcmi[1:5,6:19]-all.mi[1:5,6:19])/all.mi[1:5,6:19],2)
plot.zoo(as.vector(all.bcmi[1:5,6:19]), as.vector(all.mi[1:5,6:19]), type="l")
mp(all.bcmi)
```

Appendix E4: Codes for volatility forecasting

```
######################################SGARCH model#############################
#################Data attachement and calling of required packages into memory
attach(rates);library(rugarch)
require(forecast);require(zoo);
require(stats);require(sandwich)
require(rms);require("aod")
library(robustbase)
######################################################Creeating varaibles
usdret=diff(log(gbp))[2:1819]
USDRET<-usdret[1:1818]
extret=abs(diff(log(mwk)))[1:1818]
direct1<-ifelse(extret[-1]>extret,1, 0)
exvaret<-cbind(extret,direct1)
############################## Estimating parameters of Augmented SGARCH
usdspec<-ugarchspec(variance.model = list(model = "sGARCH",submodel="NULL",
```

```
                  garchOrder = c(1,1)), mean.model = list(armaOrder = c(0,0),
                    include.mean = TRUE),distribution.model='jsu')
usdfit<-ugarchfit(usdspec,usdret, solver = 'hybrid',out.sample=440)
 show(usdfit)
 plot(usdfit,which = "all")
######################### Estimating parameters of non-Augmented SGARCH
usdspec1<-ugarchspec(variance.model = list(model = "sGARCH",submodel="NULL",
          garchOrder = c(1, 1),variance.targeting=TRUE),
          mean.model = list(armaOrder = c(0,0),include.mean = TRUE),distribution ='jsu')
usdfit1 <-ugarchfit(usdspec1,usdret, solver = 'hybrid',out.sample=440)
          show(usdfit1)
########################################################Model statistics
################################################# Miscellaneous Statistics
uncvariance(usdfit);uncvariance(usdfit1)
halflife(usdfit);halflife(usdfit1)
persistence(usdfit);persistence(usdfit1)
infocriteria(usdfit);infocriteria(usdfit1)
#################################### Calculating root mean and absolute error
error<-abs(usdfit@fit$residuals)-sigma(usdfit)
error1<-abs(usdfit1@fit$residuals)-sigma(usdfit1)
mean(abs(error));mean(abs(error1))
sqrt(mean(error^2));sqrt(mean(error1^2))
###################################################T-tests for predicted errors
cor(error,error1); cor(error,error1)
(t.test( abs(as.vector(error)),abs(as.vector(error1)),alternative =  "greater", mu=0,
         paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
(t.test(as.vector(error^2),as.vector(error1^2),alternative =  "greater", mu=0,
        paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
################################################### Minzer-Zarnowit R-square
summary(lmrob(abs(usdfit@fit$residuals)~sigma(usdfit)))[13]
summary(lmrob(abs(usdfit1@fit$residuals)~sigma(usdfit1)))[13]
```

```
###Forecasting (following horizons were used: n=1, n=180 and n=440, to get the forecasts for
#other horizons, assign the approapriate horizon to the nuum variable)
numm=440; j=0
bboot = ugarchforecast(usdfit, n.ahead=numm,n.roll=j,out.sample=0, external.forecasts =
list(vregfor = exvaret))
bboot1 = ugarchforecast(usdfit1, n.ahead=numm,n.roll=j,out.sample=0)
plot(bboot,which = 3)
############################################################ Forecast Statistics
xm<-mean(USDRET[1379:(1378+numm)])
forecasty=USDRET[1379:(1378+numm)]-xm
Error<-abs(forecasty)-sigma(bboot)[,j+1]
Error1<-abs(forecasty)-sigma(bboot1)[,j+1]
mean(abs(Error))*100000;mean(abs(Error1))*100000
sqrt(mean(Error^2))*100000;sqrt(mean(Error1^2))*100000
####################################################### T-test for forecasted Errors
cor(sqrt(Error1^2),sqrt(abs(Error^2)))
(t.test( abs(Error),abs(Error1),alternative =  "two.sided", mu=0,
         paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
(t.test(Error^2,Error1^2,alternative =  "two.sided", mu=0,
        paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
######################################################## Minzer-Zarnowit R-square
us.lm<-lmrob(abs(forecasty)~sigma(bboot)[,j+1])
us.lm1<-lmrob(abs(forecasty)~sigma(bboot1)[,j+1])
summary(us.lm)[12];summary(us.lm1)[12]
mean(abs(us.lm$residuals))*100000;mean(abs(us.lm1$residuals))*100000
sqrt(mean(us.lm$residuals^2))*100000;sqrt(mean(us.lm1$residuals^2))*100000
######################################################################## T-test
(t.test( abs(us.lm$residuals),abs(us.lm1$residuals),alternative =  "two.sided", mu=0,
         paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
(t.test(us.lm$residuals^2,us.lm1$residuals^2,alternative =  "two.sided", mu=0,
        paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
#################################################################### Wald test
wald.test(vcov(us.lm), coefficients(us.lm),Terms = 2, L = NULL, H0 = 1,
```

```
        df = df.residual(us.lm), verbose = FALSE)
wald.test(vcov(us.lm1), coefficients(us.lm1),Terms = 2, L = NULL, H0 = 1,
        df =df.residual(us.lm1), verbose =FALSE)
##################################################################GMMT test
z = residuals(usdfit)/sigma(usdfit)
skew = dskewness("jsu",skew = coef(usdfit)["skew"], shape= coef(usdfit)["shape"])
# add back 3 since dkurtosis returns the excess kurtosis
kurt = 3+dkurtosis("jsu",skew = coef(usdfit)["skew"], shape= coef(usdfit)["shape"])
z1 = residuals(usdfit1)/sigma(usdfit1)
skew1 = dskewness("jsu",skew = coef(usdfit1)["skew"], shape= coef(usdfit1)["shape"])
# add back 3 since dkurtosis returns the excess kurtosis
kurt1 = 3+dkurtosis("jsu",skew = coef(usdfit1)["skew"], shape= coef(usdfit1)["shape"])
print(GMMTest(z, lags = 3, skew=skew, kurt=kurt))
print(GMMTest(z1, lags = 2, skew=skew1, kurt=kurt1))
##########################################################Unconditional var plots
par(mfrow=c(1,1))
sgfor<-xts(x=sigma(bboot), order.by=seq(as.Date("2015-04-14"),length=440, by="days"))
sgfor1<-xts(x=sigma(bboot1), order.by=seq(as.Date("2015-04-14"),length=440, by="days"))
combsgfor<-cbind(sgfor,sgfor1)
plot.zoo(cbind(sigma(bboot),uv1,abs(USDRET[1379:(1378+numm)])),
        col=c("blue","red"), cex.main=1,xlab=expression(Days), ylab="sigma",
  main="Upper column is for Augmented model while lower column is for Non-augmented
model")
##################################NAGARCH model##########################
#########################################################Calling and attachment
attach(rates);library(rugarch)
require(forecast);require(zoo);
require(stats);require(sandwich)
require(rms);require("aod")
library(robustbase)
############################################################# Creating regressors
usdret<-diff(log(bwp))[2:1819]
```

```
USDRET<-usdret[1:1818]
extret<-abs(diff(log(inr)))[1:1818]
direct1<-ifelse(extret[-1]>extret, 1, 0)
exvaret<-cbind(extret,direct1)
######################################## Estimating parameters of Augmented NAGARCH
usdspec<-ugarchspec(variance.model = list(model = "fGARCH",submodel="NAGARCH",
    garchOrder = c(3,0),external.regressors=exvaret,variance.targeting=TRUE),
    mean.model = list(armaOrder = c(2,1),include.mean = TRUE),distribution.model='jsu')
usdfit<-ugarchfit(usdspec,usdret, solver = 'hybrid',out.sample=440)
    show(usdfit)
    plot(usdfit,which = "all")
#################################### Estimating parameters of non-Augmented NAGARCH
usdspec1<-ugarchspec(variance.model = list(model = "fGARCH",submodel="NAGARCH",
     garchOrder = c(3, 0),variance.targeting=TRUE),
     mean.model = list(armaOrder = c(2, 1),include.mean = TRUE),distribution ='jsu')
usdfit1 <-ugarchfit(usdspec1,usdret, solver = 'hybrid',out.sample=440)
     show(usdfit1)
#######################################################################Model statistics
############################################################## Miscellaneous Statistics
uncvariance(usdfit);uncvariance(usdfit1)
halflife(usdfit);halflife(usdfit1)
persistence(usdfit);persistence(usdfit1)
infocriteria(usdfit);infocriteria(usdfit1)
################################################ Calculating root mean and absolute error
error<-(usdfit@fit$residuals)^2-sigma(usdfit)^2
error1<-(usdfit1@fit$residuals)^2-sigma(usdfit1)^2
mean(abs(error));mean(abs(error1))
(mean(error^2));(mean(error1^2))
############################################################# T-tests for predicted errors
 cor(error,error1)
cor(error,error1)
(t.test( abs(as.vector(error1)),abs(as.vector(error)),alternative =  "two.sided", mu=0,
```

```
                paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
(t.test(as.vector(error1^2),as.vector(error^2),alternative =  "two.sided", mu=0,
               paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
######################################################3 Minzer-Zarnowit R-square
summary(lmrob(usdfit@fit$residuals^2~sigma(usdfit)))$r.square
summary(lmrob(usdfit1@fit$residuals^2~sigma(usdfit1)))$r.square
###Forecasting (following horizons were used: n=1, n=180 and n=440, to get the forecasts for
#other horizons, assign the approapriate horizon to the nuum variable)
par(mfrow=c(1,1))
numm=440
j=0
bboot = ugarchforecast(usdfit, n.ahead=numm,n.roll=j,out.sample=0, external.forecasts =
list(vregfor = exvaret))
bboot1 = ugarchforecast(usdfit1, n.ahead=numm,n.roll=j,out.sample=0)
plot(bboot,which = 3)
################################################################# Forecast Statistics
xm<-mean(USDRET[1379:(1378+numm)])
forecasty=USDRET[1379:(1378+numm)]-xm
Error<-abs(forecasty)-sigma(bboot)[,j+1]
Error1<-abs(forecasty)-sigma(bboot1)[,j+1]
mean(abs(Error))*100000;mean(abs(Error1))*100000
sqrt(mean(Error^2))*100000;sqrt(mean(Error1^2))*100000
########################################################## T-test for forecasted Errors
cor(sqrt(Error1^2),sqrt(abs(Error^2)))
(t.test( abs(Error),abs(Error1),alternative =  "two.sided", mu=0,
         paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
(t.test(Error^2,Error1^2,alternative =  "two.sided", mu=0,
        paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
########################################################## Minzer-Zarnowit R-square
us.lm<-lmrob(abs(forecasty)~sigma(bboot)[,j+1])
us.lm1<-lmrob(abs(forecasty)~sigma(bboot1)[,j+1])
summary(us.lm)[12];summary(us.lm1)[12]
mean(abs(us.lm$residuals))*100000;mean(abs(us.lm1$residuals))*100000
```

```
sqrt(mean(us.lm$residuals^2))*100000;sqrt(mean(us.lm1$residuals^2))*100000
###################################################################### T -test
(t.test( abs(us.lm$residuals),abs(us.lm1$residuals),alternative =  "two.sided", mu=0,
        paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
(t.test(us.lm$residuals^2,us.lm1$residuals^2,alternative =  "two.sided", mu=0,
       paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
##################################################################### Wald test
wald.test(vcov(us.lm), coefficients(us.lm),Terms = 2, L = NULL, H0 = 1,
         df = df.residual(us.lm), verbose = FALSE)
wald.test(vcov(us.lm1), coefficients(us.lm1),Terms = 2, L = NULL, H0 = 1,
         df =df.residual(us.lm1), verbose =FALSE)
################################################################### GMMT test
z = residuals(usdfit)/sigma(usdfit)
skew = dskewness("jsu",skew = coef(usdfit)["skew"], shape= coef(usdfit)["shape"])
# add back 3 since dkurtosis returns the excess kurtosis
kurt = 3+dkurtosis("jsu",skew = coef(usdfit)["skew"], shape= coef(usdfit)["shape"])
z1 = residuals(usdfit1)/sigma(usdfit1)
skew1 = dskewness("jsu",skew = coef(usdfit1)["skew"], shape= coef(usdfit1)["shape"])
# add back 3 since dkurtosis returns the excess kurtosis
kurt1 = 3+dkurtosis("jsu",skew = coef(usdfit1)["skew"], shape= coef(usdfit1)["shape"])
print(GMMTest(z, lags = 2, skew=skew, kurt=kurt))
print(GMMTest(z1, lags = 2, skew=skew1, kurt=kurt1))
#########################################################Unconditional var plots
par(mfrow=c(1,2))
k<-c(1:365)
omega<-coef(usdfit)[length(coef(usdfit))]
OMEGA<-coef(usdfit1)[length(coef(usdfit1))]
P<-persistence(usdfit)
P1<-persistence(usdfit1)
uv<-sigma(bboot)-(omega/(1-P)*(1-P^(k-1)))
uv1<-sigma(bboot1)-(OMEGA/(1-P1)*(1-P1^(k-1)))
plot.zoo(x = cbind(uv,uv1),
```

```
        xlab="Time/horizon", ylab = "sigma",
        main =list("Unconditional forecasted sigma (NAGARCH)",font = 1,cex=0.9),
        col=c(2,3), lty=c(1,5), screens = 1)
lines(x = sigma(usdfit),
        xlab="Time/horizon", ylab = "sigma",
        main =list("Unconditional forecasted sigma (NAGARCH)",font = 1,cex=0.9),
        col=c(5,3), lty=c(1,5), screens = 1)
legend(200, 0.0083,legend=c("sigma 1","sigma 2"),lty = c(1,5),col=c(2,3),cex=0.7)
############################################################ forecasted sigma plot
plot.zoo(x = cbind(abs(forecasty),sigma(bboot),sigma(bboot1)),
        xlab="Time/horizon", ylab = "sigma",
        main=list("|returns| versus Forecasted sigma (NAGARCH)",font = 1,
                cex=0.9),
        col=c(8,2,3), lty=c(5,1,4),lwd=c(1,1,2), screens = 1)
legend(1, 0.059,legend=c("|returns|","sigma 1", "sigma 2"),lty = c(5,1,4),col=c(8,2,3),cex=0.7)
########################################## TGARCH model #####################
#############################################################Calling and attachment
attach(rates);library(rugarch)
require(forecast);require(zoo);
require(stats);require(sandwich)
require(rms);require("aod")
library(robustbase)
############################################################## Creating regressors
usdret=diff(log(ils))[2:1819]
USDRET<-usdret[1:1818]
extret=abs(diff(log(inr)))[1:1818]
direct1<-ifelse(extret[-1]>extret, 1, 0)
exvaret<-cbind(direct1)
##################### Estimating parameters of Augmented TGARCH#################
usdspec<-ugarchspec(variance.model = list(model = "fGARCH",submodel="TGARCH",
       garchOrder = c(2,2),external.regressors=exvaret,variance.targeting=TRUE),
       mean.model = list(armaOrder = c(1,1),include.mean = TRUE),distribution.model='nig')
```

```
usdfit<-ugarchfit(usdspec,usdret, solver = 'hybrid',out.sample=440)
        show(usdfit)
        plot(usdfit,which = "all")
        plot.zoo(usdfit@fit$residuals)
################ Estimating parameters of non-Augmented TGARCH#################
usdspec1<-ugarchspec(variance.model = list(model = "fGARCH",submodel="TGARCH",
          garchOrder = c(2,2),variance.targeting=TRUE),
          mean.model = list(armaOrder = c(2,1),include.mean = TRUE),distribution ='nig')
usdfit1 <-ugarchfit(usdspec1,usdret, solver = 'hybrid',out.sample=440)
          show(usdfit1)
############################################################### Model statistics
######################################################## Miscellaneous Statistics
uncvariance(usdfit);uncvariance(usdfit1)
halflife(usdfit);halflife(usdfit1)
persistence(usdfit);persistence(usdfit1)
infocriteria(usdfit);infocriteria(usdfit1)
########################################## Calculating root mean and absolute error
error<-(usdfit@fit$residuals)^2-sigma(usdfit)^2
error1<-(usdfit1@fit$residuals)^2-sigma(usdfit1)^2
mean(abs(error));mean(abs(error1))
mean(abs(error^2));mean(abs(error1^2))
###################################################### T-tests for predicted errors
cor(error,error1)
cor(error,error1)
(t.test( abs(as.vector(error1)),abs(as.vector(error)),alternative =  "two.sided", mu=0,
         paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
(t.test(as.vector(error1^2),as.vector(error^2),alternative =  "two.sided", mu=0,
        paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
###################################################### Minzer-Zarnowit R-square
summary(lmrob(abs(usdfit@fit$residuals)~sigma(usdfit)))[13]
summary(lmrob(abs(usdfit1@fit$residuals)~sigma(usdfit1)))[13]
```

```
###Forecasting (following horizons were used: n=1, n=180 and n=440, to get the forecasts for
#other horizons, assign the approapriate horizon to the nuum variable)
par(mfrow=c(1,1))
numm=440
j=0
bboot = ugarchforecast(usdfit, n.ahead=numm,n.roll=j,out.sample=0, external.forecasts =
list(vregfor = exvaret))
bboot1 = ugarchforecast(usdfit1, n.ahead=numm,n.roll=j,out.sample=0)
plot(bboot,which = 3)
############################################################ Forecast Statistics
xm<-mean(USDRET[1379:(1378+numm)])
forecasty=USDRET[1379:(1378+numm)]-xm
Error<-abs(forecasty)-sigma(bboot)[,j+1]
Error1<-abs(forecasty)-sigma(bboot1)[,j+1]
mean(abs(Error))*100000;mean(abs(Error1))*100000
sqrt(mean(Error^2))*100000;sqrt(mean(Error1^2))*100000
#cbind(as.matrix((abs(Error/abs(forecasty)))*100),
    #as.matrix((abs(Error1/abs(forecasty)))*100))
######################################################## T-test for forecasted Errors
cor(abs(Error1),abs(Error))
cor(sqrt(Error1^2),sqrt(abs(Error^2)))
(t.test( abs(Error1),abs(Error),alternative =  "two.sided", mu=0,
        paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
(t.test(Error1^2,Error^2,alternative =  "two.sided", mu=0,
       paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
#########################################################Minzer-Zarnowit R-square
us.lm<-lmrob(abs(forecasty)~sigma(bboot)[,j+1])
us.lm1<-lmrob(abs(forecasty)~sigma(bboot1)[,j+1])
summary(us.lm)[12];summary(us.lm1)[12]
mean(abs(us.lm$residuals))*100000;mean(abs(us.lm1$residuals))*100000
sqrt(mean(us.lm$residuals^2))*100000;sqrt(mean(us.lm1$residuals^2))*100000
###################################################################### Wald test
wald.test(vcov(us.lm), coefficients(us.lm),Terms = 2, L = NULL, H0 = 1,
```

```
        df = df.residual(us.lm), verbose = FALSE)
wald.test(vcov(us.lm1), coefficients(us.lm1),Terms = 2, L = NULL, H0 = 1,
        df =df.residual(us.lm1), verbose =FALSE)
####################################################################GMMT test
z = residuals(usdfit)/sigma(usdfit)
skew = dskewness("nig",skew = coef(usdfit)["skew"], shape= coef(usdfit)["shape"])
# add back 3 since dkurtosis returns the excess kurtosis
kurt = 3+dkurtosis("nig",skew = coef(usdfit)["skew"], shape= coef(usdfit)["shape"])
z1 = residuals(usdfit1)/sigma(usdfit1)
skew1 = dskewness("nig",skew = coef(usdfit1)["skew"], shape= coef(usdfit1)["shape"])
# add back 3 since dkurtosis returns the excess kurtosis
kurt1 = 3+dkurtosis("nig",skew = coef(usdfit1)["skew"], shape= coef(usdfit1)["shape"])
print(GMMTest(z, lags = 2, skew=skew, kurt=kurt))
print(GMMTest(z1, lags = 2, skew=skew1, kurt=kurt1))
############################################################Unconditional var plots
par(mfrow=c(1,1))
k<-c(1:365)
omega<-coef(usdfit)[length(coef(usdfit))]
OMEGA<-coef(usdfit1)[length(coef(usdfit1))]
P<-persistence(usdfit)
P1<-persistence(usdfit1)
uv<-sigma(bboot)-(omega/(1-P)*(1-P^(k-1)))
uv1<-sigma(bboot1)-(OMEGA/(1-P1)*(1-P1^(k-1)))
plot.zoo(x = cbind(uv,uv1),
        xlab="Time/horizon", ylab = "sigma",
 main =list("Unconditional forecasted sigma(TGARCH)",font = 1,cex=0.9),
 col=c(2,3), lty=c(1,5), screens = 1)
legend(100, 0.003,legend=c("sigma 1","sigma 2"),lty = c(1,5),col=c(2,3),cex=0.7)
#############################################################forecasted sigma plot
plot.zoo(x = cbind(abs(forecasty),sigma(bboot),sigma(bboot1)),
        xlab="Time/horizon", ylab = "sigma",
        main=list("|returns| versus Forecasted sigma (TGARCH)",font = 1,
```

```
            cex=0.9),
      col=c(8,2,3), lty=c(5,1,4),lwd=c(1,1,2), screens = 1)
legend(200, 0.06,legend=c("|returns|","sigma 1", "sigma 2"),lty = c(5,1,4),col=c(8,2,3),cex=0.7)
##################################GJRGARCHmodel ########################
##########################################################Calling and attachment
attach(rates);library(rugarch)
require(forecast);require(zoo);
require(stats);require(sandwich)
require(rms);require("aod")
library(robustbase)
############################################################## Creating regressors
usdret=diff(log(ils))[2:1819]
USDRET<-usdret[1:1818]
extret=abs(diff(log(inr)))[1:1818]
direct1<-ifelse(extret[-1]>extret, 1, 0)
exvaret<-cbind(direct1)
################## Estimating parameters of Augmented GJRGARCH#################
usdspec<-ugarchspec(variance.model = list(model = "fGARCH",submodel="TGARCH",
     garchOrder = c(2,2),external.regressors=exvaret,variance.targeting=TRUE),
     mean.model = list(armaOrder = c(1,1),include.mean = TRUE),distribution.model='nig')
usdfit<-ugarchfit(usdspec,usdret, solver = 'hybrid',out.sample=440)
     show(usdfit)
     plot(usdfit,which = "all")
############ Estimating parameters of non-Augmented GJRGARCH#################
usdspec1<-ugarchspec(variance.model = list(model = "fGARCH",submodel="TGARCH",
      garchOrder = c(2,2),variance.targeting=TRUE),
      mean.model = list(armaOrder = c(2,1),include.mean = TRUE),distribution ='nig')
usdfit1 <-ugarchfit(usdspec1,usdret, solver = 'hybrid',out.sample=440)
      show(usdfit1)
      plot(usdfit1,which = 9)
################################################################Model statistics
########################################################## Miscellaneous Statistics
```

```
uncvariance(usdfit);uncvariance(usdfit1)
halflife(usdfit);halflife(usdfit1)
persistence(usdfit);persistence(usdfit1)
infocriteria(usdfit);infocriteria(usdfit1)
############################################## Calculating root mean and absolute error
error<-(usdfit@fit$residuals)^2-sigma(usdfit)^2
error1<-(usdfit1@fit$residuals)^2-sigma(usdfit1)^2
mean(abs(error));mean(abs(error1))
(mean(error^2));(mean(error1^2))
########################################################## T-tests for predicted errors
 cor(error,error1)
cor(error,error1)
(t.test( abs(as.vector(error1)),abs(as.vector(error)),alternative =  "two.sided", mu=0,
                paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
(t.test(as.vector(error1^2),as.vector(error^2),alternative =  "two.sided", mu=0,
               paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
########################################################### Minzer-Zarnowit R-square
summary(lmrob(usdfit@fit$residuals^2~sigma(usdfit)))$r.square
summary(lmrob(usdfit1@fit$residuals^2~sigma(usdfit1)))$r.square
###Forecasting (following horizons were used: n=1, n=180 and n=440, to get the forecasts for
#other horizons, assign the approapriate horizon to the nuum variable)
par(mfrow=c(1,1))
numm=440
j=0
bboot = ugarchforecast(usdfit, n.ahead=numm,n.roll=j,out.sample=0, external.forecasts =
list(vregfor = exvaret))
bboot1 = ugarchforecast(usdfit1, n.ahead=numm,n.roll=j,out.sample=0)
plot(bboot,which = 3)
################################################################ Forecast Statistics
xm<-mean(USDRET[1379:(1378+numm)])
forecasty=USDRET[1379:(1378+numm)]-xm
Error<-abs(forecasty)-sigma(bboot)[,j+1]
Error1<-abs(forecasty)-sigma(bboot1)[,j+1]
```

```
mean(abs(Error))*100000;mean(abs(Error1))*100000
sqrt(mean(Error^2))*100000;sqrt(mean(Error1^2))*100000
########################################################## T-test for forecasted Errors
cor(sqrt(Error1^2),sqrt(abs(Error^2)))
(t.test( abs(Error),abs(Error1),alternative =  "two.sided", mu=0,
         paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
(t.test(Error^2,Error1^2,alternative =  "two.sided", mu=0,
        paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
########################################################## Minzer-Zarnowit R-square
us.lm<-lmrob(abs(forecasty)~sigma(bboot)[,j+1])
us.lm1<-lmrob(abs(forecasty)~sigma(bboot1)[,j+1])
summary(us.lm)[12];summary(us.lm1)[12]
mean(abs(us.lm$residuals))*100000;mean(abs(us.lm1$residuals))*100000
sqrt(mean(us.lm$residuals^2))*100000;sqrt(mean(us.lm1$residuals^2))*100000
############################################################################ T -test
(t.test( abs(us.lm$residuals),abs(us.lm1$residuals),alternative =  "two.sided", mu=0,
         paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
(t.test(us.lm$residuals^2,us.lm1$residuals^2,alternative =  "two.sided", mu=0,
        paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
########################################################################## Wald test
wald.test(vcov(us.lm), coefficients(us.lm),Terms = 2, L = NULL, H0 = 1,
          df = df.residual(us.lm), verbose = FALSE)
wald.test(vcov(us.lm1), coefficients(us.lm1),Terms = 2, L = NULL, H0 = 1,
          df =df.residual(us.lm1), verbose =FALSE)
######################################################################### GMMT test
z = residuals(usdfit)/sigma(usdfit)
skew = dskewness("nig",skew = coef(usdfit)["skew"], shape= coef(usdfit)["shape"])
# add back 3 since dkurtosis returns the excess kurtosis
kurt = 3+dkurtosis("nig",skew = coef(usdfit)["skew"], shape= coef(usdfit)["shape"])
z1 = residuals(usdfit1)/sigma(usdfit1)
skew1 = dskewness("nig",skew = coef(usdfit1)["skew"], shape= coef(usdfit1)["shape"])
# add back 3 since dkurtosis returns the excess kurtosis
```

```
kurt1 = 3+dkurtosis("nig",skew = coef(usdfit1)["skew"], shape= coef(usdfit1)["shape"])
print(GMMTest(z, lags = 2, skew=skew, kurt=kurt))
print(GMMTest(z1, lags = 2, skew=skew1, kurt=kurt1))
##################################################################Unconditional var plots
par(mfrow=c(1,1))
k<-c(1:365)
omega<-coef(usdfit)[length(coef(usdfit))]
OMEGA<-coef(usdfit1)[length(coef(usdfit1))]
P<-persistence(usdfit)
P1<-persistence(usdfit1)
uv<-sigma(bboot)-(omega/(1-P)*(1-P^(k-1)))
uv1<-sigma(bboot1)-(OMEGA/(1-P1)*(1-P1^(k-1)))
plot.zoo(x = cbind(uv,uv1),
        xlab="Time/horizon", ylab = "sigma",
 main =list("Unconditional forecasted sigma(TGARCH)",font = 1,cex=0.9),
 col=c(2,3), lty=c(1,5), screens = 1)
legend(100, 0.003,legend=c("sigma 1","sigma 2"),lty = c(1,5),col=c(2,3),cex=0.7)
################################################################## forecasted sigma plot
plot.zoo(x = cbind(abs(forecasty),sigma(bboot),sigma(bboot1)),
        xlab="Time/horizon", ylab = "sigma",
        main=list("|returns| versus Forecasted sigma (TGARCH)",font = 1,
              cex=0.9),
        col=c(8,2,3), lty=c(5,1,4),lwd=c(1,1,2), screens = 1)

legend(200, 0.06,legend=c("|returns|","sigma 1", "sigma 2"),lty = c(5,1,4),col=c(8,2,3),cex=0.7)
############################################EGARCH model#####################
##################################################################Calling and attachment
attach(rates);library(rugarch)
require(forecast);require(zoo);
require(stats);require(sandwich)
require(rms);require("aod")
library(robustbase)
```

```
################################################################### Creating regressors
x=inr
y=sek
## Creeating regressors
usdret=diff(log(y))[2:1819]
USDRET<-usdret[1:1818]
extret=abs(diff(log(x)))[1:1818]
direct1<-ifelse(extret[-1]>extret, 1, 0)
exvaret<-cbind(extret,direct1,abs(diff(log(nok)))[1:1818])
################### Estimating parameters of Augmented NAGARCH#################
usdspec<-ugarchspec(variance.model = list(model = "eGARCH",submodel="NULL",
            garchOrder = c(1,1),external.regressors=exvaret),
                mean.model = list(armaOrder = c(1,1),include.mean = TRUE),
                 distribution.model ='sged')
usdfit<-ugarchfit(usdspec,usdret, solver = 'hybrid',out.sample=440)
show(usdfit)
plot(usdfit,which = "all")
################ Estimating parameters of non-Augmented NAGARCH#################
usdspec1<-ugarchspec(variance.model = list(model = "eGARCH",submodel="NULL",
              garchOrder = c(1, 1),variance.targeting=FALSE),
                  mean.model = list(armaOrder = c(1, 1),
                  archm = FALSE,include.mean = TRUE),distribution ='sged')
usdfit1 <-ugarchfit(usdspec1,usdret, solver = 'hybrid',out.sample=440)
show(usdfit1)
##################################################################### Model statistics
############################################################# Miscellaneous Statistics
uncvariance(usdfit);uncvariance(usdfit1)
halflife(usdfit);halflife(usdfit1)
persistence(usdfit);persistence(usdfit1)
infocriteria(usdfit);infocriteria(usdfit1)
################################################# Calculating root mean and absolute error
error<-(usdfit@fit$residuals)^2-sigma(usdfit)^2
```

```
error1<-(usdfit1@fit$residuals)^2-sigma(usdfit1)^2
mean(abs(error));mean(abs(error1))
(mean(error^2));(mean(error1^2))
############################################################ T-tests for predicted errors
 cor(error,error1)
cor(error,error1)
(t.test( abs(as.vector(error1)),abs(as.vector(error)),alternative =  "two.sided", mu=0,
                 paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
(t.test(as.vector(error1^2),as.vector(error^2),alternative =  "two.sided", mu=0,
                paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
########################################################### Minzer-Zarnowit R-square
summary(lmrob(usdfit@fit$residuals^2~sigma(usdfit)))$r.square
summary(lmrob(usdfit1@fit$residuals^2~sigma(usdfit1)))$r.square
###Forecasting (following horizons were used: n=1, n=180 and n=440, to get the forecasts for
#other horizons, assign the approapriate horizon to the nuum variable)
par(mfrow=c(1,1))
numm=440
j=0
bboot = ugarchforecast(usdfit, n.ahead=numm,n.roll=j,out.sample=0, external.forecasts =
list(vregfor = exvaret))
bboot1 = ugarchforecast(usdfit1, n.ahead=numm,n.roll=j,out.sample=0)
plot(bboot,which = 3)
################################################################### Forecast Statistics
xm<-mean(USDRET[1379:(1378+numm)])
forecasty=USDRET[1379:(1378+numm)]-xm
Error<-abs(forecasty)-sigma(bboot)[,j+1]
Error1<-abs(forecasty)-sigma(bboot1)[,j+1]
mean(abs(Error))*100000;mean(abs(Error1))*100000
sqrt(mean(Error^2))*100000;sqrt(mean(Error1^2))*100000
########################################################### T-test for forecasted Errors
cor(sqrt(Error1^2),sqrt(abs(Error^2)))
(t.test( abs(Error),abs(Error1),alternative =  "two.sided", mu=0,
         paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
```

```
(t.test(Error^2,Error1^2,alternative =  "two.sided", mu=0,
        paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
############################################################ Minzer-Zarnowit R-square
us.lm<-lmrob(abs(forecasty)~sigma(bboot)[,j+1])
us.lm1<-lmrob(abs(forecasty)~sigma(bboot1)[,j+1])
summary(us.lm)[12];summary(us.lm1)[12]
mean(abs(us.lm$residuals))*100000;mean(abs(us.lm1$residuals))*100000
sqrt(mean(us.lm$residuals^2))*100000;sqrt(mean(us.lm1$residuals^2))*100000
############################################################################ T -test
(t.test( abs(us.lm$residuals),abs(us.lm1$residuals),alternative =  "two.sided", mu=0,
         paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
(t.test(us.lm$residuals^2,us.lm1$residuals^2,alternative =  "two.sided", mu=0,
        paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
######################################################################### Wald test
wald.test(vcov(us.lm), coefficients(us.lm),Terms = 2, L = NULL, H0 = 1,
          df = df.residual(us.lm), verbose = FALSE)
wald.test(vcov(us.lm1), coefficients(us.lm1),Terms = 2, L = NULL, H0 = 1,
          df =df.residual(us.lm1), verbose =FALSE)
####################################################################### GMMT test
z = residuals(usdfit)/sigma(usdfit)
skew = dskewness("sged",skew = coef(usdfit)["skew"], shape= coef(usdfit)["shape"])
# add back 3 since dkurtosis returns the excess kurtosis
kurt = 3+dkurtosis("sged",skew = coef(usdfit)["skew"], shape= coef(usdfit)["shape"])
z1 = residuals(usdfit1)/sigma(usdfit1)
skew1 = dskewness("sged",skew = coef(usdfit1)["skew"], shape= coef(usdfit1)["shape"])
# add back 3 since dkurtosis returns the excess kurtosis
kurt1 = 3+dkurtosis("sged",skew = coef(usdfit1)["skew"], shape= coef(usdfit1)["shape"])
print(GMMTest(z, lags = 2, skew=skew, kurt=kurt))
print(GMMTest(z1, lags = 2, skew=skew, kurt=kurt))
############################################################3 forecasted sigma plot
par(mfrow=c(1,1))
plot.zoo((abs(forecasty)),col="8", lty=5,xlab="Time/horizon",ylab=" sigma",
```

```
  main=list("Absolute returns versus Forecasted sigma for SGARCH",font = 1, cex=0.9))
plot.zoo(sigma(bboot)[,1],col="4",lty=1)
lines(sigma(bboot1)[,1],col="1",lty=2)
legend(1, 0.06,legend=c("|returns|","sigma 1", "sigma 2"),lty = c(5,1,4),col=c(8,4,1),cex=0.7)
```

Appendix E5: Codes for Value-at-Risk forecasting

```
#####################################SGARCH ################################
######################################################### calling and attachement
attach(rates);library(rugarch)
require(forecast);require(zoo);
require(stats);require(sandwich)
require(rms);require("aod")
library(robustbase)
require(MCS)
############################################################## Creating regressors
USDRET<-usdret[1:1818]
extret=abs(diff(log(mwk)))[1:1818]
extret=abs(diff(log(mwk)))[1:1818]
direct1<-ifelse(extret[-1]>extret,1, 0)
exvaret<-cbind(extret,direct1)
########################################## 99% VaR ########################
### Volatility specification for Augmented and non-Augmented SGARCH###########
usdspec<-ugarchspec(variance.model = list(model = "sGARCH",submodel="NULL",
  garchOrder = c(2,1),external.regressors=exvaret,variance.targeting=TRUE),
  mean.model = list(armaOrder = c(5,4),include.mean = TRUE),distribution.model='jsu')
usdspec1<-ugarchspec(variance.model = list(model = "sGARCH",submodel="NULL",
    garchOrder = c(2,1),variance.targeting=TRUE),
      mean.model = list(armaOrder = c(5,4),include.mean = TRUE),distribution ='jsu')
### Estimating parameters of Augmented and non-Augmented SGARCH###########
cl = makePSOCKcluster(10)
varmod = ugarchroll(usdspec, data = usdret, n.ahead = 1,forecast.length=440,
               refit.every =100, refit.window = "moving",
```

```
                solver = "hybrid", calculate.VaR = TRUE,VaR.alpha = c(0.01, 0.005,0.01),
                cluster = cl, keep.coef = TRUE)
show(varmod)
stopCluster(cl)
cl1 = makePSOCKcluster(10)
varmod1 = ugarchroll(usdspec1, data = usdret, n.ahead = 1,forecast.length=440,
               refit.every =100, refit.window = "moving",
                solver = "hybrid", calculate.VaR = TRUE,VaR.alpha = c(0.01, 0.005,0.01),
                cluster = cl1, keep.coef = TRUE)
show(varmod1)
stopCluster(cl1)
###################VAR REPORTS#############################################
report(varmod, type="VaR", VaR.alpha = 0.01, conf.level = 0.99)
report(varmod1, type="VaR", VaR.alpha = 0.01, conf.level = 0.99)
report(varmod, type="fpm")
report(varmod1, type="fpm")
plot(varmod,which=4)
plot(varmod1,which=4)
############# MCS Ranking procedure and other comparison tools #####################
realize1 <- as.data.frame(varmod)$Realized
evaluate1<-varmod@forecast$VaR[,1]
realize2 <- as.data.frame(varmod1)$Realized
evaluate2<-varmod1@forecast$VaR[,1]
mean(evaluate1);mean(evaluate2)
Loss1<-LossVaR(realize1, evaluate1, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
Loss2<-LossVaR(realize2, evaluate2, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
LOOS<-cbind(Loss1,Loss2)
MCSprocedure(Loss=LOOS,alpha=0.01,B=10000,statistic='Tmax',verbose=TRUE,set.seed(16))
mean(Loss1); mean(Loss2)
max(Loss1);max(Loss2)
cor(Loss1,Loss2)
(t.test(as.vector(Loss1),as.vector(Loss2),alternative =  "less", mu=0,
```

```
    paired = TRUE, var.equal = FALSE,conf.level = 0.95))[3]
#################### MAD and maxAD ##########################################
a=ifelse(realize1<evaluate1,(evaluate1),0)
b=ifelse(realize2<evaluate2,(evaluate2),0)
mean(abs(a[a!=0]-mean(a[a!=0])))
mean(abs(b[b!=0]-mean(b[b!=0])))
max(abs(a[a!=0]-mean(a[a!=0])))
max(abs(b[b!=0]-mean(b[b!=0])))
############################## Expected Shortfall###############################
################################### Computations ##############################
actual=as.data.frame(varmod)$Realized
VaR=varmod@forecast$VaR[,1]
f = function(x) qdist("jsu", p=x, mu = 0, sigma = 1,
    skew  = coef(varmod)[[5]][2][[1]][16], shape=coef(varmod)[[5]][2][[1]][17])
actual1=as.data.frame(varmod1)$Realized
VaR1=varmod1@forecast$VaR[,1]
f1 = function(x) qdist("jsu", p=x, mu = 0, sigma = 1,
    skew  = coef(varmod1)[[5]][2][[1]][15], shape=coef(varmod1)[[5]][2][[1]][16])
P=0.05
ES = varmod@forecast$density$Mu + varmod@forecast$density$Sigma*integrate(f, 0,
P)$value/P
ES1 = varmod1@forecast$density$Mu + varmod1@forecast$density$Sigma*integrate(f, 0,
P)$value/P
print(ESTest(0.05, actual, ES, VaR, boot = TRUE))
print(ESTest(0.05, actual1, ES1, VaR1, boot = TRUE))
mean(ES)
mean(ES1)
mean((actual- ES)^2)
mean((actual1- ES1)^2)
######################## Computing mean-daily capital requirement #################
x=cbind(varmod@forecast$VaR[,1],(3+1)*movavg(varmod@forecast$VaR[,1], 60, type="s"))
y=cbind(varmod1@forecast$VaR[,1],(3+1)*movavg(varmod1@forecast$VaR[,1], 60,
    type="s"))
```

```
abs(mean(rowMax(x, which = FALSE, ignore.zero = TRUE)))
abs(mean(rowMax(y, which = FALSE, ignore.zero = TRUE)))
######################### 95% VaR ########################
### Estimating parameters of Augmented and non-Augmented SGARCH###########
cl = makePSOCKcluster(10)
```

```
varmod = ugarchroll(usdspec, data = usdret, n.ahead = 1,forecast.length=440,  solver = "hybrid",
                 refit.every =100, refit.window = "moving", calculate.VaR = TRUE,
                 VaR.alpha = c(0.01, 0.025,0.05), cluster = cl, keep.coef = TRUE)
show(varmod)
stopCluster(cl)
cl1 = makePSOCKcluster(10)
varmod1 = ugarchroll(usdspec1, data = usdret, n.ahead = 1,forecast.length=440,
                  refit.every =100, refit.window = "moving", solver = "hybrid", calculate.VaR =
                TRUE,VaR.alpha =c(0.01, 0.025,0.05), cluster = cl1, keep.coef = TRUE)
show(varmod1)
stopCluster(cl1)
#####################VAR REPORTS############################################
report(varmod, type="VaR", VaR.alpha = 0.05, conf.level = 0.95)
report(varmod1, type="VaR",VaR.alpha = 0.05, conf.level = 0.95)
report(varmod, type="fpm")
report(varmod1, type="fpm")
plot(varmod,which=4)
plot(varmod1,which=4)
############ MCS Ranking procedure and other comparison tools #####################
realize1 <- as.data.frame(varmod)$Realized
evaluate1<-varmod@forecast$VaR[,1]
realize2 <- as.data.frame(varmod1)$Realized
evaluate2<-varmod1@forecast$VaR[,1]
evaluate2<-varmod1@forecast$VaR[,1]
mean(evaluate1);mean(evaluate2)
Loss1<-LossVaR(realize1, evaluate1, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
Loss2<-LossVaR(realize2, evaluate2, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
```

```
LOOS<-cbind(Loss1,Loss2)
MCSprocedure(Loss=LOOS,alpha=0.01,B=10000,statistic='Tmax',verbose=TRUE,set.seed(17))
mean(Loss1); mean(Loss2)
max(Loss1);max(Loss2)
##################### MAD and maxAD ############################################
```

```
 a=ifelse(realize1<evaluate1,(evaluate1),0)
b=ifelse(realize2<evaluate2,(evaluate2),0)
mean(abs(a[a!=0]-mean(a[a!=0])))
mean(abs(b[b!=0]-mean(b[b!=0])))
max(abs(a[a!=0]-mean(a[a!=0])))
max(abs(b[b!=0]-mean(b[b!=0])))
##############################NAGARCH #####################################
##################################### calling and attachement######################
attach(rates);library(rugarch)
require(forecast);require(zoo);
require(stats);require(sandwich)
require(rms);require("aod")
library(robustbase)
require(MCS)
##################################### Creating regressors #########################
usdret<-diff(log(bwp))[2:1819]
USDRET<-usdret[1:1818]
extret<-abs(diff(log(brl)))[1:1818]
direct1<-ifelse(extret[-1]>extret, 1, 0)
exvaret<-cbind(extret,direct1)
################################### 99% VaR ##################################
### Volatility specification for Augmented and non-Augmented NAGARCH###########
usdspec<-ugarchspec(variance.model = list(model = "fGARCH",submodel="NAGARCH",
        garchOrder = c(3,0),external.regressors=exvaret,variance.targeting=TRUE),
  mean.model = list(armaOrder = c(2,1),include.mean = TRUE),distribution.model='jsu')
usdspec1<-ugarchspec(variance.model = list(model = "fGARCH",submodel="NAGARCH",
          garchOrder = c(3, 0),variance.targeting=TRUE),
```

```
        mean.model = list(armaOrder = c(2, 1),include.mean = TRUE),distribution ='jsu')
### Estimating parameters of Augmented and non-Augmented NAGARCH############
cl = makePSOCKcluster(10)
varmod = ugarchroll(usdspec, data = usdret, n.ahead = 1,forecast.length=440,
                  refit.every =100, refit.window = "moving",
                  solver = "hybrid", calculate.VaR = TRUE,VaR.alpha = c(0.01, 0.005,0.01),
                  cluster = cl, keep.coef = TRUE)
show(varmod)
stopCluster(cl)
cl1 = makePSOCKcluster(10)
varmod1 = ugarchroll(usdspec1, data = usdret, n.ahead = 1,forecast.length=440,
                 refit.every =100, refit.window = "moving",
                  solver = "hybrid", calculate.VaR = TRUE,VaR.alpha = c(0.01, 0.005,0.01),
                  cluster = cl1, keep.coef = TRUE)
show(varmod1)
stopCluster(cl1)
####################VAR REPORTS###########################################
par(mfrow=c(2,2))
report(varmod, type="VaR", VaR.alpha = 0.01, conf.level = 0.99)
report(varmod1, type="VaR", VaR.alpha = 0.01, conf.level = 0.99)
report(varmod, type="fpm")
report(varmod1, type="fpm")
plot(varmod,which=4)
plot(varmod1,which=4)
############# MCS Ranking procedure and other comparison tools #####################
realize1 <- as.data.frame(varmod)$Realized
evaluate1<-varmod@forecast$VaR[,1]
realize2 <- as.data.frame(varmod1)$Realized
evaluate2<-varmod1@forecast$VaR[,1]
mean(evaluate1);mean(evaluate2)
Loss1<-LossVaR(realize1, evaluate1, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
Loss2<-LossVaR(realize2, evaluate2, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
```

```
LOOS<-cbind(Loss1,Loss2)
MCSprocedure(Loss=LOOS,alpha=0.01,B=10000,statistic='Tmax',verbose=TRUE,set.seed(16))
mean(Loss1); mean(Loss2)
max(Loss1);max(Loss2)
##################### MAD and maxAD ##############################################
a=ifelse(realize1<evaluate1,(evaluate1),0)
b=ifelse(realize2<evaluate2,(evaluate2),0)
mean(abs(a[a!=0]-mean(a[a!=0])))
mean(abs(b[b!=0]-mean(b[b!=0])))
max(abs(a[a!=0]-mean(a[a!=0])))
max(abs(b[b!=0]-mean(b[b!=0])))
#################################### Expected Shortfall############################
######################################### Computations ############################
actual=as.data.frame(varmod)$Realized
VaR=varmod@forecast$VaR[,1]
f = function(x) qdist("jsu", p=x, mu = 0, sigma = 1,
       skew  = coef(varmod)[[5]][2][[1]][13], shape=coef(varmod)[[5]][2][[1]][14])
actual1=as.data.frame(varmod1)$Realized
VaR1=varmod1@forecast$VaR[,1]
f1 = function(x) qdist("jsu", p=x, mu = 0, sigma = 1,
         skew  = coef(varmod1)[[5]][2][[1]][11], shape=coef(varmod1)[[5]][2][[1]][12])
P=0.05
ES=varmod@forecast$density$Mu+varmod@forecast$density$Sigma*integrate(f, 0, P)$value/P
ES1=varmod1@forecast$density$Mu+varmod1@forecast$density$Sigma*integrate(f, 0, P) $
            value/P
print(ESTest(0.05, actual, ES, VaR, boot = TRUE))
print(ESTest(0.05, actual1, ES1, VaR1, boot = TRUE))
mean(ES)
mean(ES1)
mean((actual- ES)^2)
mean((actual1- ES1)^2)
########### Computing mean-daily capital requirement################################
```

```
x=cbind(varmod@forecast$VaR[,1],(3+0)*movavg(varmod@forecast$VaR[,1], 60, type="s"))
y=cbind(varmod1@forecast$VaR[,1],(3+0)*movavg(varmod1@forecast$VaR[,1], 60,
type="s"))
abs(mean(rowMax(x, which = FALSE, ignore.zero = TRUE)))
abs(mean(rowMax(y, which = FALSE, ignore.zero = TRUE)))
############################################# 95% VaR ########################
####### Estimating parameters of Augmented and non-Augmented NAGARCH############
cl = makePSOCKcluster(10)
varmod = ugarchroll(usdspec, data = usdret, n.ahead = 1,forecast.length=440,
                    refit.every =100, refit.window = "moving",
                    solver = "hybrid", calculate.VaR = TRUE,VaR.alpha = c(0.01, 0.025,0.05),
                    cluster = cl, keep.coef = TRUE)
show(varmod)
stopCluster(cl)
cl1 = makePSOCKcluster(10)
varmod1 = ugarchroll(usdspec1, data = usdret, n.ahead = 1,forecast.length=440,
                     refit.every =100, refit.window = "moving",
                     solver = "hybrid", calculate.VaR = TRUE,VaR.alpha =c(0.01, 0.025,0.05),
                     cluster = cl1, keep.coef = TRUE)
show(varmod1)
stopCluster(cl1)
###################################VAR REPORTS################################
report(varmod, type="VaR", VaR.alpha = 0.05, conf.level = 0.95)
report(varmod1, type="VaR",VaR.alpha = 0.05, conf.level = 0.95)
report(varmod, type="fpm")
report(varmod1, type="fpm")
plot(varmod,which=4)
plot(varmod1,which=4)
################### MCS Ranking procedure and other comparison tools ###############
realize1 <- as.data.frame(varmod)$Realized
evaluate1<-varmod@forecast$VaR[,1]
realize2 <- as.data.frame(varmod1)$Realized
```

```
evaluate2<-varmod1@forecast$VaR[,1]
evaluate2<-varmod1@forecast$VaR[,1]
###################### meqn VAR
mean(evaluate1);mean(evaluate2)
Loss1<-LossVaR(realize1, evaluate1, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
Loss2<-LossVaR(realize2, evaluate2, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
LOOS<-cbind(Loss1,Loss2)
MCSprocedure(Loss=LOOS,alpha=0.01,B=10000,statistic='Tmax',verbose=TRUE,set.seed(17))
mean(Loss1); mean(Loss2)
max(Loss1);max(Loss2)
#################### MAD and maxAD ##########################################
 a=ifelse(realize1<evaluate1,(evaluate1),0)
b=ifelse(realize2<evaluate2,(evaluate2),0)
mean(abs(a[a!=0]-mean(a[a!=0])))
mean(abs(b[b!=0]-mean(b[b!=0])))
max(abs(a[a!=0]-mean(a[a!=0])))
max(abs(b[b!=0]-mean(b[b!=0])))
#########################TGARCH ##########################################
##################calling and attachement#########################################
attach(rates);library(rugarch)
require(forecast);require(zoo);
require(stats);require(sandwich)
require(rms);require("aod")
library(robustbase)
require(MCS)
################################### Creating regressors#########################
usdret=diff(log(ils))[2:1819]
USDRET<-usdret[1:1818]
extret=abs(diff(log(mwk)))[1:1818]
extret1=abs(diff(log(brl)))[1:1818]
direct1<-ifelse(extret[-1]>extret, 1, 0)
exvaret<-cbind(direct1,extret1)
```

```
########################################### 99% VaR ##############################
######### Volatility specification for Augmented and non-Augmented TGARCH###########
usdspec<-ugarchspec(variance.model = list(model = "fGARCH",submodel="TGARCH",
  garchOrder = c(2,2),external.regressors=exvaret,variance.targeting=TRUE),
  mean.model = list(armaOrder = c(1,1),include.mean = TRUE),distribution.model='nig')
usdspec1<-ugarchspec(variance.model = list(model = "fGARCH",submodel="TGARCH",
    garchOrder = c(2,2),variance.targeting=TRUE),
      mean.model = list(armaOrder = c(1,1),include.mean = TRUE),distribution ='nig')
############## Estimating parameters of Augmented and non-Augmented TGARCH########
cl = makePSOCKcluster(10)
varmod = ugarchroll(usdspec, data = usdret, n.ahead = 1,forecast.length=440,
                 refit.every =100, refit.window = "moving",
                 solver = "hybrid", calculate.VaR = TRUE,VaR.alpha = c(0.01, 0.005,0.01),
                 cluster = cl, keep.coef = TRUE)
show(varmod)
stopCluster(cl)
cl1 = makePSOCKcluster(10)
varmod1 = ugarchroll(usdspec1, data = usdret, n.ahead = 1,forecast.length=440,
                refit.every =100, refit.window = "moving",
                 solver = "hybrid", calculate.VaR = TRUE,VaR.alpha = c(0.01, 0.005,0.01),
                 cluster = cl1, keep.coef = TRUE)
show(varmod1)
stopCluster(cl1)
########################################VAR REPORTS############################
report(varmod, type="VaR", VaR.alpha = 0.01, conf.level = 0.99)
report(varmod1, type="VaR", VaR.alpha = 0.01, conf.level = 0.99)
report(varmod, type="fpm")
report(varmod1, type="fpm")
plot(varmod,which="all")
plot(varmod1,which="all")
############## MCS Ranking procedure and other comparison tools #####################
realize1 <- as.data.frame(varmod)$Realized
```

```
evaluate1<-varmod@forecast$VaR[,1]
realize2 <- as.data.frame(varmod1)$Realized
evaluate2<-varmod1@forecast$VaR[,1]
evaluate2<-varmod1@forecast$VaR[,1]
mean(evaluate1);mean(evaluate2)
Loss1<-LossVaR(realize1, evaluate1, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
Loss2<-LossVaR(realize2, evaluate2, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
LOOS<-cbind(Loss1,Loss2)
MCSprocedure(Loss=LOOS,alpha=0.2,B=10000,statistic='Tmax',verbose=TRUE)
mean(Loss1); mean(Loss2)
max(Loss1);max(Loss2)
#################### MAD and maxAD ##########################################
a=ifelse(realize1<evaluate1,(evaluate1),0)
b=ifelse(realize2<evaluate2,(evaluate2),0)
mean(abs(a[a!=0]-mean(a[a!=0])))
mean(abs(b[b!=0]-mean(b[b!=0])))
max(abs(a[a!=0]-mean(a[a!=0])))
max(abs(b[b!=0]-mean(b[b!=0])))
##################################### Expected Shortfall######################
### Computations
actual=as.data.frame(varmod)$Realized
VaR=varmod@forecast$VaR[,1]
f = function(x) qdist("nig", p=x, mu = 0, sigma = 1,
    skew  = coef(varmod)[[5]][2][[1]][12], shape=coef(varmod)[[5]][2][[1]][13])
actual1=as.data.frame(varmod1)$Realized
VaR1=varmod1@forecast$VaR[,1]
f1 = function(x) qdist("nig", p=x, mu = 0, sigma = 1,
      skew  = coef(varmod1)[[5]][2][[1]][10], shape=coef(varmod1)[[5]][2][[1]][11])
P=0.01
ES = varmod@forecast$density$Mu +
                   varmod@forecast$density$Sigma*integrate(f, 0, P)$value/P
ES1 = varmod1@forecast$density$Mu +
```

```
                              varmod1@forecast$density$Sigma*integrate(f, 0, P)$value/P
print(ESTest(0.01, actual, ES, VaR, boot = TRUE, n.boot = 1000))
print(ESTest(0.01, actual1, ES1, VaR1, boot = TRUE, n.boot = 1000))
mean(ES)
mean(ES1)
mean((actual- ES)^2)
mean((actual1- ES1)^2)
 ##################### Computing mean-daily capital requirement#####################
x=cbind(varmod@forecast$VaR[,1],(3+1)*movavg(varmod@forecast$VaR[,1], 60, type="s"))
y=cbind(varmod1@forecast$VaR[,1],(3+1)*movavg(varmod1@forecast$VaR[,1], 60,
type="s"))
abs(mean(rowMax(x, which = FALSE, ignore.zero = TRUE)))
abs(mean(rowMax(y, which = FALSE, ignore.zero = TRUE)))
###################################### 95% VaR ##############################
######### Estimating parameters of Augmented and non-Augmented TGARCH###########
cl = makePSOCKcluster(10)
varmod = ugarchroll(usdspec, data = usdret, n.ahead = 1,forecast.length=440,
                    refit.every =100, refit.window = "moving",
                    solver = "hybrid", calculate.VaR = TRUE,VaR.alpha = c(0.01, 0.025,0.05),
                    cluster = cl, keep.coef = TRUE)
show(varmod)
stopCluster(cl)
cl1 = makePSOCKcluster(10)
varmod1 = ugarchroll(usdspec1, data = usdret, n.ahead = 1,forecast.length=440,
                     refit.every =100, refit.window = "moving",
                     solver = "hybrid", calculate.VaR = TRUE,VaR.alpha =c(0.01, 0.025,0.05),
                     cluster = cl1, keep.coef = TRUE)
show(varmod1)
stopCluster(cl1)
##################VAR REPORTS############################################
report(varmod, type="VaR", VaR.alpha = 0.05, conf.level = 0.95)
report(varmod1, type="VaR",VaR.alpha = 0.05, conf.level = 0.95)
```

```
report(varmod, type="fpm")
report(varmod1, type="fpm")
plot(varmod,which=4)
plot(varmod1,which=4)
############# MCS Ranking procedure and other comparison tools #####################
realize1 <- as.data.frame(varmod)$Realized
evaluate1<-varmod@forecast$VaR[,1]
realize2 <- as.data.frame(varmod1)$Realized
evaluate2<-varmod1@forecast$VaR[,1]
evaluate2<-varmod1@forecast$VaR[,1]
mean(evaluate1);mean(evaluate2)
Loss1<-LossVaR(realize1, evaluate1, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
Loss2<-LossVaR(realize2, evaluate2, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
LOOS<-cbind(Loss1,Loss2)
MCSprocedure(Loss=LOOS,alpha=0.01,B=10000,statistic='Tmax',verbose=TRUE,set.seed(19))
mean(Loss1); mean(Loss2)
max(Loss1);max(Loss2)
#################### MAD and maxAD ########################################
a=ifelse(realize1<evaluate1,(evaluate1),0)
b=ifelse(realize2<evaluate2,(evaluate2),0)
mean(abs(a[a!=0]-mean(a[a!=0])))
mean(abs(b[b!=0]-mean(b[b!=0])))
max(abs(a[a!=0]-mean(a[a!=0])))
max(abs(b[b!=0]-mean(b[b!=0])))
###################################GJRGARCH ################################
###################### calling and attachement##################################
attach(rates);library(rugarch)
require(forecast);require(zoo);
require(stats);require(sandwich)
require(rms);require("aod")
library(robustbase)
require(MCS)
```

```
############################## Creating regressors#####################################
USDRET<-usdret[1:1818]
extret=abs(diff(log(mwk)))[1:1818]
extret=abs(diff(log(mwk)))[1:1818]
direct1<-ifelse(extret[-1]>extret,1, 0)
exvaret<-cbind(extret,direct1)
######################################## 99% VaR ##################################
### Volatility specification for Augmented and non-Augmented GJRGARCH############
usdspec<-ugarchspec(variance.model = list(model = "sGARCH",submodel="NULL",
  garchOrder = c(2,1),external.regressors=exvaret,variance.targeting=TRUE),
  mean.model = list(armaOrder = c(5,4),include.mean = TRUE),distribution.model='jsu')
usdspec1<-ugarchspec(variance.model = list(model = "sGARCH",submodel="NULL",
    garchOrder = c(2,1),variance.targeting=TRUE),
      mean.model = list(armaOrder = c(5,4),include.mean = TRUE),distribution ='jsu')
###### Estimating parameters of Augmented and non-Augmented GJRGARCH############
cl = makePSOCKcluster(10)
varmod = ugarchroll(usdspec, data = usdret, n.ahead = 1,forecast.length=440,
                refit.every =100, refit.window = "moving",
                solver = "hybrid", calculate.VaR = TRUE,VaR.alpha = c(0.01, 0.005,0.01),
                cluster = cl, keep.coef = TRUE)
show(varmod)
stopCluster(cl)
cl1 = makePSOCKcluster(10)
varmod1 = ugarchroll(usdspec1, data = usdret, n.ahead = 1,forecast.length=440,
               refit.every =100, refit.window = "moving",
                solver = "hybrid", calculate.VaR = TRUE,VaR.alpha = c(0.01, 0.005,0.01),
                cluster = cl1, keep.coef = TRUE)
show(varmod1)
stopCluster(cl1)
########################################VAR REPORTS############################
report(varmod, type="VaR", VaR.alpha = 0.01, conf.level = 0.99)
report(varmod1, type="VaR", VaR.alpha = 0.01, conf.level = 0.99)
```

```
report(varmod, type="fpm")
report(varmod1, type="fpm")
plot(varmod,which=4)
plot(varmod1,which=4)
############# MCS Ranking procedure and other comparison tools #####################
realize1 <- as.data.frame(varmod)$Realized
evaluate1<-varmod@forecast$VaR[,1]
realize2 <- as.data.frame(varmod1)$Realized
evaluate2<-varmod1@forecast$VaR[,1]
mean(evaluate1);mean(evaluate2)
Loss1<-LossVaR(realize1, evaluate1, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
Loss2<-LossVaR(realize2, evaluate2, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
LOOS<-cbind(Loss1,Loss2)
MCSprocedure(Loss=LOOS,alpha=0.01,B=10000,statistic='Tmax',verbose=TRUE,set.seed(16))
mean(Loss1); mean(Loss2)
max(Loss1);max(Loss2)
################### MAD and maxAD ##########################################
a=ifelse(realize1<evaluate1,(evaluate1),0)
b=ifelse(realize2<evaluate2,(evaluate2),0)
mean(abs(a[a!=0]-mean(a[a!=0])))
mean(abs(b[b!=0]-mean(b[b!=0])))
max(abs(a[a!=0]-mean(a[a!=0])))
max(abs(b[b!=0]-mean(b[b!=0])))
############################## Expected Shortfall############################
################################ Computations##############################
actual=as.data.frame(varmod)$Realized
VaR=varmod@forecast$VaR[,1]
f = function(x) qdist("jsu", p=x, mu = 0, sigma = 1,
      skew  = coef(varmod)[[5]][2][[1]][16], shape=coef(varmod)[[5]][2][[1]][17])
actual1=as.data.frame(varmod1)$Realized
VaR1=varmod1@forecast$VaR[,1]
f1 = function(x) qdist("jsu", p=x, mu = 0, sigma = 1,
```

```
     skew  = coef(varmod1)[[5]][2][[1]][15], shape=coef(varmod1)[[5]][2][[1]][16])
P=0.05
ES = varmod@forecast$density$Mu + varmod@forecast$density$Sigma*integrate(f, 0,
P)$value/P
ES1 = varmod1@forecast$density$Mu + varmod1@forecast$density$Sigma*integrate(f, 0,
P)$value/P
print(ESTest(0.05, actual, ES, VaR, boot = TRUE))
print(ESTest(0.05, actual1, ES1, VaR1, boot = TRUE))
mean(ES)
mean(ES1)
mean((actual- ES)^2)
mean((actual1- ES1)^2)
###################### Computing mean-daily capital requirement#####################
x=cbind(varmod@forecast$VaR[,1],(3+1)*movavg(varmod@forecast$VaR[,1], 60, type="s"))
y=cbind(varmod1@forecast$VaR[,1],(3+1)*movavg(varmod1@forecast$VaR[,1], 60,
type="s"))
abs(mean(rowMax(x, which = FALSE, ignore.zero = TRUE)))
abs(mean(rowMax(y, which = FALSE, ignore.zero = TRUE)))
################################# 95% VaR #################################
########Estimating parameters of Augmented and non-Augmented GJRGARCH###########
cl = makePSOCKcluster(10)
varmod = ugarchroll(usdspec, data = usdret, n.ahead = 1,forecast.length=440,
               refit.every =100, refit.window = "moving",
               solver = "hybrid", calculate.VaR = TRUE,VaR.alpha = c(0.01, 0.025,0.05),
               cluster = cl, keep.coef = TRUE)
show(varmod)
stopCluster(cl)
cl1 = makePSOCKcluster(10)
varmod1 = ugarchroll(usdspec1, data = usdret, n.ahead = 1,forecast.length=440,
                refit.every =100, refit.window = "moving",
                solver = "hybrid", calculate.VaR = TRUE,VaR.alpha =c(0.01, 0.025,0.05),
                cluster = cl1, keep.coef = TRUE)
show(varmod1)
```

```
stopCluster(cl1)
######################################VAR REPORTS###########################
report(varmod, type="VaR", VaR.alpha = 0.05, conf.level = 0.95)
report(varmod1, type="VaR",VaR.alpha = 0.05, conf.level = 0.95)
report(varmod, type="fpm")
report(varmod1, type="fpm")
plot(varmod,which=4)
plot(varmod1,which=4)
############################MCS Ranking procedure and other comparison tools ########
realize1 <- as.data.frame(varmod)$Realized
evaluate1<-varmod@forecast$VaR[,1]
realize2 <- as.data.frame(varmod1)$Realized
evaluate2<-varmod1@forecast$VaR[,1]
evaluate2<-varmod1@forecast$VaR[,1]
mean(evaluate1);mean(evaluate2)
Loss1<-LossVaR(realize1, evaluate1, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
Loss2<-LossVaR(realize2, evaluate2, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
LOOS<-cbind(Loss1,Loss2)
MCSprocedure(Loss=LOOS,alpha=0.01,B=10000,statistic='Tmax',verbose=TRUE,set.seed(17))
mean(Loss1); mean(Loss2)
max(Loss1);max(Loss2)
#################### MAD and maxAD ###########################################
 a=ifelse(realize1<evaluate1,(evaluate1),0)
b=ifelse(realize2<evaluate2,(evaluate2),0)
mean(abs(a[a!=0]-mean(a[a!=0])))
mean(abs(b[b!=0]-mean(b[b!=0])))
max(abs(a[a!=0]-mean(a[a!=0])))
max(abs(b[b!=0]-mean(b[b!=0])))
#########################EGARCH #######################
######################################calling and attachement##################
attach(rates);library(rugarch)
require(forecast);require(zoo);
```

```
require(stats);require(sandwich)
require(rms);require("aod")
library(robustbase)
require(MCS)
######################################## Creating regressors######################
x=inr
y=sek
## Creeating regressors
usdret=diff(log(y))[2:1819]
USDRET<-usdret[1:1818]
extret=abs(diff(log(x)))[1:1818]
direct1<-ifelse(extret[-1]>extret, 1, 0)
exvaret<-cbind(extret,direct1,abs(diff(log(nok)))[1:1818])
########################### 99% VaR ######################################
##############Volatility specification for Augmented and non-Augmented EGARCH#######
usdspec<-ugarchspec(variance.model = list(model = "eGARCH",submodel="NULL",
                        garchOrder = c(1,1),external.regressors=exvaret),
        mean.model = list(armaOrder = c(1,1),include.mean = TRUE), distribution.model
                   ='sged')
usdspec1<-ugarchspec(variance.model = list(model = "eGARCH",submodel="NULL",
                        garchOrder = c(1, 1),variance.targeting=FALSE),
          mean.model = list(armaOrder = c(1, 1), include.mean = TRUE),distribution ='sged')
####### Estimating parameters of Augmented and non-Augmented EGARCH############
cl = makePSOCKcluster(10)
varmod = ugarchroll(usdspec, data = usdret, n.ahead = 1,forecast.length=440,
        refit.every=100, refit.window="moving",solver="hybrid", calculate.VaR = TRUE,
        VaR.alpha = c(0.01, 0.005,0.01), cluster = cl, keep.coef = TRUE)
show(varmod)
stopCluster(cl)
cl1 = makePSOCKcluster(10)
varmod1=ugarchroll(usdspec1,data = usdret, n.ahead = 1,forecast.length=440,
        refit.every =100,refit.window="moving",solver="hybrid", calculate.VaR = TRUE,
```

```
            VaR.alpha = c(0.01, 0.005,0.01), cluster = cl1, keep.coef = TRUE)
show(varmod1)
stopCluster(cl1)
####################VAR REPORTS############################################
report(varmod, type="VaR", VaR.alpha = 0.01, conf.level = 0.99)
report(varmod1, type="VaR", VaR.alpha = 0.01, conf.level = 0.99)
report(varmod, type="fpm")
report(varmod1, type="fpm")
plot(varmod,which="all")
plot(varmod1,which="all")
############# MCS Ranking procedure and other comparison tools #####################
realize1 <- as.data.frame(varmod)$Realized
evaluate1<-varmod@forecast$VaR[,1]
realize2 <- as.data.frame(varmod1)$Realized
evaluate2<-varmod1@forecast$VaR[,1]
evaluate2<-varmod1@forecast$VaR[,1]
mean(evaluate1);mean(evaluate2)
Loss1<-LossVaR(realize1, evaluate1, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
Loss2<-LossVaR(realize2, evaluate2, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
LOOS<-cbind(Loss1,Loss2)
MCSprocedure(Loss=LOOS,alpha=0.01,B=10000,statistic='Tmax',verbose=TRUE,set.seed(1))
mean(Loss1); mean(Loss2)
max(Loss1);max(Loss2)
#################### MAD and maxAD ##########################################
a=ifelse(realize1<evaluate1,(evaluate1),0)
b=ifelse(realize2<evaluate2,(evaluate2),0)
mean(abs(a[a!=0]-mean(a[a!=0])))
mean(abs(b[b!=0]-mean(b[b!=0])))
max(abs(a[a!=0]-mean(a[a!=0])))
max(abs(b[b!=0]-mean(b[b!=0])))
### Expected Shortfall
############################## Computations ###################################
```

```
actual=as.data.frame(varmod)$Realized
VaR=varmod@forecast$VaR[,1]
f = function(x) qdist("sged", p=x, mu = 0, sigma = 1,
        skew  = coef(varmod)[[5]][2][[1]][11], shape=coef(varmod)[[5]][2][[1]][12])
actual1=as.data.frame(varmod1)$Realized
VaR1=varmod1@forecast$VaR[,1]
f1 = function(x) qdist("sged", p=x, mu = 0, sigma = 1,
    skew  = coef(varmod1)[[5]][2][[1]][8], shape=coef(varmod1)[[5]][2][[1]][9])
P=0.05
ES=varmod@forecast$density$Mu +varmod@forecast$density$Sigma*integrate(f, 0, P)$value/P
ES1=varmod1@forecast$density$Mu+varmod1@forecast$density$Sigma*integrate(f, 0, P)$value/P
print(ESTest(0.05, actual, ES, VaR, boot = TRUE))
print(ESTest(0.05, actual1, ES1, VaR1, boot = TRUE))
mean(ES)
mean(ES1)
mean((actual- ES)^2)
mean((actual1- ES1)^2)
###################### Computing mean-daily capital requirement######################
x=cbind(varmod@forecast$VaR[,1],(3+1)*movavg(varmod@forecast$VaR[,1], 60, type="s"))
y=cbind(varmod1@forecast$VaR[,1],(3+1)*movavg(varmod1@forecast$VaR[,1], 60, type="s"))
abs(mean(rowMax(x, which = FALSE, ignore.zero = TRUE)))
abs(mean(rowMax(y, which = FALSE, ignore.zero = TRUE)))
########################## 95% VaR ##########################################
######## Estimating parameters of Augmented and non-Augmented SGARCH###########
cl = makePSOCKcluster(10)
varmod = ugarchroll(usdspec, data = usdret, n.ahead = 1,forecast.length=440,
          refit.every =100, refit.window = "moving",
          solver = "hybrid", calculate.VaR = TRUE,VaR.alpha = c(0.01, 0.025,0.05),
                cluster = cl, keep.coef = TRUE)
show(varmod)
```

```
stopCluster(cl)
cl1 = makePSOCKcluster(10)
varmod1 = ugarchroll(usdspec1, data = usdret, n.ahead = 1,forecast.length=440,
                 refit.every =100, refit.window = "moving",
                 solver = "hybrid", calculate.VaR = TRUE,VaR.alpha =c(0.01, 0.025,0.05),
                 cluster = cl1, keep.coef = TRUE)
show(varmod1)
stopCluster(cl1)
##################VAR REPORTS##############################################
report(varmod, type="VaR", VaR.alpha = 0.05, conf.level = 0.95)
report(varmod1, type="VaR",VaR.alpha = 0.05, conf.level = 0.95)
report(varmod, type="fpm")
report(varmod1, type="fpm")
plot(varmod,which="all")
plot(varmod1,which="all")
############# MCS Ranking procedure and other comparison tools #####################
realize1 <- as.data.frame(varmod)$Realized
evaluate1<-varmod@forecast$VaR[,1]
realize2 <- as.data.frame(varmod1)$Realized
evaluate2<-varmod1@forecast$VaR[,1]
evaluate2<-varmod1@forecast$VaR[,1]
mean(evaluate1);mean(evaluate2)
Loss1<-LossVaR(realize1, evaluate1, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
Loss2<-LossVaR(realize2, evaluate2, which = 'asymmetricLoss', type = 'normal',delta=25,0.99)
LOOS<-cbind(Loss1,Loss2)
MCSprocedure(Loss=LOOS,alpha=0.01,B=10000,statistic='Tmax',verbose=TRUE)
mean(Loss1); mean(Loss2)
max(Loss1);max(Loss2)
##################### MAD and maxAD ##########################################
a=ifelse(realize1<evaluate1,(evaluate1),0)
b=ifelse(realize2<evaluate2,(evaluate2),0)
mean(abs(a[a!=0]-mean(a[a!=0])))
```

```
mean(abs(b[b!=0]-mean(b[b!=0])))
max(abs(a[a!=0]-mean(a[a!=0])))
max(abs(b[b!=0]-mean(b[b!=0])))
```