An ICA-GARCH Approach to Computing Portfolio VaR with Applications to South African Financial Markets

Victor Mombeyarara

Supervisor: Dr Blessing Mudavanhu

Master of Management in Finance & Investment

Faculty of Commerce Law and Management

Wits Business School

University of The Witwatersrand

2016

Thesis submitted in fulfilment of the requirements for the degree of

Master of Management in Finance and Investment

FACULTY OF COMMERCE LAW AND MANAGEMENT

WITS BUSINESS SCHOOL

UNIVERSITY OF THE WITWATERSRAND

Supervisor: Signature:

DECLARATION

I, ***Victor Mombeyarara***, solemnly declare that the research work reported in this thesis is my own, except where otherwise indicated and acknowledged. It is submitted for the degree of Master of Management in Finance and Investment at the University of the Witwatersrand, Johannesburg. This thesis has not, either in whole or in part, been submitted for a degree or diploma to any other university.

Victor Mombeyarara

Signed at-------------------------------------On the-----------Day of----------------------------2017

Abstract

The Value-at-Risk (VaR) measurement – which is a single summary, distribution independent statistical measure of losses arising as a result of market movements – has become the market standard for measuring downside risk. There are some diverse ways to computing VaR and with this diversity comes the problem of determining which methods accurately measure and forecast Value-at-Risk. The problem is two-fold. First, what is the distribution of returns for the underlying asset? When dealing with *linear* financial instruments – where the relationship between the return on the financial asset and the return on the underlying is linear– we can assume normality of returns. This assumption becomes problematic for *non-linear* financial instruments such as options. Secondly, there are different methods of measuring the volatility of the underlying asset. These range from the univariate GARCH to the multivariate GARCH models. Recent studies have introduced the Independent Component Analysis (ICA) GARCH methodology which is aimed at computational efficiency for the multivariate GARCH methodologies. In our study, we focus on *non-linear* financial instruments and contribute to the body of knowledge by determining the optimal combination for the measure for volatility of the underlying (univariate-GARCH, EWMA, ICA-GARCH) and the distributional assumption of returns for the financial instrument (assumption of normality, the Johnson translation system). We use back-testing and out-of-sample tests to validate the performance of each of these combinations which give rise to six different methods for value-at-risk computations.

Dedication

This research report is dedicated to

Tinaye Gabriel and Takudzwa Michael, my first-born twins…grow to be as diligent and intelligent as your father!!

My Father: Maita Sinyoro, Zvaitwa Muroro!!

The Mombeyarara family: You inspired me to reach for the stars, if I fall I will land on the clouds!!

Acknowledgements

It is said that, "in the end you start thinking about the beginning". I look back with pride at the decision I made to study for the Master of Management in Finance and Investment (MMFI) programme. Now I have not simply become better educated, but better rounded – a "whole" person – and I am inspired to change my community by more than accumulating knowledge but applying it.

Particularly, I would like to thank the following people

Dr. Blessing Mudavanhu, my supervisor, who inspired me to look at simple but efficient solutions to solving complex problems. I wish we could share your simplifying view of the mathematical world with the younger generation. That way we could have more people prepared and ready to tackle complex problems in the world.

My Family; all I can say is that I wouldn't have made it without your continued encouragement. You guys are the best, no man ever had a more loving family.

My friends. I met some wonderful people in the MMFI class of 2016. You have all inspired me to carry the burden of solving the world's problems with pride. If we don't it, no one else will. It's been an amazing journey.

Table of Contents

List of Tables

List of Figures

List of Acronyms

ARCH Auto-Regressive Conditional Heteroscedasticity

AUD Australian Dollar

BEKK Baba-Engle-Kraft-Kroner

BRW (Boudoukh, Richardson, Whitelaw)

CVAR Conditional Value-at-Risk

DCC Dynamic Conditional Correlation

EGARCH Exponential Generalised Auto-Regressive Conditional Heteroscedasticity

ES Expected Shortfall

EVD Eigenvalue Decomposition

EVT Extreme Value Theory

EWMA Exponentially Weighted Moving Average

FHS (filtered historical simulations)

FTSE Financial Times Stock Exchange

GARCH Generalised Auto-Regressive Conditional Heteroscedasticity

ICA Independent Component Analysis

IGARCH Integrated Generalised Auto-Regressive Conditional Heteroscedasticity

JSE Johannesburg Stock Exchange

LTCM Long Term Capital Management

MGARCH Multivariate Generalised Auto-Regressive Conditional Heteroscedasticity

MMFI Master of Management in Finance and Investment

P/L Profit or Loss

PBH Pickans-Balkema-de Haan

PCA Principal Component Analysis

SWARCH Switching Regime Auto-Regressive Conditional Heteroscedasticity

TVaR Tail Value-at-Risk

USD United States Dollar

VaR Value-at-Risk

VEC Vector Error Correction

ZAR South African Rand

1 Introduction

The Value at Risk (VaR) measurement is a risk assessment technique which generalises the likelihood of underperforming by providing a statistical measure of downside risk. VaR assesses the potential losses on a portfolio over a given future period of time with a given degree of confidence. (ActEd Financial Economics Notes, 2016).

The recent past has seen regulatory requirements being geared almost exclusively toward a Value at Risk (VaR) concept as a measure of downside market risk. (Xu and Wirjanto, 2013). This has led to an increasing need for more efficient and more accurate methods of measuring and forecasting VaR for investment portfolio risk management.

1.1 Objectives of the Research

Previous research focusing on VaR modelling in South African financial markets (McMillan and Thupayagale, 2010) focused on the univariate Generalised Auto-Regressive Heteroscedasticity (GARCH) models for modelling and forecasting Value-at-Risk. A univariate model takes into account only one variable and ignores the temporal dependence of that particular variable to other variables. In reality, financial volatilities move together over time across different assets and markets.

Multivariate GARCH (MGARCH) models step away from the more simplified univariate GARCH models so as to model volatility and correlation transmission as well as spill over effects. (Silvennoinen & Terasvirta, 2008). However, the multivariate GARCH models – vector error correction (VEC) GARCH model (Bollerslev, Engle and Wooldridge, 1988); Baba-Engle-Kraft-Kroner (BEKK) model; Matrix Exponential GARCH model (Kawakatsu, 2006) – contain a large number of parameters rendering them computationally intensive and therefore less tractable.

Independent Component Analysis (Hyvärinen, 1999), can be used to transform the observed multidimensional financial time series vector into components that are statistically as independent from each other as possible. In previous research (Oja, Kiviluoto and Malaroiu, 2000; Wu, YU and LI, 2006; Xu and Wirjanto, 2013) VaR was modelled using the ICA-GARCH approach for *linear* asset portfolios. In their experimental results, Wu, YU and LI, 2006, show that the ICA-GARCH models are more effective at modelling Value-at-Risk for risk management purposes than existing methods, including Dynamic Conditional Correlation (DCC), Principal Component Analysis (PCA-GARCH), and Exponentially Weighted Moving Average (EWMA)

When dealing with *linear* financial instruments – where the change in the return on the financial asset and the change in the return on the underlying are linearly related – we can assume normality of returns as in the work of Wu, YU and LI, (2006). This assumption becomes problematic for *non-linear* financial instruments such as options. In this research, we step away from the *linear* asset portfolio to model the multivariate portfolio VaR for *non-linear* assets using the ICA-GARCH approach. We look at the case where we do not assume a distribution for the returns of the financial asset but use Johnson (1949)'s translation system to determine the distribution from the first four moments of the returns. In doing so, our objective is to analyse the performance of the ICA-GARCH approach to measuring multivariate portfolio VaR for non-linear assets. We carry out back-testing and out-of-sample tests of the performance of the ICA-GARCH model for VaR estimation and compare this with the Risk-Metrics (1996) Exponentially Weighted Moving Average (EWMA) approach as well as the univariate GARCH approach of McMillan and Thupayagale (2010).

1.2 Research Questions

If applied to *non-linear* financial assets in South African financial markets, does the ICA-GARCH approach to computing *multivariate* portfolio VaR where the underlying distribution is estimated using the Johnson's distribution lead to better performing estimate of VaR and more quickly converging VaR computations as well as more accurate VaR estimates and forecasts than the *univariate* GARCH and EWMA approaches?

1.3 Brief Literature Review

Value at risk (VaR) estimation falls into one of three approaches: historical simulations, Monte Carlo simulation and parametric approaches. With parametric approaches, we mostly rely on both an approximation of the portfolio and strong assumptions about the distribution of the risk factors' returns (usually that the risk factors are jointly normally distributed). VaR is then computed by using the standard deviations (s.d.) and correlations ρ of financial returns under the assumption that these returns are normally distributed. (RiskMetrics, 1996)

In practice, the assumption of return normality has proven to be extremely risky. This was the biggest mistake that Long Term Capital Management (LTCM) made and as a result underestimating their portfolio risks at the extremes Jorion (2000). If we are to proceed without making any assumptions about the distribution of the underlying financial returns then we can make use of the Johnson (1949) translation system to estimate the distribution of returns.

The other issue concerns the method used to compute the standard deviations of financial returns. In the original RiskMetrics framework, the risk factors' log-returns were assumed to be *conditionally* normally distributed (having a *multivariate* normal distribution), the conditionality being on the variance-covariance matrix of returns. RiskMetrics (1996) mainly focus on the

Exponential Weighted Moving Average (EWMA) to forecast the parameters of the multivariate conditional normal distribution.

However, there are now more accurate methods of estimating the standard deviations and correlations for VaR purposes. Some of these methods are Extreme Value Techniques (Parkinson, 1980), Two-Step Regression Analysis (Davidian and Carroll, 1987), GARCH (Bollerslev, 1986), Stochastic Volatility (Harvey et. al, 1994) and Applications of Chaotic Dynamics (LeBaron, 1994).

The importance of the multivariate approaches is that a univariate model takes into account only one variable and ignores the temporal dependence of that particular variable to other variables. In reality, financial volatilities move together over time across different assets and markets.

Multivariate GARCH (MGARCH) models step away from the more simplified univariate GARCH models so as to model volatility and correlation transmission as well as spill over effects (Silvennoinen & Terasvirta, 2008). The first multivariate GARCH model was proposed by Bollerslev, Engle and Wooldridge, 1988 as an extension of the univariate GARCH model. However, the number of parameters to estimate in a typical multivariate model are often very large, and the restrictions to guarantee the positive definiteness of the conditional covariance matrix are often difficult to enforce in practice. (Xu and Wirjanto, 2013). Wu, YU and LI, 2006 propose the use Independent Component Analysis (ICA-GARCH) models which are computationally more efficient for estimating the multivariate volatilities as compared to the multivariate GARCH model proposed by Bollerslev, Engle and Wooldridge, 1988

We can then combine the distribution derived using the Johnson translation system and the volatility estimated using the ICA-GARCH approach to *possibly* come up with a method that performs better than other existing methodologies such as the EWMA and the univariate-GARCH.

1.4 Research Approach

For our study, we will be analysing the performance of the following methods for computing the VaR for non-linear financial assets:

i. Univariate GARCH assuming normality of returns

ii. Univariate GARCH using Johnson's Translation System

iii. EWMA assuming normality of returns

iv. EWMA using Johnson's Translation System

v. ICA-GARCH assuming normality of returns

vi. ICA-GARCH using Johnson's Translation System

For the simple VaR methodology where we assume normality of returns, Value at Risk can be computed using 1.6449 multiplied by the standard deviation of $\widehat{r_{p,t}}$ (the return on the portfolio). 1.6449 is the 5th percentile of the standard normal distribution. The standard deviation will be calculated using the three methodologies (ICA-GARCH, univariate GARCH & EWMA)

For the VaR where we use Johnson's Translation System we follow the following steps:

Step 1. We estimate delta ($\tilde{\delta}_i$), the rate of change of the value of the financial instrument with respect to the changes in the underlying's price, gamma ($\tilde{\Gamma}_i$), the rate of change in the delta with respect to the change in the underlying' price, theta ($\tilde{\theta}_i$) (the sensitivity of the value of the financial instrument to time. These parameters are derived from the Black-Scholes formula for options. We also calculate ${\sigma^2}_{i,t}$ where ${\sigma^2}_{i,t}$ is the volatility of the underlying asset, *i* at time *t* and is calculated using the ICA-GARCH, univariate GARCH & EWMA methods.

Step 2. We calculate the mean, variance, skewness coefficient and kurtosis coefficient of the non-linear financial assets.

Step 3. We then use Slifker and Shapiro (1980)'s selection criteria to determine the distribution and estimates of the parameters γ, δ, ξ and λ for the asset. γ and δ are shape parameters, λ is a scale parameter and ξ is a location parameter for the normalising transformation.

Step 4. We compute the percentiles of $\widehat{r_{\iota,t}}$ (the return on the non-linear financial asset) distributions using some transformations.

Step 5. Use this percentile for VaR calculations.

Finally, we will use back-testing and out-of-sample tests to validate the performance of each of the six combinations and methods for value-at-risk computations.

1.5 Data Requirements

Our portfolio will be made up of is made up of 3 non-linear financial assets as follows:

i. A long call on the FTSE/JSE TOP40 Index
ii. A long put option on the USD/ ZAR currency exchange rate.
iii. A short call on the gold price

The dataset is from the periods of January 1, 2011 to December 31, 2016 representing 1302 daily observations. For model comparison with the univariate GARCH and EWMA, we divide the dataset into two parts. The first 1102 observations are for model training while the remaining 200 observations make up the out-of-sample dataset for the evaluation of forecasting precision.

1.6 Conclusion

In this chapter, we have laid down the framework within which we will be operating to solve the research problem. In the following chapter, we have lay down in greater detail the method of computing VaR that we shall be using and the steps we will be following. We have also briefly laid down our research approach as well as the data requirement necessary for our research study.

2 On Value-at-Risk and Volatility Measures

In this chapter, we highlight and explore the literature relevant to our work by first looking at VaR (Value-at-Risk). Value-at Risk is first defined and some of the methods used in practice to measure it are briefly explored. Much focus is given on VaR measurement using the RiskMetrics approach before introducing the ICA and ICA-GARCH concepts.

2.1 Defining VaR

Perhaps the risk measure of choice used in the financial industry due to its simplicity of computation and interpretation, value-at-risk (VaR) is a single summary, distribution independent statistical measure of losses arising as a result of "typical" market movements (Danı´elsson. 2011). The importance of Value-at-Risk lies in the fact that it measures the loss on a portfolio in such a way that we can attach a probability p of losses being equal to or exceeding VaR and a probability $(1-p)$ of these losses being lower than VaR which makes it a relatively easy risk measure to understand.

In statistical terms, if we define Q as a random variable representing the distribution of the profit or loss (P/L) on a portfolio, and if we also define *q* as a particular realisation of Q, VaR can be represented statistically as:

$$\Pr[Q \leq -VaR(p)] = p \tag{2.1}$$

or

$$p = \int_{-\infty}^{-VaR(p)} f_q(x)dx \tag{2.2}$$

where; $f_q(\cdot)$ is the probability density function of the profit or loss (P/L) function.

Graphically, we can represent VaR as follows:

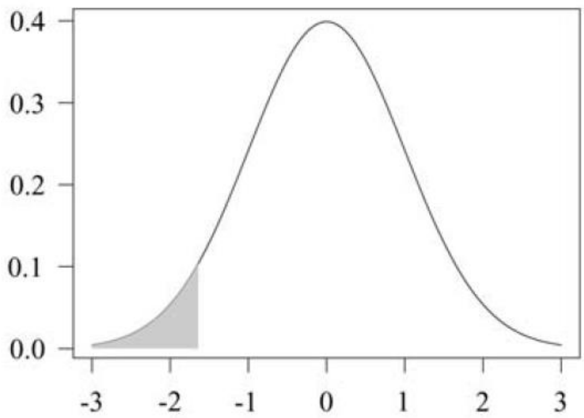


Figure 1: P/L density, $f_q(\cdot)$, and VaR

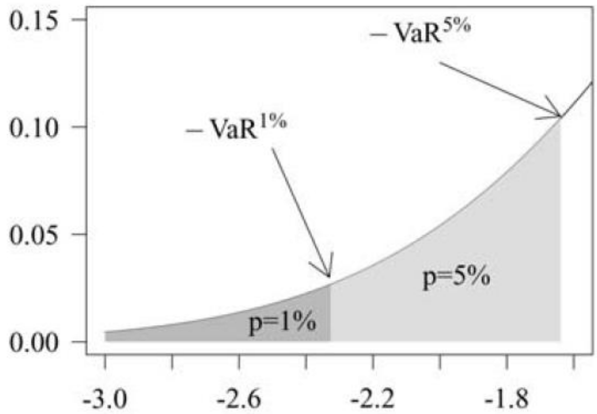


Figure 2: Left tail of $f_q(\cdot)$ and VaR

Figure 1 shows the entire density of the P/L function whereas Figure 2 zooms in on the left tail.

The shaded areas identify the 1% and 5% probabilities. (Danı´elsson. 2011)

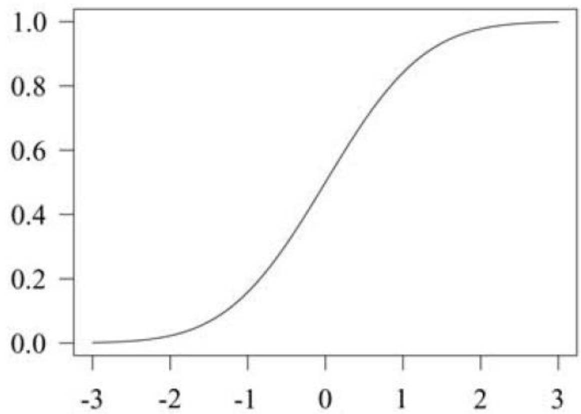


Figure 3: P/L distribution, $F_q(\cdot)$, and VaR

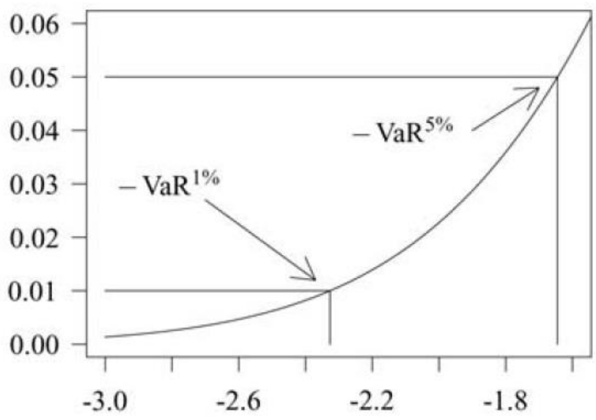


Figure 4: Tail of $F_q(\cdot)$ and VaR

Figure 3 shows the entire distribution of the P/L function while Figure 4 shows the left part of the distribution. (Danı´elsson. 2011)

So important is VaR that it has become the new regulatory framework's yardstick for quantifying investment portfolio risk but problems sometimes arise with the way VaR is used to assess and measure an investment portfolio's risk. The case of Long Term Capital Management (LTCM)'s failure has been widely ascribed to the way the hedge fund used Value at Risk (VaR) (Jorion, 2000). Jorion (2000) finds that in the case of LTCM, VaR itself was not the culprit but the way it was parameterised. One example of the inappropriate parameters used was the 10-day horizon used to set the amount of equity capital needed. Typically, the horizon must be related to the liquidity of the assets or alternatively the time it would take to raise additional funds or implement corrective action. 10 days is adequate for a commercial bank as it is assumed that investors/depositors take on average 10 days to liquidate their assets but is insufficient for a hedge fund where investors get in and out of positions more frequently and where there is more wide use of leverages. This was clearly so in the case of LTCM.

Another mistake that LTCM made was to assume return normality and as a result underestimated their portfolio risks at the extremes. As a result of this shortcoming, the concept of VaR is often times supplemented by the use of some other more rigorous risk measures such as Conditional

Value-at-Risk (CVAR) also referred to as the Expected Shortfall (ES), Tail Value-at-Risk (TVaR) or Extreme Value Theory (EVT) which are better able to model the tails of the P/L distribution.

Extreme Value Theory (EVT) is a branch of statistics whose main result is to model the distribution of values above a given threshold. The Pickans-Balkema-de Haan (PBH) theorem describes the distribution of these observations above a particular high threshold as a generalized Pareto distribution. (Levine, 2009). Artzner, Delbaen, Eber and Heath, (1999) argue that VaR leads to Pareto-inferior allocations where there are extreme deviations from the median of the probability distribution.

In addition, VaR can also fail to appropriately account for portfolio risk diversification. (Artzner et al, 1999). VaR is also known to violate the sub-additivity hypothesis of the so-called coherent risk measures due to the fact that VaR does not reflect the entire tail of the P/L distribution. (Gu´eant, n.d.)

However, despite its shortcomings, regulatory requirements have been geared almost exclusively toward a Value at Risk (VaR) concept as a measure of downside market risk. (Xu and Wirjanto, 2013). In practice, most banks and insurance companies use Value at Risk for regulatory purposes (in particular to measure and quantify their regulatory capital) as well as also using it as an internal risk measure. However, due to the drawbacks outlined above, many practitioners supplement VaR figures with Conditional Value-at-Risk (CVAR) or Expected Shortfall (ES) indices, but Value-at-Risk remains a critical component for internal risk quantification and management. (Gu´eant, n.d.)

2.2 Overview of Value-at-Risk Estimation

In this section, we give a high-level laydown for Value-at-Risk Estimation. The aim is to introduce the methodologies for VaR estimation without getting into much mathematical rigour.

Later on, in Section 2.4, we give granular details of VaR estimation with more detailed explanations of the steps involved.

2.2.1 VaR Computation for Linear Positions

VaR can be easily estimated by assuming a linear relationship between the value of an asset and the value of its underlying. This is true of the assets themselves (where the relationship is 1-to-1 between the value of the asset and the value of the underlying) and derivatives such as forwards and futures where the relationship between the value of the derivative and the underlying is linear. Value at risk (VaR) estimation for linear positions falls into one of three methodologies: historical simulations, Monte Carlo simulation and parametric approaches.

Historical Simulations

These are non-parametric approaches where we make an assumption that future behaviour of some risk factor will *replay* its past behaviour during a certain period of time. Some of the methodologies used in practice under the historical simulations approach are the Boudoukh, Richardson, Whitelaw (BRW), the Hull and White approach or the filtered historical simulations (FHS) approach. Extreme Value Theory is also used to provide better estimates for the extreme quantiles. (Gu´eant, n.d.)

Monte Carlo simulation.

The Monte-Carlo simulations use the same type of simulations as in historical approaches but the samples here are not assumed to be based on past realisations of the risk factors but rather rely on calibrated distributions of the risk factors and draw scenarios from this joint distribution. A direct consequence of this is that any distribution can be carved for the risk factors with the obvious disadvantage that a lot of parameters will need to be estimated.

Analytical/Parametric Approaches

Also known as the Linear VaR or the Variance-Covariance VaR, this is the simplest VaR method and is the most commonly used in practice. Most parametric approaches rely on both an approximation of the portfolio and *strong* assumptions about the distribution of the risk factors' returns (usually that the risk factors are *jointly* normally distributed). VaR is then computed by using the standard deviations (s.d.) and correlations ρ of financial returns under the assumption that these returns are normally distributed. (RiskMetrics, 1996)

Gaussian assumptions about the risk factors are the most commonly used but these do not necessarily hold in practice and can lead to underestimation of the tail losses (Gu´eant, n.d.). Portfolio exposures are assumed to be linear and since the portfolio return is a linear combination of normal variables, it is itself normally distributed. Thus, the portfolio volatility can be calculated by using the covariance matrix and weight vector easily.

2.2.2 VaR Computation for Non-Linear Positions

In the standard parametric methods outlined above, an assumption was made that portfolio exposures are linear in nature. This becomes an inaccurate assumption when dealing with non-linear financial instruments such as options. (RiskMetrics, 1996) outline two methods to measure VaR for non-linear positions; analytical approximations and structured Monte Carlo Simulation. These two methods differ in how the value of the portfolio changes with market movements. While the analytical approach approximates changes in value, the structured Monte Carlo approach fully re-values portfolios under different scenarios.

Structured Monte Carlo Simulation

Structured Monte Carlo simulation involves creating a large number of possible scenarios and revaluing the asset under each of these scenarios. VaR is then approximated by defining it as, for example, the 5th percentile of the distribution of value changes. Due to the required revaluations, this approach is much more computationally more intensive than the method we present below.

Analytical Approximations

Here we approximate the non-linear relationship using a mathematical expression that relates the return on the position to the return on the underlying. The methods are based on approximations of the portfolio using a Taylor series expansion and thus relies on the "Greeks" of the assets in the portfolio. (Gu´eant, n.d.). Here the change in the value of the instrument is approximated not just by the delta but also by the gamma (which measures the curvature of changes in the value) as well as the other Greeks; vega, rho, and theta. These can also be used to enhance the accuracy of the approximation. Two common types used here are the delta and delta-gamma approximations. (RiskMetrics, 1996). Other common approaches are the Greek Normal VaR, Delta Normal VaR or Delta-Gamma Normal VaR

A Simple Example:

Consider the following Portfolio example set out in (RiskMetrics, 1996):

Asset 1: A 1-year zero-coupon bond of AUD 1 million to be received in one year's time. The spot 1-year AUD rate is an effective interest rate of 10% per annum so that the current market value of the instrument is: $AUD1{,}000{,}000/1.1 = AUD909{,}090.9$.

Asset 2: An at-the-money AUD put/ZAR call option with a contract size of AUD 1 million and expiration date one month in the future. The premium of the option is 0,0105 and the spot exchange rate at which the contract was concluded is 1,538 AUD/ZAR. The implied volatility at which this option is priced is 14% p.a.

Of course, the value of this portfolio is dependent on the AUD/ZAR exchange rate and the 1-year AUD bond price.

We also make the following assumptions:

i. Our risk horizon is 5 days

ii. The daily volatilities of these two assets are;

a. $\sigma_{1y} = 0{,}08\%$

b. $\sigma_{\frac{AUD}{ZAR}} = 0{,}42\%$

c. $\rho_{1y,\frac{AUD}{ZAR}} = -0{,}17$

σ is the volatility of the currencies and ρ is the correlation coefficient of the two assets' returns.

Here we are going to focus on price risk alone (delta) and ignore the other risks (vega, rho, theta).

Solution 1: Delta Normal VaR Approximation

The simplest approach is to estimate the changes in the option value via a linear model; the delta approximation. We can calculate the delta for the option to be -0,4919 in this example.

The first step is to write down the return on the portfolio whose VaR we are trying to estimate. The return on this portfolio, denoted by r_p, consisting of the zero-coupon bond and the put on the AUD/call on the ZAR is:

$$r_p = r_{1y} + r_{\frac{AUD}{ZAR}} + \delta r_{\frac{AUD}{ZAR}} \tag{2.3}$$

where: r_{1y} = the price return on the 1-year AUD interest rates, $r_{\frac{AUD}{ZAR}}$ = the return on the AUD/ZAR exchange rate and δ = the delta of the option = -0,4919.

The idea here is to incorporate the return of the option into the return of the portfolio via the sensitivity of the option to the sensitivity of the underlying. The delta of an option is the rate of change of the value of the option with respect to the changes in the underlying share's price.

Later on, in Section 2.4.2 we will be demonstrating was to compute the portfolio return by calculating the actual return on an option using the Delta-Gamma-Theta methodology.

Under the assumption that the portfolio returns are normally distributed, VaR at the 95% confidence level is then given by:

$$VaR = 1.6449\sqrt{\sigma_{1y}^2 + (1+\delta)^2\sigma_{\frac{AUD}{ZAR}}^2 + 2(1+\delta)\rho_{1y,\frac{AUD}{ZAR}}\sigma_{1y}\sigma_{\frac{AUD}{ZAR}}} \qquad (2.4)$$

Here, 1.6449 is simply the 5% point of the standard normal distribution, in other words, 5% of the values from a standard normal distribution are greater than 1.6445 or equivalently are greater than 1.6445 standard deviations when approaching the mean from above:

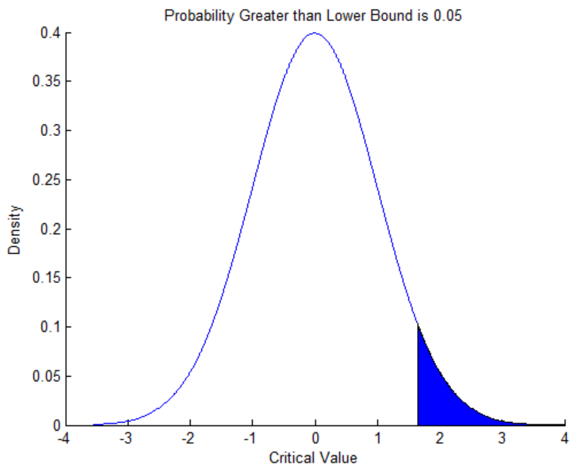


Figure 5: 5% point for the standard normal distribution

Using the volatilities and correlations given above as well as the value for the δ of the option, and scaling the 1 year AUD rate to a weekly rate (using the square root of 50) the weekly VaR using the delta equivalent approach is given by:

	Market value in ZAR	VaR(1w)
1-yr DEM cash flow	R591 086	R1 745
FX position - FX hedge	R300 331	R4 654
Diversified Portfolio	R891 417	R4 684

Table 1: Total Portfolio VaR

The portfolio's VaR is therefore given by R4 684. This value also demonstrates the impact of diversification on VaR as it is lower than the VaR values simply added together.

Solution 2: Delta-Gamma Normal VaR Approximation

This approach is more accurate than the earlier method above. However, this accuracy is eroded in extreme movements in the value of the exchange rates. This is for the simple reason that the delta is a linear approximation of a non-linear relationship

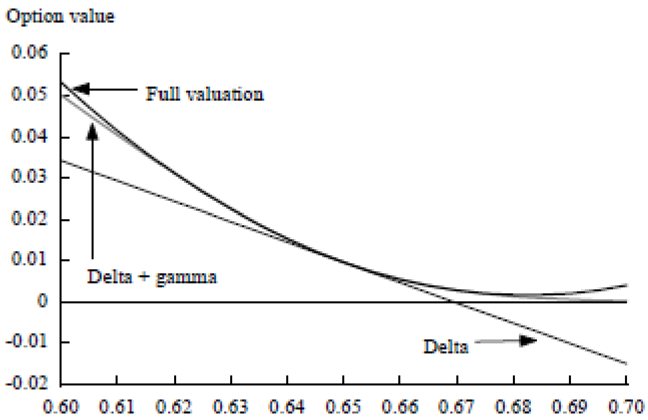


Figure 6: AUD/ZAR exchange rate

By including the gamma term which accounts for non-linear effects (i.e. squared returns) of changes in the exchange rates, we improve this approximation.

The expression for the portfolio return is now given by:

$$r_p = r_{1y} + r_{\frac{AUD}{ZAR}} + \delta r_{\frac{AUD}{ZAR}} + 0{,}5\Gamma P_{\frac{AUD}{ZAR}}\left(r_{\frac{AUD}{ZAR}}\right)^2 \tag{2.5}$$

where: r_{1y} = the price return on the 1-year AUD interest rates, $r_{\frac{AUD}{ZAR}}$ = the return on the AUD/ZAR exchange rate, $P_{\frac{AUD}{ZAR}}$ the value of the AUD/ZAR exchange rate when the VaR forecast is made, δ = the delta of the option = - 0,4919 and Γ = the gamma of the option=15,14.

The gamma term here introduces skewness into the distribution of r_p. This means that the assumption of normality is now violated and as such we can no longer use the same approach as in 2.4 above. The new approach involves computing the first 4 moments of r_p and finding a suitable distribution whose first four moments match those of r_p. We then compute the 5th percentile of $r_p$based on this distribution. We explain this methodology in greater detail in Section 2.4.2 below.

Solution 3: Structured Monte-Carlo Simulation

Given the last limitation outlined above, i.e. where the P/L distribution may not necessarily be normally distributed, one way of working around this problem is to use the Monte Carlo Methodology which instead of estimating changes in the value of the portfolio using the product of a rate change (σ) and sensitivity (δ, Γ) rather focuses on revaluing positions at changed rate levels.

2.3 Volatility Estimation and Forecasting

The general VaR formula can be written as:

$$VaR^{p\%} = \sigma \times p\% \; point \; for \; the \; P/L \; distribution \times \vartheta \times \sqrt{T} \tag{2.6}$$

where σ is the volatility of the returns for which we are computing VaR, $p\% \; point \; for \; P/L \; distribution$is would be 1.6449 the P/L distribution is normal and $p\% = 5\%$, ϑ is the value of the asset and $\sqrt{T}$ is necessary if the holding period is different to the period for which σ applies to otherwise it will just be 1.

It is therefore important to have an estimate of σ in order for us to be able to calculate VaR.

In the original RiskMetrics framework, the risk factors' log-returns were assumed to be *conditionally* normally distributed (having a *multivariate* normal distribution), the conditionality being on the variance-covariance matrix of returns. The returns themselves may not necessarily be normally distributed and have fatter tails than is otherwise predicted by the normal distribution.

RiskMetrics (1996) mainly focus on the Exponential Weighted Moving Average (EWMA) to forecast the parameters of the multivariate conditional normal distribution. However, there are now more accurate methods of estimating the standard deviations and correlations for VaR purposes. Some of these methods are Extreme Value Techniques (Parkinson, 1980), Two-Step Regression Analysis (Davidian and Carroll, 1987), GARCH (Bollerslev, 1986), Stochastic Volatility (Harvey et. al, 1994) and Applications of Chaotic Dynamics (LeBaron, 1994). GARCH-type models are the most commonly used in practice. Tests of GARCH-type models on foreign exchange and stock markets have showed that these are better approaches to estimating volatility than moving averages in particular over shorter time horizons such as a day or a week. These models are numerous but some of the more common models used in practice are the generalised ARCH (GARCH), Integrated GARCH (IGARCH), Exponential GARCH (EGARCH) and Switching Regime ARCH (SWARCH).

2.3.1 Univariate vs Multivariate Models

A univariate model takes into account only one variable and ignores the temporal dependence of that particular variable to other variables. In reality, financial volatilities move together over time across different assets and markets.

An Example: Covariance Structures

Let us assume our data consists of the gold prices g_i and exchange rates x_i over several years t_i. The following separate regressions represent two univariate models:

$$x_i = \beta_{x_0} + \beta_{x_1} t_i + \varepsilon_{x_i} \tag{2.7a}$$

$$g_i = \beta_{g_0} + \beta_{g_1} t_i + \varepsilon_{g_i} \tag{2.7b}$$

In the univariate case, no information about the gold prices flows through to the model about the exchange rates and vice-verse. Analysis would be carried out on each of the models without regard of the relationship between the two variables. In a multivariate setting, however the gold prices and exchange rates would be modelled jointly, such as:

$$\mathbf{Y}_i = \begin{bmatrix} x_i \\ g_i \end{bmatrix} = \mathbf{W}\boldsymbol{\beta} + \begin{bmatrix} \varepsilon_{x_i} \\ \varepsilon_{g_i} \end{bmatrix}$$

$$= \mathbf{W}\boldsymbol{\beta} + \boldsymbol{\varepsilon_i}$$

$$\boldsymbol{\varepsilon_i} \sim \left(\mathbf{0}, \begin{bmatrix} \boldsymbol{\sigma_1^2} & \boldsymbol{\sigma_{12}} \\ \boldsymbol{\sigma_{12}} & \boldsymbol{\sigma_2^2} \end{bmatrix} \right)$$

The vectors $\mathbf{Y}_i$ and $\boldsymbol{\varepsilon_i}$ capture the responses and errors for the two observations that belong to the same subject. The errors for financial returns on the gold would now have the correlation given by,

$$Corr\left[\varepsilon_{g_i}, \varepsilon_{x_i}\right] = \frac{\sigma_{12}}{\sqrt{\sigma_1^2 \sigma_2^2}}$$

where σ_{12} is the covariance between the gold prices and exchange rates, σ_1^2 and σ_2^2 are the variances of the gold prices and exchange rates.

It is then through this correlation that information about the gold prices flows through to the exchange rates and vice versa. The RiskMetrics VaR approach used a univariate model of volatility in the form of the Exponential Weighted Moving Average (EWMA) model. Even the further enhancements we pointed out earlier: generalised ARCH (GARCH), Integrated GARCH (IGARCH), Exponential GARCH (EGARCH) and Switching Regime ARCH (SWARCH) are all univariate models and suffer from this shortfall of not being able to capture the inter-dependencies between the particular portfolio VaR in which we are interested in and the VaR from other portfolios and assets. It is this weakness that has led to the development of the multivariate GARCH models.

2.3.2 Multivariate GARCH Models

As we pointed out earlier, understanding the co-movements of financial time series is of great importance. Multivariate GARCH (MGARCH) models step away from the more simplified univariate GARCH models so as to model volatility and correlation transmission as well as spill over effects. (Silvennoinen & Terasvirta, 2008).

The first GARCH model for the conditional covariance matrix was the vector error correction (VEC-GARCH) model proposed by Bollerslev, Engle and Wooldridge (1988). This also had an ARCH version in Engle, Granger and Kraft (1984) One of the major pitfalls of this model is that imposing positive definiteness of the conditional covariance matrix in this model is difficult.

The $\mathrm{VEC}-\mathrm{GARCH}(1,1)$ model is given by:

$$\mathrm{vech}(\boldsymbol{H}_t) = c + \boldsymbol{A}_1 \mathrm{vech}(r_{t-1} r'_{t-1}) + \boldsymbol{B}_1 \mathrm{vech}(\boldsymbol{H}_{t-1}) \tag{2.8a}$$

The more general $\mathrm{VEC}-\mathrm{GARCH}\ (\mathrm{p},\mathrm{q})$ model can be written as:

$$\mathrm{vech}(\boldsymbol{H}_t) = c + \sum_{j=1}^{q} \boldsymbol{A}_j \mathrm{vech}\left(r_{t-j} r'_{t-j}\right) + \sum_{j=1}^{p} \boldsymbol{B}_j \mathrm{vech}\left(\boldsymbol{H}_{t-j}\right) \tag{2.8b}$$

where: $\mathrm{vech}(\cdot)$ is an operator that stacks the columns of the lower triangular part of its argument square matrix, *c* is an $N(N+1)/2 \text{ x } 1$ vector and $\boldsymbol{A}_j$ and $\boldsymbol{B}_j$ are $N(N+1)/2 \text{ x } N(N+1)/2$ parameter matrices. (Silvennoinen & Terasvirta, 2008)

The main problem arises when estimating the parameters of a VEC model since this model is very much computationally demanding. The number of parameters that needs to be estimated equals

$$(p+q)(N(N+1)/2)^2 + N(N+1)/2 \tag{2.9}$$

This is very large except for the case where N is small. Bollerslev, Engle and Wooldridge (1988) then proposed a restricted version of the VEC model above where $\boldsymbol{A}_j$ and $\boldsymbol{B}_j$ are diagonal matrices. However, (Silvennoinen & Terasvirta, 2008) argues that this model is too restrictive since no interaction is allowed between the conditional variances and co-variances.

Another model that can be viewed as a restricted version of the VEC-GARCH model is the Baba-Engle-Kraft-Kroner (BEKK) model taking the form $BEKK(1,1,1)$

$$\boldsymbol{H}_t = CC' + \boldsymbol{A}_{1,1}{}' r_{t-1} r_{t-1}' \boldsymbol{A}_{1,1} + \boldsymbol{B}_{1,1}{}' \boldsymbol{H}_{t-1} \boldsymbol{B}_{1,1} \tag{2.10a}$$

This can be generalised to the $BEKK(p,k,q)$ model as:

$$\boldsymbol{H}_t = CC' + \sum_{j=1}^{q} \sum_{k=1}^{K} \boldsymbol{A}_{kj}{}' \, r_{t-j} r_{t-j}' \boldsymbol{A}_{kj} + \sum_{j=1}^{p} \sum_{k=1}^{K} \boldsymbol{B}_{kj}{}' \boldsymbol{H}_{t-j} \boldsymbol{B}_{kj} \tag{2.10b}$$

where $\boldsymbol{A}_{kj}, \boldsymbol{B}_{kj}$ and C are $N \times N$ parameter matrices and C is lower triangular (Engle and Kroner, 1995). Interpretation of the parameters of this model is not easy. Estimation of the BEKK model still involves heavy computations due to several matrix inversions. The number of parameters $(p+q)KN^2 + N(N+1)/2$ in the full BEKK model is still large.

Kawakatsu (2006) proposed the Matrix Exponential GARCH model which is a generalisation of the univariate Exponential (EGARCH) model of Nelson (1991). This Matrix Exponential $\mathrm{EGARCH}(1,1)$ model can be written as:

$$\text{vech}(\ln \boldsymbol{H}_t - C) = \boldsymbol{A}_1\eta_{t-1} + \boldsymbol{F}_1(|\eta_{t-1}|-E|\eta_{t-1}|) + \boldsymbol{B}_i\text{vech}(\ln \boldsymbol{H}_t - C) \qquad (2.11a)$$

This can also be written more generally as the EGARCH (p, q) model given by:

$$\text{vech}(\ln \boldsymbol{H}_t - C) = \sum_{i=i}^{q} \boldsymbol{A}_i\eta_{t-i} + \sum_{i=1}^{q} \boldsymbol{F}_i\,(|\eta_{t-i}|-E|\eta_{t-i}|) + \sum_{i=1}^{p} \boldsymbol{B}_i\text{vech}\,(\ln \boldsymbol{H}_t - C) \quad (2.11b)$$

where *C* is a symmetric N x N matrix and $\boldsymbol{A}_i, \boldsymbol{B}_i$ and $\boldsymbol{F}_i$ are parameter matrices of sizes $N(N+1)/2$ x N, $N(N+1)/2$ x $N(N+1)/2$, and $N(N+1)/2$ x N, respectively. This model still contains a large number of parameters. Let us now look at a method of estimating the multivariate GARCH model that is much more computationally less demanding.

2.3.3 Independent Component Analysis

We have discussed how some of the multivariate GARCH models – vector error correction (VEC) GARCH model (Bollerslev, Engle and Wooldridge, 1988); Baba-Engle-Kraft-Kroner (BEKK) model; Matrix Exponential GARCH model (Kawakatsu, 2006) – contain a large number of parameters rendering them computationally intensive and therefore less tractable.

We now take a more detailed look at an alternative method, ICA (Hyvärinen, 1999), which is a statistical technique for transforming an observed multidimensional random vector into components that are statistically as independent from each other as possible. We shall show later that this method is more computationally tractable compared to the models outlined above.

Let us assume we start with a realisation of m continuous valued scalar *multivariate* random variables $x_1, x_2, \dots, x_m$. We then arrange the x_i observed scalar *multivariate* variables into an m-dimensional random vector $\mathbf{x} = (x_1, x_2, \dots, x_m)^T$. The observed *m* scalar *multivariate* random variables $x_1, x_2, \dots, x_m$ are assumed to be a linear combination of *n* unknown *independent components* $s_1, s_2, \dots, s_n$. These unknown independent components are mutually statistically independent with zero mean. These elements are also non-Gaussian. Also, we must assume that $n \leq m$ (Hyv¨arinen, Karhunen, and Oja, 2001)

If we also arrange the component variables s_i into a vector $\mathbf{s} = (s_1, s_2, \dots, s_n)^T$, the relationship can be written as:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{2.12}$$

where **A** is an unknown m x n full rank matrix called the *mixing matrix*. The mixing coefficients or elements of matrix **A** are unknown. The problem of ICA is to estimate the matrix **A** from which we can obtain **W** as the (pseudo) inverse of the estimate of the matrix **A.** (Hyvärinen, 1999). The independent components are then obtained using the relationship:

$$\mathbf{s} = \mathbf{W}\mathbf{x} \tag{2.13}$$

The one restriction of the model is that we can only estimate non-Gaussian independent component except in the case where only one of the independent components is Gaussian. (Hyvärinen and Oja, 2000)

An Example

Consider two independent components s_1 and s_2 that have the following uniform distributions:

$$p(s_i) = \begin{cases} \frac{1}{2\sqrt{3}} & \text{if } |s_i| \leq \sqrt{3} \\ 0 & \text{otherwise} \end{cases} \tag{2.14}$$

The joint density of s_1 and s_2 is then uniform on a square. This follows from the definition of statistical independence where $p(s_1, s_2) = p(s_1)p(s_2)$

Let us then mix these two components using the following mixing matrix:

$$A_0 = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}$$

From this we obtain two mixed variables x_1 and x_2. The mixed data has a uniform distribution but the mixed random variables x_1 and x_2 are not independent anymore. To illustrate this last point, if x_1 attains its minimum or maximum values then this completely determines the value taken by x_2.

The problem is estimating the date model of Independent Component Analysis is to estimate the matrix A_0 using only information contained in the mixtures x_1 and x_2

Why Restrict to Non-Gaussian Variables

In the case of Gaussian variables, we are only able to estimate the ICA model to an orthogonal transformation. The matrix **A** is not identifiable for Gaussian independent components. (Hyvärinen and Oja, 2000)

Assume that the mixing matrix is orthogonal and that the s_i are Gaussian. The variables x_1 and x_2 are Gaussian, uncorrelated and have unit variance. Their joint density function is given by:

$$p(x_1, x_2) = \frac{1}{2\pi} \exp\left(-\frac{{x_1}^2 + {x_2}^2}{2}\right) \tag{2.15}$$

This distribution is of course symmetric and therefore does not contain any information about the directions of the columns of the mixing matrix **A**. In simple terms, one can prove that the distribution of any orthogonal transformation of the Gaussian (x_1, x_2) has the exact same distribution as (x_1, x_2) and also x_1 and x_2 are independent.

Pre-Processing

Pre-processing the data before carrying out ICA not only simplifies the ICA algorithm but also reduces the number of parameters to be estimated.

i. Centering

We need to first make the x a zero-mean variable by subtracting its mean vector $\mathbf{m} = E\{\mathrm{x}\}$. This automatically implies that s is zero-mean as well. This step is just to simplify the ICA algorithms. If we so desire, after estimating the mixing matrix A, we can add the mean vector of s (given by $\mathbf{A}^{-1}\mathbf{m}$) back to the centred estimates of s.

ii. Sphering or Whitening

Here (after centering) we transform the observed vector **x** *linearly* so that we have a new vector $\tilde{\mathbf{x}}$ whose components are uncorrelated and has equal unit variances. In other words, the covariance matrix of $\tilde{\mathbf{x}}$ equals the identity matrix I.

$$E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^{\mathrm{T}}\} = \mathbf{I} \tag{2.16}$$

To achieve whitening, we can use the method of eigenvalue decomposition (EVD) of the covariance matrix:

$$E\{\mathbf{x}\mathbf{x}^{\mathrm{T}}\} = \mathbf{E}\mathbf{D}\mathbf{E}^{\mathrm{T}} \tag{2.17}$$

where: $\mathbf{E} =$ the orthogonal matrix of eigenvectors of $E\{\mathbf{x}\mathbf{x}^{\mathrm{T}}\}$ and **D** = is the diagonal matrix of its eigenvalues = diag($d_1, d_2, \dots, d_n$)

Also, $E\{\mathbf{x}\mathbf{x}^{\mathrm{T}}\}$ can be estimated in a standard way from the available sample $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$

Sphering can now be achieved by setting:

$$\tilde{\mathbf{x}} = \mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^{\mathrm{T}}\mathbf{x} \tag{2.18}$$

where the matrix $\mathbf{D}^{-1/2}$ is obtained through a component-wise operation $\mathbf{D}^{-1/2} = \mathrm{diag}({d_1}^{-1/2}, {d_2}^{-1/2}, \dots, {d_n}^{-1/2})$

From (2.12) and (2.18) we now have:

$$\tilde{\mathbf{x}} = \mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^{\mathrm{T}}\mathbf{A}\mathbf{s} = \widetilde{\mathbf{A}}\mathbf{s} \tag{2.19}$$

Sphering has now the mixing matrix $\mathbf{A}$ into a new one $\widetilde{\mathbf{A}}$. The new mixing matrix $\widetilde{\mathbf{A}}$ is orthogonal. That is:

$$E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^{\mathrm{T}}\} = \widetilde{\mathbf{A}}E\{\mathrm{ss}^{\mathrm{T}}\}\widetilde{\mathbf{A}}^{\mathrm{T}} = \widetilde{\mathbf{A}}\widetilde{\mathbf{A}}^{\mathrm{T}} = \mathbf{I} \tag{2.20}$$

The ultimate result is that sphering reduces the number of parameters to be estimated. Because an orthogonal matrix contains $n(n-1)/2$ degrees of freedom, we only need to estimate the new, orthogonal mixing matrix $\widetilde{\mathbf{A}}$. Since an orthogonal matrix contains half of the number of parameters of an arbitrary matrix, we can say that we have now solved half of the ICA problem.

Algorithms for Independent Component Analysis

To optimise the problem set out above so as to determine the mixing matrix, there are various algorithms as set out in Hyvärinen (1999). The choice of a suitable algorithm depends on the stability of the algorithm, it's convergence speed as well as the memory requirements.

Some of the algorithms available are: Jutten-Hérault algorithm, Non-linear decorrelation algorithms, Algorithms for maximum likelihood or infomax estimation, Non-linear PCA algorithms, Neural one-unit learning rules, Other neural (adaptive) algorithms, Tensor-based algorithms, Weighted covariance methods, The FastICA algorithm

i. The FastICA Algorithm

The FastICA algorithm is a computationally efficient algorithm for performing ICA estimation. This batch (block) algorithm uses a fixed-point iteration that has been found to be 10-100 times faster at converging than the conventional gradient descent methods for ICA. (Hyvärinen, 1999) Here we do not delve much into the construction of the of the FastICA algorithm but list some of the properties of this algorithm that make it more desirable compared to other methods for solving the ICA problem as set out in Hyvärinen and Oja, (2000):

i. It has a much faster convergence.
ii. The algorithm is easy to use and apply in practice.
iii. The algorithm finds directly independent components of any non-Gaussian distribution. For other algorithms, some estimate of the probability distribution function has to be first available before the non-linearity is chosen.
iv. One can obtain algorithms that are robust and/or of minimum variance
v. The independent components can be estimated one by one
vi. The algorithm is computationally simple, and requires little memory space.

2.3.4 ICA-GARCH

Independent Component Analysis (ICA) can be applied to model multivariate asset return volatilities as a linear combination of several univariate Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models. Here we start with multivariate GARCH models and use Independent Component Analysis (ICA) to decompose these multivariate time series into statistically independent time series. (Wu, YU and LI, 2006). The resulting ICA-GARCH models are then used to estimate the multivariate volatilities for VaR estimation. This approach is much more computationally tractable (Wu, YU and LI, 2006) as compared to the multivariate GARCH model proposed by Bollerslev, Engle and Wooldridge (1988) for which we have already shown earlier that as the number of dimensions increase, the number parameters to be estimated also increase substantially which becomes computationally prohibitive.

2.4 Computing VaR: Detailed Analysis

In this section, we now go into much more detail about VaR computations. The choice of approach is dependent on the type of positions that are "at risk". In our study, we carry out VaR computations using two analytical approaches:

i. Simple VaR for linear instruments

ii. Delta-Gamma VaR for non-linear instruments

As stated earlier, the terms linear and non-linear describe the relationship that exists between a position's underlying returns to that position's relative change in value.

2.4.1 Simple VaR Calculations

The example below deals with VaR estimation at the 95% confidence level. Let's consider a portfolio consisting of *N* positions and that each of these positions consists of one cashflow on which we have both volatility and correlation forecasts. We denote the relative change in the value of the n^{th} position at time t by $\widehat{r_{n,t}}$. From this we can write the change in the value of the portfolio, $\widehat{r_{p,t}}$ as

$$\widehat{r_{p,t}} = \sum_{n=1}^{N} \omega_n \widehat{r_{n,t}} = \sum_{n=1}^{N} \omega_n \delta_n r_{n,t} \tag{2.21}$$

where: ω_n is the total nominal amount (as opposed to the weight) that is invested in the n^{th} position.

The VaR on a portfolio of simple linear instruments can be computed using 1.6449 multiplied by the standard deviation of $\widehat{r_{p,t}}$. 1.6449 is the 5th percentile of the standard normal distribution. If we are calculating the one day VaR (in other words when the VaR forecast horizon is one day), this standard deviation is calculated one day ahead.

The VaR estimate is given by:

$$VaR_t = \sqrt{\vec{\sigma}_{t|t-1} R_{t|t-1} \vec{\sigma}_{t|t-1}^{\,T}} \tag{2.22}$$

where:

$$\vec{\sigma}_{t|t-1} = \begin{bmatrix} 1.65\sigma_{1,t|t-1}\omega_1\delta_1 & 1.65\sigma_{2,t|t-1}\omega_2\delta_2 & \dots & 1.65\sigma_{N,t|t-1}\omega_N\delta_N \end{bmatrix} \tag{2.23a}$$

is the individual VaR 1 x N vector and

$$R_{t|t-1} = \begin{bmatrix} 1 & \rho_{12,t|t-1} & \cdots & \rho_{1N,t|t-1} \\ \rho_{21,t|t-1} & 1 & \vdots & \vdots \\ \dots & \dots & \vdots & \dots \\ \rho_{N1,t|t-1} & \cdots & \cdots & 1 \end{bmatrix} \tag{2.23b}$$

is the $N \times N$ correlation matrix of the returns on the underlying cashflows. Here the fundamental assumption is that the portfolio return follows a conditional normal distribution.

Important Point to Note on Equity Positions

To calculate the VaR for an equity given the returns on the market index, we can use the equation:

$$VaR_t = V_t \cdot \beta_t \cdot 1.65\sigma_{m,t} \tag{2.24}$$

where: $1.65\sigma_{m,t}$ = The VaR estimate of the appropriate stock index and β_t = the sensitivity of the stock to changes in the value for the index.

Fixed income instruments

With a portfolio of fixed income instruments, RiskMetrics, (1996) mention two issues that arise around:

i. The correct variable to use for measuring volatility and correlations that is should we use the price or yields for computations.

ii. Incorporating the "roll down" and "pull-to-par" effects of bonds in the Value-at-Risk calculations.

To deal with Point i, the RiskMetrics, (1996) approach computes the **price** volatilities and correlations on fixed income instruments by first computing the zero rates for all instruments with a maturity of over a year and then constructing prices from these series using the expression:

$$P_t = e^{-y_t N} \tag{2.25}$$

where: y_t is the current yield on the *N*-period zero-coupon bond. For money market rates, prices are constructed from the formula:

$$P_t = \frac{1}{(1 + y_t)^N} \tag{2.26}$$

However, practitioners like to think of volatilities on fixed income in terms of the yield and as such in terms of yield volatility. From (2.25) we have the price return calculation given by:

$$r_t = \ln(P_t / P_{t-1}) = N(y_{t-1} - y_t) \tag{2.27}$$

Therefore, the standard deviation of price returns is given by:

$$\sigma_t = N\sigma(y_{t-1} - y_t) \tag{2.28}$$

where: $\sigma(y_{t-1} - y_t)$ is the standard deviation of $y_{t-1} - y_t$.

What the equation above is saying is that to get the price return volatility we must multiply the terms to maturity of the underlying instrument by the standard deviation of the absolute changes in the yields.

In a similar way, from 2.26 we can also obtain:

$$r_t = \ln(P_t/P_{t-1}) = N\sigma\left[ln\left(\frac{1+\ y_{t-1}}{1+y_t}\right)\right] \tag{2.29}$$

We can then obtain the standard deviation of price returns as:

$$\sigma_t = N\sigma\left[ln\left(\frac{1+\ y_t}{1+y_{t-1}}\right)\right] \tag{2.30}$$

where: $\sigma\left[ln\left(\frac{1+y_t}{1+y_{t-1}}\right)\right]$ is the standard deviation of $ln\left(\frac{1+y_t}{1+y_{t-1}}\right)$

To deal with Point 2., Finger, (1996) points out that in the real world, the bond's market value systematically increases toward it's par value (the so-called "pull to par effect") and its daily volatility decreases as it moves closer to par (the so-called "roll down" effect). This is opposed to the RiskMetrics, (1996) assumptions: 1) there is no expected change in the market value of the bond and 2) the volatility of the bond's market value scales up with the square root of the time horizon, assumptions which effectively mean that cash are as if the maturity of the bond will always be the same.

The correct methodology (Fisher, 1966) for measuring VaR for cashflows that occur in *T* days over a forecast horizon of *t* days *(t<T)* is given by:

i. First use the *T-t* period rate, y_{T-t}, to discount the cashflows that occur in T days' time. We denote this discounted value by V_{T-t}
ii. Then compute VaR as $V_{T-t}(\sigma_{T-t}\sqrt{t})$

These steps effectively address the pull to par and roll down effects. (RiskMetrics, 1996)

2.4.2 Delta-Gamma Normal VaR Calculations (Portfolios with Non-Linear Positions)

This method allows for more accurate VaR computation for portfolios containing options. This can be viewed as an extension of the Delta Normal VaR methodology except that here we incorporate the delta, gamma and theta of individual options in the VaR calculations.

We start with the case of a single option. Here we assume that each option is a function of one cashflow and write the return on the option as:

$$\widehat{r_{i,t}} = \tilde{\delta}_i r_{i,t} + 0.5\tilde{\Gamma}_i r_{i,t}^2 + \tilde{\theta}_i n \tag{2.31}$$

where: $\tilde{\delta}_i = \eta_i \delta_i$, $\tilde{\Gamma}_i = \eta_i P_{i,t} \Gamma_i$, $\eta = \left(\frac{V_t}{P_t}\right)$, $\tilde{\theta}_i = \theta_i / V_i$, $n =$ VaR forecast horizon and $V_i =$ option's premium

Derivation of equation (2.31) has been presented in the Appendix.

It's important to note that equation (2.31) is a reasonable approximation when the Greeks δ and Γ are stable as the underlying price changes. If small changes in the underlying causes large changes in δ and Γ then the delta-gamma approach doesn't perform well (RiskMetrics, 1996).

Determining the Distribution of the Option's Returns

The next task is to determine the distribution of these option returns given by $\widehat{r_{i,t}}$. To do this we need to compute the numerical values of the moments of $\widehat{r_{i,t}}$ (recall that this is the returns on the option as opposed to $r_{i,t}$, the returns on the underlying asset). Also, RiskMetrics (1996) assume that the returns of the underlying are normally distributed with mean 0 and variance $\sigma^2{}_{i,t}$). RiskMetrics (1996) present the table below comparing the statistical features of an option and its underlying return:

Statistical parameter	Option	Underlying
Return	$\widehat{r_{i,t}}.$	$r_{i,t}$
Mean	$0.5\tilde{\Gamma}\sigma^2{}_{i,t} + \tilde{\theta}_i n$	0
Variance	$\tilde{\delta}_i^2\sigma^2{}_{i,t} + 0.5\tilde{\Gamma}^2\sigma^4{}_{i,t}$	$\sigma^2{}_{i,t}$
Skewness	$3\tilde{\delta}_i^2\tilde{\Gamma}_i\sigma^4{}_{i,t} + \tilde{\Gamma}_1^3\sigma^6{}_{i,t}$	0
Kurtosis	$12\tilde{\delta}_i^2\tilde{\Gamma}_1^2\sigma^6{}_{i,t} + 3\tilde{\Gamma}_1^4\sigma^8{}_{i,t} + 3\sigma^4{}_{i,t}$	$3\sigma^4{}_{i,t}$

Table 2: Statistical Features of an Option and its Underlying Return.

Again, $\tilde{\delta}_i = \eta_i\delta_i$, $\tilde{\Gamma}_i = \eta_i P_{i,t}\Gamma_i$, $\eta = \left(\frac{V_t}{P_t}\right)$, $\tilde{\theta}_i = \theta_i/V_i$, $n =$ VaR forecast horizon and $V_i =$ option's premium

To determine these numerical values, we need the estimates of $\tilde{\delta}_i$, $\tilde{\Gamma}_i$, $\tilde{\theta}_i$ and $\sigma^2{}_{i,t}$. Estimates of the first three are easily found by applying a Black-Scholes type valuation. In our study, we use the ICA-GARCH approach to estimate the variance $\sigma^2{}_{i,t}$ of the underlying asset.

Having obtained these first four moments of $\widehat{r_{i,t}}$'s distribution, we then find a distribution that has the same moments for which we know what it is exactly. To do this we need to apply the Johnson Translation System to match the moments of $\widehat{r_{i,t}}$'s distribution to one of a set of possible distributions called Johnson Distributions.

The Johnson Translation System

If we have a continuous variable *X* whose distribution is unknown and we wish to approximate, Johnson (1949) proposed three normalizing transformations having the general form:

$$Z = \gamma + \delta f\left(\frac{X - \xi}{\lambda}\right) \tag{2.32}$$

where: $f(\,)$ is a monotonic transformation function, Z is a standard normal variable, γ and δ are shape parameters, λ is a scale parameter and ξ is a location parameter.

Further, it is assumed that $\delta > 0$ and $\lambda > 0$

The transformations (George and Ramachandran, 2011) proposed by Johnson (1949) are as:

i. Lognormal system of distributions denoted by S_L:

$$Z = \gamma + \delta ln\left(\frac{X - \xi}{\lambda}\right), X > \xi \qquad (2.33a)$$

$$= \gamma^* + \delta ln(X - \xi), X > \xi \qquad (2.33b)$$

This system covers the Lognormal distribution

ii. The bounded system of distributions denoted by S_B

$$Z = \gamma + \delta ln\left(\frac{X - \xi}{\xi + \lambda - X}\right), \xi < X < \xi + \lambda \qquad (2.34)$$

This system covers the Gamma, Beta and many other distributions that are bounded on the lower end, upper end or both.

iii. The unbounded system of distributions denoted by S_U

$$Z = \gamma + \delta ln\left[\left(\frac{X - \xi}{\lambda}\right) + \left\{\left(\frac{X - \xi}{\lambda}\right)^2 + 1\right\}^{1/2}\right], -\infty < X < \infty \qquad (2.35a)$$

$$Z = \gamma + \delta \sinh^{-1}\left(\frac{X - \xi}{\lambda}\right) \qquad (2.35b)$$

This covers the *t*, normal and other distributions that are unbounded.

After transformation of (2.32), Z follows a standard normal distribution and as such, the probability density function (pdf) of each of the equations in the Johnson family can be derived.

In general, if X follows the Johnson distribution and

$$Y = \left(\frac{X - \xi}{\lambda}\right) \qquad (2.36)$$

The pdf of X (George and Ramachandran, 2011) is given by:

$$p(x) = \frac{\delta}{\lambda\sqrt{2\pi}} \times g'\left(\frac{X-\xi}{\lambda}\right) \times \exp\left\{-\frac{1}{2}\left[\gamma + \delta \cdot g\left(\frac{X-\xi}{\lambda}\right)\right]^2\right\} \qquad (2.37)$$

for all $x \in H$, where

$H = [\xi, +\infty)$ for the S_L family of distributions

$H = [\xi, \xi + \lambda]$ for the S_B family of distributions

$H = (-\infty, +\infty)$ for the S_U family of distributions

Also:

$g(y) = \ln(y)$ for S_L family of distributions

$g(y) = \ln(y/(1-y))$ for S_B family of distributions

$g(y) = \ln[y + \sqrt{y^2 + 1}$ for S_U family of distributions

and as such

$g'(y) = \frac{1}{y}$ for S_L family

$g'(y) = \frac{1}{[y(1-y)]}$ for S_B family

$g'(y) = \frac{1}{\sqrt{y^2+1}}$ for S_U family

Now, to find the estimates of γ, δ, ξ and λ we use percentile matching of the Johnson system which involves estimating *k* required parameters by matching *k* selected quantiles of the standard normal distribution with the corresponding quantile estimates of the target population (George and Ramachandran, 2011).

Slifker and Shapiro (1980) introduced a selection rule to give estimates of the Johnson parameters γ, δ, ξ and λ. This rule uses a function of four percentiles for selecting one of the 3 families (Log-normal, Unbounded and Bounded).

The rule works as follows;

Choose any fixed value *z* (0<*z*<1) of a standard normal variate. Determine the percentile P_ζ corresponding to ζ = *-3z, -z, z, 3z:*

For example, with *z = 0.5*

$$P_{-1,5} = (1 - 0{,}93319) * 100 = 6{,}681\%$$

$$P_{-0,5} = (1 - 0{,}69146) * 100 = 30{,}854\%$$

$$P_{0,5} = 0{,}69146 * 100 = 69{,}146\%$$

$$P_{1,5} = 0{,}93319 * 100 = 93{,}319\%$$

From the data, let $x_{-3z}, x_{-z}, x_z, x_{3z}$ be the percentiles of data values corresponding to the four selected percentiles of the normal distribution above.

The type of Johnson distribution (Slifker and Shapiro, 1980) chosen is based on the value of the discriminant *d* calculated as:

$$d = \frac{mn}{p^2} \tag{2.38}$$

where: $p = x_z - x_{-z}$, $m = x_{3z} - x_z$ and $n = x_{-z} - x_{-3z}$

If the calculated discriminant is given by:

$d > 1.001 \Leftrightarrow$ Unbounded

$d < 0.999 \Leftrightarrow$ Bounded

$0.999 \leq d \leq 1.001 \Leftrightarrow$ Lognormal

The parameter estimates for the Johnson S_U distribution are:

$$\hat{\delta} = \frac{2z}{\cosh^{-1}\left[\frac{1}{2}\left(\frac{m}{p} + \frac{n}{p}\right)\right]} \tag{2.39}$$

$$\hat{\gamma} = \hat{\delta}\sinh^{-1}\left[\frac{\frac{n}{p} - \frac{m}{p}}{2\left(\frac{m}{p}\frac{n}{p} - 1\right)^{1/2}}\right] \tag{2.40}$$

$$\hat{\lambda} = \frac{2p\left(\frac{m}{p}\frac{n}{p} - 1\right)^{1/2}}{\left(\frac{m}{p} + \frac{n}{p} - 2\right)\left(\frac{m}{p} + \frac{n}{p} + 2\right)^{1/2}} \tag{2.41}$$

and

$$\xi = \frac{x_z + x_{-z}}{2} + \frac{p\left(\frac{n}{p} - \frac{m}{p}\right)}{2\left(\frac{m}{p} + \frac{n}{p} - 2\right)} \tag{2.42}$$

The parameter estimates for the Johnson S_B distribution are:

$$\hat{\delta} = \frac{z}{\cosh^{-1}\left(\frac{1}{2}\left[\left(1 + \frac{p}{m}\right)\left(1 + \frac{p}{n}\right)\right]^{1/2}\right)}; \delta > 0 \tag{2.43}$$

$$\hat{\gamma} = \hat{\delta}\sinh^{-1}\left[\frac{\left(\frac{p}{n} - \frac{p}{m}\right)\left[\left(1 + \frac{p}{m}\right)\left(1 + \frac{p}{n}\right) - 4\right]^{1/2}}{2\left(\frac{p}{m}\frac{p}{n} - 1\right)}\right] \tag{2.44}$$

$$\hat{\lambda} = \frac{p\left[\left\{\left(1 + \frac{p}{m}\right)\left(1 + \frac{p}{n}\right) - 2\right\}^{2} - 4\right]^{1/2}}{\frac{p}{m}\frac{p}{n} - 1} \tag{2.45}$$

and

$$\xi = \frac{x_z + x_{-z}}{2} - \frac{\lambda}{2} + \frac{p\left(\frac{p}{n} - \frac{p}{m}\right)}{2\left(\frac{p}{m}\frac{p}{n} - 1\right)} \tag{2.46}$$

The parameter estimates for the Johnson S_L distribution are:

$$\hat{\delta} = \frac{2z}{\ln\left(\frac{m}{p}\right)} \tag{2.47}$$

$$\hat{\gamma}^{*} = \hat{\delta}\ln\left[\frac{\frac{m}{p} - 1}{p\left(\frac{m}{p}\right)^{1/2}}\right];\ \gamma^{*} is\ for\ both\ \gamma\ and\ \lambda \tag{2.48}$$

and

$$\xi = \frac{x_z + x_{-z}}{2} - \frac{p}{2}\frac{\frac{m}{p} + 1}{\frac{m}{p} - 1} \tag{2.49}$$

Given these estimates, we can then calculate any percentile of $\widehat{r_{i,t}}$'s distribution. This approximate percentile is then used in the VaR calculation.

Once we have our parameters, make the transformation from the percentile of the returns on the underlying to the percentile of the distribution of the returns using the transformation:

$$\widehat{r_{i,t}} = \sinh\left(\frac{(r_{i,t} - \gamma)}{\delta}\right) \cdot \lambda + \xi \Leftrightarrow \text{Unbounded} \tag{2.50}$$

$$\widehat{r_{i,t}} = \frac{\beta(\xi + \lambda) + \xi}{(1 + \beta)}; \beta = e^{\left(\frac{(r_{i,t} - \gamma)}{\delta}\right)} \Leftrightarrow \text{Bounded} \tag{2.51}$$

$$\widehat{r_{i,t}} = \lambda e^{\left(\frac{(r_{i,t} - \gamma)}{\delta}\right)} + \xi \Leftrightarrow \text{Lognormal} \tag{2.52}$$

Computing the VaR

Recalling that the general VaR formula (Equation 2.6) can be written as:

$$VaR^{p\%} = \sigma \times p\%\; point\; for\; the\; P/L\; distribution\; \times \vartheta \times \sqrt{T}$$

where σ is the volatility of the returns for which we are computing VaR, up until now, what we have done is to calculate the $p\%\; point\; for\; P/L\; distribution$ using the Johnson Translation System. We already know that, ϑ is the value of the asset and $\sqrt{T}$ is necessary if the holding period is different to the period for which σ applies to otherwise it will just be 1.

For a single option, σ is calculated using the formula defined in Table 2 above and as such we can compute our VaR estimates. However, portfolio of options below, we proceed as follows:

Calculating $\boldsymbol{p\%\; point}$ *and* $\boldsymbol{\sigma}$ *for a Portfolio Containing Options*

For the *p% point*, we find the 5th percentile of $\widehat{r_{p,t}}$'s distribution the same way we found the 5th percentile of $\widehat{r_{i,t}}$'s distribution, as shown previously. The only difference is in the formal and we show the formulas shortly.

To find σ, Let's consider $\widehat{r_{p,t}}$ the portfolio return for a portfolio made up of 3 options given by:

$$\widehat{r_{p,t}} = \omega_1 \widehat{r_{1,t}} + \omega_2 \widehat{r_{2,t}} + \omega_3 \widehat{r_{3,t}} \tag{2.53}$$

where

$$\omega_i = \frac{V_i}{\sum_{i=1}^{3} V_i} \tag{2.53a}$$

To compute the moments of $\widehat{r_{p,t}}$ we need the covariance matrix $\sum$ of the underlying returns $\{r_{1,t}, r_{2,t}, r_{3,t}\}$ and the δ, Γ and θ cashflow vectors as:

$$\tilde{\delta} = \begin{bmatrix} \tilde{\delta}_1 \\ \tilde{\delta}_2 \\ \tilde{\delta}_3 \end{bmatrix}, \tilde{\Gamma} = \begin{bmatrix} \tilde{\Gamma}_1 & 0 & 0 \\ 0 & \tilde{\Gamma}_2 & 0 \\ 0 & 0 & \tilde{\Gamma}_3 \end{bmatrix}, \text{and } \tilde{\theta} = \begin{bmatrix} \tilde{\theta}_1 \\ \tilde{\theta}_2 \\ \tilde{\theta}_3 \end{bmatrix} \tag{2.54}$$

Once again, we use the ICA-GARCH method to determine $\sum$. To find the 5th percentile of $\widehat{r_{p,t}}$'s distribution, we apply the same method as we did to find the 5th percentile of $\widehat{r_{\iota,t}}$'s distribution as shown in an earlier section.

The expressions for the first two moments is now given by:

$$\mu_{p,t} = 0.5 \cdot \text{trace}\left[\tilde{\Gamma}\sum\right] + \sum_{i=1}^{3} \tilde{\theta}_i \tag{2.55}$$

$$\sigma_{p,t}^2 = \hat{\delta}^T \sum \hat{\delta} + 0.5 \cdot \text{trace}\left[\left(\tilde{\Gamma}\sum\right)^2\right] \tag{2.56}$$

Here we add an adjustment factor compared to the linear cases where $\Gamma = 0$. The trace of the matrix $\tilde{\Gamma}\sum$ is the sum of the *N* eigenvalues of $\tilde{\Gamma}\sum$. The trace of $\left(\tilde{\Gamma}\sum\right)^2$ is the sum of the squared eigenvalues of $\tilde{\Gamma}\sum$ and so forth (Pichler & Selitsch, 1999).

If we standardise the portfolio returns by letting:

$r_{p,t} = \frac{\widehat{r_{p,t}} - \mu_{p,t}}{\sigma_{p,t}}$, the higher moments of $r_{p,t}$ with $\kappa \geq 3$ are given by:

$$E\left(r_{p,t}{}^{k}\right) = \frac{\frac{1}{2}\kappa!\,\hat{\delta}^{T}\sum[\Gamma\sum]^{\kappa-2}\hat{\delta} + \frac{1}{2}(\kappa-1)!\cdot \mathrm{trace}[\Gamma\sum]^{\kappa}}{\sigma_{p,t}{}^{\kappa/2}} \tag{2.57}$$

$\kappa = 3$ gives the skewness and $\kappa = 4$ gives the kurtosis (Pichler & Selitsch, 1999).

So, to summarise, the steps to follow to compute the VaR for a non-linear asset are:

Step 1. We estimate $\tilde{\delta}_i$, $\tilde{\Gamma}_i$, $\tilde{\theta}_i$ for each option (from the Black-Scholes formula) and $\sigma^2{}_{i,t}$(from the volatility modelling methodologies, GARCH, EWMA and ICA-GARCH methodologies)

Step 2. We calculate the mean, variance, skewness coefficient and kurtosis coefficient using the formulae in Table 3.1

Step 3. We then use Slifker and Shapiro (1980)'s selection criteria to determine the distribution and estimates of the parameters γ, δ, ξ and λ. γ and δ are shape parameters, λ is a scale parameter and ξ is a location parameter for the normalising transformation.

Step 4. Compute the percentiles of $\widehat{r_{i,t}}$'s distributions based on the transformations in (2.50), (2.51) and (2.52)

Step 5. Calculate the σ for the portfolio

Step 5. Compute the VaR calculation.

2.5 Research Instruments

We use the following Applications for our analysis:

EViews 8: For regression analysis and estimating our volatility models.

MATLAB: For running the fast ICA algorithm. We used the robust fast ICA algorithm based on Hyvärinen and Oja, (2000) to estimate the Independent Components for our data.

MS Excel: For building our VaR computations. This also include coding in VBA for iterations that calculate the trace function as necessitated by equations in Section 2.4.2 above.

2.6 Data Requirements

We computed VaR calculations on a portfolio made up of 3 linear financial assets as follows:

i. A long call option on the GOLDS Index
i. A long put option on the FTSE/JSE TOP 40 Index
ii. A long call option on the ZAR/USD currency exchange rate.

These were largely obtained from the Bloomberg terminal.

Our period of observations is the period from December 31, 2010 to December 31, 2016. We further split the time series further into an in-sample period for model training and out-of-sample data for the evaluation of forecasting precision. In essence, we split our data as follows:

	Dates	Observations	Proportions
In-Sample Period	2010/12/31 to 2015/12/31	1305	83 %
Out-Of-Sample Period	2015/12/31 to 2016/12/31	261	17%
Total	2010/12/31 to 2016/12/31	1565	100%

Table 3: In and Out-of-Sample Split

We obtained data from the Bloomberg Terminal using the following Bloomberg tickers:

i. Equities – FTSE/JSE Africa Top40 Tradeable Index (TOP40)
ii. Commodities – Gold Spot $/OZ Commodity Index (GOLDS)
iii. Foreign Exchange – USD/ZAR Exchange Rates

2.7 Pre-processing

We lay-out in detail the steps we followed to make sure that the data is the right structure for our analysis in Section 4.1 below. However, at a glance, we first test the raw price data for stationarity. For this we use the Augmented Dickey-Fuller test performed on the price data for the different indices. The hypothesis is as follows,

$H_0: y_t \sim I(1)$ Null hypothesis series is non-stationery - has a unit root

$H_1: y_t \sim I(0)$ Alternative hypothesis the series is stationary - does not have a unit root

If we establish that the data is stationary, we transform it into another form by some sort of transformation which here will be differencing or taking logs before differencing. By stationary, we mean here that the statistical properties of the data such as its joint probability distribution (strict stationarity) as well as its first two moments (weak stationarity) remain constant over time.

Further, we use log-linear interpolation to deal with discontinuities in the data which could possibly give rise to discontinuities. Before we perform the Independent Component Analysis, we also center and whiten the data as laid out in Section 2.3.3 above. Once we have carried out these tasks and made one final test for stationarity on the processed data, we will then continue with our analysis.

Further, we use the following system of equations to fit the EWMA and GARCH model onto the data:

Univariate GARCH(1,1)

In order to fit the $GARCH(1,1)$ model to our data, we use the following model:

$$\sigma_t^2 = \gamma V_L + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2$$

Or equivalently:

$$\sigma_t^2 = \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2$$

where: σ_t^2 is the conditional variance, V_L is the long-term average volatility, u_{t-1}^2 is the square if the previous period residual, σ_{t-1}^2 is the fitted variance from the model during the previous period and γ, α & β are the parameters of the model

The log-likelihood function for the $GARCH(1,1)$ model with normal distribution becomes:

$$L_N(\theta) = ln \prod_t \frac{1}{\sqrt{(2\pi\sigma_t^2)}} e^{-\frac{u_t^2}{2\sigma_t^2}}$$

$$= -\frac{1}{2}\sum_t \left[ln(2\pi) + ln(\sigma_t^2) + \frac{u_t^2}{\sigma_t^2}\right]$$

So, in Excel we will use Solver to solve the following optimisation problem:

$$\min L_N(\theta) = \frac{1}{2}\sum_t \left[ln(2\pi) + ln(\sigma_t^2) + \frac{u_t^2}{\sigma_t^2}\right]$$

subject to

$$\sigma_t^2 - \omega - \alpha u_{t-1}^2 - \beta \sigma_{t-1}^2 = 0$$

$$-\omega \leq 0$$

$$-\alpha \leq 0$$

$$-\beta \leq 0$$

$$\alpha + \beta - 1 \leq 0$$

The optimisation problem above leads to estimates of γ, α & β

Exponentially Weighted Moving Average (EWMA)

The exponentially weighted moving average (σ_t) is calculated as:

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)r_{t-1}^2$$

where: r_t is the value of the time series at time *t* and λ is the smoothing parameter (a non-negative constant between 0 and 1)

Now, the EWMA is a special case of the $GARCH(1,1)$ and the $GARCH(1,1)$ is a generalised case of the EWMA. The main difference between these two models is that GARCH includes the additional term for mean reversion, the γV_L term. To go from $GARCH(1,1)$ to EWMA, consider the model below:

$$GARCH(1,1) = \sigma_t^2 = \omega + \alpha u_{t-1}^2 + \beta\sigma_{t-1}^2$$

Letting $\omega = 0$ and $(\alpha + \beta) = 1$, the above expression simplifies to:

$$GARCH(1,1) = \sigma_t^2 = \alpha u_{t-1}^2 + (1-\alpha)\sigma_{t-1}^2$$

This is now equivalent to the formula for the EWMA:

$$EWMA = \sigma_t^2 = \alpha u_{t-1,t}^2 + (1-\alpha)\sigma_{t-1}^2$$

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)u_{t-1,t}^2$$

The λ is the "decay" parameter. RiskMetrics (1996) proposed a value for the decay factor λ of 0,94 for daily data at a level of tolerance (γ_{tol}) of 0,01. The relationship linking the decay parameter, level of tolerance and required number of historical returns is given by:

$$1 - \lambda^n = (1 - \gamma_{tol})$$

We can then re-write the EWMA model as:

$$\sigma_t^2 = 0{,}94\sigma_{t-1}^2 + (0{,}06)u_{t-1}^2$$

ICA-GARCH(1,1)

The procedure for the ICA-GARCH(1,1) is the same as that for the Univariate-GARCH(1,1) except that now we run the FastICA algorithm on the data to obtain the Independent Components before fitting a GARCH(1,1) onto these Independent Components.

2.8 Back testing

Here we present a simple method of determining the appropriateness of the models. We construct 1-day VaR forecasts over a selected period say 6-months period. We then compare these forecasts to the actual realized profits or losses represented here by the 1-day returns.

Recall that our portfolio return is calculated as:

$$r_{p,t} = \sum_{i=1}^{2} w_i r_{i,t}$$

where w_i is the weight of each asset and $r_{i,t}$ is the return of each asset in the portfolio.

The Value-at-Risk was computed and presented in the preceding sections.

Assessing Model Performance

The simplest measure of performance which we employ here is a count of the number of times that the VaR estimate falls short in predicting future losses/gains. In other words, "under estimates" future losses/gains. Here we assume that on each particular day there is a 5% chance that the observed loss exceeds the forecast VaR

To give more perspective, let us start with a random variable *X(t)* on any day t such that *X(t)* = 1 if the realised loss is greater than the forecast VaR and *X(t)* = 0 otherwise.

The distribution of *X(t)* can be thought of as a Bernoulli distribution written in the form:

$$f(X(t)|0{,}05) = \begin{cases} 0{,}05^{X(t)}(1-0{,}05)^{1-X(t)} & X(t) = 0 \; or \; 1 \\ 0 & Otherwise \end{cases}$$

Let's suppose we observe *X(t)* for a total of *T* days, *t=1,2,3,…,T*. The random variable *X(t)* has an expected value of 0,05 (from the mean of the Bernoulli distribution). The total number of violations of VaR over this period of time is given by

$$X_T = \sum_{t=1}^{T} X(t)$$

Given our level of confidence *a*, the expected value of X_T which is the expected number of violations of VaR violations over T Days is given by

$$T \times a$$

So, in the case where we are dealing with say $a = 5\%$ over 260 days, the expected number of violations of VaR is 13. Therefore, we expect to observe 1 violation of VaR every 20 days.

The importance of this simplified method lies in the fact that the probability of observing violations of VaR over *T* days is the same as the probability of observing violations of VaR at any point in time *t*.

2.9 Conclusion

In this chapter, we have explored some of the key literature and concepts for the topics relevant to our research. We have looked at the methods of estimating Value-at Risk and also set the tone on the use of univariate vs multivariate volatility estimates. Whereas multivariate volatilities are more accurate, the main challenge with their estimation is the ease of computation. Independent Component Analysis (ICA) from statistics and signal processing can be applied to model multivariate asset return volatilities as a linear combination of several univariate Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models. We have also laid down in greater detail the method of computing VaR that we used and the steps we followed together with our research methodology as well as the data analysis methods and tools we used in our research

3 Results

In this chapter, we assess the performance of the different methods for computing portfolio Value-at-Risk. First, we analyse the three volatility models under study (univariate GARCH, ICA-GARCH and EWMA). We then go on to use these volatility models to compute the VaR estimates for our portfolios and present the results. Finally, we use back-testing methods to evaluate the performance of these models.

3.1 Volatility Models

We used three methods for computing the volatility namely, the univariate-GARCH, EWMA and the ICA-GARCH approach. Before carrying out our analysis, we pre-processed the data using EViews 8 as follows:

3.1.1 Data Pre-Processing

Our stock price data for the Golds Index, Top40 Index and USDZAR exchange rate is in South African Rand and as such rebasing is not necessary.

Descriptive Statistics and Time Series Analysis:

Before carrying out any empirical project, it is necessary to perform a descriptive analysis of the data in order to note patterns, unusual behaviours and trends in the data. The observations noted will be of great use when later analysing and interpreting the results of empirical analysis.

The graphs below are the time series plots for the price data for the Golds Index, Top40 Index and USDZAR for the period 2010/12/31 to 2016/12/31 in EViews 8. Note the fall in prices of the GOLDS index from 2012 when would commodity prices fell:

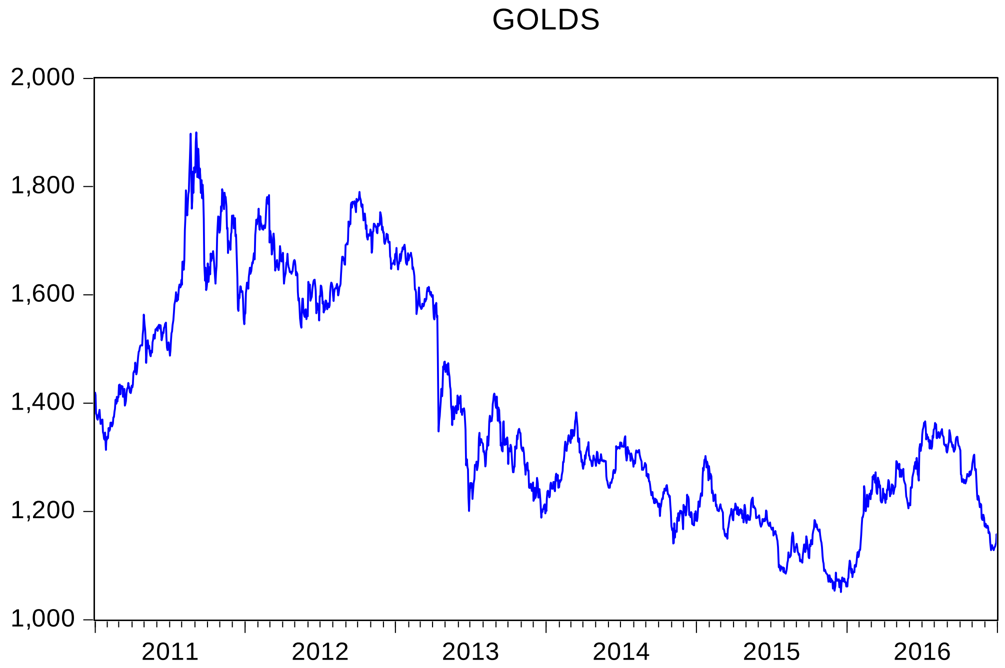


Figure 7: GOLS Index

In contrast, the TOP40 index has been steadily increasing recovering from the 2008 financial crisis:

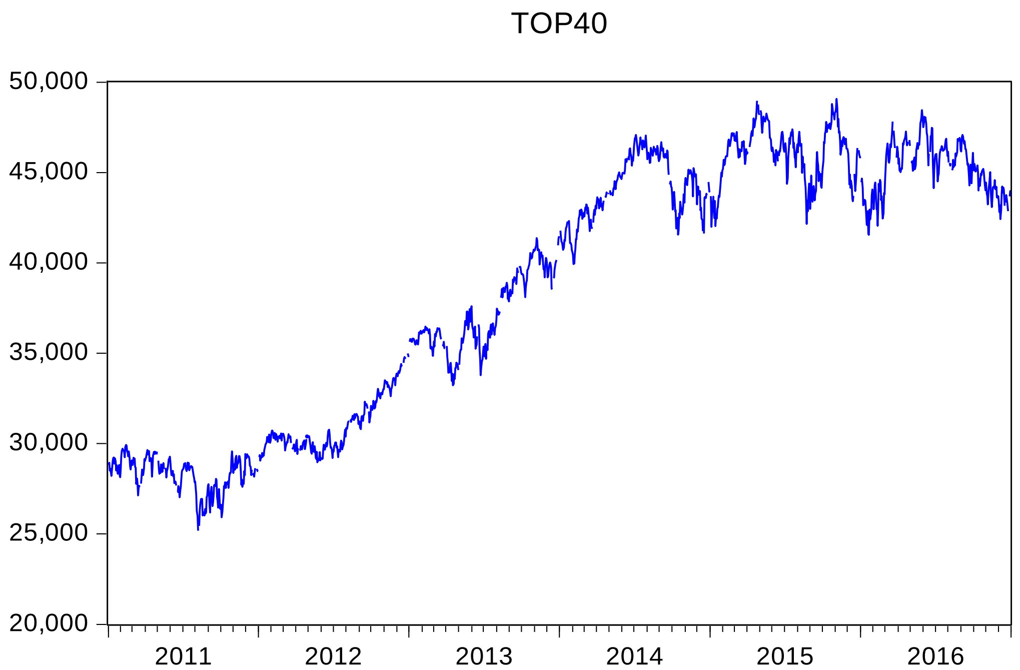


Figure 8: TOP40 Index

In the same way as the fall in commodity prices, the rand was also depreciating over the same period losing its value against the US dollar:

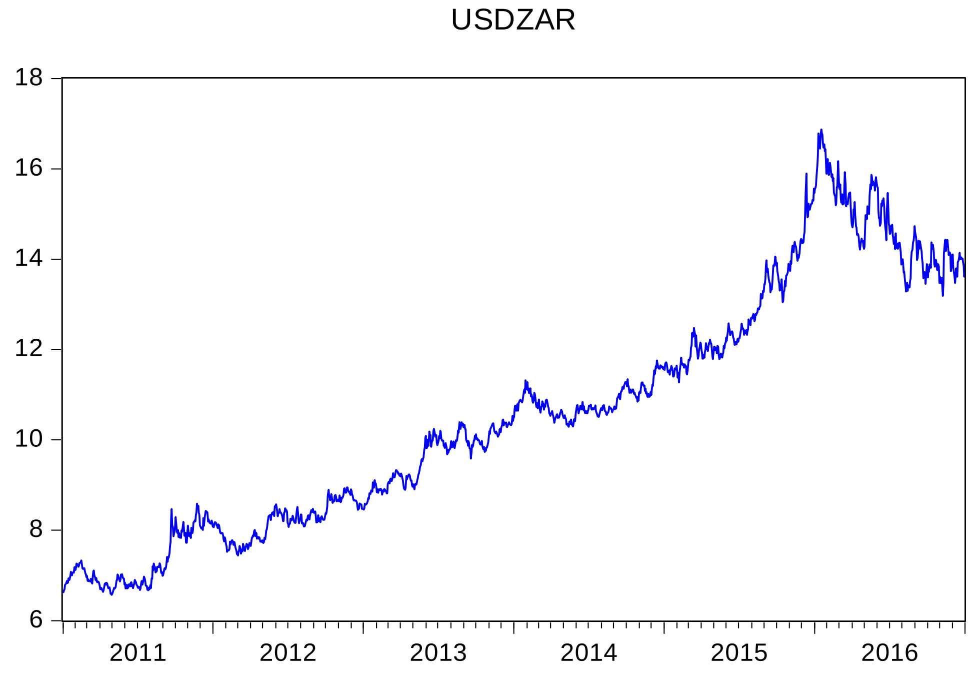


Figure 9: USDZAR Exchange Rate

There is strong evidence of trends in the price data. The impulsiveness in prices is seen with price increases rather than with price decreases. This is to say that for all data sets there are more significant spikes observed in the increase in prices and less significant spikes observed in the decline in prices. There are also numerous positive outliers for price increase spikes. The plots of the stock prices characterise the stylised features of price data.

Stationarity Test

Finally, we have to check whether our series are stationary or not. This is very important, for the stationarity or otherwise of a series can strongly influence its behaviour and properties. The hypothesis being tested is whether the series is non-stationary that is, contains a unit root.

$H_0: y_t \sim I(1)$ Null hypothesis series is non-stationery - has a unit root

$H_1: y_t \sim I(0)$ Alternative hypothesis the series is stationary - does not have a unit root

The table below extracted from EViews 8 shows the test statistics and p-values for the Augmented Dickey-Fuller test performed on the price data for different indices:

Null Hypothesis: GOLDS has a unit root
Number of Lags: 12
Method: Least Squares

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1,2318	0,6627
Test critical values:	1% level	-3,4343	
	5% level	-2,8632	
	10% level	-2,5677	

*MacKinnon (1996) one-sided p-values.

Table 4: ADF Test on Golds Index

Null Hypothesis: TOP40 has a unit root
Number of Lags: 12
Method: Least Squares

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1,0164	0,7493
Test critical values:	1% level	-3,4350	
	5% level	-2,8635	
	10% level	-2,5679	

*MacKinnon (1996) one-sided p-values.

Table 5: ADF test on Top40 Index

Null Hypothesis: USDZAR has a unit root
Number of Lags: 12
Method: Least Squares

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1,0923	0,7209
Test critical values:	1% level	-3,4343	
	5% level	-2,8632	
	10% level	-2,5677	

*MacKinnon (1996) one-sided p-values.

Table 6: ADF on USDZAR

The test statistics in all 3 cases are less negative than the test critical values and hence there is no sufficient evidence to reject the null hypothesis of a unit root in the stock prices. Therefore, the stock prices themselves are not stationary.

Dealing with Non-Stationarity

We then took log percentage returns of the prices for the three indices for the period 2010/12/31 to 2016/12/31 to create three new series. These series were created in the form of three new series dgolds, dtop40 and dusdzar in EViews 8 using the following formulae:

d*Indexname* = 100*dl*Indexname* = 100*(l*Indexname* – l*Indexname* (-1)) where l*Indexname* = *log(Indexname)* and *Indexname* = {GOLDS, TOP40 and USDZAR}

The table below shows the test statistics and p-values for the Augmented Dickey-Fuller test performed on the log percentage returns for our indices:

Null Hypothesis: DGOLDS has a unit root
Number of Lags: 12
Method: Least Squares

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-39,4363	0,0000
Test critical values:	1% level	-3,4343	
	5% level	-2,8632	
	10% level	-2,5677	

*MacKinnon (1996) one-sided p-values.

Table 7: ADF Test on DGolds Index

Null Hypothesis: DTOP40 has a unit root
Number of Lags: 12
Method: Least Squares

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-28,7226	0,0000
Test critical values:	1% level	-3,4350	
	5% level	-2,8635	
	10% level	-2,5679	

*MacKinnon (1996) one-sided p-values.

Table 8: ADF test on DTop40 Index

Null Hypothesis: DUSDZAR has a unit root
Number of Lags: 12
Method: Least Squares

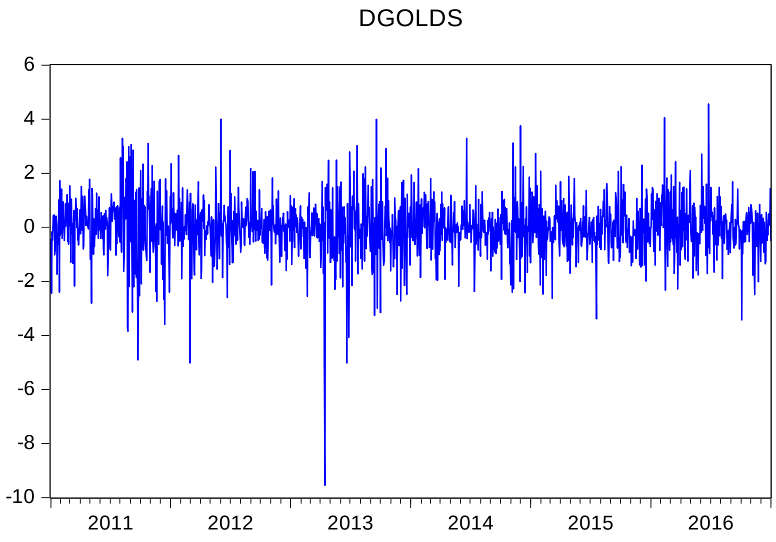
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-39,4056	0,0000
Test critical values:	1% level	-3,4343	
	5% level	-2,8632	
	10% level	-2,5677	

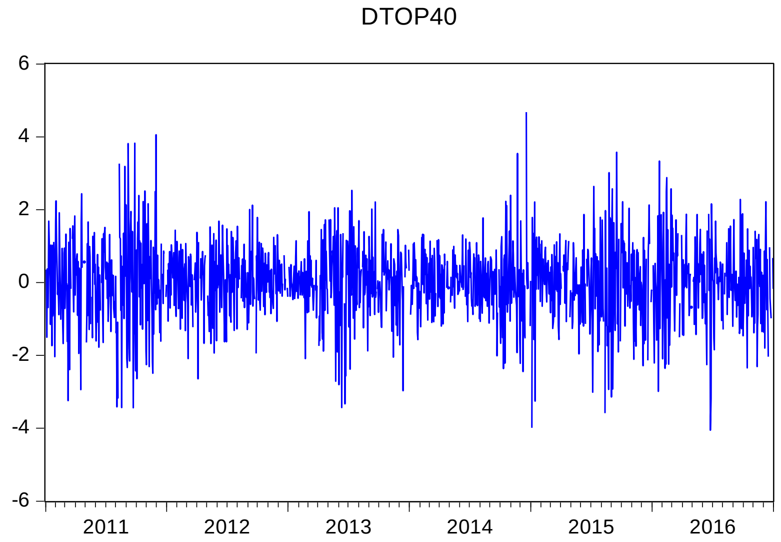
*MacKinnon (1996) one-sided p-values.

Table 9: ADF on DUSDZAR

The test statistics in all three cases are more negative than the test critical values and hence the null hypothesis of a unit root in the log percentage returns is convincingly rejected. The log percentage returns are stationary and as such, our analysis is going to be based on the log percentage returns rather than the stock prices themselves.

A plot of these three new series generated for the log percentage returns is shown in the figure below:





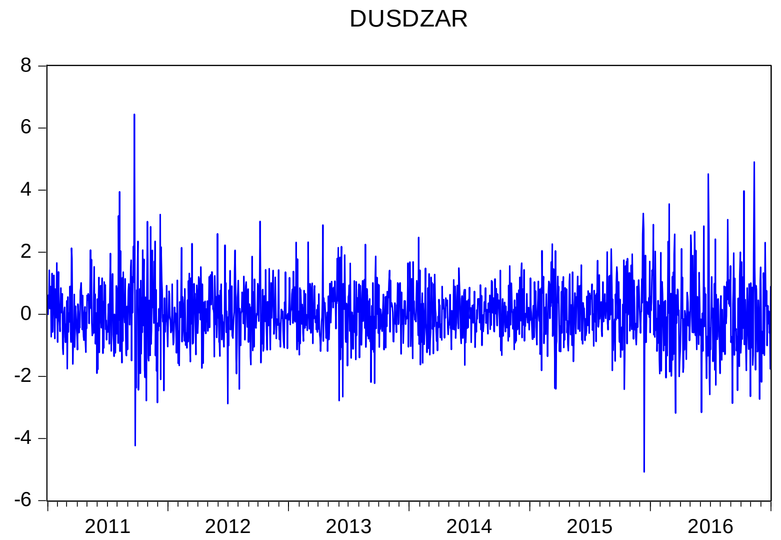


Figure 10: Daily log percentage returns for GOLDS, TOP40 and USDZAR

Note the existence of outliers in all three plots. The time series plots show volatility clustering in other words, the current level of volatility seems to be positively correlated with its level during the preceding periods.

Test for Normality

Below is a histogram of the log percentage returns of the stocks for all three indices for the periods 2010/12/31 to 2016/12/31together with the descriptive statistics. These graphs were extracted from EViews 8.

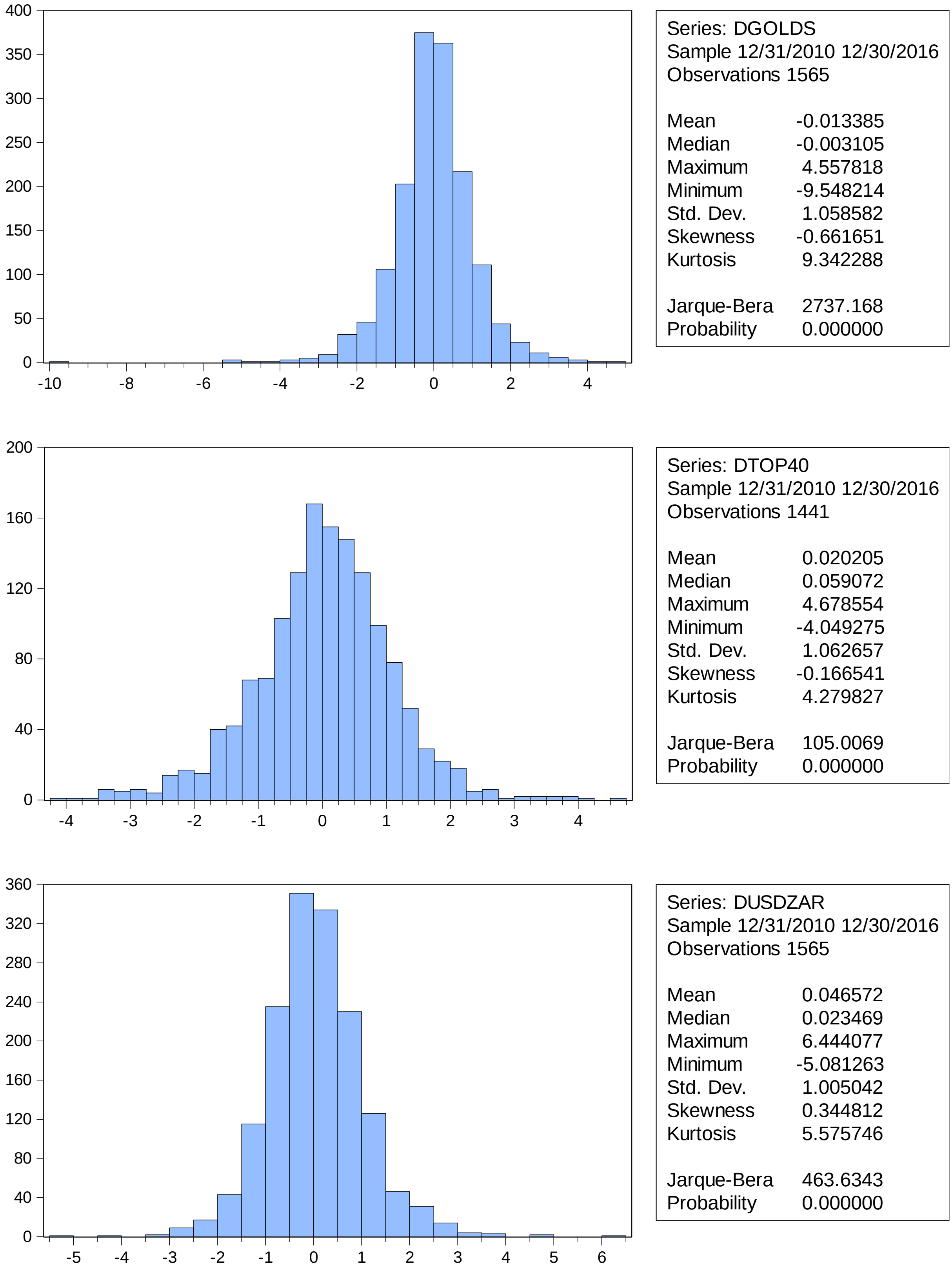


Figure 11: Histogram and stats for the indices

From the descriptive statictics for the log percentage returns for each stock price data set, both series are not normally ditsibuted. The skewness for all three log percentage returns is not 0 and are thus asymmetric about their mean value and have an excess kurtoses of at least 4 for all three cases. This however is one of the sylised fetures of financial time series data. Jarque-Bera p-values of 0 shows strong evidence against normality.

Dealing with missing values

Any GARCH analysis works with continuous data. If there are breaks in the time series due to holidays, EViews will throw errors and the GARCH analysis will fail. To eliminate this problem, we carried out a log-linear interpolation on the TOP40 index to create the TOP40b index which is a continuous series. The graph below is the time series plot for the price data for the Top40 Index for the period 2010/12/31 to 2016/12/31:

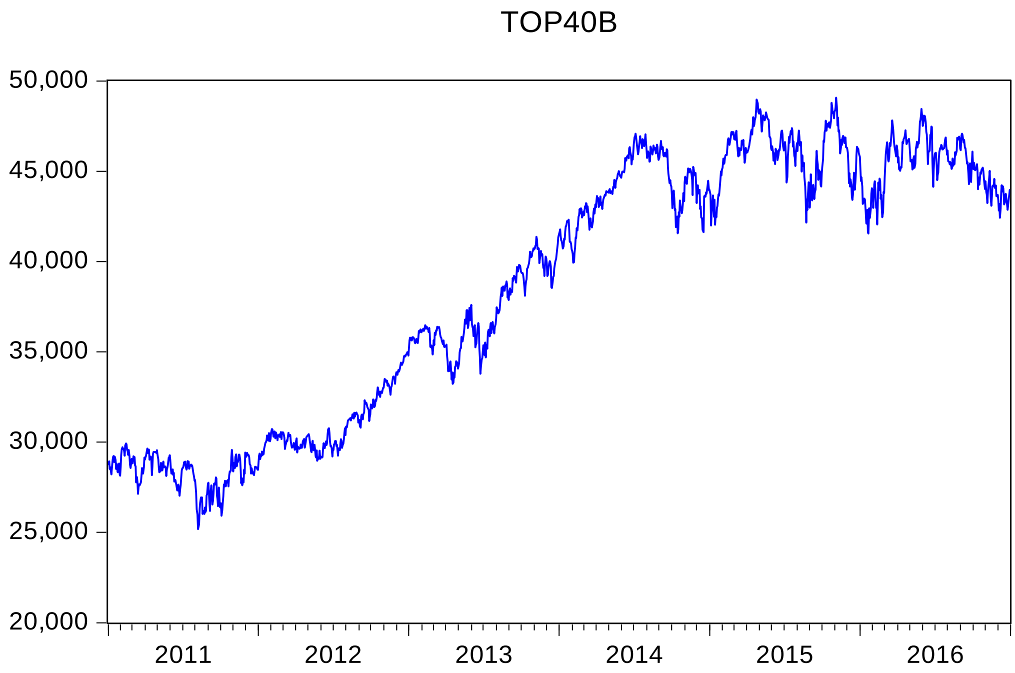


Figure 12: TOP40B Index

Compared to the TOP40 index we plotted earlier,

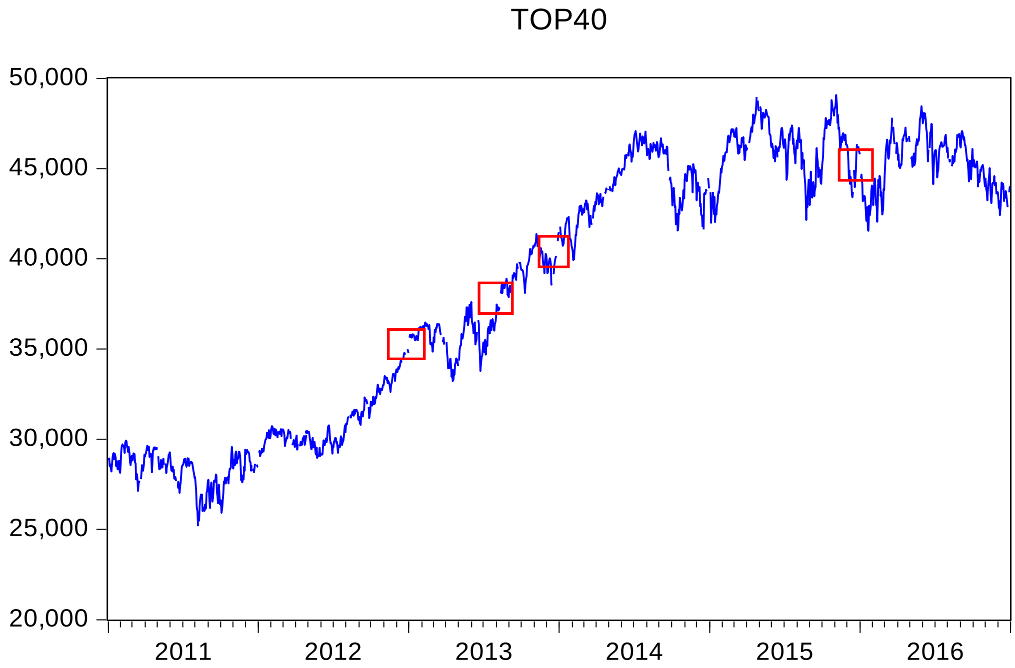


Figure 13: TOP40 Index

the plots are the same except that we have now dealt with the discontinuities highlighted in red above.

Our analysis of the top 40 index was thus carried out on the TOP40B index.

Independent Component Analysis

To carry out the ICA, we exported the data from EViews into MATLAB where we used the FastICA algorithm to carryout Independent Component Analysis on our data:

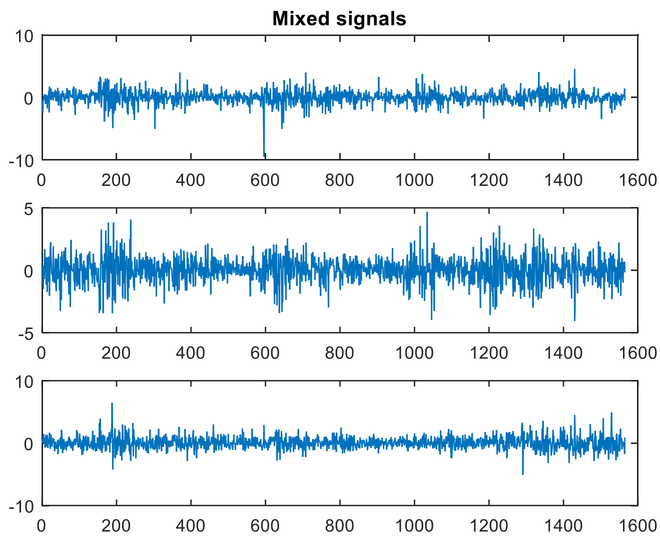


Figure 14: Plot of the mixed signals.

These mixed signals (from top to bottom) are the original DGOLDS, DTOP40b and USDZAR indices before we have carried out the ICA. Notice that this is the same as the plots in Figure 10 above. Because we can't work with the data as time series in the FastICA algorithm, we have observation number and not necessarily the date on the horizontal axis. We will recombine the series with the dates after carrying out the Independent Component Analysis.

Earlier we mentioned that the pre-processing of data for ICA involves centering followed by Sphering or Whitening. In the figure below we have plotted a graph of the whitened signals:

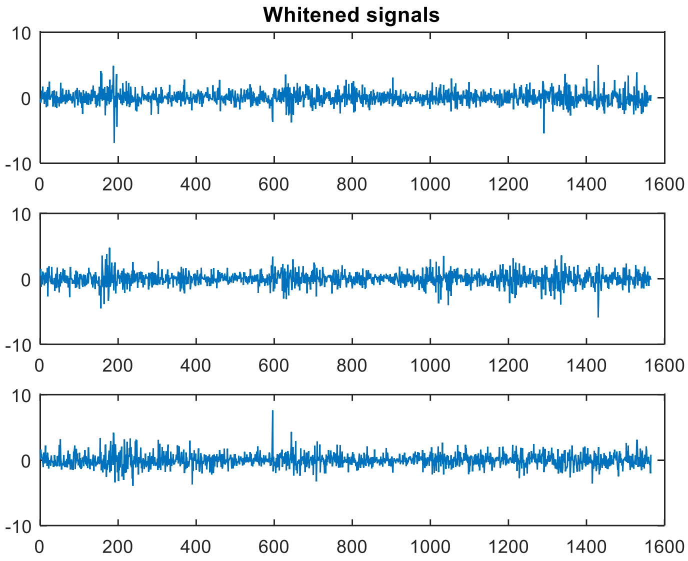


Figure 15: Plot of the whitened signals

The figure below is a plot of the Independent Components:

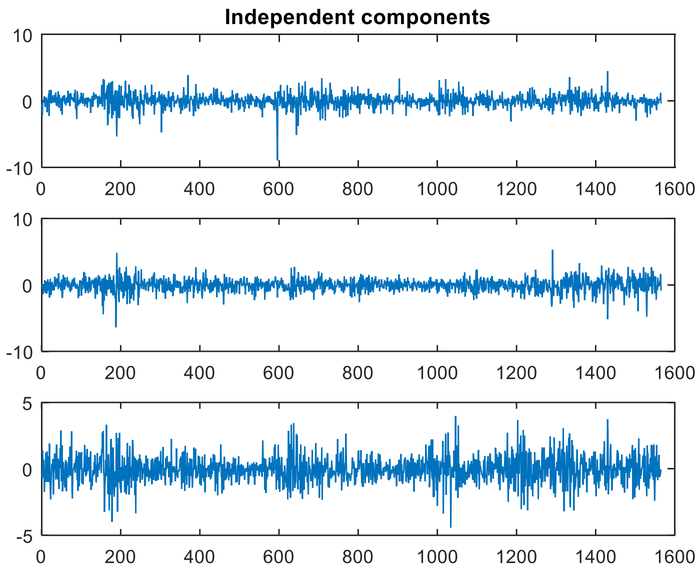


Figure 16: Plot of the Independent Components

The next step after carrying out the ICA and obtaining out Independent Components is to export the data from MATLAB back into EViews for further analysis.

To put things into further perspective, we have plotted below the graphs of each of the mixed signals superimposed on that of the Independent Components:

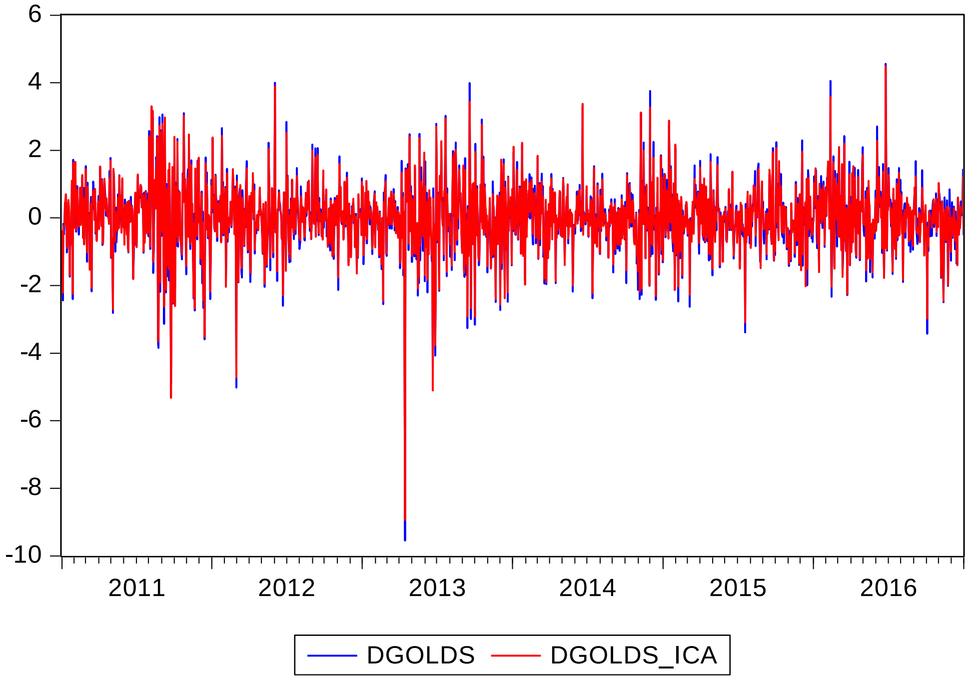


Figure 17: Comparing the GOLDS Mixed Signals to the Independent Components

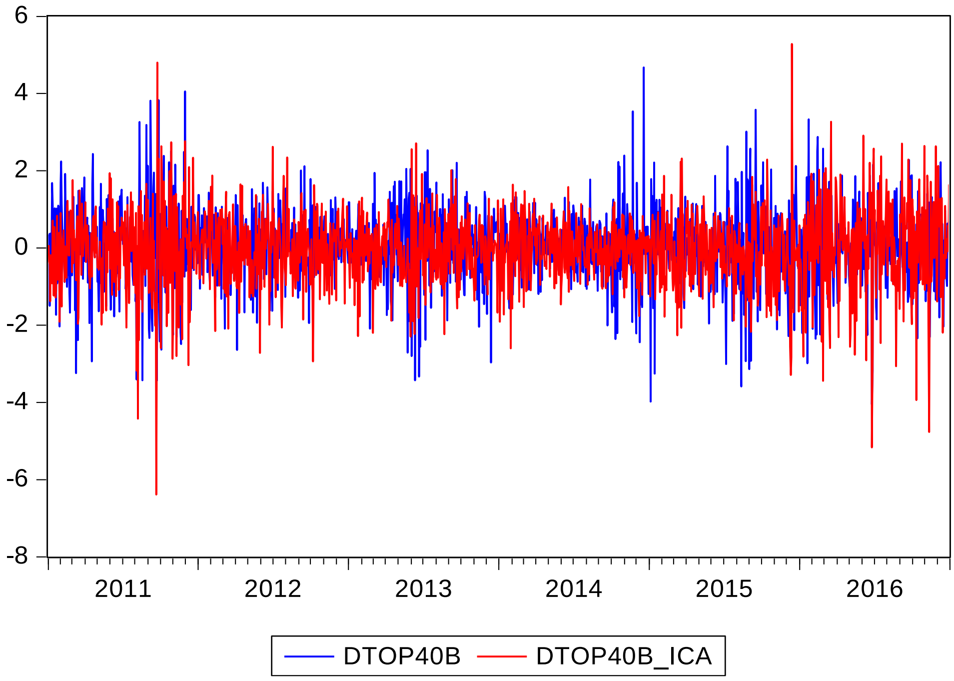


Figure 18: Comparing the TOP40B Mixed Signals to the Independent Components

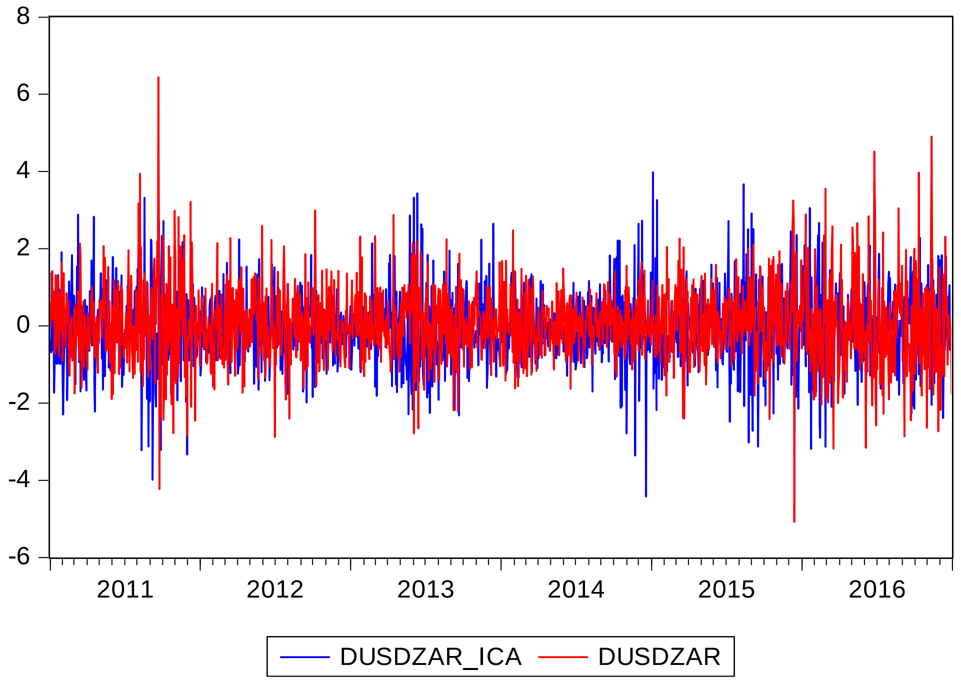


Figure 19: Comparing the USDZAR Mixed Signals to the Independent Components

Parameter Estimation of the Volatility Models

The next step involves Parameter Estimation

We began by estimating the parameters for the Univariate-GARCH(1,1) model. For parameter estimation, we fitted the GARCH(1,1) model on the data for the period 2010/12/21 to 2015/12/31 leaving out 2016 for the out of sample tests. The results are as follows:

Dependent Variable: DGOLDS
Method: ML - ARCH (Marquardt) - Normal distribution
Sample (adjusted): 1/03/2011 12/31/2015
Variance Equation
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0,0578	0,0104	5,5695	0,0000
RESID(-1)^2	0,0784	0,0063	12,4192	0,0000
GARCH(-1)	0,8711	0,0133	65,2792	-

R-squared	-0,0000
Adjusted R-squared	-0,0000
Akaike info criterion	2,8754
Schwarz criterion	2,8873

Table 10: GARCH(1,1) on Golds Index

Dependent Variable: DTOP40B
Method: ML - ARCH (Marquardt) - Normal distribution
Sample (adjusted): 1/03/2011 12/30/2015

Variance Equation

GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0,0156	0,0050	3,1076	0,0019
RESID(-1)^2	0,0743	0,0136	5,4810	0,0000
GARCH(-1)	0,9124	0,0151	60,6037	-

R-squared	-0,0002
Adjusted R-squared	-0,0002
Akaike info criterion	2,7364
Schwarz criterion	2,7523

Table 11: GARCH(1,1) on Top40 Index

Dependent Variable: DUSDZAR
Method: ML - ARCH (Marquardt) - Normal distribution
Sample (adjusted): 1/03/2011 12/30/2015

Variance Equation

GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0,0468	0,0148	3,1573	0,0016
RESID(-1)^2	0,0945	0,0147	6,4447	0,0000
GARCH(-1)	0,8521	0,0288	29,6236	0,0000

R-squared	-0,0001
Adjusted R-squared	-0,0001
Akaike info criterion	2,5968
Schwarz criterion	2,6126

Table 12: GARCH(1,1) on USDZAR

Next we estimated the parameters for the ICA-GARCH(1,1) model. For parameter estimation, we fitted the ICA-GARCH(1,1) model on the data for the period 2010/12/21 to 2015/12/31 leaving out 2016 for the out of sample tests. It's important to note that the procedure is essentially the same as the Univariate-GARCH(1,1) model except that here we are now fitting the GARCH(1,1) model on our independent components generated by the FastICA algorithm in Matlab. The results are as follows:

Dependent Variable: DGOLDS_ICA
Method: ML - ARCH (Marquardt) - Normal distribution
Sample (adjusted): 1/03/2011 12/30/2015
Variance Equation
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

Variable	**Coefficient**	**Std. Error**	**z-Statistic**	**Prob.**
C	0,0421	0,0072	5,8244	0,0000
RESID(-1)^2	0,0698	0,0055	12,7650	0,0000
GARCH(-1)	0,8895	0,0104	85,8801	-

R-squared	-0,0000
Adjusted R-squared	-0,0000
Akaike info criterion	2,7703
Schwarz criterion	2,7763

Table 13: ICA-GARCH(1,1) on Golds Index

Dependent Variable: DTOP40B_ICA
Method: ML - ARCH (Marquardt) - Normal distribution
Sample (adjusted): 1/03/2011 12/30/2015
Variance Equation
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

Variable	**Coefficient**	**Std. Error**	**z-Statistic**	**Prob.**
C	0,0518	0,0151	3,4217	0,0006
RESID(-1)^2	0,1090	0,0162	6,7226	0,0000
GARCH(-1)	0,8306	0,0307	27,0448	0,0000

R-squared	-0,0001
Adjusted R-squared	-0,0001
Akaike info criterion	2,5644
Schwarz criterion	2,5803

Table 14: ICA-GARCH(1,1) on Top40 Index

Dependent Variable: DUSDZAR_ICA
Method: ML - ARCH (Marquardt) - Normal distribution
Sample (adjusted): 1/03/2011 12/30/2015
Variance Equation
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

Variable	**Coefficient**	**Std. Error**	**z-Statistic**	**Prob.**
C	0,0169	0,0051	3,3197	0,0009
RESID(-1)^2	0,0805	0,0144	5,5784	0,0000
GARCH(-1)	0,9036	0,0161	56,2267	-

R-squared	-0,0001
Adjusted R-squared	-0,0001
Akaike info criterion	2,6638
Schwarz criterion	2,6796

Table 15: ICA-GARCH(1,1) on USDZAR

The parameter estimates for the EWMA model are already specified as GARCH 0.94 and ARCH 0.06 as derived in Section 3.6 above and as such we do not carry out any parameter estimation here and go ahead and forecast using our models estimated above

3.2 VaR Forecasts

After the volatility modelling and forecasting, we went ahead and built our VaR models. For this we created a portfolio of options as follows:

ii. A long call option on the GOLDS Index

iii. A long put option on the FTSE/JSE TOP 40 Index

iv. A long call option on the ZAR/USD currency exchange rate.

All 3 were 6 month options. The steps followed can be summarised as follows:

Step 1. Estimate $\tilde{\delta}_i, \tilde{\Gamma}_i, \tilde{\theta}_i$ for each option (from the Black-Scholes formula) and we have ${\sigma^2}_{i,t}$(from the volatility modelling methodologies, GARCH, EWMA and ICA-GARCH methodologies above)

Step 2. Compute the portfolio return based on the system of equations we specified earlier in Section 2.4.2

Step 3. Calculate the mean, variance, skewness coefficient and kurtosis coefficient for our portfolio return.

Step 4. Use Slifker and Shapiro (1980)'s selection criteria to determine the distribution and estimates of the parameters γ, δ, ξ and λ. γ and δ are shape parameters, λ is a scale parameter and ξ is a location parameter for the normalising transformation.

Step 5. Compute the percentiles of $\widehat{r_{i,t}}$'s distributions based on the Johnson Translation

Step 6. Calculate the σ for the portfolio from the portofolio variance computed in Step 3.

Step 7. Use the percentile in Step 4 and the σ in Step 5 to compute the VaR calculation.

As we outlined in Section 3.9, we assess the appropriateness of a VaR model based on the total number of violations of VaR over a specified period of time. In the following section, we present the results for the out-of-sample tests. In all cases, the strike price is chosen such that the options are at the money at the start of the forecast period. We now present our results below:

Results Excluding Theta (Delta-Gamma Method)

We now show the results of the number of VaR violations for the Delta VaR compared to the Delta-Gamma-(Theta)-VaR when we exclude Theta from the calculations:

	Number of Violations	**Expected No of Violations**	**Percentage of Violations**
Delta-Gamma VaR (Normal)	55	7	42%
Delta-Gamma VaR (Translated)	58	7	45%

Table 16: Results Comparison for the Univariate-GARCH Model

	Number of Violations	**Expected No of Violations**	**Percentage of Violations**
Delta-Gamma VaR (Normal)	55	7	42%
Delta-Gamma VaR (Translated)	58	7	45%

Table 17: Results Comparison for the EWMA Model

	Number of Violations	**Expected No of Violations**	**Percentage of Violations**
Delta-Gamma VaR (Normal)	51	7	39%
Delta-Gamma VaR (Translated)	51	7	39%

Table 18: Results Comparison for the ICA-GARCH Model

Results Including Theta (Delta-Gamma-Theta)

The number of VaR violations for the Delta VaR compared to the Delta-Gamma-(Theta)-VaR when we include Theta in the calculations are as follows:

	Number of Violations	Expected No of Violations	Percentage of Violations
Delta-Gamma-Theta VaR (Normal)	3	7	2%
Delta-Gamma-Theta VaR (Translated)	-	7	-

Table 19: Results Comparison for the Univariate-GARCH Model

	Number of Violations	Expected No of Violations	Percentage of Violations
Delta-Gamma-Theta VaR (Normal)	4	7	3%
Delta-Gamma-Theta VaR (Translated)	-	7	-

Table 20: Results Comparison for the EWMA Model

	Number of Violations	Expected No of Violations	Percentage of Violations
Delta-Gamma-Theta VaR (Normal)	16	7	12%
Delta-Gamma-Theta VaR (Translated)	5	7	4%

Table 21: Results Comparison for the ICA-GARCH Model

From the results above, we see that for the Delta-Gamma method (excluding theta) we can still use the Normal Distribution percentile for the VaR computations as this leads to better VaR forecasts than the Translated Percentiles. However, both sets of estimates are still providing poor VaR forecasts with VaR violations exceeding 30% on average.

When we include the higher moment, theta to get the Delta-Gamma-Theta forecasts, the Normal Distribution percentile is now producing weak forecasts with the Translated Percentiles giving 0 violations in all three cases (Univariate-GARCH, EWMA and ICA-GARCH). On the whole, the Delta-Gamma-Theta methodology is a better VaR forecasting technique than the Delta-Gamma method.

We also see that the ICA-GARCH methodology failed to produce better VaR forecasts for our sample period as it produced more VaR violations that the other 2 (Univariate-GARCH and EWMA).

We now look at the translated percentiles from the Univariate-GARCH, EWMA and ICA-GARCH under the Delta-Gamma-Theta methodology to compare the "strictness" of the forecasts

We now rank the absolute values of these translated percentiles:

	Translated Percentile		
	Delta	**Delta-Gamma**	**Delta-Gamma-Theta**
Univariate-GARCH	0,04	0,76	24,50
EWMA	0,05	1,33	22,06
ICA-GARCH	0,04	0,77	4,32

Table 22: Absolute values of the Translated 5th Percentiles

	Translated Percentile		
	Delta	**Delta-Gamma**	**Delta-Gamma-Theta**
Univariate-GARCH	8	6	1
EWMA	7	4	2
ICA-GARCH	9	5	3

Table 23: Ranking the absolute values of the Translated 5th Percentiles

The tables above show that the 5th percentiles of the different methodologies. The Univariate-GARCH Delta-Gamma-Theta methodology provides the best VaR estimates based on the critical values of the translated percentiles.

Additionally, we now show the different VaR estimates superimposed on the same graph as the portfolio return:

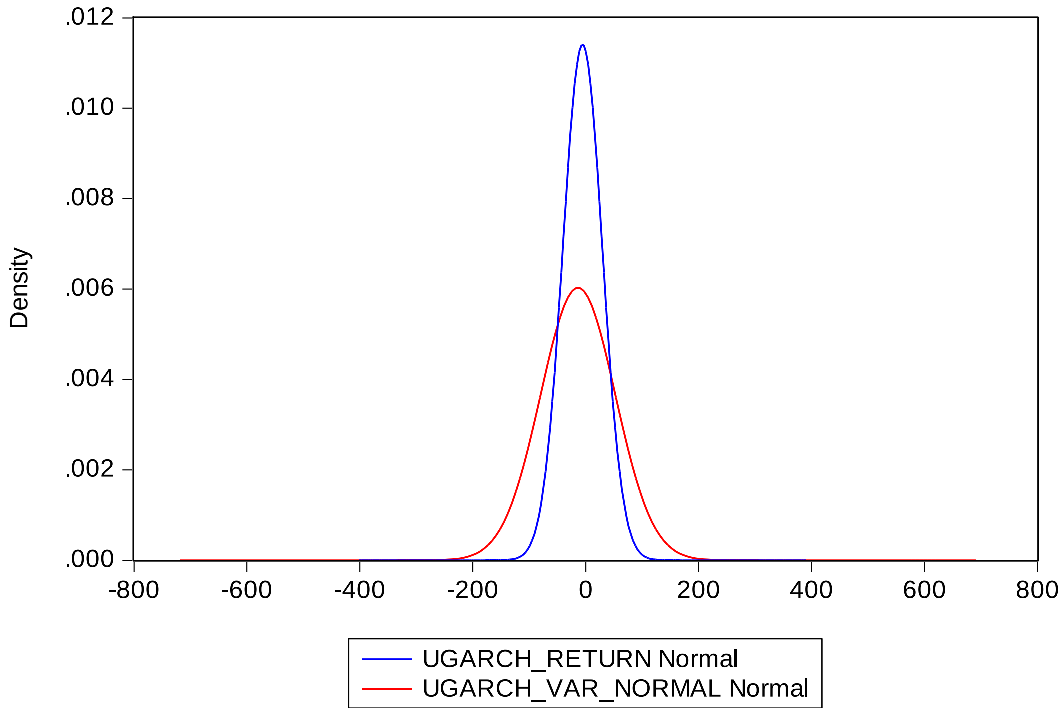


Figure 20: Univariate-GARCH Delta-Gamma-Theta Normal Estimate

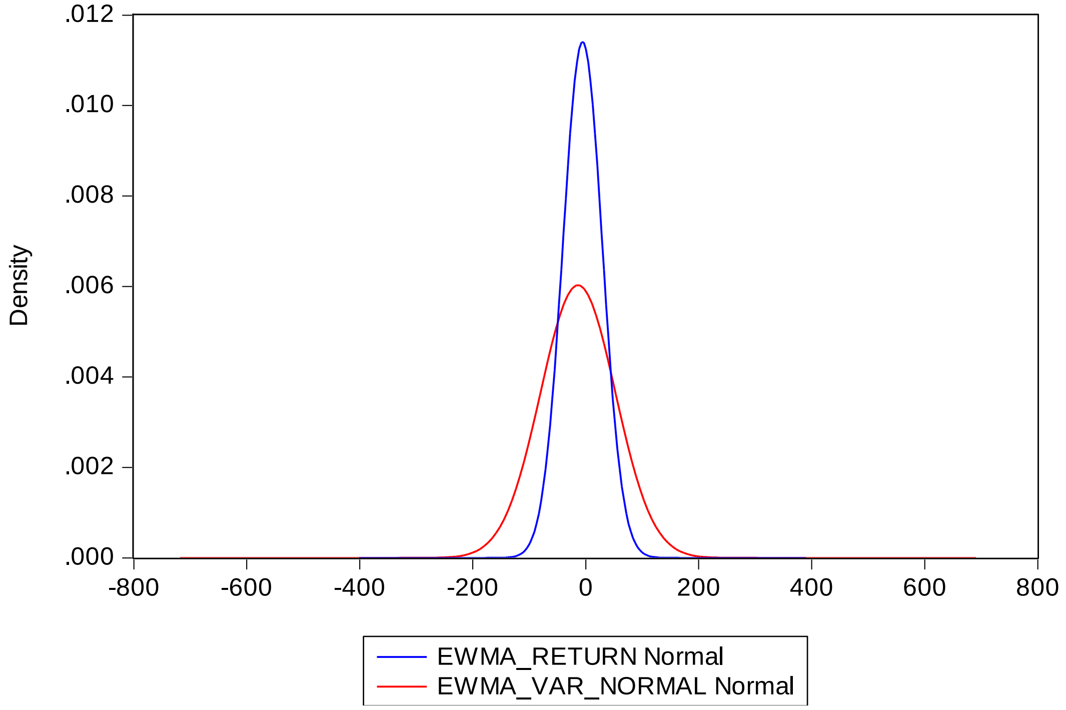


Figure 21: EWMA Delta-Gamma-Theta Normal Estimate

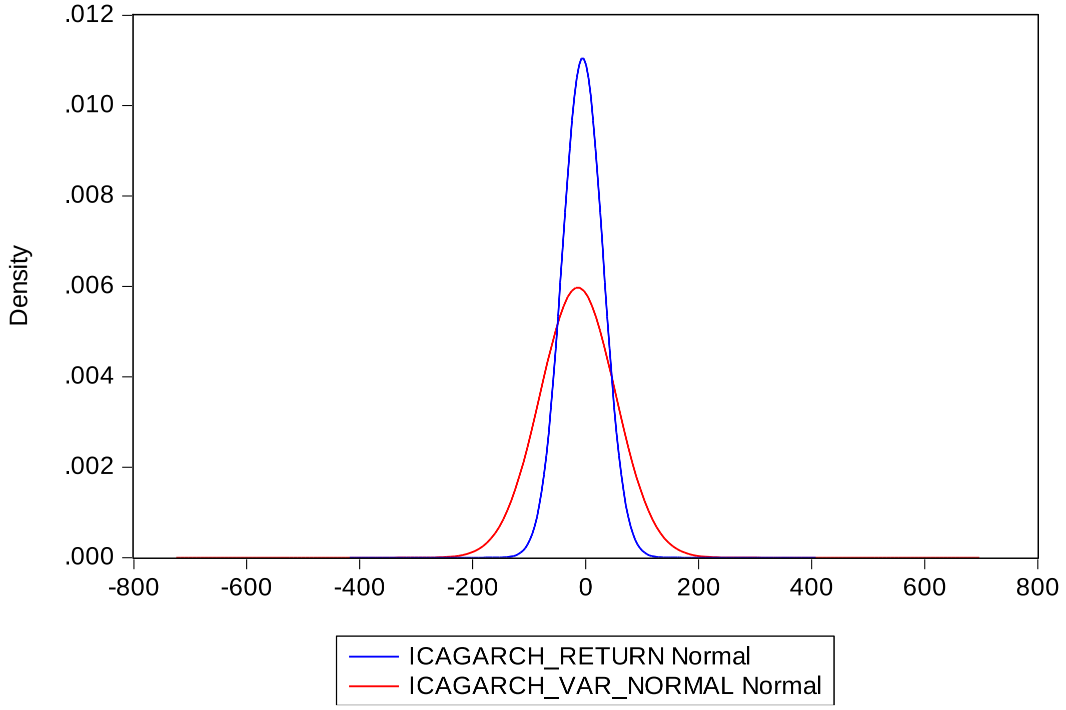


Figure 22: ICA-GARCH Delta-Gamma-Theta Normal Estimate

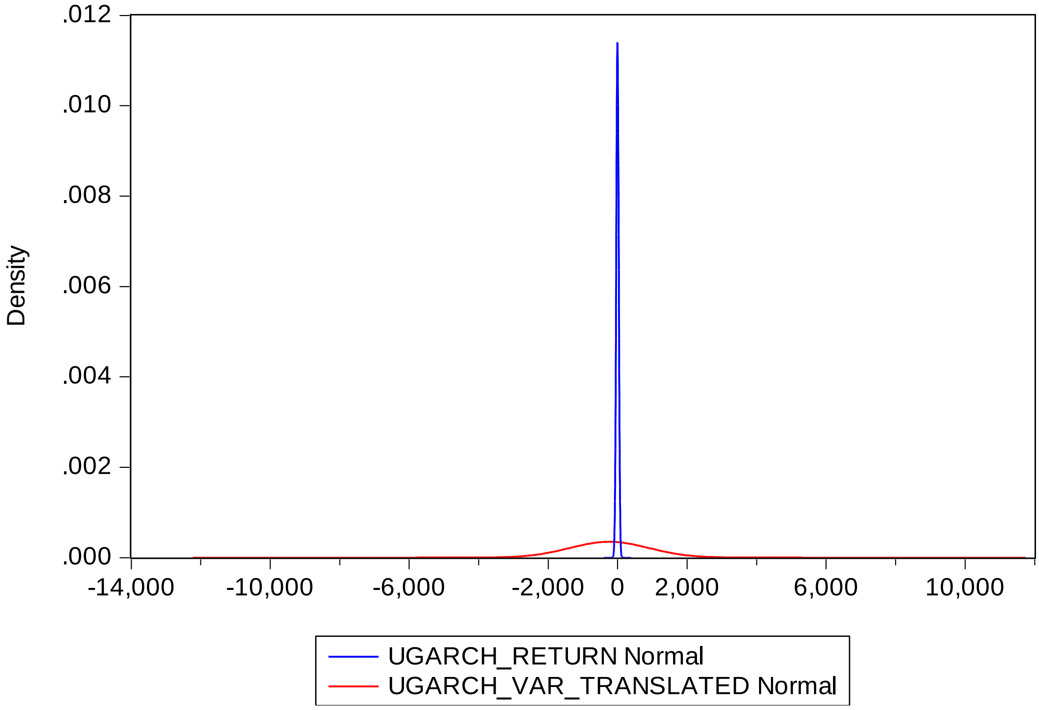


Figure 23: Univariate Delta-Gamma-Theta Translated Estimate

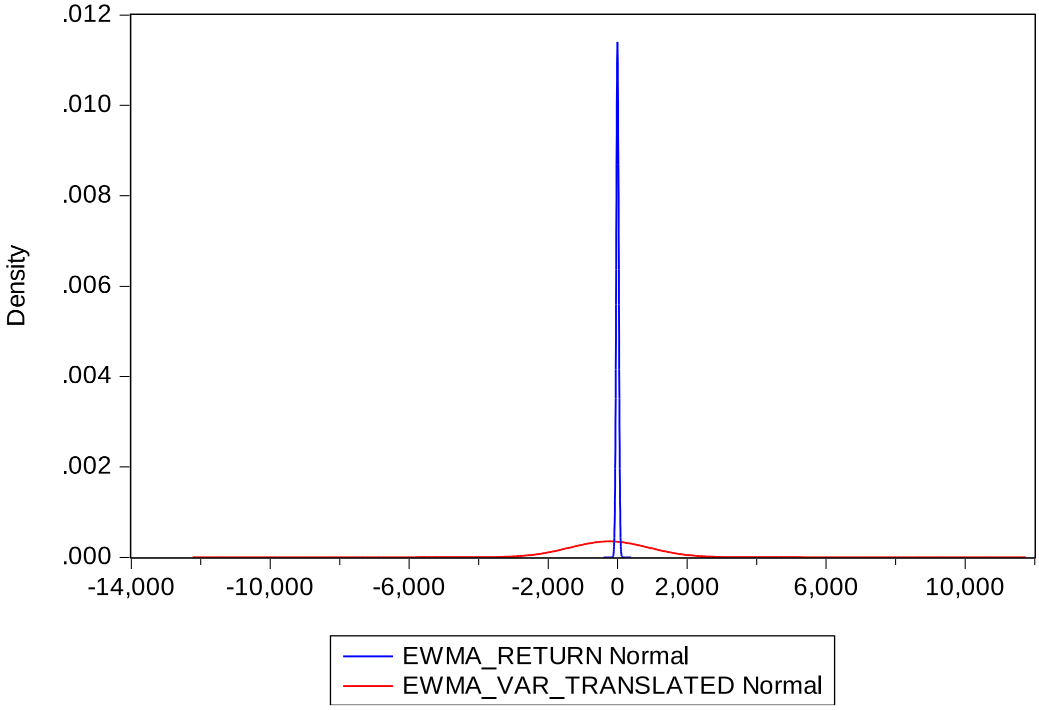


Figure 24: EWMA Delta-Gamma-Theta Translated Estimate

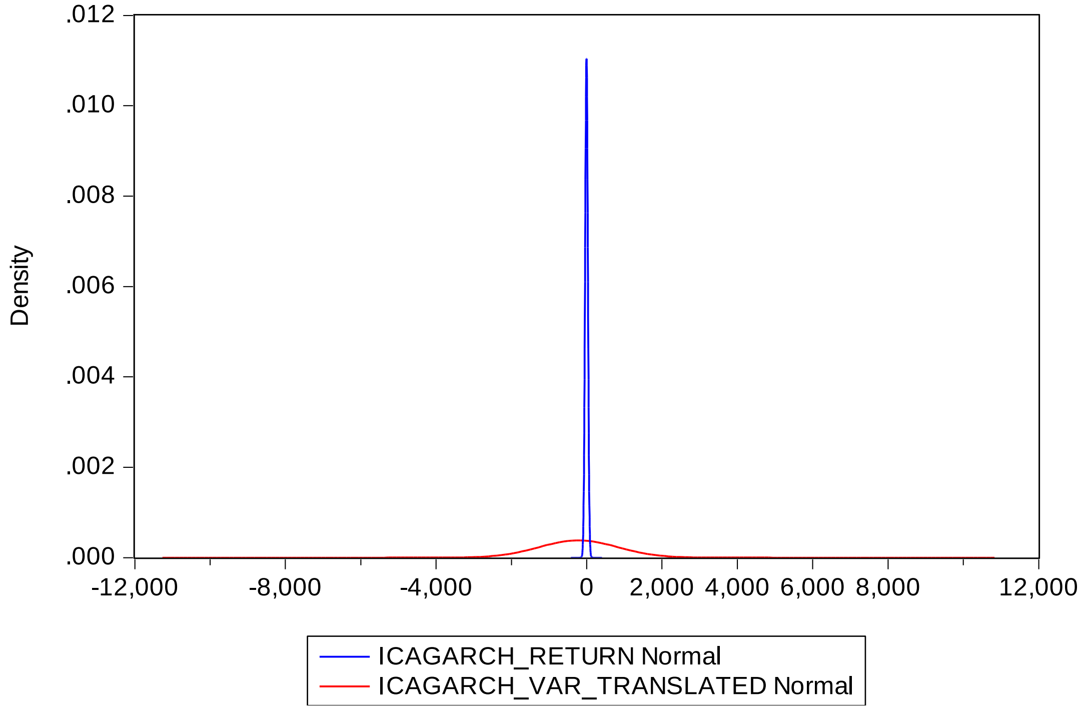


Figure 25: ICA-GARCH Delta-Gamma-Theta Translated Estimate

In all cases, we see that the Delta-Gamma-Theta VaR estimates for the translated distribution give better VaR forecasts than their Normal Distribution counterparts.

3.3 Conclusion

In this chapter, we have presented our results and findings on the different VaR computation methodologies we were studying. The results pointed towards the Delta-Gamma-Theta methodology being a superior VaR technique compared to the Delta-Gamma and subsequently the Delta methodology. Within the Delta-Gamma-Theta methodology itself, the Univariate-GARCH methodology provided better estimates of VaR for the period chosen giving better forecasts than the ICA-GARCH and EWMA methodologies. In the next chapter, we conclude our study and also give recommendations for further research.

4 Final Remarks and Recommendations

The main purpose of this study was to answer the research question: If applied to *non-linear* financial assets in South African financial markets, does the ICA-GARCH approach to computing *multivariate* portfolio VaR where the underlying distribution is estimated using the Johnson's distribution lead to better performing estimate of VaR and more quickly converging VaR computations as well as more accurate VaR estimates and forecasts than the *univariate* GARCH and EWMA approaches.

We created a portfolio of three options, a call option on the GOLDS index, a put option on the TOP40 Index and a call option on the USDZAR exchange rate. All three options were at-the-money at the beginning of the forecast period. For simplicity, we assumed that they were only comprised of a single cashflow which is on exercise date. From this we computed the Greeks and subsequently the first four moments of the portfolio.

The first VaR estimates we computed were for the Delta-VaR method. The predictions were very poor with our VaR estimates failing to capture many of the losses suffered by the portfolio. We then looked at the Delta-Gamma and Delta-Gamma-Theta methodologies. The results showed that the Delta-Gamma-Theta methodology is a superior VaR technique compared to the Delta-Gamma and subsequently the Delta methodology. The addition of the higher order theta leads to more accurate projections.

In addition, within the Delta-Gamma-Theta methodology, the ICA-GARCH approach for computing the volatilities of the different assets in the portfolio did not lead to much better VaR forecasts for our sample and the chosen period. The Univariate-GARCH method provided the most accurate forecasts of all the three methodologies.

It should be noted that several factors could have influenced the outcome of our research. These include the choice of the number of assets in the portfolio. While we worked with three assets, for multivariate volatilities this could be considered too small a number. Further studies could be carried out to determine the results once we are dealing with a large portfolio with say more than 20 options. Another case to consider would be the results when we have a mix of long and short positions in the options. This could produce different results for the correlations and as such different results for the portfolio variance.

It is also interesting to note that here we used a period of 6 months from 2016/07/01 to 2017/12/31. Since here we are incorporating theta into the calculations, it would be interesting to carry-out scenario analysis to determine if the results will be consistent over different forecast horizons.

References

ActEd Financial Economics Notes (2016). Actuarial Education Company 2016

Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, 9, 203–228. doi:10.1111/1467-9965.00068

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* (31)3, 307-327 https://doi.org/10.1016/0304-4076(86)90063-1

Bollerslev, T., R. F. Engle, & J. M. Wooldridge (1988). A capital asset pricing model with time-varying covariances. *The Journal of Political Economy*, 96, 116–131.

Danı´elsson, J. (2011). *Financial Risk Forecasting*. United Kingdom, UK: Wiley Finance Series

Davidian, M. & Carroll, R. (1987). Variance function estimation. *JASA,* 82, 1079-1091

Engle, R. and Kroner, K. (1995). Multivariate Simultaneous GARCH. *Econometric Theory*, 11, 122-150. http://dx.doi.org/10.1017/S0266466600009063

Engle, R. F., C. W. J. Granger, and D. Kraft (1984). Combining competing forecasts of inflation using a bivariate ARCH model. *Journal of Economic Dynamics and Control*, 8, 151–165.

Finger, C. C. (1996) Accounting for the "pull to par" and "roll down" for RiskMetrics cashflows *RiskMetrics Monitor*, (September 16, 1996).

Fisher, L. (1996). Some New Stock-Market Indexes. *Journal of Business*, 39, (1966), 191–225.

George, Florence and Ramachandran, K. M. (2011) Estimation of Parameters of Johnson's System of Distributions, *Journal of Modern Applied Statistical Methods*, (10)2, Article 9. Available at: http://digitalcommons.wayne.edu/jmasm/vol10/iss2/9

Gu´eant, O. (n.d.) Computing the Value at Risk of a Portfolio: Academic literature and Practitioners' response.

Harvey, A.C., Ruiz, E. and Shephard, N.G. (1994). Multivariate Stochastic Variance Models. *Rev Econ Stud* 61(2), 247-264. DOI: https://doi.org/10.2307/2297980 http://www.jstor.org/stable/2352357

Hyvärinen, A., Karhunen and Oja, E. (2001). Independent Component Analysis. *New York, J. Wiley.*

Hyvärinen, A. (1999) Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. *IEEE Trans. on Neural Networks*, 10(3), 626-634. https://www.cs.helsinki.fi/u/ahyvarin/papers/TNN99new.pdf

Hyvärinen, A. and E. Oja (1997). A Fast Fixed-Point Algorithm for Independent Component Analysis. *Neural Computation*, 9, 1483-1492.

Hyvärinen, A. and Oja, E. (2000). Independent Component Analysis - Algorithms and Applications. *Neural Networks,* (13), 411–430.

Johnson, N. L. (1949). Systems of Frequency Curves Generated by Methods of Translation. *Biometrika,* 36 (1/2), 149–176. doi:10.2307/2332539. JSTOR 2332539

Jorion, P. (2000). Risk management lessons from Long-Term Capital Management. *European Financial Management*, 6, 277–300. doi:10.1111/1468-036X.00125

Kawakatsu, H. (2006). Matrix exponential GARCH. *Journal of Econometrics* 134(1), 95–128 http://dx.doi.org/10.1016/j.jeconom.2005.06.023

LeBaron, B. (1994). Chaos and nonlinear forecastability in economics and finance. *Philosophical Transactions of the Royal Society of London A* 348(1686), 397-404. DOI: 10.1098/rsta.1994.0099

Levine, D. (2009). Modelling Tail Behavior with Extreme Value Theory. Risk Management. September 2009 – Issue 17

McMillan, D. & Thupayagale, P. (2010). Evaluating Stock Index Return Value-at-Risk Estimates in South Africa: Comparative Evidence for Symmetric, Asymmetric and Long Memory GARCH Models. *Journal of Emerging Market Finance* 2010 9, 325 http://emf.sagepub.com/content/9/3/325

Miles, M.B. and Huberman, A.M. (1994) *Qualitative Data Analysis (2nd edition).* p. 278-280. Thousand Oaks, CA: Sage Publications

Moody, D. (2002). *Empirical Research Methods Research Methods Class*, March 8, 15 & 22, 2002

Mouton, J. (1996). *Understanding Social Research.* Van Schaik Publishers, 1996. p. 175

Myers, M.D., and Newman, M. (2006). The qualitative interview in IS research: Examining the craft *Information and Organization.* pp 2-26.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica* 59, 347–370.

Oja, E., Kiviluoto, K., and Malaroiu, S., (2000). Independent component analysis for financial time series. *Proceedings of the IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium* (AS-SPCC '00), 111–116.

Parkinson, M. (1980) The Extreme Value Method for Estimating the Variance of the Rate of Return. *The Journal of Business*, (53)1, 61-65, Stable URL:

Pichler, S. and Selitsch, K. (1999): A Comparison of Analytical VaR Methodologies for Portfolios That Include Options, *Working Paper TU Wien*, Available at: http://www.gloriamundi.org

Risk-Metrics (1996) *RiskMetrics™ - Technical Document Fourth Edition*

Silvennoinen, A. & Terasvirta, T. (2008). Multivariate GARCH Models. *CREATES Research Paper* 2008-6. Available at SSRN: https://ssrn.com/abstract=1148139 or http://dx.doi.org/10.2139/ssrn.1148139

Slifker, J. F. and Shapiro, S. S. (1980) The Johnson System: Selection and Parameter Estimation. *Technometrics*, 22(4), 239–246.

Wu, YU and LI, (2006). Value At Risk Estimation Using Independent Component Analysis-Generalized Autoregressive Conditional Heteroscedasticity (ICA-GARCH) Models. *International Journal of Neural Systems,* 16(5), 371–382.

Xu, D. & Wirjanto, T. S. (2013). Computation of Portfolio VaRs with GARCH-Type Volatility. Available at SSRN: https://ssrn.com/abstract=2242929

Yin, R. K. (2003). *Case study research: Design and methods (3rd Ed.).* Thousand Oaks, CA: Sage.

Zito, T., Wilbert, N., Wiskott, L. & Berkes, P. (2009). Modular toolkit for Data Processing (MDP): a Python data processing frame work. *Frontiers in Neuroinformatics*. (2008) 2(8). doi:10.3389/neuro.11.008.2008.

Appendix: Derivation of The Equation for the Return on an Option

We start with the value of the option at time *t+n* given a value at time t as well as the changes in the prices of the underlying:

$$V_{t+n} = V_t + \delta \cdot (P_{t+n} - P_t) + 0.5 \cdot \Gamma \cdot (P_{t+n} - P_t)^2 + \theta \cdot (\tau_{t+n} - \tau_t) \qquad (A.1)$$

This expression can be re-written as:

$$V_{t+n} - V_t = \delta \cdot (P_{t+n} - P_t) + 0.5 \cdot \Gamma \cdot (P_{t+n} - P_t)^2 + \theta \cdot (\tau_{t+n} - \tau_t) \qquad (A.2)$$

We can re-express this equation as:

$$V_t \cdot \left(\frac{V_{t+n} - V_t}{V_t}\right) = \delta \cdot P_t \cdot \left(\frac{(P_{t+n} - P_t)}{P_t}\right) + 0.5 \cdot \Gamma \cdot {P_t}^2 \cdot \left(\frac{(P_{t+n} - P_t)}{P_t}\right)^2 + \theta \cdot (\tau_{t+n} - \tau_t) (A.3)$$

Dividing through by P_t gives:

$$\left(\frac{V_t}{P_t}\right) \cdot \left(\frac{V_{t+n} - V_t}{V_t}\right) = \delta \cdot \left(\frac{(P_{t+n} - P_t)}{P_t}\right) + 0.5 \cdot \Gamma \cdot P_t \cdot \left(\frac{(P_{t+n} - P_t)}{P_t}\right)^2 + \left(\frac{\theta}{P_t}\right) \cdot (\tau_{t+n} - \tau_t) (A.4)$$

We now define the following terms:

$$R_V = \left(\frac{V_{t+n} - V_t}{V_t}\right)$$

$$R_P = \left(\frac{(P_{t+n} - P_t)}{P_t}\right)$$

$$n = (\tau_{t+n} - \tau_t)$$

$$\eta = \left(\frac{V_t}{P_t}\right)$$

We can then re-write A.4. as:

$$R_V = \eta\delta R_P + 0.5(\alpha\Gamma P_t)(R_P)^2 + \left(\frac{\theta}{V_t}\right)n \qquad (A.5)$$

$$R_V = \tilde{\delta}R_P + 0.5\tilde{\Gamma}(R_P)^2 + \tilde{\theta}(\tau_{t+n} - \tau_t) \qquad (A.6)$$