UNIVERSITEIT·STELLENBOSCH·UNIVERSITY
jou kennisvennoot • your knowledge partner

Goodness of fit for different dynamic hedging strategies

Marguerite Bezuidenhout

Assignment presented in partial fulfilment
of the requirements for the degree of
Hons BComm (Financial Risk Management)
at the Stellenbosch University

Supervisor: Mr C.J. van der Merwe

Degree of confidentiality: A

October 2015

PLAGIARISM DECLARATION

1. Plagiarism is the use of ideas, material and other intellectual property of another's work and to present it as my own.
2. I agree that plagiarism is a punishable offence because it constitutes theft.
3. I also understand that direct translations are plagiarism.
4. Accordingly all quotations and contributions from any source whatsoever (including the internet) have been cited fully. I understand that the reproduction of text without quotation marks (even when the source is cited) is plagiarism.
5. I declare that the work contained in this assignment, except otherwise stated, is my original work and that I have not previously (in its entirety or in part) submitted it for grading in this module/assignment or another module/assignment.

Student number	**Signature**
16947681	M.B.
Name and surname	**Date**
Marguerite Bezuidenhout	20 October 2015

Acknowledgements

The author would like to thank the following people for their contribution towards this project:

- My supervisor for this research assignment, Mr. Carel van der Merwe, for all his support and guidance throughout the project. Always giving new ideas and different ways to think about the problem at hand.

- My parents and family for their continual support and love.

- Prof. W.J. Conradie and Dr. J.D. van Heerden for their contribution and leadership during my studies.

Abstract

Hedging refers to a strategy used to lower the overall risk of a portfolio. There exists a broad selection of models available in the market to implement a dynamic hedging strategy. In this research assignment the Constant Conditional Correlation (CCC), Dynamic Conditional Correlation (DCC) and the Copula Based-GARCH models are constructed and executed to evaluate the most appropriate model for the selected data. Mili and Abid (2004:659) argues that a dynamic hedging strategy leads to greater risk reduction as opposed to a static hedging strategy. Various summary statistics are calculated to detect if there is any autocorrelations present in the data, this is used for the testing of Autoregressive Conditional Heteroskedastic (ARCH) effects. Some goodness-of-fit (GOF) tests are calculated to interpret the data and to decide on appropriate distributions, this is followed by different information criterions to use in the process for model selection. Finally, the parameters of the three models will be estimated through the process of maximum-likelihood. The results of each model fit are discussed and relevant comparisons are made.

There are two hedging strategies implemented in this study, the first is the Top 40 index that will be directly hedged with its own futures, and secondly using foreign currency futures to cross hedge the currency exposure of holding foreign equity. It was found that the Copula-Based GARCH model, which permits nonlinear and asymmetric dependence between the two assets in the cross-hedge portfolio, results in the most appropriate model fit. However, comparing the DCC and CCC-GARCH models, the conclusion is that the DCC-GARCH model is more appropriate than the CCC-GARCH model, therefore implying that the more dynamic a model is the better.

Key words:

Constant Conditional Correlation, Dynamic Conditional Correlation, Copula, GARCH, ARCH

Opsomming

Verskansing verwys na ‘n strategie wat gebruik word om die algehele risiko van ‘n portefeulje te verlaag. Daar is ‘n wye keuse van modelle beskikbaar om ‘n dinamiese verskansings strategie te implimenteer. In hierdie navorsings opdrag gaan die Konstante Voorwaardelike Korrelasie (KVK), Dinamiese Voorwaardelike Korrelasie (DVK) en die Koepel gebaseerde-GARCH modelle gebou en uitgevoer word om die mees geskikte model vir die geselekteerde data te vind. Mili en Abid(2004:659) argumenteer dat ‘n dinamiese verskansings strategie lei tot meer risiko vermindering teenoon ‘n statiese verskansings strategie. Verskeie opsommings statistieke word bereken om vas te stel of daar enige outokorrelasies teenwoordig is in die data, met ander woorde die toetsing vir Outoregressiewe Voorwaardelike Heteroskedastic (ORVH) effekte. ‘n Paar passingstoetse word bereken om die data te verstaan en om gepaste verdelings te selekteer, gevolg deur verskillende inligtings kriteria wat gebruik word in die proses van model seleksie. Ten slotte, sal die parameters van die drie modelle beraam word deur die proses van maksimum-waarskynlikheid. Die resultate van elke model word bespreek en relevante vergelykings word gemaak.

Daar is twee verskansings strategieë wat in hierdie studie geïmplementeer word, die eerste is die Top 40 indeks wat direk verskans word met sy eie termynkontrakte, en tweedens die gebruik van buitelandse valuta termynkontrakte om te kruis verskans deur die valuta blootstelling van die hou van buitelandse aandele. Daar is gevind dat die Koepel gebaseerde GARCH model, wat nie-lineêre en asimmetriese afhanklikheid tussen die twee bates toelaat in die kruisverskansings portefeulje, lei tot die mees geskikte model passing. Maar in die vergelyking van die DVK en KVK-GARCH modelle, is die gevolgtrekking dat die DVK-GARCH model meer toepaslik is as die KVK-GARCH model en dus kan die veronderstelling gemaak word dat hoe meer dinamies ‘n model is hoe beter is die model.

Sleutel woorde:

Konstante Voorwaardelike Korrelasie, Dinamiese Voorwaardelike Korrelasie, Koepels, ORVH

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

List of abbreviations and/or acronyms

ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criterion
ARCH	Autoregressive Conditional Heteroskedastic
ARMA	Autoregressive Moving-average
BEKK	Baba, Engle, Kraft and Kroner
BIC	Bayesian Information Criterion
CCC	Constant Conditional Correlation
DCC	Dynamic Conditional Correlation
DF	Dickey-Fuller
EK	Excess Kurtosis
GARCH	Generalised Autoregressive Conditional Heteroskedasticity
GJR	Glosten, Jagannathan, and Runkle
GOF	Goodness-of-Fit
HQC	Hannan–Quinn Information Criterion
JB	Jarque-Bera
LBP	Ljung-Box Pierce
LM	Lagrange Multiplier
MSCI	Morgan Stanley Capital International
USD	United States Dollar
VAR	Vector Auto-Regression
ZAR	South African Rand

CHAPTER 1
INTRODUCTION

1.1 INTRODUCTION

Portfolio protection is often as important as portfolio appreciation. Hull (2012:63) argues that an equity portfolio is hedged to minimize systematic risk. This also enables an investor to step outside of the market for a small period of time, since hedging might be more cost effective that selling the portfolio and buying it back at a later stage. In this research assignment, dynamic hedging will refer to the hedging of equity positions with relevant equity futures.

Short-dated derivative instruments held in a portfolio are used to hedge long-dated derivatives to ensure against undesirable losses. Limitation to doing dynamic hedging involves being very careful when one hedges the higher order Greeks like gamma and vega, because transaction costs will increase (Greenbaum & Ravindrant, 2002:5). Kotzé (2010:1) states that in practice, hedging is often carried out using a position in futures rather than one in the underlying security.

Dynamic hedging, as opposed to static hedging, refers to the continuous buying or selling of the underlying security to maintain a delta of zero. Static hedging refers to the process where one hedges on inception and then to leave it as it is. A static hedge, therefore, does not need to be rebalanced. Mili and Abid (2004:657) describes an optimal hedge ratio as the proportion of the cash position that should be covered with an opposite position on a futures market.

Known literature places great emphasis on the estimation of a static hedge ratio by the use of the ordinary least-squares techniques. However, recent studies employ different bivariate conditional volatility models to enable an estimation of a time-varying hedge ratio. The advantage of the time-varying hedge ratio is that it takes into account the continuous changes in the joint distribution of spot and futures returns (Hsu, Tseng & Wang, 2008:1096). Whereas a constant hedge ratio does not change as time passes.

Using a static approach to hedge has the drawback that, when estimating the optimal hedge ratio, it overlooks the fact that two series are co-integrated. Most of the regression models already developed assumes that the variances and covariance of futures and spot returns are constant, and therefore obtaining a constant hedge ratio.

Copula functions are used to construct multivariate distributions and to investigate the dependence structure between random variables. A copula is the joint distribution of a vector of uniform random variables. Szego (2004:1) states that it is possible to map any vector of random variables into a vector of random variables with uniform margins, which enables the split of the margins of that vector to grasp the dependence. Using copula functions such as the Gumbel and Clayton, rather than the usual Gaussian assumption, can produce a richer dependence.

Most dynamic hedging models assume that the futures and spot returns follow a multivariate normal distribution with linear dependence. In the absence of the multivariate normality assumption, the joint distribution is decomposed into its marginal distributions and a copula, which can be considered both individually and simultaneously. The copula refers to the dependence structure between the spot and the futures returns. The dependence parameters in the copula function represent time-varying processes that seek to find possible dynamic and nonlinear relations between the spot and futures returns (Hsu, Tseng & Wang, 2008:1097).

Mili and Abid (2004:659) argues that the Generalised Autoregressive Conditional Heteroskedasticity (GARCH) time-varying hedge ratio has the advantage above the constant hedge ratio that it ultimately leads to larger risk reduction. The construction of a time varying optimal hedge ratio that depends on the conditional variances and covariance between the spot and futures returns, implements a dynamic hedge ratio that takes the rapid stock of information accessible in both spot and futures markets into account.

1.2 PROBLEM STATEMENT

Earlier studies have already shown that the traditional regression-based static method is unsuitable for hedging using futures, and therefore a variety of different dynamic hedging strategies have emerged in the market. In a dynamic hedging strategy, delta neutrality is key (i.e. the delta of the portfolio equals zero). This stems from the idea that gamma is an indicator of how fast the delta of a portfolio will change.

An approach where a time-varying hedge ratio is present is a better method to implement to ensure a reduction of risk in a portfolio. A competitive time-varying dynamic hedging strategy has to be estimated to replace that of the static hedge ratio in the market. A model needs to be developed that takes account of the fact that two separate series can be co-integrated, since a time-varying hedge ratio is dependent on the conditional variances and covariance between the spot and futures returns. In addition, the joint distribution of spot and futures returns are continuously changing.

1.3 RESEARCH QUESTION

Is it possible to improve the goodness of fit of a model by implementing a copula-based GARCH model as opposed to the Constant Conditional Correlation and Dynamic Conditional Correlation-GARCH models?

1.4 RESEARCH OBJECTIVES

This research will cover a class of new copula-based GARCH models, to test whether they have a better fit compared to the constant conditional correlation (CCC) GARCH and the dynamic conditional correlation (DCC) GARCH models.

The Gaussian copula function will be implemented in the copula-based GARCH model to estimate the possible dependence structure between spot and futures returns, and to show their joint distribution with full flexibility. Two more copula functions will be described in the literature review. Various summary statistics will be computed to understand the data and to apply certain tests to enable the fit of a proper model. Parameter estimation for the different models will be done though maximum likelihood.

Goodness-of-fit tests are implemented to test how well the selected data fits a specific model. The best model will be chosen on the basis of the different information criterions. In this research assignment, the information criterions to be implemented are known as the Akaike, Bayesian, Shibata and Hannan-Quinn information criterions. The models will be implemented for two hedging scenarios. First, the Top 40 Index will be directly hedged with its own futures, and secondly using foreign currency futures to cross hedge the currency exposure of holding foreign equity.

1.5 IMPORTANCE/BENEFITS OF THE STUDY

The use of hedge funds in personal financial portfolios has increased dramatically since the beginning of the 21st century. Typically, hedge funds are only open to a limited range of professional or wealthy investors.

Smith and Stulz (1985:111) suggested that firms hedge to reduce the probability of facing bankruptcy costs. Companies transfer risk to better risk bearers that are diversified and that have better access to capital markets. Hedging also leads to increased debt capacity. Since every company has an optimal mix of debt and equity financing, hedging transfers some of the company's debt to firms outside the company. This enables the firm to undertake a greater amount of debt. Companies hedge for a lower tax liability. Lastly, firms in need of financing may choose to rely on internal funds only. In order to do so, they have to implement hedging policies that would smooth out their cash flows to meet future funding needs.

Hedging is an important procedure that is starting to become an integral part in today's businesses. For this reason investors and companies have to implement the correct hedging strategy for a specific market and product at the correct time. Careful time and consideration has to be taken before a hedging decision is made, since every hedge has its cost. The benefits of hedging has to outweigh the costs incurred in order to be successful.

The copula-based GARCH and DCC GARCH models are implemented to take into consideration that the hedge is constantly changing as time passes. By making this adjustment, individuals and companies implementing a dynamic hedging strategy should reduce their potential losses if unfavourable events occur. By the use of copula functions the specifications of the joint distribution of assets become more precise and the risk exposures of the portfolios, more manageable.

1.6 PROJECT OUTLINE

Chapter 2 will be an in depth literature review on the construction and derivation of the different hedging strategies such as the CCC GARCH, DCC GARCH and the copula-based GARCH-models that are used for futures hedging. The third chapter consists of the research methodology, that will explain how the different methodologies will be implemented. Chapter 4 follows with the application of the above literature. In other words each strategy will be applied and the necessary comparisons will be made. This includes the computations and estimation of the copula function. The corresponding parameters of each proposed model will be estimated by the use of maximum likelihood.

The following statistics will be computed to enable comparisons: mean, standard deviation, skewness, kurtosis, Jarque-Bera statistic, Ljung-Box test statistics and lastly the Lagrange multiplier statistic. Dickey-Fuller and Johansen tests will also be obtained and the various information criterions will be computed. In the final chapter all the conclusions and open ended questions will be stated.

CHAPTER 2
LITERATURE REVIEW

2.1 INTRODUCTION

Hedging refers to the reduction in risk by exploiting the correlation amongst various risky instruments. A range of financial instruments exist such as insurance policies, forward contracts, swaps, options, over-the-counter products, derivative products, and futures contracts, which can all be used for hedging purposes.

Dynamic replication for hedging strategies makes use of a more delicate procedure than that used for static replication. Sharma, Vaish, Pandey and Gupta (2010:139) argues that for the process of dynamic replication, there exists a trading desk that deals with the different transaction costs, liquidity constraints, choosing price development models and all the uncertainties that follow.

Hsu, Tseng and Wang (2008:1098) states that a crucial input in the hedging of risks is the optimal hedge ratio. Hedging risks has become an essential issue. It is vital to determine the optimal amount of hedging instruments. Therefore, the calculation of the optimal hedge ratio plays an important role in the hedging process. The optimal hedge ratio is defined as the ratio of futures holdings to a spot position that minimises the risk of the hedged portfolio. Let s_t and f_t be the respective changes in the spot and futures prices at time t. If the joint distribution of spot and futures returns remains the same over time, then the conventional risk-minimising hedge ratio δ^*, is given by the following equation:

$$\delta^* = \frac{cov(s_t, f_t)}{var(f_t)}$$

An estimation of this static hedge ratio is simply undertaken from the least squares regression of s_t on f_t. Since the expected relationship between economic or financial variables may be better captured by a time varying parameter model as opposed to a fixed coefficient model, the optimal hedge ratio can therefore be one that is time varying rather than static. Hatemi-J and Roca (2006:295) states that with the arrival of new information, the joint distribution of these assets may be time-varying, in which case the static hedging strategy is not suitable for an extension to multi-period futures hedging. Hsu, Tseng and Wang (2008:1098) argues that conditional on the information set at time $t-1$, the optimal time-varying hedge ratio is obtained by minimising the risk of the hedged return $s_t - \delta_{t-1} f_t$ or as follows:

$$\delta_{t-1}^* = \frac{cov_{t-1}(s_t, f_t)}{var_{t-1}(f_t)}$$

In the following section various summary statistics that can be used to summarise data, will be discussed. Next follows an in depth justification of the three different hedging models namely the CCC, DCC and the copula based-GARCH models. The sub sections thereafter entails the clarification of parameter estimation, various goodness of fit tests and the different information criterions.

2.2 SUMMARY STATISTICS

Engle (1982:987) states that it is possible that an uncorrelated time series can still be serially dependent due to a dynamic conditional variance process. Therefore, if there is any autocorrelation present in the autoregressive conditional heteroskedastic (ARCH) effects, the series can be serially dependent. Engle's ARCH test is a Lagrange multiplier test to assess the significance of the ARCH effects. A time series is stationary if its statistical properties do not vary with time. The Augmented Dickey-Fuller test, tests whether there is a unit root present in the data, that is, it tests for non-stationarity. The Johansen test, tests for co-integration in the data. It is an important test to implement since ignoring co-integration aspects in a time series may lead to a spurious regression problem, which occurs if arbitrarily trending or non-stationary series are regressed on each other.

2.2.1 Lagrange Multiplier Statistic

Trench (1987:2) argues that Lagrange multipliers enable the capability to maximise or minimise functions with the limitation that only points on a certain surface is considered. To find the critical points of a function $f(x,y,z)$ on a level surface $g(x,y,z) = C$, the following system of simultaneous equations has to be solved:

$$\nabla f(x,y,z) = \lambda \nabla \mathrm{g}(\mathrm{x},\mathrm{y},\mathrm{z})$$
$$g(x,y,z) = C$$

Remembering that ∇f and ∇g represents vectors, it can be written as a collection of four equations in the four unknowns x, y, z and λ:

$$f_x(x,y,z) = \lambda g_x(x,y,z)$$
$$f_y(x,y,z) = \lambda g_y(x,y,z)$$
$$f_z(x,y,z) = \lambda g_z(x,y,z)$$
$$g(x,y,z) = C$$

The variable λ is a dummy variable known as the Lagrange multiplier, however only the values of x, y and z are actually important. Once all the critical point values are obtained, substitute them into f to see where the maxima and minima are located. The critical points where f is the greatest are called the maxima and the critical points where f is smallest are known as the minima.

2.2.2 Dickey-Fuller Test

Cryer and Chan (2010:201) states that a Augmented Dickey–Fuller test (ADF) is a test for a unit root in a time series sample. There are three methods for testing for a unit root in a time series.

1. Testing the model without a constant term
2. Testing the model with a constant term
3. Testing the model with a linear time trend

Consider the following hypothesis for the Augmented Dickey-Fuller t-test

$\boldsymbol{H_0}$: $\theta = 1$

$\boldsymbol{H_1}$: $\theta < 1$

The Dickey-Fuller test statistic is then defined as:

$$DF \equiv t - ratio = \frac{\hat{\theta} - 1}{std(\hat{\theta})}$$

The null hypothesis of a unit root is rejected when the test statistic DF is lower than the corresponding critical value. However if the null hypothesis cannot be rejected the series should be differenced.

2.2.3 Johansen Test

A Johansen Test refers to a procedure for testing co-integration in several vectors of variables that are integrated of order one, generally denoted as $I(1)$. This test allows more than one co-integrating relationship making it more generally appropriate (Hjalmarsson & Österholm, 2007:4).

Johansen's methodology takes its starting point in a vector autoregression (VAR) of order p given by:

$$y_t = \mu + A_1 y_{t-1} + \cdots + A_p y_{t-p} + \varepsilon_t$$

where, y_t denotes an n×1 vector of variables that are integrated of order one, and ε_t denotes an n×1 vector of innovations. Now VAR can be rewritten as follows:

$$\Delta y_t = \mu + \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t$$

where, $\Pi = \sum_{i=1}^{p} A_i - I$ and $\Gamma_i = -\sum_{j=i+1}^{p} A_j$.

If the coefficient matrix Π has a reduced rank of $r < n$, then there exists $n \times r$ matrices α and β each with rank r such that $\Pi = \alpha\beta'$ and $\beta' y_t$ is stationary. Here r denotes the number of co-integrating relationships, the elements of α are known as the adjustment parameters and each column of β is a co-integrating vector (Hjalmarsson & Österholm, 2007:4).

There are two kinds of Johansen tests:

1. The trace test:

$$J_{trace} = -T \sum_{i=r+1}^{n} \ln(1 - \hat{\lambda}_i)$$

2. Maximum eigenvalue test:

$$J_{max} = -T ln(1 - \hat{\lambda}_{r+1})$$

where, T denotes the sample size and $\hat{\lambda}_i$ denotes the i^{th} largest canonical correlation. The trace test tests the null hypothesis of r co-integrating vectors against the alternative hypothesis of n co-integrating vectors. The maximum eigenvalue test, conversely, tests the null hypothesis of r co-integrating vectors against the alternative hypothesis of $r + 1$ co-integrating vectors. Neither of these test statistics follows a chi-square distribution in general, therefore asymptotic critical values can be found in Johansen and Juselius (1990). Subsequently the critical values used for the maximum eigenvalue and trace test statistics are grounded on a pure unit-root assumption, and they will therefore no longer be correct when the variables in the system are near-unit-root processes (Hjalmarsson & Österholm, 2007:5).

2.3 GARCH - GENERAL AUTOCORRELATED CONDITIONAL HETEROSKEDASTICITY

Peters (2008:4) argues that in the financial industry volatility is an imperative concept that measures the state of uncertainty in returns. It is known that volatility fluctuates over time and has a tendency to cluster in times of large volatility and times of low volatility. This phenomenon is known as heteroskedasticity. Another aspect to take account for is that volatility has made known to be autocorrelated, that is, that today's volatility depends on the volatilities in the past. Bearing in mind that volatility is not directly observable in the market the necessity for a good model to help estimate and forecast is vital. A type of model that captures the above properties is known as a GARCH -General Autocorrelated Conditional Heteroskedasticity model. This model has proven to be successful in estimating and predicting volatility changes. Now follows the three different GARCH models that will be considered in this study.

2.3.1 Constant Conditional Correlation (CCC)-GARCH Model

Peters (2008:5) argues that the conditional correlations were anticipated to be constant and that only the conditional variances were time varying. Meanwhile correlation in practice for many assets changes as time moves forward, therefore assumption that the conditional correlation is constant over time is not convincing.

The CCC GARCH model is a model in which the volatility of an asset is defined only through lagged squared innovations and volatility of its own. Nakatani and Teräsvirta (2009:147) argues that the investigation of interdependence in volatility is imperative for portfolio risk management on the one hand, and is necessary for research on the degree of market integration on the other. A great number of researchers have found sufficient evidence that the conditional variances of financial time series are interacting. In the CCC GARCH model, co-movements between heteroskedastic time series are modelled by permitting each series to follow a separate GARCH process while limiting the conditional correlations between the GARCH processes to be constant (Teräsvirta, 2012:1).

Although the CCC GARCH model has clear computational advantages over the multivariate GARCH (BEKK) model of Engle and Kroner (1995), the correlation structure between the spot and futures markets is quite restricted.

Kroner and Sultan (1993) proposed the following bivariate error-correction model of s_t and f_t with a constant correlation GARCH(1,1) structure for the estimation of δ_t^*:

$$S_t = \alpha_{0s} + \alpha_{1s}(S_{t-1} - \lambda F_{t-1}) + \varepsilon_{st}$$

$$f_t = \alpha_{0f} + \alpha_{1f}(S_{t-1} - \lambda F_{t-1}) + \varepsilon_{ft}$$

$$\begin{bmatrix} \varepsilon_{st} \\ \varepsilon_{ft} \end{bmatrix} \Big| \boldsymbol{\Psi}_{t-1} \sim N(0, H_t)$$

$$H_t = \begin{bmatrix} h_{s,t}^2 & h_{sf,t} \\ h_{sf,t} & h_{f,t}^2 \end{bmatrix} = \begin{bmatrix} h_{s,t} & 0 \\ 0 & h_{f,t} \end{bmatrix} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \begin{bmatrix} h_{s,t} & 0 \\ 0 & h_{f,t} \end{bmatrix} = D_t R D_t$$

$$h_{s,t}^2 = c_s + a_s \varepsilon_{s,t-1}^2 + b_s h_{s,t-1}^2$$

$$h_{f,t}^2 = c_f + a_f \varepsilon_{f,t-1}^2 + b_f h_{f,t-1}^2$$

Where S_{t-1} and F_{t-1} represents the corresponding spot and futures prices, and $S_{t-1} - \lambda F_{t-1}$ denotes the error-correction term, $\boldsymbol{\Psi}_{t-1}$ denotes the information set at time $t-1$, and the disturbance term is given by the following equation $\varepsilon_t = (\varepsilon_{st}, \varepsilon_{ft})'$ follows a bivariate normal distribution with zero mean and a conditional covariance matrix H_t with a constant correlation ρ.

The GARCH term allows the hedge ratio to be time-varying, while the error-correction term characterises the long-run connection between the spot and futures prices. The theory of a constant correlation model may be too restrictive to be in line with actuality. The DCC GARCH model was introduced by Engle and Sheppard (2001) and Engle (2002) to liberate this restriction and improve the elasticity of the hedging models (Hsu, Tseng & Wang, 2008:1096).

2.3.2 Dynamic Conditional Correlation (DCC)-GARCH Model

Engle and Sheppard (2001) introduced the dynamic conditional correlation (DCC) model as a means of considering the flexible correlation structure between assets. Engle (2002:1) states that these models have the flexibility of univariate GARCH models together with parsimonious parametric models for the correlations.

The DCC model has a two-step algorithm to estimate the parameters that makes the model relatively simple to implement in practice. In the first step, the conditional variances are estimated via a univariate GARCH model. In the second step the parameters for the conditional correlation given the parameters from the first step are estimated. Peters (2008:5) argues that this method makes it possible to estimate covariances of a large amount of assets without too much challenging computations. Finally, the DCC model includes conditions that make the covariance matrix positive definite at all points in time.

However, the DCC model assumes that the spot and futures returns follow a multivariate normal distribution with linear dependence. Hsu, Tseng and Wang (2008:1096) argues that numerous empirical studies have shown that numerous financial asset returns are skewed, leptokurtic, and asymmetrically dependent, which contradicts the statement above. For example, the volatility of an assets return refers to the standard deviation of the changes in value during a specific time horison. In the long run returns tend to move towards a mean value (mean reverting). The changes in value that appears during these times are both positive and negative (asymmetric), mostly close to the mean value but some changes obtain extreme values (leptokurtic). As mentioned above, the volatility of today's returns is conditional on the past volatility and tends to cluster (Peters, 2008:6). Hence, these characteristics should be considered in the specifications of any effective hedging model.

In contrast with the CCC GARCH model, the DCC GARCH model allows the correlation R to be time-varying:

From (1) it follows, that:

$$H_t = D_t R_t D_t = D_t J_t Q_t J_t D_t$$

where D_t represents the diagonal matrix of conditional standard deviations from univariate GARCH models, $Q_t = (q_{ij,t})_{2\times 2}$ is a positive definite matrix, $J_t = diag\{q_{s,t}^{-\frac{1}{2}}, q_{f,t}^{-\frac{1}{2}}\}$, and Q_t satisfies the following:

$$Q_t = (1 - \theta_1 - \theta_2)\,\overline{Q} + \theta_1\, \zeta_{t-1}\, \zeta_{t-1} + \theta_2\, Q_{t-1}$$

Where, θ_1 and θ_2 are non-negative parameters that satisfy the following restriction $\theta_1 + \theta_2 < 1$.

2.3.3 Copula-Based GARCH Model

Hsu, Tseng and Wang (2008:1096) states that a new copula-based GARCH model will be presented for the estimation of an optimal hedge ratio. Without the assumption of multivariate normality, the joint distribution can be split into its marginal distributions and a copula.

The marginal distributions can be any non-elliptical distribution, while the copula function describes the dependence structure between the spot and futures returns. The proposed hedging model uses the GJR-skewed-t (Glosten, Jagannathan, and Runkle) specification for the marginal distributions. Including the three different copulas namely the Gaussian, Gumbel, and Clayton for the joint distribution to allow a wide range of possible dependence structures. The dependence parameters in these copulas are modelled as time-varying processes to capture possible dynamic and nonlinear relationships between the spot and futures returns.

Hsu, Tseng and Wang (2008:1100) argues that a copula function enables the consideration of the marginal distributions and the dependence structures both separately and simultaneously. Therefore, the joint distribution can be specified with full elasticity with respect to asset returns, which will correspond to a more reasonable outcome. Glosten, Jagannathan, and Runkle (1993) and Hansen (1994), specifies the GJR-skewed-t models for shocks in the spot and futures returns. The model specifications are as follows:

The conditional variance for asset $i,\ i = s, f$, is given by

$$h_{i,t}^2 = c_i + b_i h_{i,t-1}^2 + a_{i,1}\varepsilon_{i,t-1}^2 + a_{i,2}k_{i,t-1}\varepsilon_{i,t-1}^2$$

$$\varepsilon_{i,t}\big|\Psi_{t-1} = \mathrm{h}_{i,t}z_{i,t} \qquad \text{where } z_{i,t} \sim \mathrm{skewed} - \mathrm{t}(z_i|\eta_i, \phi_i)$$

with $k_{i,t-1} = 1$ when $\varepsilon_{i,t-1}$ is negative, else $k_{i,t-1} = 0$.

The density function of the skewed-t distribution is

$$skewed - t(z|\eta,\phi) = \begin{cases} bc\left(1 + \frac{1}{\eta-2}\left(\frac{bz+a}{1-\phi}\right)^2\right)^{-\eta+\frac{1}{2}}, & z < -\frac{a}{b} \\ bc\left(1 + \frac{1}{\eta-2}\left(\frac{bz+a}{1+\phi}\right)^2\right)^{-\eta+\frac{1}{2}} & , z \geq -\frac{a}{b} \end{cases}$$

The values of a, b, and c are defined as

$$a \equiv 4\,\phi\, c\, \frac{\eta-2}{\eta-1} \qquad b \equiv 1 + 3\phi^2 - a^2 \qquad c \equiv \frac{\Gamma(\eta+\frac{1}{2})}{\sqrt{\pi(\eta-2)\Gamma(\frac{\eta}{2})}}$$

Where, η denotes the kurtosis parameter and ϕ is the asymmetry parameter. These parameters are limited to the following restriction $4 < \eta < 30$ and $-1 < \phi < 1$. Therefore, the particular marginal distributions of spot and futures returns are asymmetric, fat-tailed and non-Gaussian.

Assume that the conditional cumulative distribution functions of z_s and z_f are $G_{s,t}(z_{s,t}|\Psi_{\mathrm{t-1}})$ and $G_{f,t}(z_{f,t}|\Psi_{t-1})$, respectively. The conditional copula function, denoted as $C_t(u_t, v_t|\Psi_{t-1})$, is defined by the two time-varying cumulative distribution functions of random variables $u_t = G_{s,t}(z_{s,t}|\Psi_{t-1})$ and $v_t = G_{f,t}(z_{f,t}|\Psi_{t-1})$. Let φ_t be the bivariate conditional cumulative distribution functions of $z_{s,t}$and $z_{f,t}$.

Hsu, Tseng and Wang (2008:1101) states that using Sklar's theorem, the following is obtained

$$\begin{aligned}\Phi_t(z_{s,t}, z_{f,t}|\Psi_{t-1}) &= C_t(u_t, v_t|\Psi_{t-1})\\ &= C_t(G_{s,t}(z_{s,t}\big|\Psi_{t-1}),\ G_{f,t}(z_{f,t}|\Psi_{t-1})|\Psi_{t-1})\end{aligned}$$

The bivariate conditional density function of $z_{s,t}$ and $z_{f,t}$ can be constructed as

$$\begin{aligned}\varphi_t(z_{s,t}, z_{f,t}\big|\Psi_{t-1}) = &\ c_t(G_{s,t}(z_{s,t}\big|\Psi_{t-1}), G_{f,t}(z_{f,t}|\Psi_{t-1})|\Psi_{t-1})\\ &\times g_{s,t}(z_{s,t}|\Psi_{t-1}) \times g_{f,t}(z_{f,t}|\Psi_{t-1})\end{aligned} \tag{5}$$

Where $c_t(u_t, v_t\left|\Psi_{t-1}) = \frac{\partial^2 C_t(u_t,v_t|\Psi_{t-1})}{\partial u_t,\partial v_t, g_{s,t}(z_{s,t}|\Psi_{t-1})}\right.$ is the conditional density of $z_{s,j}$, and $g_{f,t}(z_{f,t}|\Psi_{t-1})$ is the conditional density of $z_{f,t}$.

2.3.3.1 Copulas

Patton (2007:2) states that the variance of the return on a portfolio of risky assets depends on the variances of the individual assets and on the linear correlation amongst the assets in the portfolio. In general, the distribution of the return on a portfolio depends on the univariate distributions of the individual assets in the portfolio and on the dependence between each of the assets, which is captured by a function known as a copula.

A copula is a function that links together univariate distribution functions to form a multivariate distribution function. If all the variables are continuously distributed, then their copula is merely a multivariate distribution function with $Uniform(0,1)$ univariate marginal distributions. Consider a random vector, $\boldsymbol{X} = [X_1, X_2, \dots, X_n]'$, with joint distribution $\boldsymbol{F}$ and marginal distributions $F_1, F_2, \dots, F_n$. Sklar's theorem provides the mapping from the individual distribution functions to the joint distribution function:

$$F(x) = C\big(F_1(x_1), F_2(x_2), \dots, F_n(x_n)\big), \quad \forall\ x \in \mathbb{R}^n \tag{1}$$

From any multivariate distribution, $\boldsymbol{F}$, the marginal distributions, F_i, can be extracted as well as the copula, $\boldsymbol{C}$. Given any set of marginal distributions $(F_1, F_2, \dots, F_n)$ and any copula $\boldsymbol{C}$, equation (1) can be used to find a joint distribution with the given marginal distributions. A significant feature of this outcome is that the marginal distributions do not need to be similar in any way to one another, neither is the choice of the copula constrained by the choice of the marginal distributions.

Since each marginal distribution, F_i, holds all of the univariate information on the individual variable X_i, while the joint distribution $\boldsymbol{F}$ holds all the univariate and multivariate information, it is clear that the information contained in the copula $\boldsymbol{C}$ must be all of the dependence information between the $X_i's$. Therefore, copulas are occasionally also known as dependence functions. Also note that if U_i is defined as the probability integral transform of X_i, i.e. $U_i \equiv F_i(X_i)$, then $U_i \sim Uniform(0,1)$. It can be shown that $\boldsymbol{U} = [U_{1,}U_2, \ldots, U_n]' \sim \boldsymbol{C}$, the copula of $\boldsymbol{X}$. If the joint distribution function is n-times differentiable, then taking the n^{th} cross partial derivative of equation (1) the following is obtained:

$$\begin{aligned}
\boldsymbol{f}(x) &= \frac{\partial^n}{\partial_{x_1}\partial_{x_2} \ldots \partial_{x_n}} \boldsymbol{F}(x) \\
&= \prod_{i=1}^{n} f_i(x_i) \cdot \frac{\partial^n}{\partial_{u_1}\partial_{u_2} \ldots \partial_{u_n}} \boldsymbol{C}\left(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)\right) \\
&= \prod_{i=1}^{n} f_i(x_i) \cdot \boldsymbol{c}\left(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)\right)
\end{aligned}$$

From the above it is clear that the joint density is equal to the product of the marginal densities and the copula density denoted by $\boldsymbol{c}$. This also indicates that the joint log-likelihood is simply the sum of the univariate log-likelihoods and the copula log-likelihood, which is essential in the estimation of copula-based models:

$$\log \boldsymbol{f}(x) = \sum_{i=1}^{n} \log f_i(x_i) + \log \boldsymbol{c}\ (F_1(x_1), F_2(x_2), \ldots, F_n(x_n))$$

Patton (2007:3) states that the decomposition of a joint distribution into its marginal distributions and a copula allows a great deal of flexibility in identifying a model for the joint distribution. This is an advantage when the shape and goodness-of-fit of a model for the joint distribution is of primary interest. When accumulated knowledge is available about the distributions of specific variables and using that knowledge in constructing a joint distribution, then copulas also play a valuable part.

2.3.3.2 Applications of copulas in finance

The incentive for the use of copulas in finance originates from the increased empirical evidence that the dependence between important asset returns is non-normal. An example of non-normal dependence is when two asset returns show greater correlation during market declines than during market expansions. Asset returns exhibit non-normal dependence when the dependence is not consistent with a normal copula. The suggestion that dependence between asset returns are non-normal has wide-ranging implications for financial decision making, in risk management, multivariate option pricing, portfolio decisions and credit risk (Patton, 2007:11).

2.3.3.3 Gaussian Copula

Aas (2004:2) states that implicit copulas do not have a simple closed form, but are inferred by using well-known multivariate distribution functions. An example of an implicit copula is the Gaussian copula. The Gaussian copula is a distribution over the unit cube $[0,1]^d$. It is constructed from a multivariate normal distribution over $\mathbb{R}^d$ by making use of the probability integral transform.

The Gaussian copula follows as:

$$C_\rho(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(u)} \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \, exp\left\{-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right\} dxdy$$

where ρ denotes the parameter of the copula, and $\Phi^{-1}(\cdot)$ denotes the inverse of the standard univariate Gaussian distribution function.

2.3.3.4 Gumbel Copula

Aas (2004:3) argues that there are a number of copulas which are not derived from multivariate distribution functions, but do have simple closed forms. Well-known explicit copulas are the Clayton and Gumbel copulas. The Gumbel copula is used to model asymmetric dependence in data. This copula is well-known for its capability to capture strong upper tail dependence and weak lower tail dependence. If the outcomes are likely to be greatly correlated at high values but less correlated at low values, then the Gumbel copula is a suitable choice (Mahfoud, 2012:19).

Aas (2004:3) claims that the Gumbel copula shows greater dependence in the positive tail than in the negative. This copula is given as follows:

$$C_\delta(u,v) = \exp(-[(-\log u)^\delta + (-\log v)^\delta]^{\frac{1}{\delta}})$$

where, $0 < \delta \leq 1$ is a parameter controlling the dependence. Perfect dependence is achieved if $\delta \to 0$, while $\delta = 1$ implies independence.

2.3.3.5 Clayton Copula

Mahfoud (2012:18) states that Clayton first introduced the Clayton copula in 1978. It is typically used to study correlated risks because of their ability to capture lower tail dependence. Aas

(2004:3) argues that the Clayton copula, which is an asymmetric copula, shows greater dependence in the negative tail than in the positive. This copula is given by the following:

$$C_\delta(u,v) = (u^{-\delta} + v^{-\delta} - 1)^{-\frac{1}{\delta}}$$

Where, $0 < \delta < \infty$ is a parameter controlling the dependence. Perfect dependence is obtained if $\delta \ \rightarrow \infty$, while $\delta = 0$ implies independence.

2.4 PARAMETER ESTIMATION

Hsu, Tseng and Wang (2008:1102) states that the model parameters are estimated by use of the maximum likelihood method. Thus at time t, the log-likelihood function can be derived by taking the logarithm of equation (5) in section 2.3.3.

$$log\varphi_t = logc_t + log\ g_{s,t} + \log g_{f,t^*}$$

Let the parameters in $g_{s,t}$ and $g_{f,t}$ be denoted as θ_s and θ_f, whereas the parameters in c_t are denoted as θ_c. These parameters can therefore be estimated by maximising the following log-likelihood function:

$$L_{s,f}(\theta) = L_s(\theta_s) + L_f\left(\theta_f\right) + L_c(\theta_c)$$

where, $\theta = (\theta_s, \theta_f, \theta_c)$ and L_k represents the sum of the log-likelihood function values across observations of the variable k (Hsu, Tseng & Wang, 2008:1101).

As the dimensions of the estimated equation may be quite large, it is difficult in practice to achieve a simultaneous maximisation of $L_{s,t}(\theta)$ for all of the parameters. To solve this problem successfully, the two-step estimation procedure is used.

In the first step, the parameters of the marginal distribution are estimated from the univariate time series by: $\hat{\theta}_S = \sum_{t=1}^{T} \log g_{s,t}(z_{s,t}|\Psi_{t-1};\theta_S)$

$$\hat{\theta}_f = \sum_{t=1}^{T} \log g_{f,t}(z_{f,t}|\Psi_{t-1};\theta_f)$$

In the second step, given the marginal estimates obtained above, the dependence parameters are estimated by: $\hat{\theta}_c = \sum_{t=1}^{T} \log c_t(\Psi_{t-1};\hat{\theta}_S,\hat{\theta}_f,\theta_c)$

2.5 GOODNESS-OF-FIT TESTS

Olivares and Garcia-Forero (2010:190) states that the goodness-of-fit (GOF) of a statistical model describes how well it fits into a set of observations. GOF indices summarise the discrepancy between the observed values and the values expected under a statistical model. GOF statistics are GOF indices with known sampling distributions, usually obtained using asymptotic methods that are used in statistical hypothesis testing. This section introduces two GOF tests to consider in the modelling of the data. The Jarque–Bera tests for normality. And as an alternative to Engle's ARCH test, serial dependence (ARCH effects) in a residual series can be tested by conducting a Ljung-Box Pierce tests on the first m lags of the residual series and the squared residual series.

2.5.1 Jarque-Bera Test

The Jarque–Bera test is one of the various goodness-of-fit tests. It tests whether the selected sample data has a skewness and kurtosis value corresponding to that of the normal distribution. The Jarque-Bera test for normality is defined as follows:

Consider testing the following hypothesis

$$\boldsymbol{H_0}\text{: } Normal\ Distribution \quad (Skewness\ and\ Excess\ Kurtosis\ equals\ zero)$$

$$\boldsymbol{H_1}\text{: } Non-normal\ Distribution$$

The Jarque-Bera test statistic is then defined as:

$$JB = n \cdot \left[\frac{S^2}{6} + \frac{(EK)^2}{24}\right]$$

where, n denotes the number of observations in the dataset, S represents the sample skewness and EK indicates the excess kurtosis:

$$Sample\ Skewness = \ S = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2)^{\frac{3}{2}}}$$

$$Sample\ Kurtosis = \ K = \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2)^2}$$

$$Excess\ Kurtosis = \ EK = K - 3$$

In the equations above, $\hat{\mu}_3$ and $\hat{\mu}_4$ denotes the estimates of third and fourth central moments, respectively, $\bar{x}$ denotes the sample mean, and $\hat{\sigma}^2$ indicates the estimate of the variance. If the data originates from a normal distribution, the JB test statistic asymptotically has a chi-squared distribution with two degrees of freedom. Thus the JB –statistic can be used to test the hypothesis of whether the data originates from a normal distribution. The null hypothesis is a joint hypothesis of the skewness and the excess kurtosis equaling zero. Samples from a normal distribution have an expected skewness and an expected excess kurtosis equal to zero. Therefore the null hypothesis of normality is rejected if the calculated test statistic exceeds the critical value from a $\chi_2^2 \sim$ distribution (Domański, 2010:76).

2.5.2 Ljung-Box-Pierce test

A set of m autocorrelations can be tested at once by making use of the Ljung-Box-Pierce test. Consider the following hypothesis test:

$$\boldsymbol{H_0}\!: \ \rho_k = 0 \quad for\ all\ k \leq m$$

$$\boldsymbol{H_1}\!: \ \rho_k \neq 0 \quad for\ some\ value\ of\ k\ where\ k \leq m$$

The Ljung-Box-Pierce test statistic is then defined as:

$$Q = T(T+2)\sum_{k=1}^{m} \frac{r_k^2}{T-k} \sim \chi_m^2$$

for each m, where m denotes the number of autocorrelations to be tested, T is the number of observations in a time series and r_k denotes the observed autocorrelations. The null hypothesis is rejected if the test statistic Q exceeds the critical value from a $\chi_m^2 \sim$ distribution. If the null hypothesis is rejected the conclusion is that the series is not random, in other words it is not a white noise process. However if the null hypothesis is not rejected there is some support that the series fits a white-noise model.

Autocorrelation is defined as the cross-correlation of an observation with itself at different points in time. Informally, it is the similarity between observations as a function of the time lag between them. It is a mathematical tool for discovering repeating patterns (Cryer & Chan, 2010:150).

2.6 INFORMATION CRITERION

Information criterions refer to a measure of the relative quality of a statistical model for a given set of data. Given a collection of models for the data, the various information criterions estimate the quality of each model, relative to each of the other models. Hence, information criterions are tools for appropriate model selection.

2.6.1 Akaike Information Criteria

Cavanaugh (2012:13-16) argues that the Akaike information criterion (AIC) represents a measure of the quality of statistical models for a given dataset. AIC is a means for model selection and it originated through information theory. Assume that there is a statistical model of some data. Let L denote the maximised value of the likelihood function of the model and k the number of estimated parameters in the model. The AIC of the model then follows as:

$$AIC = 2k - 2\ln(L)$$

Given a set of possible models for the given data, the preferred model is then the one with the smallest AIC value. Thus AIC rewards goodness of fit, but it also contains a penalty that is an increasing function of the number of estimated parameters, in other words the penalty discourages overfitting.

2.6.2 Bayes Information Criteria

Bayesian information criterion (BIC) or also known as the Schwarz criterion denotes a criterion for model selection between a finite set of models. Cavanaugh (2012:1) claims that the model with the smallest BIC value will be preferred. It is based on the likelihood function. The BIC of the model then follows as:

$$BIC = -2 \cdot \ln(L) + k \cdot ln(n)$$

where, $L \equiv$ maximised value of the likelihood function

$k \equiv$ the number of free parameters to be estimated

$n \equiv$ the number of observations

2.6.3 Shibata Information Criteria

Kadilar and Erdemir (2002:128) claims that Shibata investigated the asymptotic properties of Akaike's estimate and proved that the AIC does not produce a consistent estimate of the order of an autoregressive model. The Shibata Criterion of the model then follows as:

$$S_k = (1 + 2 \cdot \frac{dk+1}{n})^d \cdot |\textstyle\sum_k \cdot|$$

where, $\sum_k \cdot \equiv covariance\ matrix\ obtained\ from\ the\ least\ squares\ residuals$

$n \equiv$ the number of observations

$d \equiv$ dimension

$k \equiv$ number of orders of the series

2.6.4 Hannan-Quinn Information Criteria

Tu and Xu (2012:1) states that the Hannan–Quinn information criterion (HQC) denotes a criterion for model selection. The HQC of the model then follows as:

$$HQC = -2 \cdot L + 2 \cdot k \cdot log \cdot \log(n)$$

where, $L \equiv log - likelihood$

$k \equiv number\ of\ parameters$

$n \equiv number\ of\ observations$

2.7 SUMMARY

In this chapter, introductory concepts that are used in hedging strategies were discussed. The different hedging models were explained as well as a brief explanation of different summary statistics. Followed by two different goodness-of-fit tests and the different information criterions. In the chapters that follow, the literature of this chapter will be implemented. Chapter 3 entails the explanation of the research methodology, followed by the empirical results in Chapter 4. Chapter 5 concludes the research assignment and finishes off with some open questions.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 INTRODUCTION

This chapter entails the research methodology implemented to calculate the goodness-of-fit for the three different hedging models. The focus is primarily on the fitting of the hedging models and the estimation of their parameters, which enables the calculation of the respective information criterions for model selection. The proposed methodologies are applied to calculate the results indicated in Chapter 4.

3.2 DESCRIPTION OF DATA

In the sections that follow, the GOF of the various hedging models will be examined for a stock index and for currency futures. The stock index that will be considered is known as the Top 40 index. To emphasize the effectiveness of the copula-based GARCH model, it may be more interesting to consider an application for which the futures are less correlated with the underlying asset. Therefore, the performance of the various models are compared by cross hedging the currency exposure of holding the MSCI-Rand index with USD/ZAR futures.

All the data used in this research assignment was obtained from Bloomberg and Capital Synergy. The Top 40 data and its futures runs from 4 January 2010 to 31 December 2014. The MSCI-Rand and USD/ZAR data runs from 16 September 2013 to 19 August 2015. The futures comprise of 6-month futures contracts and are rolled over. The asset returns are represented by the changes in the logarithm of the daily closing prices.

3.3 R – PACKAGES USED IN THIS STUDY

The core packages necessary to conduct this research assignment are the following: zoo, rmgarch, timeSeries, fAssets, xlsx, FinTS, fGarch, MTS, DistributionUtils, copula, tseries, moments and urca. Next, an explanation of the essential packages are discussed.

3.3.1 Package: rmgarch

The rmgarch package provides various multivariate GARCH models with methods for fitting, filtering, forecasting and simulating. It consists of feasible multivariate GARCH models including DCC-GARCH, GO-GARCH and Copula-GARCH models. At present, the Dynamic Conditional Correlation (with multivariate Normal, Laplace and Student distribution) models are fully implemented, with methods for spec, fit, filter, forecast, simulation, and rolling estimation and forecasting. The Copula-GARCH model is implemented with the multivariate Normal and Student t distributions, with dynamic and static estimation of the correlation.

3.3.2 Package: fGarch

This package consists of econometric functions for modelling GARCH processes. Generalized Autoregressive Conditional Heteroskedastic (GARCH), models have become important in the analysis of time series data. For this purpose, the family of GARCH functions offers functions for simulating, estimating and forecasting various univariate GARCH-type time series models in the conditional variance and an Autoregressive Moving-average (ARMA) specification in the conditional mean. The function garchFit is a numerical implementation of the maximum log-likelihood approach under different assumptions like a normal or student-t distribution. Several diagnostic analysis tools check the parameter estimates.

3.3.3 Package: copula

The copula package provides classes of commonly used elliptical, archimedean, extreme value and other copula families. It also includes methods for density, distribution, random number generation, and plots. Fitting copula models and goodness-of-fit tests. Independence and serial (univariate and multivariate) independence tests, and other copula related tests.

3.3.4 Package: tseries

This package conducts time series analysis and computational finance. The Augmented Dickey–Fuller tests in this package are constructed by making use of the general regression equation that incorporates a constant and a linear trend and the t-statistic for a first order autoregressive coefficient that equals one is computed. The number of lags used in the regression is k. If the computed statistic is outside the table of critical values, then a warning message is generated. Missing values are not allowed.

3.3.5 Package: moments

This package contains the function to perform the Jarque-Bera test on a given data sample to determine if the data are samples drawn from a normal population. The Jarque-Bare test works as follows, under the hypothesis of normality, data should be symmetrical (i.e. skewness should be equal to zero) and have skewness close to three. Then the Jarque-Bera statistic is chi-square distributed with two degrees of freedom. Pearson’s measure of kurtosis is also available in this package as well as the function that computes the skewness of a given dataset.

3.3.6 Package: urca

This package conducts unit root and co-integration tests for time series data. It includes the Johansen test procedure for a given data set. The "trace" or "eigen" statistics are reported and the matrix of eigenvectors as well as the loading matrix.

3.4 SUMMARY STATISTICS, UNIT ROOT AND CO-INTEGRATION TESTS

Summary statistics like the Lagrange multiplier is calculated to detect any autocorrelation present in the data in order to select the appropriate ARMA model. The Ljung-Box Pierce test is implemented to check for ARCH effects and the Jarque-Bera test checks for the normality of the data. The Dicky-Fuller test indicates if there is a unit root present and therefore indicates whether the series is stationary or non-stationary. The Johansen test, tests for co-integration between the datasets. In this research assignment, the Johansen test is conducted between the Top 40 Index and its corresponding futures and between the MSCI-Rand Index and the USD/ZAR futures.

3.5 THE THREE MODELS AND THE INFORMATION CRITERIONS

The Constant Conditional Correlation-, Dynamic Conditional Correlation- and Copula Based-GARCH models are constructed by mainly using the rmgarch package. It also gets incorporated into the relevant model whether the conditional variances are constant or dynamic. In the case of the Copula-Based GARCH model the Gaussian copula will be used. However, two other copula functions are described in detail in the literature review that could be considered as opposed to using the Gaussian copula function.

Parameter estimation for each of the models are obtained through the process of maximum-likelihood. The various information criterions are also calculated for each model to use in the process of model selection. Lastly, for interest the copula functions are displayed in a 3-dimentional graph to see how using different copulas can make a difference in the results.

3.6 SUMMARY

This chapter discussed the methodology to implement the theory discussed in Chapter 2. The method to construct each model such as the CCC-GARCH, DCC-GARCH and Copula-Based GARCH models was discussed. Followed by the different information criterions used to compare the models with one another. In the next chapter the results will be represented and discussed.

CHAPTER 4
RESULTS

4.1 INTRODUCTION

This chapter represents and describes the results. The results represent the outcomes after the implementation of the research methodology captured in Chapter 3. In the following sections, the estimated parameters of the CCC-GARCH, DCC-GARCH and copula-based GARCH models are calculated. The summary statistics concerning each of the models are calculated, as well as the goodness-of-fit tests to enable comparisons.

4.2 SUMMARY STATISTICS AND GRAPHICAL PRESENTATION OF THE LOG-RETURN DATA

4.2.1 Summary Statistics

In this subsection the different summary statistics are computed to understand the data in such a way as to enable the estimation and selection of the most appropriate model.

Table 4.1 reports the results of the different summary statistics.

Summary Statistics				
	Assets			
	Top 40		MSCI-Rand	USD/ZAR
Statistics	**Stock**	**Futures**	**Spot**	**Futures**
Mean	0.000442322	0.000441907	0.000395164	0.00033063
Standard Deviation	0.01047332	0.01084127	0.009979288	0.00740623
Minimum	-0.03835729	-0.04276186	-0.0358143	-0.0263537
Maximum	0.04678554	0.04234158	0.05818463	0.02766695
Skewness	-0.04562985	-0.1224396	0.1178092	0.1984872
Kurtosis	4.518223	4.326269	5.749292	3.24209
Jarque-Bera	120.87*	95.04*	23.364*	12.757*
Q(24)	28.098	32.705	24.84	22.362
Q^2(24)	376.41*	378.29*	54.856*	63.621*
ARCH(5)	74.221*	81.365*	5.4051	81.365*

Note that the,∗, indicates significance at a 1% level. Statistical significance is attained when the p-value is less than the significance level α. The p-value is the probability of obtaining at least as extreme results given that the null hypothesis is true whereas alpha is the probability of rejecting the null hypothesis given that it is true.

Table 4. 1: Summary Statistics

There is some relatively small positive skewness in the Top 40 data and small negative skewness in the MSCI-Rand index and the USD/ZAR futures. The kurtosis for the Top 40 data and the MSCI-Rand is far from the value three indicating fatter tails than the normal distribution.

However, the kurtosis of the USD/ZAR futures is close to three. The Jarque-Bera statistics are all significant indicating that the data does not originate from a normal distribution. Therefore, concluding that the unconditional distributions of the spot and futures returns are asymmetric, fat-tailed, and non-Gaussian. The Ljung–Box tests indicate that there are no serial correlation present in any of the Top 40, USD/ZAR and MSCI-Rand returns. However, both the Ljung-Box test and the Lagrange multiplier statistics for the ARCH effects present strong autocorrelations in the squared returns for all assets. Thus, the data contains ARCH effects and has to be considered in the study, therefore the modelling of GARCH is selected.

Note that, $Q(24)$ denotes the Ljung–Box test statistic up to the 24th-order serial correlation in the returns, $Q^2(24)$ denotes the Ljung–Box test statistic for the serial correlations in the squared returns and ARCH(5) denotes the Lagrange multiplier test up to the fifth-order ARCH effects.

Table 4.2 reports the results of the unit roots and co-integration tests.

Unit Root and Co-integration Tests				
	Assets			
	Top 40		**MSCI-Rand**	**USD/ZAR**
Statistics	**Stock**	**Futures**	**Spot**	**Futures**
ADF (Price)	-2.7932	-2.8009	-2.4376	-2.9057
ADF (Returns)	-11.4*	-11.165*	-8.1512*	-7.7854*
Trace	80.9708	-	7.0161	-
λ	0.2308	-	2.3998	-

Note that the,∗, indicates significance at a 1% level. Statistical significance is attained when the p-value is less than the significance level α. The p-value is the probability of obtaining at least as extreme results given that the null hypothesis is true whereas alpha is the probability of rejecting the null hypothesis given that it is true.

Table 4. 2: Unit root and co-integration

The Augmented Dickey–Fuller tests illustrate that the spot and futures prices have a unit root, but taking a first-difference leads to stationarity, this is seen by the significance of the ADF(Returns) in the table above. The Johansen trace statistics shows that the spot and futures prices for stock markets, the MSCI-Rand index and the USD/ZAR futures are not co-integrated and therefore the error-correction term ought to be omitted. Knowing that the data is not co-integrated means there will not be any spurious problems when fitting the models and estimating their parameters.

The ADF tests are applied to test the null hypothesis of a unit root for the spot and futures prices and the returns. Trace denotes the Johansen trace test, with the null hypothesis being that there is no co-integration present in the data. The λ parameter denotes the estimated co-integration parameter.

4.2.2 Graphical Presentation of the data

The two graphs below illustrate the various indices and futures with their corresponding returns over the sample period. This is just to obtain a visual look at what the data looks like when the daily log-returns are plotted against time.

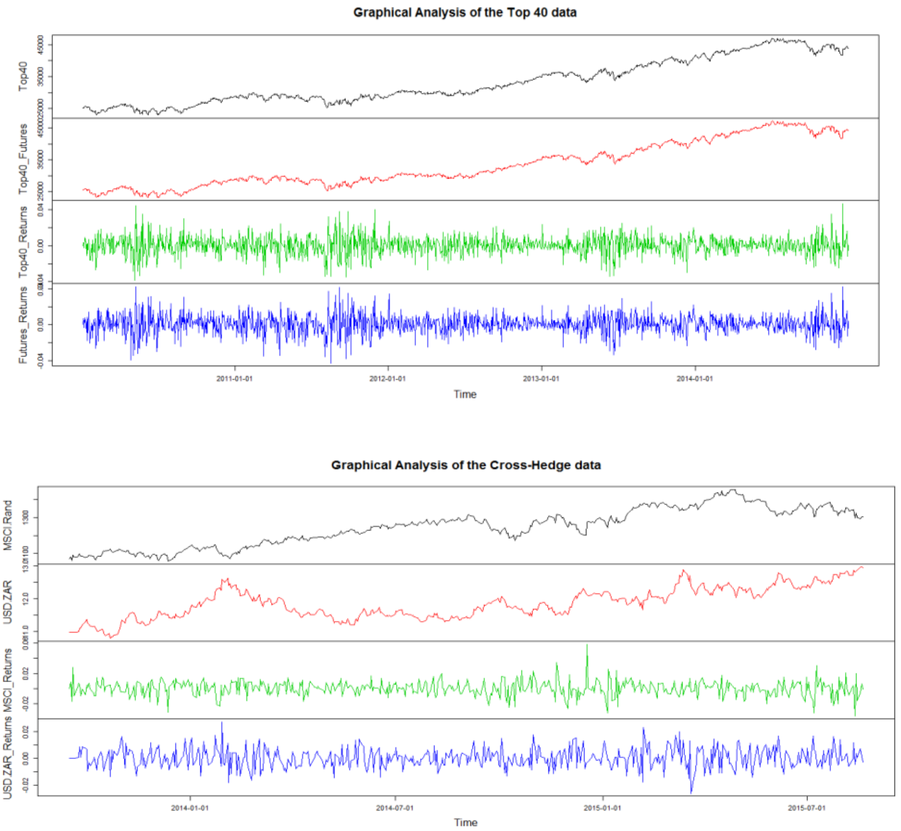

Figure 4. 1: Graphical analysis

The next graphs represents the scatter plots between the Top 40 index and the Top 40 futures and their returns. The graphs show that hedging a stock with its own futures directly represents a straight line, because spot and futures returns in the direct hedge are tied closely together by the no-arbitrage condition. Thus, concluding that the correlation between the Top 40 index and its futures equals one. Moreover, the correlation between the returns on the Top 40 index and its futures equals a value close to one. A positive correlation means that y tends to increase as x increases.

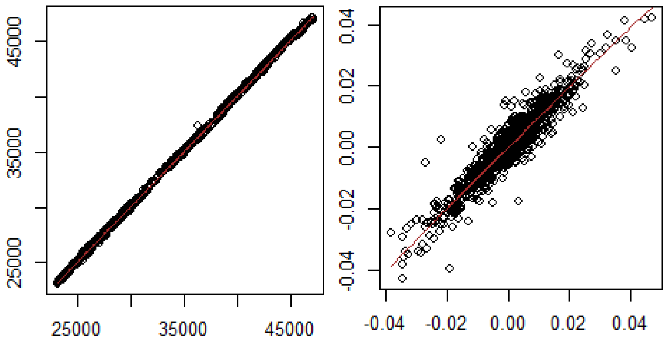

Figure 4. 2: Scatter plot of Top 40

Left: Scatter plot of the Top 40 Index and its futures

Right: Scatter plot of the returns on the Top 40 Index and its futures

Next follows the scatter plots between the MSCI-Rand index and the USD/ZAR futures. These scatter plots show however, that in the case of the cross hedge the spot and futures returns are less linearly correlated with one another. Therefore allowing asymmetric dependence, and implying dynamic hedging to be important consideration.

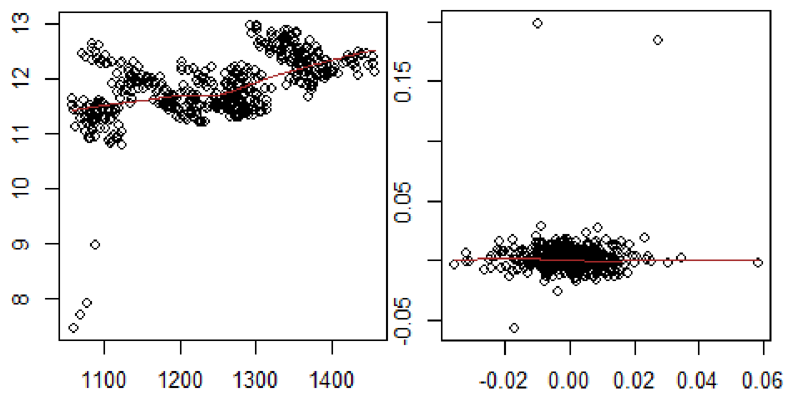

Figure 4. 3: Scatter plot of MSCI-Rand

Left: Scatter plot of the MSCI-Rand Index and the USD/ZAR futures

Right: Scatter plot of the returns on the MSCI-Rand Index and the USD/ZAR futures

4.3 ESTIMATED RESULTS FOR THE CCC-GARCH MODEL

Table 4.3 reports the results of the estimated parameters and the corresponding standard errors of the Top 40 index and its futures through the method of maximum likelihood for the Constant Conditional Correlation GARCH model.

Constant Conditional Correlation GARCH Model Estimations				
Parameters	**Estimate**	**Std. Error**	**t-value**	**Pr(> | t |)**
[Top 40].mu	0.000673	0.000264	2.550707	0.01075
[Top 40].ar1	-0.023099	0.028746	-0.803578	0.42164
[Top 40].omega	0.000002	0.000003	0.617635	0.53682

[Top 40].alpha1	0.074025	0.041479	1.784636	0.07432
[Top 40].beta1	0.908716	0.047037	19.319313	0
[Top 40 - Futures].mu	0.000678	0.000615	1.10284	0.2701
[Top 40 - Futures].ar1	-0.039339	0.02883	-1.364518	0.17241
[Top 40 - Futures].omega	0.000002	0.000024	0.071813	0.94275
[Top 40 - Futures].alpha1	0.065276	0.259414	0.251629	0.80133
[Top 40 - Futures].beta1	0.921507	0.29599	3.113306	0.00185
[Joint].C1	0.99	0.024501	40.406262	0
Log-Likelihood	8838.916			
Information Criteria				
Akaike	-14.08			
Bayes	-14.035			
Shibata	-14.08			
Hannan-Quinn	-14.063			

Table 4. 3: CCC-GARCH for Top 40 data

Important aspects to consider of the above table is the estimates of the standard errors. The smaller the standard errors of the model the better are the parameter estimations of the model. The standard error statistic measures the accuracy, dispersion and precision of the sample as an estimate of the population parameter. It is an important indicator of how reliable an estimate of the population parameter the sample statistic is.

For model selection, consult the various information criterions namely the Akaike, Bayes, Shibata and Hannan-Quinn. The largest negative value represents the more appropriate model to select. The log-likelihood values can be compared for each scenario for the different models. The larger the log-likelihood value of the model the more appropriate is that model. For the rest of the models and their parameter estimations please consult Appendix: A.

For example, investigating Table 4.3 and Table A.1 the following conclusions are made. The standard errors of each of the estimated parameters is smaller for the direct hedge using the CCC-GARCH model than that of the cross-hedge. This is also supported by the various information criterions, which are all greater for the cross-hedge model as opposed to the direct hedge model. Therefore the CCC-GARCH model is ultimately a more appropriate model for the direct-hedge.

Comparing overall the CCC-GARCH, DCC-GARCH and Copula-Based GARCH models the following conclusions are made. The standard errors of each of the estimated parameters is the smallest for the Copula-Based GARCH model with respect to the cross-hedge. However for the direct hedge the standard errors of the DCC-GARCH model is the smallest. Next, consider the information criterions of the models, the Copula-Based GARCH model is the most appropriate model with respect to the cross-hedge and the direct hedge since it has the smallest information criteria values.

In terms of model fitting, the log-likelihood functions in most of the copula-based GARCH models are greater than the log-likelihoods of the CCC-GARCH and DCC-GARCH models. In other words, allowing the dependence measure to be time-varying could be more crucial than permitting the dependence structure to be asymmetric, because spot and futures returns in the direct hedge are tied closely by their no-arbitrage condition.

However, the copula model has the highest log-likelihood for the cross hedge, in which spot and futures returns are less linearly correlated and thus allowing asymmetric dependence in dynamic hedging to be important.

4.4 COPULA GRAPHICS

The following figure illustrates the 3-dimensional view of a copula functions for the direct hedge between the Top 40 Index and its futures as well as their returns.

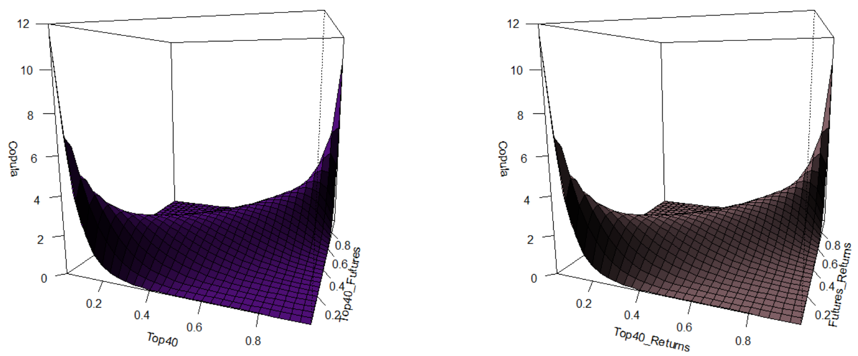

Figure 4. 4: Copula Functions

Left: Copula function for the Top 40 Index and its futures

Right: Copula function for the Top 40 Index returns and its futures returns

These graphs show that the Top 40 index is perfectly positively associated with the Top 40 futures. For further interest on 3-dimentional copula functions and the Gaussian copula, feel free to consult Appendix: C.

4.5 CONCLUSION

Taking account of the above results the Copula-Based GARCH model, which permits nonlinear and asymmetric dependence between the two assets in the cross-hedge portfolio, results in greater risk reduction. Therefore the model performs best if the futures are less correlated with the underlying asset. The outcomes for the Copula-Based GARCH model are better than the outcomes of the remaining models, but only by a slight amount thus making the choice of model depend on the specific problem at hand. The advantages the Copula-Based GARCH model has makes it a model to consider when hedging dynamically. Reflecting from the results above the DCC-GARCH model is more appropriate than the CCC-GARCH model. Concluding therefore that the more dynamic the model the better. But the outcomes of the three different hedging models are almost similar. Therefore making the choice of model depends on the specific problem and other aspects such as cost and time effectiveness of each model has to be taken into account.

CHAPTER 5
CONCLUSION AND OPEN QUESTIONS

There exists several different models and theories for implementing a dynamic hedging strategy. The Copula-Based GARCH model was implemented to verify whether tolerating a nonlinear and asymmetric dependence between two assets can be a more effective model than the CCC-GARCH and DCC-GARCH models.

Chapter 2 represented a literature review of the three models implemented, as well as some elementary discussions on some theory needed to understand the models. The three models had to be compared to one another. Therefore choosing the most appropriate model some goodness of fit tests and various information criterions were explained. Some attention was also given to different summary statistics. Chapter 3 represented the research methodology to implement the theory.

The results of the estimated parameters and standard errors of each model are represented in Chapter 4. The Copula-Based GARCH model performed the best for the cross-hedge portfolio and the DCC-GARCH model performed the best for the direct hedge portfolio. Although the results only differed slightly. Note that the statistical programming code used in this study was used and implemented independently by the author and can be viewed as ineffective programming. The code was transcribed in the statistical programming package called R and is denoted in Appendix D.

The Copula-Based GARCH model turned out to be the model most appropriate in the situation where futures are less correlated with their underlying assets. Thus leaving room for some open questions that can be asked to further future research. Is an even more dynamic approach better even if the results in this research assignment only differ slightly to those of a less dynamic approach? Is it better to implement the simplest model as opposed to the more complex models when the outcomes are very close to one another?

Concluding that the Copula-Based model does have an advantage, the fact to consider is this small advantage worth the complexity. The fact that the results differ by such a small amount between the various model makes model selection difficult. The best model suitable will depend on the specific situation at hand. It might be more cost effective and productive to implement the simpler models since the results are practically the same. Furthermore, extending the hedging procedure by making use of options can be a factor to consider in future research.

REFERENCES

Aas, K. 2004. Modelling the dependence structure of financial assets: A Survey of four copulas.

Baillie, R.T. & Myers, R.J. 1991. Bivariate GARCH estimation of the optimal commodity futures hedge. *Journal of Applied Econometrics*, 6(1):109-124.

Cavanaugh, J. 2012. The Akaike Information Criterion. Unpublished doctoral dissertation. United States: University of Lowa.

Cavanaugh, J. 2012. The Bayesian Information Criterion. Unpublished doctoral dissertation. United States: University of Lowa.

Cryer, J.D. & Chan, K. 2010. *Time Series Analysis with applications in R*. 2nd Ed. Springer.

Decarlo, L. 1997. On the meaning and use of kurtosis. *Psychological Methods*, 2(1):292-307.

Domański, C. 2010. Properties of the Jarque-Bera Test. Poland: University of Lodziensis.

Dragulescu, A. 2014. Package 'xlsx'. [Online]. Available: http://groups.google.com/group/R-package-xlsx [2015, May 30].

Engle, R. 2002. Dynamic Conditional Correlation – A Simple Class of Multivariate GARCH Models. *Forthcoming journal of Business and Economic Statistics.*

Engle, R. 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4):987-1007.

Ghalanos, A. 2015. Package 'rmgarch'. [Online]. Available: http://www.unstarched.net [2015, May 30].

Graves, S. 2009. Package 'FinTS'. [Online]. Available: http://faculty.chicagobooth.edu/ruey.tsay/teaching/bs41202/sp2009 [2015, May 30].

Greenbaum, M.C. & Ravindrant, K. 2002. Dynamic Hedging. *Colorado Springs Spring Meeting*, 1(27).

Gregory, K. B. 2011. The Gaussian Copula Model.

Groeneveld, R. & Meeden, G. 1984. Measuring Skewness and Kurtosis. *Journal of the Royal Statistical Society*, 33(1):391-399.

Hatemi-J, A. & Roca, E. 2006. Calculating the optimal hedge ratio: constant, time varying and the Kalman Filter approach. *Applied Economics Letters*, 13(1):293-299.

Hjalmarsson, E. & Ősterholm, P. 2007. Testing for Cointegration using the Johansen Methodology when variables are Near-Integrated. *IMF Working Paper*, 7(1):2-19.

Hofert, M. & Kojadinovic, I. & Maechler, M. & Yan, J. 2015. Package ‘copula’. [Online]. Available: http://nacopula.r-forge.r-project.org/ [2015, May 30].

Hsu, C. & Tseng, C. & Wang, Y. 2008. Dynamic hedging with futures: a copula-based GARCH model. *Journal of Futures Market*, 28(1):1095-1116.

Hull, J.C. 2012. *Options, Futures and other Derivatives.* 8th Ed. USA: Pearson Education, Inc.

Jouanin, J. & Riboulet, G. & Roncalli, T. & Opérationnelle, G. & Lyonnais, C. Financial Application of Copula Functions.

Kadilar, C. & Erdemir, C. 2002. Comparison of Performance among Information Criteria in VAR and Seasonal VAR models. *Journal of Mathematics and Statistics*, 31(1):127-137.

Komsta, L. & Novomestk. F. 2015. Package ‘moments’. [Online]. Available: http://www.r-project.org, http://www.komsta.net/ [2015, May 30].

Kotzé, A. A. 2010. Delta hedging: Futures vs Underlying spot*. Financial Chaos Theory.*

Koutmos, G. & Pericly, A. 1998. Dynamic Hedging of commercial paper with T-Bill futures*. The Journal of Futures Markets*, 18(18):925.

Lai, Y. & Chen, C. & Gerlach, R. 2009. Optimal dynamic hedging via copula-threshold-GARCH models. *Mathematics and Computers in Simulation*, 79(1):2609-2624.

Mahfoud, M. 2012. Bivariate Archimedean copulas: an application to two stock market indices. Unpublished doctoral dissertation. Amsterdam: Vrije Universiteit Amserdam.

Mili, M. & Abid, F. 2004. Optimal hedge ratios estimates: Static vs Dynamic hedging. *India Finance*, 18(1):655-670.

Nakatani, T & Teräsvirta, T. 2009. Testing for volatility interactions in the Constant Conditional Correlation GARCH model. *The Econometrics Journal*, 12(1):147-163.

Olivares, M & Garcia-Forero, C. 2010. Goodness-of-Fit Testing. Unpublished doctoral dissertation. Spain: University of Barcelona.

Patton, A. 2007. Copula-Based Models for Financial Time Series. Unpublished doctoral dissertation. England: University of Oxford.

Peters, T. 2008. Forecasting the covariance matrix with the DCC GARCH model. Unpublished doctoral dissertation. Stockholm: University of Stockholm.

Pfaff, B. & Stigler, M. 2013. Package ‘urca’. [Online]. Available: https://cran.r-project.org/package=urca [2015, May 30].

R Development Core Team. 2015.R. Edition 3.2.2 [Online]. Available: http://www.r-project.org/ [2015, May 30].

Rmetrics Core Team & Wuertz, D. & Setz, T. & Chalabi, Y. 2014. Package ‘fAssets’. [Online]. Available: https://www.rmetrics.org [2015, May 30].

Rmetrics Core Team. & Wuertz, D. & Setz, T. & Chalabi, Y. 2015. Package ‘timeSeries’. [Online]. Available: http://www.rmetrics.org [2015, May 30].

Scott, D. 2012. Package ‘DistributionUtils’. [Online]. Available: https://cran.r-project.org/package=DistributionUtils [2015, May 30].

Sharma, S.K. & Vaish, A.K. & Pandey, R. & Gupta, C. 2010. Dynamic Delta Hedging. *Asia-Pacific Business Review*, 6(4):139-154.

Sterman, J. 1984. Appropriate Summary Statistics for evaluating the historical fit of system dynamics models. Unpublished doctoral dissertation. Boston: Massachusetts Institute of Technology.

Szego, P.G. 2004. Risk Measures for the 21st Century. *John Wiley & Sons.*

Teräsvirta, T. 2012. Modelling conditional correlations of asset returns: A smooth transition approach. Unpublished doctoral dissertation. Denmark: University of Aarhus.

Tong, W.H. 1996. An examination of dynamic hedging. *Journal of International Money and Finance,* 15(1):19-35.

Trapletti, A. & Hornik, K. & LeBaron, B. 2015. Package ‘tseries’. [Online]. Available: https://cran.r-project.org/package=tseries [2015, May 30].

Trench, W. 1987. The Method of Lagrange Multipliers. Unpublished doctoral dissertation. Texas: University of Trinity.

Tsay, R. 2015. Package ‘MTS’. [Online]. Available: https://cran.r-project.org/web/packages/MTS/MTS.pdf [2015, May 30].

Tu, S. & Xu, L. 2012. A theoretical investigation of several model selection criteria for dimensionality reduction. Unpublished doctoral dissertation. Hong Kong: Chinese University of Hong Kong.

Weber, E. 2008. Structural Constant Conditional Correlation. *Discussion paper*, 15(1):1-12.

Wuertz, D. & Chalabi, Y. & Miklovic, M. & Boudt, C. & Chausse, P. 2013. Package ‘fGarch’. [Online]. Available: http://www.rmetrics.org [2015, May 30].

Zeileis, A. & Grothendieck, G. & Ryan, J. & Andrews, F. 2015. Package ‘zoo’. [Online]. Available: http://zoo.R-Forge.R-project.org/ [2015, May 30].

APPENDIX A: TABLES

Table A.1 reports the results of the estimated parameters and the corresponding standard errors of the MSCI-Rand index and the USD/ZAR futures through the method of maximum likelihood for the Constant Conditional Correlation GARCH model.

Constant Conditional Correlation GARCH Model Estimations				
Parameters	**Estimate**	**Std. Error**	**t-value**	**Pr(> | t |)**
[MSCI-RAND].mu	0.000512	0.000402	1.27362	0.202798
[MSCI-RAND].ar1	-0.055417	0.046761	-1.1851	0.235979
[MSCI-RAND].omega	0.000002	0.000004	0.43281	0.665152
[MSCI-RAND].alpha1	0.056765	0.049619	1.144	0.252624
[MSCI-RAND].beta1	0.928731	0.054034	17.18803	0
[USD/ZAR].mu	0.000304	0.000298	1.01988	0.307787
[USD/ZAR].ar1	-0.030701	0.045513	-0.67454	0.499969
[USD/ZAR].omega	0.000003	0.000002	1.68574	0.091845
[USD/ZAR].alpha1	0.063465	0.008365	7.58675	0
[USD/ZAR].beta1	0.890246	0.012949	68.7504	0
[Joint].C1	-0.101903	0.043386	-2.34876	0.018836
Log-Likelihood	3352.731			
Information Criteria				
Akaike	-13.421			
Bayes	-13.328			
Shibata	-13.422			
Hannan-Quinn	-13.384			

Table A. 1: CCC-GARCH Model Estimations of the MSCI-Rand index and USD/ZAR futures

Table A.2 reports the results of the estimated parameters and the corresponding standard errors of the Top 40 index and its futures through the method of maximum likelihood for the Dynamic Conditional Correlation GARCH model.

Dynamic Conditional Correlation GARCH Model Estimations				
Parameters	**Estimate**	**Std. Error**	**t-value**	**Pr(> | t |)**
[Top 40].mu	0.000673	0.000264	2.552285	0.010702
[Top 40].ar1	-0.023099	0.028869	-0.800136	0.423632
[Top 40].omega	0.000002	0.000003	0.625694	0.531516
[Top 40].alpha1	0.074025	0.041264	1.793924	0.072825
[Top 40].beta1	0.908716	0.046982	19.341997	0
[Top 40 - Futures].mu	0.000678	0.000612	1.107522	0.268068

[Top 40 - Futures].ar1	-0.039339	0.0288	-1.365917	0.171965
[Top 40 - Futures].omega	0.000002	0.000023	0.072299	0.942364
[Top 40 - Futures].alpha1	0.065276	0.25773	0.253273	0.800058
[Top 40 - Futures].beta1	0.921507	0.294337	3.130789	0.001743
[Joint].dcc(a1)	0.132187	0.029246	4.519878	0.000006
[Joint].dcc(b1)	0	0.114457	0.000001	0.999999
Log-Likelihood	9414.468			
Information Criteria				
Akaike	-14.994			
Bayes	-14.941			
Shibata	-14.995			
Hannan-Quinn	-14.974			

Table A. 2: DCC-GARCH Model Estimations for the Top 40 data

Table A.3 reports the results of the estimated parameters and the corresponding standard errors of the MSCI-Rand index and the USD/ZAR futures through the method of maximum likelihood for the Dynamic Conditional Correlation GARCH model.

Dynamic Conditional Correlation GARCH Model Estimations				
Parameters	**Estimate**	**Std. Error**	**t-value**	**Pr(> | t |)**
[MSCI-RAND].mu	0.000512	0.000401	1.27569	0.202066
[MSCI-RAND].ar1	-0.055417	0.046668	-1.18747	0.235041
[MSCI-RAND].omega	0.000002	0.000004	0.43295	0.665053
[MSCI-RAND].alpha1	0.056765	0.049597	1.14451	0.252412
[MSCI-RAND].beta1	0.928731	0.054036	17.18713	0
[USD/ZAR].mu	0.000304	0.000298	1.01905	0.308181
[USD/ZAR].ar1	-0.030701	0.044929	-0.68332	0.494405
[USD/ZAR].omega	0.000003	0.000002	1.67919	0.093115
[USD/ZAR].alpha1	0.063465	0.00838	7.5737	0
[USD/ZAR].beta1	0.890246	0.013	68.47995	0
[Joint].dcc(a1)	0.011194	0.018015	0.62138	0.534352
[Joint].dcc(b1)	0.94256	0.089328	10.5517	0
Log-Likelihood	3353.09			
Information Criteria				
Akaike	-13.414			
Bayes	-13.304			
Shibata	-13.415			
Hannan-Quinn	-13.371			

Table A. 4: DCC-GARCH Model Estimations for the MSCI-Rand index and USD/ZAR futures

Table A.4 reports the results of the estimated parameters and the corresponding standard errors of the Top 40 index and its futures through the method of maximum likelihood for the Copula-Based GARCH model.

Copula-Based GARCH Model Estimations				
Parameters	**Estimate**	**Std. Error**	**t-value**	**Pr(> | t |)**
[Top 40].mu	0.000693	0.000193	3.588822	0.000332
[Top 40].ar1	0.898576	0.034748	25.860086	0
[Top 40].omega	0.000002	0.000003	0.631603	0.527647
[Top 40].alpha1	0.07254	0.040621	1.785774	0.074136
[Top 40].beta1	0.910048	0.046285	19.661669	0
[Top 40 - Futures].mu	0.000694	0.000875	0.792859	0.42786
[Top 40 - Futures].ar1	0.895171	0.072813	12.294142	0
[Top 40 - Futures].omega	0.000002	0.000039	0.042361	0.966211
[Top 40 - Futures].alpha1	0.063581	0.427637	0.148681	0.881805
[Top 40 - Futures].beta1	0.923335	0.488083	1.891757	0.058523
[Joint].dcc(a1)	0.081923	0.070364	1.164276	0.244312
[Joint].dcc(b1)	0.021204	0.268833	0.078872	0.937134
[Joint].dcc(b1)	0.086431	0.067192	1.286341	0.198324
Log-Likelihood	9423.014			
Information Criteria				
Akaike	-15.005			
Bayes	-14.943			
Shibata	-15.005			
Hannan-Quinn	-14.982			

Table A. 5: Copula-Based GARCH Model Estimations for the Top 40 index and its futures

Table A.5 reports the results of the estimated parameters and the corresponding standard errors of the MSCI-Rand index and the USD/ZAR futures through the method of maximum likelihood for the Copula-Based GARCH model.

Copula-Based GARCH Model Estimations				
Parameters	**Estimate**	**Std. Error**	**t-value**	**Pr(> | t |)**
[MSCI-RAND].mu	0.000521	0.000402	1.27485	0.202364
[MSCI-RAND].ar1	-0.055417	0.046678	-1.18722	0.235139
[MSCI-RAND].omega	0.000002	0.000004	0.43291	0.66508
[MSCI-RAND].alpha1	0.056765	0.049614	1.14412	0.252574
[MSCI-RAND].beta1	0.928731	0.054048	17.1833	0
[USD/ZAR].mu	0.000304	0.000299	1.0184	0.308487
[USD/ZAR].ar1	-0.030701	0.045023	-0.68189	0.495311
[USD/ZAR].omega	0.000003	0.000002	1.68016	0.092926
[USD/ZAR].alpha1	0.063465	0.008376	7.57738	0
[USD/ZAR].beta1	0.890246	0.012991	68.52939	0

[Joint].dcc(a1)	0.0107	0.017208	0.62182	0.534064
[Joint].dcc(b1)	0.933018	0.148441	6.28544	0
[Joint].dcc(g1)	0.004759	0.02532	0.18797	0.850897
Log-Likelihood	3353.127			
Information Criteria				
Akaike	-13.414			
Bayes	-13.304			
Shibata	-13.415			
Hannan-Quinn	-13.371			

Table A. 6: Copula-Based GARCH Model Estimations for the MSCI-Rand index and USD/ZAR futures

APPENDIX B: COPULA GRAPHICS

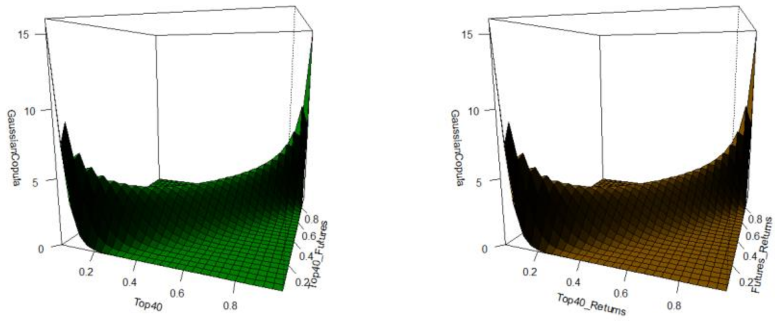

Figure B. 1: Gaussian copula functions

Left: Gaussian Copula function for the Top 40 index and its futures

Right: Gaussian Copula function for the Top 40 Index returns and its futures returns

The figures that follow illustrates the copula functions and Gaussian-copula for the cross-hedge between the MSCI-Rand Index and the USD/ZAR futures.

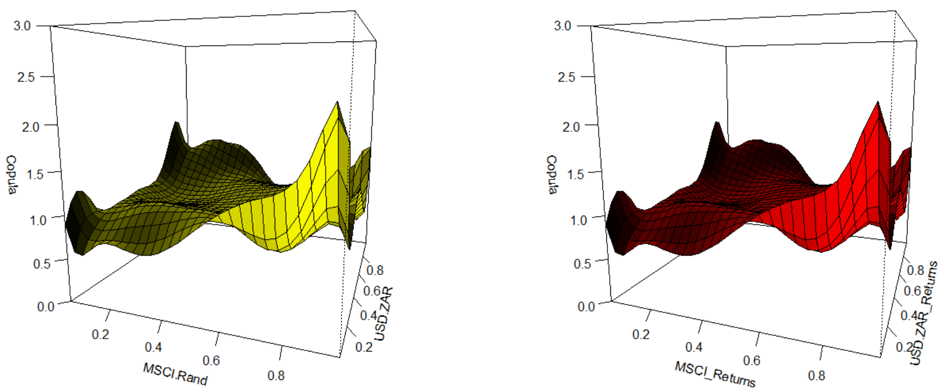

Figure B. 2: Cross-Hedge copula functions

Left: Copula function for the MSCI-Rand index and the USD/ZAR futures

Right: Copula function for the MSCI-Rand Index returns and the USD/ZAR futures returns

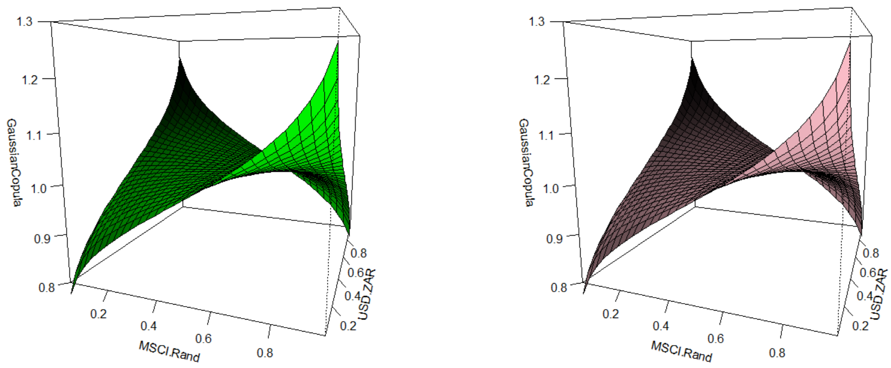

Figure B. 3: Gaussian copula for Cross-Hedge

Left: Gaussian Copula function for the MSCI-Rand index and the USD/ZAR futures

Right: Gaussian Copula function for the MSCI-Rand Index returns and the USD/ZAR futures returns

APPENDIX C: REFERENCE TO SUMMARY STATISTICS

C.1 Mean

When used without conditions, the mean refers to the arithmetic average of a data set. To calculate the mean, add all the values in the data set and divide by the number of observations.

We differentiate between a population mean and the sample mean. The population mean μ is given by the following: $\mu = \frac{1}{N}\sum x_i$

where, $\sum x_i$ denotes the sum of all the values in the population and N denotes the population size.

The sample mean $\bar{x}$ is given by: $\bar{x} = \frac{1}{n}\sum x_i$

The formulas for the population mean and the sample mean are practically identical. The only difference is whether the data that is used represents the entire population or a sample of the population. In practice, working with a sample and not the entire population is more constructive (Sterman, 1984, p.52-54).

C.2 Variance and Standard Deviation

The standard deviation is the greatest common measure of spread. A deviation represents the difference between a value and the mean.

The population variance follows as: $\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{N}$

The sample variance follows as: $s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$

In the sample variance the denominator is $n - 1$ instead of n. The reason for this is because when n is large, $n - 1 \approx n$, so the numerical outcomes from the two formulas will be similar. However, when n is small, the sample variance will provide a bigger outcome than the population variance. And this is necessary to derive an unbiased estimate for the population variance.

The sample standard deviation is the square root of the variance:

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

(Sterman, 1984, p.52-54).

C.3 Kurtosis

Kurtosis provides a measure of the thickness in the tails of a probability density function. The normal distribution has a kurtosis value equal to three (Domański, 2010, p.77). In symmetric distributions, positive kurtosis suggests heavy tails and peaked-ness, whereas negative kurtosis indicates light tails and flatness (DeCarlo, 1997, p.292). Kurtosis, defined as the standardised fourth population moment about the mean is given by:

$$\beta_2 = \frac{E(X-\mu)^4}{(E(X-\mu)^2)^2} = \frac{\mu_4}{\sigma^4}$$

where, E denotes the expectation operator, μ is the mean, μ_4 represents the fourth moment about the mean and σ is the standard deviation. The normal distribution has a kurtosis of three, and $\beta_2 - 3$ is regularly used so that the reference normal distribution has a zero kurtosis (DeCarlo, 1997, p.292). A sample equivalent to β_2 is achieved by substituting the population moments with the sample moments, which then gives:

$$b_2 = \frac{\sum(X_i - \bar{X})^4/n}{(\sum(X_i - \bar{X})^2/n)^2}$$

where, b_2 denotes the sample kurtosis, $\bar{X}$ is the sample mean and n is the number of observations.

Tailed-ness and peaked-ness are both components of kurtosis since kurtosis represents the movement of a mass that does not affect the variance. Consider the case of positive kurtosis, where a higher peak frequently accompanies heavier tails. Note that if the mass is moved from the shoulders of a distribution to its tails, then the variance will be bigger. To leave the variance unchanged, mass has to be moved from the shoulders to the centre, which gives a rewarding decrease in the variance and a peak. For negative kurtosis, the variance will be unchanged if mass is moved from the tails and centre of the distribution to its shoulders, resulting in light tails and flatness. It should be acknowledged that even though tailed-ness and peaked-ness are often both components of kurtosis, kurtosis can also imitate the effect of mainly one of these components, such as heavy tails (DeCarlo, 1997, p.292).

C.4 Skewness

Skewness provides a measure of how symmetric the observations are about the mean. The normal distribution has skewness value equal to zero. A distribution that is skewed to the right is known to have positive skewness and a distribution that is skewed to the left has negative skewness (Domański, 2010, p.77).

Skewness describes asymmetry from the normal distribution. There is positive and negative skewness, depending on whether the data points are skewed to the left or to the right of average of the dataset. Most datasets, including stock prices and asset returns, have either positive or negative skewness rather than following the balanced normal distribution, which has a skewness of zero. By knowing which way data is skewed, better estimations can be obtained of whether a data point will be more or less than the mean. The skewness of a random variable X is often measured by the standardised third central moment given by:

$$\beta_2 = \frac{\mu_3}{\sigma^3}$$

The value of this measure can become arbitrarily large, making it difficult to interpret (Groeneveld and Meeden, 1984, p.391).

APPENDIX D: R CODE

```
library(zoo)
library(rmgarch)
library(timeSeries)
library(fAssets)
library(xlsx)
library(FinTS)
library(fGarch)
library(MTS)
library(DistributionUtils)
library(copula)
library(tseries)
library(moments)
library(urca)

Return_Data <- read.table(file = "F:\\Research Assign-
ment\\DATA\\Return_Data.txt",header=TRUE)
Cross_Hedge <- read.table(file = "F:\\Research Assign-
ment\\DATA\\Cross_Hedge_Data.txt",header=TRUE)

model.1 <- as.timeSeries(Return_Data)
plot(model.1,main="Graphical Analysis of the Top 40 data")
model.2 <- as.timeSeries(Cross_Hedge)
plot(model.2,main="Graphical Analysis of the Cross-Hedge data")

par(mfrow = c(1,2))
assetsHistPlot(model.1[,1])
assetsHistPlot(model.1[,2])
assetsHistPlot(model.1[,3])
assetsHistPlot(model.1[,4])
assetsHistPlot(model.2[,1])
assetsHistPlot(model.2[,2])
assetsHistPlot(model.2[,3])
assetsHistPlot(model.2[,4])

assetsCorTestPlot(model.1[,1:2])
assetsCorTestPlot(model.1[,3:4])
```

```
assetsCorTestPlot(model.2[,1:2])
assetsCorTestPlot(model.2[,3:4])

assetsCorgramPlot(model.1[,1:2])
assetsCorgramPlot(model.1[,3:4])
assetsCorgramPlot(model.2[,1:2])
assetsCorgramPlot(model.2[,3:4])

assetsCorImagePlot(model.1[,1:2])
assetsCorImagePlot(model.1[,3:4])
assetsCorImagePlot(model.2[,1:2])
assetsCorImagePlot(model.2[,3:4])
######################################################################
# plot returns with squared and absolute returns

par(mfrow = c(1,1))
dataToPlot = cbind(Return_Data[,4], (Return_Data[,4])^2,
abs(Return_Data[,4]))
colnames(dataToPlot) = c("Returns", "Returns^2", "abs(Returns)")
plot.zoo(dataToPlot, main="Top 40 Daily Returns", col="blue")

dataToPlot = cbind(Return_Data[,5], (Return_Data[,5])^2,
abs(Return_Data[,5]))
colnames(dataToPlot) = c("Returns", "Returns^2", "abs(Returns)")
plot.zoo(dataToPlot, main="Top40 Futures Daily Returns", col="purple")

dataToPlot = cbind(Cross_Hedge[,4], (Cross_Hedge[,4])^2,
abs(Cross_Hedge[,4]))
colnames(dataToPlot) = c("Returns", "Returns^2", "abs(Returns)")
plot.zoo(dataToPlot, main="MSCI-RAND Index Daily Returns",
col="brown")

dataToPlot = cbind(Cross_Hedge[,5], (Cross_Hedge[,5])^2,
abs(Cross_Hedge[,5]))
colnames(dataToPlot) = c("Returns", "Returns^2", "abs(Returns)")
plot.zoo(dataToPlot, main="USD-ZAR Futures Daily Returns", col="red")
######################################################################
# Summary Statistics:
```

```
 Mean1 <- mean(Return_Data[,4])
 Mean2 <- mean(Return_Data[,5])
 Mean3 <- mean(Cross_Hedge[,4])
 Mean4 <- mean(Cross_Hedge[,5])

 SD1 <- sd(Return_Data[,4])
 SD2 <- sd(Return_Data[,5])
 SD3 <- sd(Cross_Hedge[,4])
 SD4 <- sd(Cross_Hedge[,5])

 Min1 <- min(Return_Data[,4])
 Min2 <- min(Return_Data[,5])
 Min3 <- min(Cross_Hedge[,4])
 Min4 <- min(Cross_Hedge[,5])

 Max1 <- max(Return_Data[,4])
 Max2 <- max(Return_Data[,5])
 Max3 <- max(Cross_Hedge[,4])
 Max4 <- max(Cross_Hedge[,5])

 Skew1 <- skewness(Return_Data[,4])
 Skew2 <- skewness(Return_Data[,5])
 Skew3 <- skewness(Cross_Hedge[,4])
 Skew4 <- skewness(Cross_Hedge[,5])

 Kurt1 <- kurtosis(Return_Data[,4])
 Kurt2 <- kurtosis(Return_Data[,5])
 Kurt3 <- kurtosis(Cross_Hedge[,4])
 Kurt4 <- kurtosis(Cross_Hedge[,5])

 JB1 <- jarque.bera.test(Return_Data[,4])
 JB2 <- jarque.bera.test(Return_Data[,5])
 JB3 <- jarque.bera.test(Cross_Hedge[,2])
 JB4 <- jarque.bera.test(Cross_Hedge[,3])

#Ljung-Box Tests:
#Top40
Box.test(Return_Data[,4], type="Ljung-Box", lag = 24)
```

```
Box.test(((Return_Data[,4])^2), type="Ljung-Box", lag = 24)
#Top40_Futures
Box.test(Return_Data[,5], type="Ljung-Box", lag = 24)
Box.test(((Return_Data[,5])^2), type="Ljung-Box", lag = 24)
#MSCI-Rand
Box.test(Cross_Hedge[,4], type="Ljung-Box", lag = 24)
Box.test(((Cross_Hedge[,4])^2), type="Ljung-Box", lag = 24)
#USD/RAND-Futures
Box.test(Cross_Hedge[,5], type="Ljung-Box", lag = 24)
Box.test(((Cross_Hedge[,5])^2), type="Ljung-Box", lag = 24)
# use ArchTest() function from FinTS package for Engle's LM test
# Top40
ArchTest(Return_Data[,4],lags=5)
#Top40-Futures
ArchTest(Return_Data[,5],lags=5)
#MSCI-RAND
ArchTest(Cross_Hedge[,4],lags=5)
#USD/RAND-Futures
ArchTest(Return_Data[,5],lags=5)
###############################################################
#Top40(price)
DF1 <- adf.test(Return_Data[,2])
ur.df(Return_Data[,2])
#Top40(returns)
DF2 <- adf.test(Return_Data[,4])
ur.df(Return_Data[,4])
#Top40_Futures(price)
DF3 <- adf.test(Return_Data[,3])
ur.df(Return_Data[,3])
#Top40_Futures(returns)
DF4 <- adf.test(Return_Data[,5])
ur.df(Return_Data[,5])
#MSCI(price)
DF5 <- adf.test(Cross_Hedge[,2])
ur.df(Cross_Hedge[,2])
#MSCI(returns)
DF6 <- adf.test(Cross_Hedge[,4])
ur.df(Cross_Hedge[,4])
```

```
#USD/ZAR(price)
DF7 <- adf.test(Cross_Hedge[,3])
ur.df(Cross_Hedge[,3])
#USD/Rnd(returns)
DF8 <- adf.test(Cross_Hedge[,5])
ur.df(Cross_Hedge[,5])
######################################################################
#Johansen Tests:

J1 <- ca.jo(Return_Data[,2:3],type = c("eigen", "trace"))
J2 <- ca.jo(Cross_Hedge[,2:3],type = c("eigen", "trace"))
######################################################################
#CCC Moddelling:
uspec.n = multispec(replicate(2, ugarchspec(mean.model =
list(armaOrder = c(1,0)))))
spec.cccn = cgarchspec(uspec.n, dccOrder = c(1, 1), distribu-
tion.model= list( copula = "mvnorm", method = "ML", time.varying =
FALSE))
fit.1 = cgarchfit(spec.cccn, data = Return_Data[,4:5], solver =
'solnp',fit.control = list(eval.se = TRUE))
resi1=residuals(fit.1)
covariance <- cov(resi1)

uspec.n = multispec(replicate(2, ugarchspec(mean.model =
list(armaOrder = c(1,0)))))
spec.cccn = cgarchspec(uspec.n, dccOrder = c(1, 1), distribu-
tion.model= list( copula = "mvnorm", method = "ML", time.varying =
FALSE))
fit.2 = cgarchfit(spec.cccn, data = Cross_Hedge[,4:5], solver =
'solnp',fit.control = list(eval.se = TRUE))
resi2=residuals(fit.2)
covariance_2 <- cov(resi2)

#DCC Moddelling:
#DCC with multivariate normal distribution:
uspec.n = multispec(replicate(2, ugarchspec(mean.model =
list(armaOrder = c(1,0)))))
spec.dccn = dccspec(uspec.n, dccOrder = c(1, 1),model="DCC", distribu-
tion ='mvnorm')
fit.1 = dccfit(spec.dccn, data = Return_Data[,4:5], solver =
'solnp',fit.control = list(eval.se = TRUE))
```

```
resi1=residuals(fit.1)
plot(fit.1)
covariance <- cov(resi1)
optimal_hedge_ratio_1 <- (0.0001073814/0.0001174076)

uspec.n = multispec(replicate(2, ugarchspec(mean.model =
list(armaOrder = c(1,0)))))
spec.dccn = dccspec(uspec.n, dccOrder = c(1, 1),model="DCC", distribu-
tion ='mvnorm')
fit.2 = dccfit(spec.dccn, data = Cross_Hedge[,4:5], solver =
'solnp',fit.control = list(eval.se = TRUE))
resi2=residuals(fit.2)
plot(fit.2)
covariance_2 <- cov(resi2)
optimal_hedge_ratio_2 <- (0.000002677817/0.0002083352)

# Copula Moddelling
uspec.n = multispec(replicate(2, ugarchspec(mean.model =
list(armaOrder = c(1,0)))))
spec.cbgn = cgarchspec(uspec.n, dccOrder = c(1, 1),asymmetric = TRUE,
distribution.model = list(copula ="mvnorm",
method ="ML", time.varying = TRUE))
fit.1 = cgarchfit(spec.cbgn, data = Return_Data[,4:5], solver =
'solnp',fit.control = list(eval.se = TRUE))
resi1=residuals(fit.1)
uspec.n = multispec(replicate(2, ugarchspec(mean.model =
list(armaOrder = c(1,0)))))
spec.cbgn = cgarchspec(uspec.n, dccOrder = c(1, 1),asymmetric = TRUE,
distribution.model = list(copula ="mvnorm",
method ="ML", time.varying = TRUE))
fit.2 = cgarchfit(spec.cbgn, data = Cross_Hedge[,4:5], solver =
'solnp',fit.control = list(eval.se = TRUE))
resi2=residuals(fit.2)

fit1 = garchFit(formula = ~arma(1,1)+garch(1,1),data=Return_Data[,4],
cond.dist ="std")
fit2 = garchFit(formula = ~ arma(1,1)+garch(1,1),data=Return_Data[,5],
cond.dist ="std")
resid1 <- residuals(fit1)
resid2 <- residuals(fit2)
cov(resid1,resid2)
```

```
var(resid2)
optimal_hedge_ratio_1 <- (0.0001064253/0.0001163693)

dat_res <- cbind(resid1,resid2)
m_res <- apply(dat_res, 2, mean)
v_res <- apply(dat_res, 2, var)
dat_res_std =cbind((dat_res[,1]-m_res[1])/sqrt(v_res[1]),(dat_res[,2]-
m_res[2])/sqrt(v_res[2]))

fit3 = garchFit(formula = ~ arma(1,1)+garch(1,1),data=Cross_Hedge[,4],
cond.dist ="std")
fit4 = garchFit(formula = ~ arma(1,1)+garch(1,1),data=Cross_Hedge[,5],
cond.dist ="std")
resid3 <- residuals(fit3)
resid4 <- residuals(fit4)
cov(resid3,resid4)
var(resid4)
optimal_hedge_ratio_2 <- (0.00000215055/0.0002103779)

dat_res1 <- cbind(resid3,resid4)
m_res1 <- apply(dat_res1, 2, mean)
v_res1 <- apply(dat_res1, 2, var)
dat_res_std1 =cbind((dat_res1[,1]-
m_res1[1])/sqrt(v_res1[1]),(dat_res1[,2]-m_res1[2])/sqrt(v_res1[2]))

fix(copula_1)
function ()
{uv = dat_res_std[,1:2]
n = nrow(uv)
uv = cbind(rank(uv[,1]),rank(uv[,2]))/(n+1)
xy = qnorm(uv)
s = 0.3
dc = Vectorize(function(x,y)
mean(dnorm(rep(qnorm(x),n),xy[,1],s)*dnorm(rep(qnorm(y),n),xy[,2],s))/
dnorm(qnorm(x))/dnorm(qnorm(y)))
vx = seq(1/30,29/30,by=1/30)
vz = outer(vx,vx,dc)
zl = c(0,12)
```

```
persp(vx,vx,vz,theta=20,phi=10,col="purple",shade=TRUE,xlab=names(Retu
rn_Data)[2],ylab=names(Return_Data)[3],zlab="Copula",ticktype="detaile
d",zlim=zl)}


fix(copula_2)
function ()
{uv = dat_res_std[,1:2]
n = nrow(uv)
uv = cbind(rank(uv[,1]),rank(uv[,2]))/(n+1)
xy = qnorm(uv)
s = 0.3
dc = Vectorize(function(x,y)
mean(dnorm(rep(qnorm(x),n),xy[,1],s)*dnorm(rep(qnorm(y),n),xy[,2],s))/
dnorm(qnorm(x))/dnorm(qnorm(y)))
vx = seq(1/30,29/30,by=1/30)
vz = outer(vx,vx,dc)
zl = c(0,12)
persp(vx,vx,vz,theta=20,phi=10,col="pink",shade=TRUE,xlab=names(Return
_Data)[4],ylab=names(Return_Data)[5],zlab="Copula",ticktype="detailed"
,zlim=zl)}


fix(copula_3)
function()
{uv = dat_res_std[,1:2]
n = nrow(uv)
uv = cbind(rank(uv[,1]),rank(uv[,2]))/(n+1)
norm.cop <- normalCopula(0.5)
norm.cop <- normalCopula(fitCopula(norm.cop,uv)@estimate)
dc = function(x,y) dCopula(cbind(x,y), norm.cop)
vx = seq(1/30,29/30,by=1/30)
vz = outer(vx,vx,dc)
zl = c(0,16)
persp(vx,vx,vz,theta=20,phi=10,col="green",shade=TRUE,xlab=names(Retur
n_Data)[2],ylab=names(Return_Data)[3],zlab="GaussianCopula",ticktype="
detailed",zlim=zl)}


fix(copula_4)
function()
{uv = dat_res_std[,1:2]
n = nrow(uv)
```

```
uv = cbind(rank(uv[,1]),rank(uv[,2]))/(n+1)
norm.cop <- normalCopula(0.5)
norm.cop <- normalCopula(fitCopula(norm.cop,uv)@estimate)
dc = function(x,y) dCopula(cbind(x,y), norm.cop)
vx = seq(1/30,29/30,by=1/30)
vz = outer(vx,vx,dc)
zl = c(0,16)
persp(vx,vx,vz,theta=20,phi=10,col="orange",shade=TRUE,xlab=names(Return_Data)[4],ylab=names(Return_Data)[5],zlab="GaussianCopula",ticktype="detailed",zlim=zl)}
######################################################################
fix(copula_1.1)
function ()
{uv = dat_res_std1[,1:2]
n = nrow(uv)
uv = cbind(rank(uv[,1]),rank(uv[,2]))/(n+1)
xy = qnorm(uv)
s = 0.3
dc = Vectorize(function(x,y)
mean(dnorm(rep(qnorm(x),n),xy[,1],s)*dnorm(rep(qnorm(y),n),xy[,2],s))/
dnorm(qnorm(x))/dnorm(qnorm(y)))
vx = seq(1/30,29/30,by=1/30)
vz = outer(vx,vx,dc)
zl = c(0,3)
persp(vx,vx,vz,theta=20,phi=10,col="yellow",shade=TRUE,xlab=names(Cross_Hedge)[2],ylab=names(Cross_Hedge)[2],zlab="Copula",ticktype="detailed",zlim=zl)}


fix(copula_2.1)
function ()
{uv = dat_res_std1[,1:2]
n = nrow(uv)
uv = cbind(rank(uv[,1]),rank(uv[,2]))/(n+1)
xy = qnorm(uv)
s = 0.3
dc = Vectorize(function(x,y)
mean(dnorm(rep(qnorm(x),n),xy[,1],s)*dnorm(rep(qnorm(y),n),xy[,2],s))/
dnorm(qnorm(x))/dnorm(qnorm(y)))
vx = seq(1/30,29/30,by=1/30)
vz = outer(vx,vx,dc)
```

```
zl = c(0,3)
persp(vx,vx,vz,theta=20,phi=10,col="red",shade=TRUE,xlab=names(Cross_Hedge)[4],ylab=names(Cross_Hedge)[5],zlab="Copula",ticktype="detailed",zlim=zl)}


fix(copula_3.1)
function()
{uv = dat_res_std1[,1:2]
n = nrow(uv)
uv = cbind(rank(uv[,1]),rank(uv[,2]))/(n+1)
norm.cop <- normalCopula(0.5)
norm.cop <- normalCopula(fitCopula(norm.cop,uv)@estimate)
dc = function(x,y) dCopula(cbind(x,y), norm.cop)
vx = seq(1/30,29/30,by=1/30)
vz = outer(vx,vx,dc)
zl = c(0.7,1.4)
persp(vx,vx,vz,theta=20,phi=10,col="green",shade=TRUE,xlab=names(Cross_Hedge)[2],ylab=names(Cross_Hedge)[3],zlab="GaussianCopula",ticktype="detailed",zlim=zl)}
```