Copula Modeling Analysis on Multi-Dimensional Portfolios with Backtesting

by

Yibo Dang

(Under the Direction of Cheolwoo Park)

ABSTRACT

In recent years, risk management has become one of the most crucial areas in the financial industry. The main interest of financial researchers is how to measure risk effectively and reliably. For this purpose, Value-at Risk (VaR) has been designed to make such measurements. However, VaR measurements are not usually accurate, especially for multi-dimensional portfolios. Because the dependency structures among each asset in a given portfolio cause the measuring problem, we need to find valid ways to overcome or avoid potential obstacles. In this thesis, Copula models are built step by step to predict VaR(α) of multi-dimensional portfolios. Additionally, our common Copula models are applied in real cases in order to see their universality, namely, to see how large the portfolio's dimensions can be before the model fails, and to see which models predict VaR(α)s more accurately. We use backtesting, a criterion with two hypothesis test called Unconditional Coverage Test and Conditional Coverage Test, for additional verification.

INDEX WORDS: Backtesting, Copula, Dependency Structures, Multi-dimensional portfolios, Value-at Risk

Copula Modeling Analysis on Multi-Dimensional Portfolios with Backtesting

by

Yibo Dang

B.S. in Mathematics, Beijing Institute of Technology, China, 2014

B.E. in Economics, Beijing Institute of Technology, China, 2014

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2016

Copula Modeling Analysis on Multi-Dimensional Portfolios with Backtesting

by

Yibo Dang

Major Professor: Cheolwoo Park

Committee: Ping Ma
Sue Ryung Chang

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2016

Copula Modeling Analysis on Multi-Dimensional Portfolios with Backtesting

Yibo Dang

March, 28, 2016

ACKNOWLEDGEMENTS

Through the period of working on this thesis, I have received a lot of help from my advisor Dr. Cheolwoo Park, without his guidance, advice, or patience, it is impossible for me to accomplish this academic work. I am also very grateful for his former instruction of my theoretical knowledge throughout the M.S. program. In addition, I appreciate Dr. Ping Ma and Dr. Sue Ryung Chang for their time and help serving as my thesis committee members. Last but not the least, I would like to thank my parents for their enduring support during my academic journey at the University of Georgia.

Table of Contents

LIST OF TABLES

Page

LIST OF FIGURES

Page

Chapter 1

Introduction

1.1 Background

In financial markets, the relationship between risk and return plays a key role when analyzing and building investment portfolios scientifically. Over the past hundred years, risk prediction and management have become more and more prevalent in modern financial markets. By applying mathematical models and ideas and quantifying the investment process, many long lasting problems in options, futures and other derivatives were no longer an issue and a prosperous financial market is a cornerstone of society. In real world investment, securitization is of particular interest because of its role in the financial crisis that started in 2007. The crisis has its origins in financial products created from mortgages in the United States, but rapidly spread from the US to other countries and from the virtual financial markets to the real economy. Many financial institutions failed. Some others had to be rescued by national governments and are still recovering from the bankruptcy crisis. The underlying cause can be dated to 1960s (John C. Hull, 2007), when US banks started to package mortgage-backed securities and sold the named portfolios to their target clients. However, this named portfolios sometimes only brought about profits for the banks, instead of investors or other financial institutions, which made the young market be filled with crisis of confidence. From that time, regulators and investors increased their research into using quantitative tools to measure risk.

One of the most popular ways to measure risk is Value-at Risk (VaR). During the 1990's, VaR was widely adopted for measuring market risk in trading portfolios. All the financial institutions released their own VaR measurements almost every day. On the basis of VaR measures, JP Morgan replaced a cumbersome system of notional market risk limits with a simple system of VaR called $\mathbf{RiskMetrics}^{TM}$, which comprised a detailed description of the techniques and a routine updated covariance matrix for several hundred key factors. They made the service free to the public which contributes in advocating VaR to the widest audience.

With the further development of VaR measurements in recent decades, regulatory building-block VaR measure and proprietary VaR measures of investment banks are challenged with applicable issues. More specifically, even if we neglect the autocorrelation among the dependence structures, the joint distribution of risk factors brings normality assumptions into the consideration. Hence, Copula methods are introduced for solving the dependence problems and making the prediction and measurements of portfolio riskiness on options and other derivatives more stable. In this thesis, we build time series models and use backtesting as a criterion to help investors manage their assets more effectively and safely.

1.2 Objective

The VaR of a multi-dimensional portfolio is usually hard to predict, because the dependency structure of the distributions exists in higher dimensional cases. Furthermore, the relationships between each asset in the portfolios also need to be specified. The objective of this thesis work is to check the variability and stability of our copula models by backtests, which examine the adequacy of VaR measures. In this way, more sophisticated and safe risk measurements can be obtained for our equity securities portfolios. We used datasets collected online through R programming from NASDQ and NYSE, and the structures of portfolios were analyzed using a financial perspective. In order to make our conclusion more general, we applied different copula methods onto different multi-dimensional stock portfolios. Gaussian Copulas, Archimedean Copulas and Hierarchical Archimedean Copulas (HAC/NAC) are covered in Chapters 2,3 and 4 respectively. Then, we used both unconditional and conditional coverage tests as criteria for checking the goodness of our models. These methods highly rely on an understanding of statistical distributions, because multivariate distributions need to be modeled and estimated by copulas for the joint distributions of asset returns. In addition, time series models like multivariate DCCGARCH, GO-GARCH were selected to predict the risks in advance and their parameters were also simulated. Although this thesis focuses on practical applications, theoretical modeling parts are also worth consideration in the preparing process. Thus, Chapters 3 and 4 are mainly restricted to statistical concepts and our inference in financial markets, in an order from VaR measurements to Copula methods. After that, Chapter 5 is focused on copula modeling process. Chapter 6 provides a detailed description of the usage in an empirical study. In the last chapter, brief conclusions are formed based on the outputting results.

Chapter 2

Brief Introduction to Value at Risk measurements

2.1 Basics

As previously stated, Value-at-Risk (VaR) is a well-known measure for qualifying and controlling the risk of a portfolio. Basically, investors will get insight of how much they may earn or lose on their investments by the statistical characteristics and the momentum of the assets. VaR, first conceived in 1993 partly in response to various financial disasters (Glyn A. Holton, 2002), is used as an internal risk management tool for many financial institutions.

Assume that L_{t+1} represents the possible profits and losses (P&L) of a portfolio in the following periods up to date t + 1, and F_{t+1} represents the distribution function of L_{t+1}. Under this denotation, mathematically, the VaR of the portfolio of the assets can be determined using a given α by:

$$VaR_t = F_{t+1}^{-1}(\alpha) = \inf\{x \in \mathbb{R};\ F_{t+1}(x) \geq \alpha\}.$$

The goal of VaR analysis is to approximate the means, variance and correlations of returns. More generally, one needs to specify the characteristic statistics for the joint assets portfolio. The change in the value of the portfolio is assumed to be of the form:

$$L_{t+1} = \sigma_t Z_{t+1}$$
$$\sigma_t^2 = \omega_t^T \Sigma_t \omega_t$$

where ω_t indicates the weight of each component of the whole assets and Z_t is an independent and identically distributed random variable (i.i.d. r.v.) which complies to N (0,1). Σ_t is of main interest in our analysis. Basically, it stands for the variance of innovations fitted by Maximum Likelihood Estimation in our time series modeling process. As for multi-dimensional portfolios, σ_t^2 replaces Σ_t with weighted vectors ω_t multiplied.

In practical application, we are going to use the former information collected for the loss function to predict the potential future movement of the P&L. However, this is difficult to fulfill and thus the above VaR equation is mostly shown here as a brief introduction to basic VaR modeling, we still need to introduce effective methods for achieving the goals we have mentioned.

2.2 Common Methodologies for Calculating VaR

2.2.1 Parametric approach

This method is also referred to as the Delta-Normal method. The initial assumption made is that the daily geometric returns of the market variables with a mean return of zero. Historical data are of high importance since σ_t^2 and the covariance matrix Σ_t both rely on the trends and distributions of them.

By this method, VaR is computed by multiplying the vector of the first derivatives of the value V of the portfolio with respect to the risk factor variables R ("delta"), and then multiplying by parameter σ and the critical value with respect to the standard normal distribution at level α. This description can be manifested as below:

$$VaR_\alpha = \Delta \times \sigma Z_\alpha dR \ .$$

This method was introduced by the JPMorgan's $\boldsymbol{RiskMetrics}^{TM}$, in which MSCI ESG research for Volkswagen Rating can be a good example. However, as from the assumption, this Delta-Normal method is valid for linear exposures but inaccurate for non-linear portfolios or for skewed distributions.

2.2.2 Historical Simulation

The Historical Simulation method directly involves using past data as a guide to what might happen in the future. The set of possible future scenarios is fully represented by the things happened over a specific historical window. In this method, we do not need to make any distributional assumption, although parameter fitting may be performed on the resulting distribution.

We apply the current weights to the historical asset returns by going back in time such as over the last year. We thus obtain a distribution of the change of portfolio returns from today's value. Usually, some of the changes will involve losses and some of them involve profits. Ordering the changes from worst to best, and then, by obtaining the 99% VaR, for example, one can say that 1% of the profits or losses are in the lowest group among the returns of the whole year, and 99% are higher ones.

2.2.3 Monte Carlo Simulation method

Monte Carlo Simulation method is by far the most powerful and flexible (Franke, 2008). No assumption of linearity or specific distribution has to be made in advance. In addition, it can incorporate most of the pragmatic distributional properties, such as fat tails and time varying volatilities. Also, longer observed periods are also valid under Monte Carlo simulation method.

This method can be briefly summarized by two steps:

a) A stochastic process is specified for the financial variables;
b) Simulate the fictitious prices for all financial variables of interest, ordering them and find the largest return as the VaR_α among the lowest group that is below the critical value α*100%.

There are many other approaches developed in recent decades, like Delta-Gamma method, and Monte- Carlo method based on it. In all, one should select the most suitable techniques in view of the basic properties of the data and the purpose willing to be achieved.

Chapter 3

Copula

3.1 Basics

Generally, the risk associated with a portfolio may be originated from market risk, credit risk and operational risk. A conventional procedure to model joint distributions of financial returns is by approximating them with multivariate normal distributions, which implies the structures returns dependency is always reduced to a fixed type. Nevertheless, empirical evidence for the normal assumptions are barely verified and an alternative model is needed, with more flexible dependency structures and arbitrary marginal distributions. Fortunately, all of the above requirements can be achieved by turning to copulas.

The definition of Copula is straightforward. The flexibility of copulas basically follows from Sklar's Theorem (1959), which says that each joint distribution can be "decomposed" into its marginal distribution accompanying with a copula "responsible" for the dependence structure:

$$\mathrm{F}(x_1, x_2, \dots, x_d) = Copula(F_1(x_1), F_2(x_2), \dots, F_d(x_d)|\theta).$$

Copula(.) is an alternative expression of the multivariate joint distribution $\mathrm{F}(x_1, x_2, \dots, x_d)$; $\theta$is a parameter of the Copula structure; and $F_i(x_i)$ is the marginal distribution function with i=1,2,3…d. As imagined, copula simplifies the estimation of a multivariate joint distribution and we accept the marginal distributions to be different from normal distribution. Moreover, copula allows the time-varying dynamic change in the model. With different methods, copula can be defined in many ways to be applied under many different situations.

3.2 Introduction to different types of Copulas

This section briefly reviews definitions of those commonly used copulas. Some important properties of copulas are also reviewed. Firstly, consider the general definition of copula:

Definition 1 (Copula)

A function mapping from $[0,1]^d\ to$ [0,1] is called function C(.) with properties:

- $\mathrm{C}(\mu_1, \mu_2, \dots, \mu_d)$ is monotone increasing when all the elements $\mu_i \in [0,1]$.
- $\mathrm{C}(1,1,1, \dots, \mu_i, 1, \dots,1) = \mu_i$ for all $\mu_i \in [0,1]$, with i = 1,2, … d.
- The copula is set to be zero if one of the μ_i is zero.
- If $\mu_i < \mu_i'$, for all $(\mu_1, \mu_2, \dots, \mu_d)$ and $(\mu_I', \dots, \mu_d') \in [0,1]^d$, the following inequality exists:

$$\sum_{i_1}^{2} \dots \sum_{i_d}^{2} (-1)^{i_1+\cdots+i_d}\, C(x_{1d}, x_{2d}) \geq 0,$$

where $x_{1d} = \mu_i$ and $x_{2d} = \mu_i'$.

The function C(.) is thus called a d dimensional copula and has a d dimensional multivariate uniform transformation. The fourth property states that any d-dimensional copula on a unit cube $[0,1]^d$ is an C-Volume of rectangle [0, μ_i] x [0, μ_i'], which is non-negative.

Accordingly, as referred to Nelson (2006), Cherubini, Luciano, and Joe (1997), if F(.) belongs to a class of elliptical distributions, then the corresponding C(.) is an elliptical copula. However, in most real situations, this type of copulas cannot be obtained explicitly since the inverse marginal distribution function $F^{-1}(.)$ and the cumulative density function are usually with integral forms which can hardly be calculated. One of the most reliable ways to overcome this drawback is using Archimedean Copula as an alternative (Ostap, Alexander, 2014), which, however, also has its internal flaws with restricted to limited dimensions of the portfolios. This issue will be discussed more in latter chapters and quantified to figure out the applicable dimensions.

By Sklar's theorem, an elliptical copula is a copula with respect to a known elliptical distribution, detailed introduction about the distribution can be found in Fang, Kotz and Ng's (1990) work. Generally, elliptical copula contains Gaussian Copula and t Copula, which are both popular used in risk management because of their easy implementation.

Definition 2 (Gaussian Copula)

The Gaussian copula represents the dependence structure of the multivariate normal distribution. This means if we have a vector $\mathrm{X} = (x_1, x_2, \dots, x_d)^T \sim N_d(\mu, \Sigma)$, where $x_i \sim N(\mu_i, \sigma_i)$, for i=1, 2, …d, then a copula exists as:

$$F_X(x_1, x_2, \dots, x_d) = C\big(F_1(x_1), F_2(x_2), \dots, F_d(x_d)\big),$$

or, equivalently, the elliptical copula determined by F is:

$$C(u_1, u_2, \dots, u_d) = F_X\left(F_1{}^{-1}(u_1), F_2{}^{-1}(u_2), \dots, F_d{}^{-1}(u_d)\right),$$

and as defined, $F_X(.)$ is the multivariate joint function for X, and $F_i(x_i)$ is its corresponding marginal distribution. In the second equation, let F_i be CDF of the ith margin and F_i^{-1} be its inverse function (Quantile function). Convenience in obtaining conditional distributions is another advantage in using them for prediction (Frees and Wang, 2005). Next, let Y be a function of X, where Y=G(X) and $G_i(X) = (x - \mu_i/\sigma_i)$. In this case, $Y_i \sim N(0,1)$, and $\mathrm{Y} = (y_1, y_2, \dots, y_d)^T \sim N_d(0, \Psi)$, which means we standardized X to comply to a multivariate normal distribution and Ψ is a correlation matrix associated with Σ.

A copula can be defined as the following:

$$F_Y(y_1, y_2, \dots, y_d) = C_{Gauss}\big(\Phi(y_1), \Phi(y_2), \dots, \Phi(y_d)\big).$$

For simplification, we use $u_i = \Phi(y_i)$ as an alternative, then the above formula is rewritten as (Franke, Hafner, 2008):

$$C_{Gauss}(u_1, u_2, \dots, u_d) = C_{Gauss}\big(\Phi(y_1), \Phi(y_2), \dots, \Phi(y_d)\big)$$

$$= F_Y(y_1, y_2, \dots, y_n) = F_Y\big(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_d)\big)$$

$$= \int_{-\infty}^{\Phi^{-1}(u_1)} \dots \int_{-\infty}^{\Phi^{-1}(u_d)} (2\pi)^{-\frac{d}{2}} |\Psi|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} r^T \Psi^{-1} r\right) dr_1 \dots dr_d.$$

Since the standardized function G(X) is increasing, according to Definition 1, it follows that C= C_{Gauss}. Through this way, the Gaussian copula is obtained.

Definition 3 (t Copula)

Similar to the derivation of Gaussian Copula, t copula is derived on $X \sim t_d(\nu, \mu, \Sigma)$, and given by:

$$C_\nu(u_1, u_2, \dots, u_d) = |\Psi|^{-\frac{1}{2}} \frac{\Gamma(\frac{\nu + d}{2})(\Gamma\left(\frac{\nu}{2}\right))^{d-1}(1 + \frac{1}{\nu}\zeta^T \Psi^{-1} \zeta)^{-\frac{\nu+d}{2}}}{(\Gamma\left(\frac{\nu + 1}{2}\right))^d \prod_{j=1}^d (1 + \frac{1}{\nu}\zeta_j^2)^{-\frac{\nu+d}{2}}}.$$

Also, Ψ is a correlation matrix associated with covariance matrix Σ, ν is the degree of freedom and $\zeta_j = t_\nu^{-1}(u_j)$ (Franke, Hafner, 2008):. Note that the Gaussian Copula and t Copula both pertain to elliptical copula family.

Definition 4 (Gumbel Copula):

The Gumbel Copula is not constructed based on Sklar's Theorem, instead, it belongs to Archimedean Copulae and is related to Laplace transforms of univariate distribution functions. Gumbel Copula is widely applied in financial market. Its generator and functions are:

$$\Phi(\mathrm{x}, \theta) = \exp\left(-x^{\frac{1}{\theta}}\right), 1 \le \theta < \infty, x \in [0, \infty),$$

$$C_\theta(u_1, u_2, \dots, u_d) = \exp\, [-\{\sum_{j=1}^d (-log u_j)^\theta\}]^{\frac{1}{\theta}}.$$

In contrary to the elliptical copula, Gumbel copula visually leads to asymmetric contour diagrams, as shown in Figure 1 (d=2). The Gumbel copula shows stronger linkage between positive values, in the meantime, it also shows more variability and more mass in the negative tail.

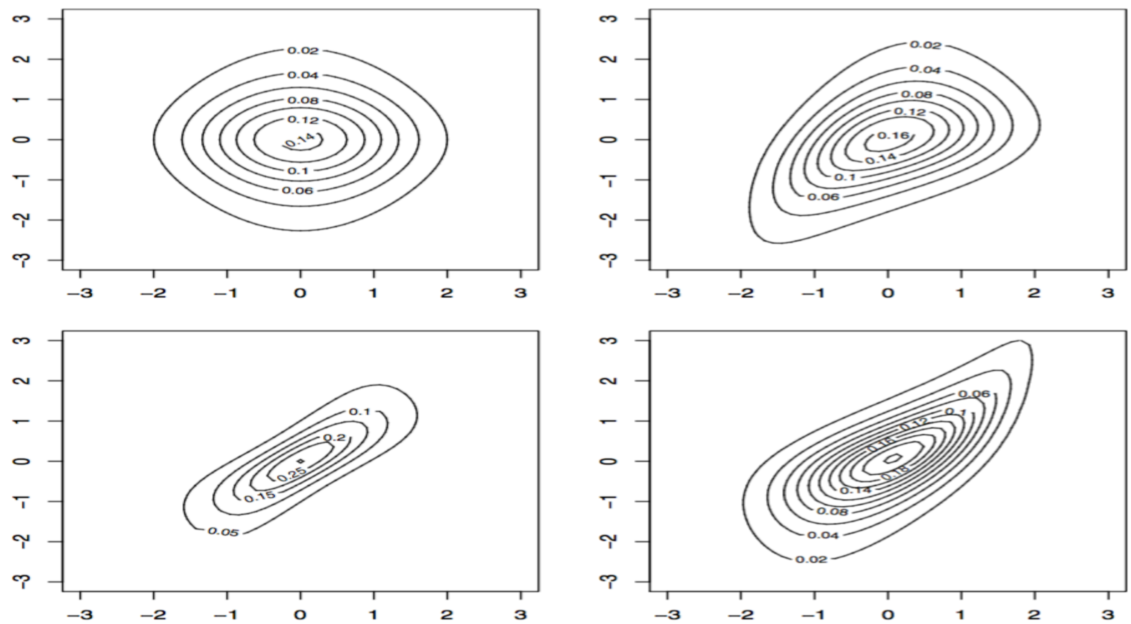

Figure 1 pdf contour plots, Gaussian Copula(top left), Clayton (top right), Frank (bottom left), Gumbel (bottom right)

Definition 5 (Clayton Copula):

$$C_\theta(u_1, u_2, \ldots, u_d) = \exp\left[\left(\sum_{j=1}^{d} {u_j}^{-\theta}\right) - d + 1\right]^{\theta}.$$

Then, for $x \in [0, \infty)$, the generator belonging to Clayton Copula is : $\Phi(x, \theta) = (\theta x + 1)^{-1/\theta}$, where $\theta \in [-1,0) \cup (0, \infty)$. While Gumbel copula puts more weights on the upper tail, Clayton Copula gives more clusters to the lower tail.

Definition 6 (Ali-Milkhail-Haq Copula)

Ali-Milkhail-Haq (AMH) Copula is also a well-known copula, whose function is denoted as following:

$$C_\theta(u_1, u_2, \dots, u_d) = \frac{\prod_{j=1}^{d} u_j}{1 - \theta\{\prod_{j=1}^{d}(1-u_j)\}}.$$

If $\theta = 0$, then AMH copula is simplified as a product copula, which is the simplest copula. If $\theta = 1$, there exists a significant lower tail dependence. What needs to be mentioned here is that AMH copula can only be effective in bivariate cases (Nelsen 1999), so it will not be considered and discussed in Chapter 4 and 6, our modeling and testing processes.

Some other Copulas, like Joe (1997) copula, Frank (1979) are either owns similar properties as copulas we have discussed or not used commonly, so we will not repeatedly consider their applications in the financial market. Summary of the common Archimedean Copulas is shown in Table 1.

Table 1: Summary of 3 Archimedean Copulas for df >2

Copula Family	**Parameter Interval**	**Generator** $\Phi(t,\theta)$	**Inverse of Generator** $\Phi^{-1}(u)$	**Frailty Distribution**
Clayton(1978)	$\theta \geq 0$	$(1+t)^{-1/\theta}$	$u^{-\theta} - 1$	Gamma
Frank(1979)	$\theta \geq 0$	$-\theta^{-1}\ln(1 + e^{-t}(e^{-\theta} - 1))$	$-\ln\frac{e^{-\theta u}-1}{e^{-\theta}-1}$	Log series
Gumbel(1960)	$\theta \geq 1$	$\exp(-u^{1/\theta})$	$(-\ln t)^{\theta}$	Positive Stable

Definition 7 (Hierarchical/ Nested Archimedean Copula)

Hierarchical Archimedean Copula is a recently developed copula method which defines the whole dependence structure in a recursive way. In other words, analysts are allowed to apply different copula measures in different levels. This provides us the possibility to arbitrarily choose a combination of copulas we want for different levels, and thus make the most perfect predictions.

The special case, so called fully nested Archimedean Copula function is:

$$C(u_1, u_2, \dots, u_d) = \phi_{d-1}[\phi_{d-1}^{-1} \circ C(\{\phi_1, \dots, \phi_{d-2}\})(u_1, u_2, \dots, u_{d-1}) + \phi_{d-1}^{-1}(u_d)] \, .$$

In this formula, ϕ_i indicates copula the generator at different copula levels.

3.3 Copula Estimation

The estimation of a copula based multivariate distribution involves both the estimation of the copula parameters θ and the estimation of the margins F_j . If we are only interested in the dependency structure, then the estimated parameter vector is independent of any predicted models for each margin. When it comes to nonparametric marginal distributions, it can be shown the consistency and normality of maximum likelihood estimators (MLE). In addition, according to Franke, Hardle and Hafner's work, the goodness and properties of $\hat{\theta}$ mainly rely on the estimation of the margins.

More specifically, suppose a d-dimensional random variable X with parametric marginal distributions $F_j(x_j; \delta_j)$, j= 1,2,…,d. where δ_j is one of the components of the parameter vector. Then according to the informal definition of copula we have discussed in the former section, the distribution of X can be decomposed as:

$$F_X(x_1, x_2, \dots, x_d) = C\{F_1(x_1; \delta_1), F_2(x_2; \delta_2), \dots, F_d(x_d; \delta_d); \theta\}$$

and the associated PDF as:

$$\mathrm{f}(x_1, x_2, \dots, x_d; \delta_1, \delta_2, \dots, \delta_d, \theta) = c\{F_1(x_1; \delta_1), F_2(x_2; \delta_2), \dots, F_d(x_d; \delta_d); \theta\} \prod_{j=1}^{d} f_j(x_j; \delta_j)$$

where $\mathrm{c}(u_1, u_2, \dots, u_d) = \frac{\partial^d C_\theta(u_1, u_2, \dots, u_d)}{\partial u_1 \dots \partial u_d}$.

Now we are interested in finding the MLE of the parameters of marginal functions and the copula function. Firstly, we can denote the parameter vector as: $\Omega = (\delta_1, \delta_2, \dots, \delta_d, \theta)^T$, where $\delta_j's$ are parameters for the ith marginal distributions, and θ belongs to the copula function.

Then the likelihood function is given by:

$$\mathrm{L}(\Omega|x_1, x_2, \dots, x_T) = \prod_{t=1}^{T} \mathrm{f}\big(x_{1,t}, x_{2,t}, \dots, x_{d,t}|\delta_1, \delta_2, \dots, \delta_d, \theta\big) .$$

Subsequently the log-likelihood function is derived as:

$$\begin{aligned} \mathrm{l}(\Omega|x_1, x_2, \dots, x_T) &= \sum_{t=1}^{T} \log\left(c\{F_1\big(x_{1,t}; \delta_1\big), F_2\big(x_{2,t}; \delta_2\big), \dots, F_d\big(x_{d,t}; \delta_d\big); \theta\}\right) \\ &+ \sum_{t=1}^{T}\sum_{j=1}^{d} log\, f_j\big(x_{j,t}; \delta_j\big). \end{aligned}$$

In the above formulas, we pick d observations for each sample $x_t = (x_{1,t}, x_{2,t}, \dots, x_{d,t})$. Here, t=1, 2, ..., T.

In order to obtain the MLE, we firstly need to estimate the parameter vector Ω. By Full Maximum Likelihood (FML) Estimation, estimating Ω is equal to solve the equation:

$$\frac{\partial l(\Omega|\mathrm{X})}{\partial \delta_i} = \frac{\partial l(\Omega|\mathrm{X})}{\partial \theta} = 0.$$

However, in practice, it costs too much time if the dimension is not moderate. Another popular estimating method is called the Inference for Margins (IFM), unlike the former one step FML method, this method estimates the parameter from the marginal distributions δ_j in the first step; secondly, it estimates the dependence parameter θ through the pseudo log likelihood function. The main advantage of IFM method is that it highly reduces the computational and notational complexity (Franke, Hardle and Hafner, 2008).

Next estimating method is the Canonical Maximum Likelihood, differing from FML and

IFM method, it estimates parameters through an empirical distribution function $\widehat{F_j(x)} = \frac{1}{T+1}\sum_{t=1}^{T} l(x_{j,t} < x)$. After margins are estimated, they are collected to form a pseudo sample and are used to estimate the copula. Notice that the estimation for marginal distributions is nonparametric.

With margins prepared, we are ready to estimate the required parameters of copula.

In terms of elliptical copulas, we usually have two estimating methods, one is Gaussian Copula estimation and the other is t-Copula estimation. Once giving the copula density function, we are able to estimate the covariance matrix Σ by the pseudo log-likelihood function. In this way, we can get the estimator of copula parameters.

As for t-copula estimation, one possible method is based on the estimation of Kendall's τ with methods of moments.

Each element φ_{ij} of the correlation matrix is estimated by:

$$\widehat{\varphi_{ij}} = \sin\left(\frac{\pi}{2}\widehat{\rho_\tau}(u_i, u_j)\right).$$

Then if the estimated correlation matrix $\widehat{\Psi}$ is fixed, we are able to estimate the degree of freedom ν through ML. In order to obtain the value of $\widehat{\varphi_{ij}}$, we also need to estimate the Kendall's τ coefficient for each pair of observation i and j. For pseudo sample $\{u_t\}_{t=1}^{T}$, we have each $u = (u_1, u_2, \ldots, u_d)^T \in [0,1]^d$. Then, the estimator of this coefficient is as following:

$$\widehat{\rho_\tau}(u_i, u_j) = \binom{T}{2}^{-1} \sum_{1 \le t_1 \le t_2 \le T} sign(u_{i,t_1} - u_{i,t_2})(u_{j,t_1} - u_{j,t_2}).$$

In this way, $\hat{\nu}$ can be estimated with the arg max of the pseudo log-likelihood function.

On basis of the theoretical results, regular copula estimation methods contain the following ones: Kendall's tau, Blomqvist's beta, minimum distance estimators, the maximum-likelihood estimator, a simulated maximum-likelihood estimator, CML, FML and IFM. All of these methods can be utilized under either known or unknown margins. Although the maximum likelihood method requires the simulated residuals for copula fitting process should be within the interval [0,1], it does not mean that we cannot make use of ML method if residuals are not satisfied with the constraint. In fact, pseudo-observations are generated here in R to solve the problem.

On the one hand, under known margins, maximum-likelihood estimator provides us the most precise results (Hofert, M., Mächler, M., & Mcneil, A. J., 2012). In addition, in programming software like R, if we consider running time and numerical stability during backstage calculation, ML estimator is also desirable to the extent of current dimensions. On the other hand, if pseudo-observations have to be produced, there has barely no gap between ML estimator and other estimators. No matter how accurate the estimating method is, we still cannot simulate copula precisely without a proper marginal distribution. As for how to choose the margins when simulating copula, we will discuss it in the following part.

3.4 Defining marginal distributions in multi-dimensional copulas

As discussed, copula methods are introduced to decompose the continuous joint distribution of dependence structure into separate marginal distributions. Commonly used dispersion structures include: autoregressive of order 1, exchangeable, Toeplitz, and unstructured. The corresponding positive definite correlation matrices are listed one by one below, in the case that the degree of freedom equals to 5:

$$\begin{pmatrix} 1 & \rho_1 & \rho_1^2 & \rho_1^3 & \rho_1^4 \\ \rho_1 & 1 & \rho_1 & \rho_1^2 & \rho_1^3 \\ \rho_1^2 & \rho_1 & 1 & \rho_1 & \rho_1^2 \\ \rho_1^3 & \rho_1^2 & \rho_1 & 1 & \rho_1 \\ \rho_1^4 & \rho_1^3 & \rho_1^2 & \rho_1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1 & \rho_1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_1 & \rho_1 & \rho_1 \\ \rho_1 & \rho_1 & 1 & \rho_1 & \rho_1 \\ \rho_1 & \rho_1 & \rho_1 & 1 & \rho_1 \\ \rho_1 & \rho_1 & \rho_1 & \rho_1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1 & \rho_2 & \rho_3 & \rho_4 \\ \rho_1 & 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_2 & \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_3 & \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_4 & \rho_3 & \rho_2 & \rho_1 & 1 \end{pmatrix}, \text{and}$$

$$\begin{pmatrix} 1 & \rho_1 & \rho_2 & \rho_3 & \rho_4 \\ \rho_1 & 1 & \rho_3 & \rho_4 & \rho_5 \\ \rho_2 & \rho_3 & 1 & \rho_1 & \rho_2 \\ \rho_3 & \rho_4 & \rho_1 & 1 & \rho_3 \\ \rho_4 & \rho_5 & \rho_2 & \rho_3 & 1 \end{pmatrix}$$

where ρ_j's are dispersion parameters.

Some practitioners used one of the correlation matrices (usually autoregressive of order 1 and simple exchangeable Archimedean (Jun Yan, 2007)) as a direct proxy for multivariate dependency. However, it is widely acknowledged that prices, returns, and other financial parameters and variables are not always normally distributed. In fact, many of them have fat tails and exhibit the so called "tail dependence". (see Figure 2 for the visualized proofs). In Figure 2, we selected one of the most active securities on NASDQ, the Alphabet Inc., which is quoted as GOOG, as our example. As known, it is a public company with the highest market value on the planet. From the curves shown in this figure, skewed t distribution fits the real log-return density function the best, while others have more or less deviations. Although skewed t gives a better fit, it underestimates the peak point as others do. If time permits, we can find infinite univariate case like this.

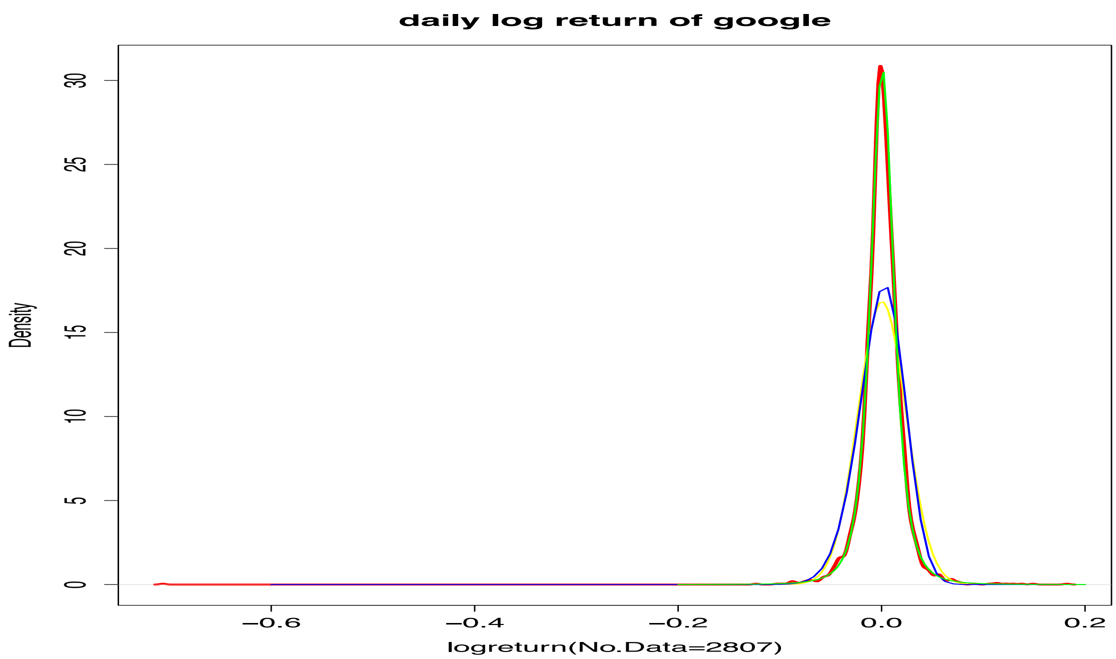

Figure 2 Density function of log-return of Google (red) with fitted normal distribution (yellow), skewed normal (blue) and skewed t (green)

Therefore, there is of necessity to build an accurate multivariate distribution before modeling a reliable copula. This raises a concern of how to choose a proper marginal distribution, and, once chosen, how to calibrate the given data to fit it.

Now, let us take a close look at the definition of skewed t distribution in the form of normal mixture. This concept prepares us with the knowledge of how to select marginal distributions believable.

Definition 8 (Normal Mean-Variance Mixture Representation of Skewed t Distribution)

Let μ and φ be two parameters in multivariate space $\mathbb{R}^n$, and Σ be a d x d real positive semi-definite matrix, and the degree of freedom $\nu > 2$. Then the d dimensional skewed t distributed random variable X can be denoted as:

$$X \sim ST_d(\upsilon, \mu, \Sigma, \varphi),$$

Given this denotation, X is a multivariate normal mean-variance mixture variable with its distribution given by:

$$X \stackrel{\text{def}}{=} \mu + W\varphi + \sqrt{W}Z,$$

where

1. $Z \sim N_d(\mu, \Sigma)$, a d dimensional multivariate normal distribution with mean 0 and covariance matrix Σ.

2. W ~ InverseGamma($\upsilon/2, \upsilon/2$), with pdf: $f(x) = \frac{\beta^{\alpha} x^{-\alpha-1} e^{-\frac{\beta}{x}}}{\Gamma(\alpha)}$. It is worth noting that W and Z are mutually independent.

Then, with W given, the distribution of X can be rewritten as:

$$X|W \sim N_d(\mu + W\nu, W\Sigma),$$

which complies a d dimensional multivariate normal distribution, with EX=$\mu + \varphi \frac{\nu}{\nu-2}$ and $\mathrm{Cov}(X) = E(W)\Sigma + \mathrm{var}(W)\nu\nu'$.

If the degree of freedom ν is fixed, it is possible to calibrate elliptical distribution to suit our financial data. Many researchers like McNeil et al. (2005) and Hu (2005) have conducted algorithms to achieve the goal. Generally, algorithms, like Expectation Maximization (EM), approximately conduct estimations of Maximum Likelihood method.

From these researches, it can be learnt that calibrating the joint distribution is much faster and mathematically solid than modifying existing copulas. From Definition 8, this normal mean-variance mixture distribution has more flexibility to calibrate parameters to suit different type of portfolios.

In short, from the definition of the normal form of the skewed t distribution, since Σ is positive semi-definite, it has non-zero tail dependence. Furthermore, adopting skewed t distribution gives heavier tails at either upper or lower part for the margins. Therefore, in real equity swaps practitioners prefer skewed t distribution to skewed normal as the margins to fit copulas. In the latter modeling chapter, we will also utilize this marginal distribution for parametric simulation.

Chapter 4

Modeling Log-Returns with Multivariate Time Series and Copula Methods

4.1 Expression of the portfolio's Log-Returns

In this section, we will build a reliable model to predict VaR on a profit and loss view of a linear portfolio. Traditionally, the return of an asset is usually specified by its log transformation since raw data vibrated drastically from one period to the other, which makes it hard to grab their properties. Statistically, this log transformation will make the time-varying data much more stable in terms of variance. In financial time series, log-returns is basically defined as:

$$\text{LogRet}_{t+1} = \log\left[\frac{P_{t+1} - P_t}{P_t}\right] = \log(P_{t+1} - P_t)\ ,$$

where P_t is the time point price (market value of the aimed asset) at t and LogRet is our interested profits or losses.

Contrary to the previous static process, log-returns are predicted through a time-varying model and assumed to be time dependent.

As usual, the Value-at-Risk at level α of a portfolio is defined as the α-quantile value from the CDF of profit & loss function. Here, some concepts need to be clarified in preparation. Profit & Loss function:

$$L_{t+1} = (V_{t+1} - V_t) = \sum_{i=1}^{d} w_i S_{i,t}\{\exp\,(X_{i,t+1} - 1\}\ ,$$

where L_{t+1} is the profit & loss of a linear portfolio, V_t is the value of the portfolio at time point t. $X_{i,t+1}$ is our target log-returns of ith asset, w_i is the weight of ith asset in our established portfolio and S is its corresponding "market value".

Then, given a so called safety threshold x, the CDF of L without the time index can be expressed as:

$$F_L(x) = P(L \leq x)\,.$$

Hereby, the nominal VaR is given by:

$$VaR(\alpha) = F_L^{-1}(\alpha).$$

From these two equations, it can be summarized that the distribution of L depends on the distribution of log-returns. Therefore, modeling the distributions of them is the first thing to do before extracting the αth quantile from F_L. As mentioned, log-returns of portfolios are assumed to be time-dependent, thus the target process $\{X_t\}$ can be modeled as:

$$X_{i,t} = \mu_{i,t} + \sigma_{i,t}\epsilon_{i,t}\,, i = 1,2, \ldots, d\,.$$

where $\epsilon_{i,t}'s$ are mutually i.i.d innovations produced by copula modeling processes. E($\epsilon_{i,t}$)=0 and Var($\epsilon_{i,t}$)=1. With known historical information of the stochastic process, we have $\mu_{i,t} = E(X_{i,t}|\mathcal{F}_{t-1})$ and $\sigma_{i,t}^2 = E\left[\left(X_{i,t} - \mu_{i,t}\right)^2|\mathcal{F}_{t-1}\right]$. Next, for the innovation vector $\varepsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_d)$, its marginal distribution are estimated from copula family, which is shown below:

$$F_\varepsilon(\epsilon_1, \epsilon_2, \ldots, \epsilon_d) = C_\theta\left(F_1(\varepsilon_1), \ldots, F_d(\varepsilon_d)\right).$$

Now, all of the parameters in the log-return model are able to be obtained. In order to achieve the prediction, following steps need to be proceeded in order:

1. Estimation of the residuals $\widehat{\varepsilon_t}$.
2. Specification and Simulation of marginal distributions.
3. Decomposition of the dependency structure.
4. Estimation of the supposed mean and variance vectors.

4.2 Time Series Modeling Process

This section will focus on accomplishing the three steps listed above by using multivariate time series models. The basic idea of time series models is predicting the current or uncertainty behaviors with known information. Since financial products, like equity securities, options, securitized financial derivatives, etc., being swapped on second markets satisfy the constraints of stochastic processes (see Figure 3 as an example), it is valid to utilize the time series models to predict parameters we are interested.

4.2.1 Brief Introduction to Univariate GARCH model

Modeling risks had been a key issue for many financial institutions for several years, because the variances of data change constantly. Thanks to the development of the generalized autoregressive conditional heteroscedasticity (GARCH) model, it is possible to model the time-varying variance of oscillated economic data. GARCH models are tools for forecasting and analyzing the volatility of time series that varies over time. Therefore, to predict the final Log-Returns, using GARCH models should be a wise choice.

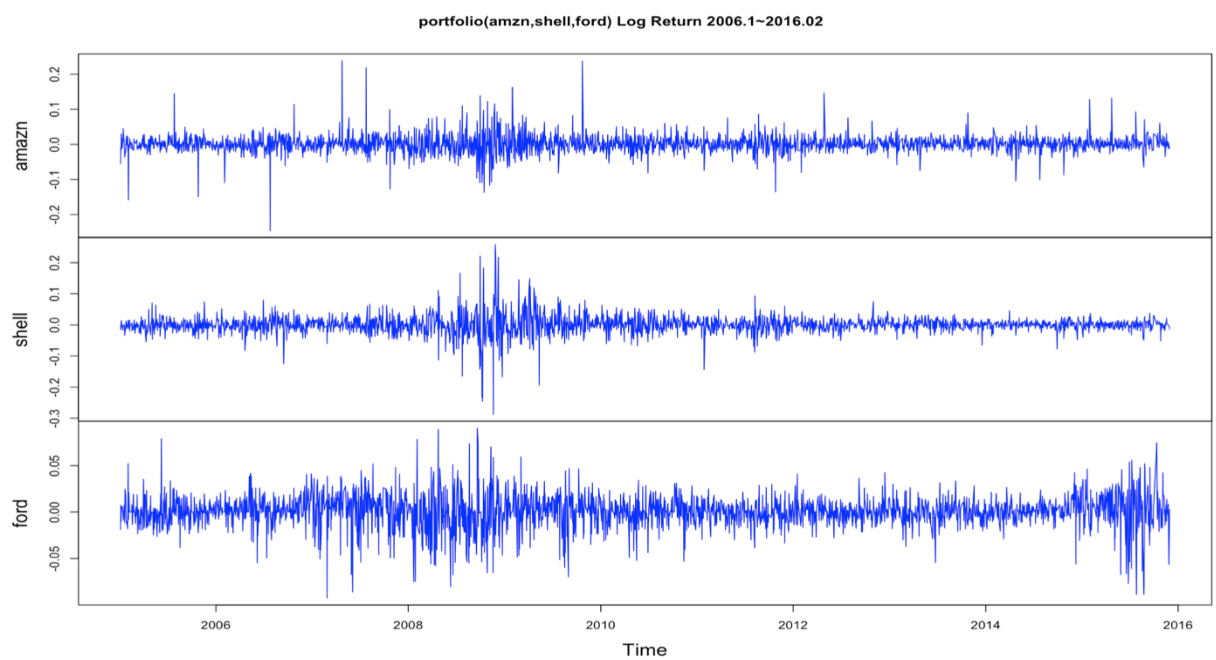

Figure 3 Log-Returns of our 3-dimensional portfolio, amazon, shell and ford.

To get a deeper understanding of Univariate GARCH model, we firstly need to know more about ARMA model. ARMA is a model combined autoregressive (AR) process with moving average (MA) process. General ARMA model was firstly proposed by Peter Whittle in 1951. The so called ARMA(p,q) model then can be written as:

$$X_t = \phi + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \cdots + \alpha_p X_{t-p} + \beta_1 \varepsilon_{t-1} + \cdots + \beta_q \varepsilon_{t-q} + \varepsilon_t ,$$

where ϕ denotes the grand mean of the interested variable X_t (Whittle, 1951). $\alpha_1, \ldots, \alpha_p$ are parameters for AR(p) process, while $\beta_1, \ldots, \beta_q$ are parameters belonging to MA(q) process. $\varepsilon_t' s$ are explained as the given historical log-returns, which are white noise terms with $\varepsilon_t \sim N(0, \sigma^2)$.

Since GARCH model is designed to predict time-varying variance, σ_t^2 is set to be the target variable and inert into ARMA(p,q) model given above. By replacing X_t in ARMA model with σ_t^2, Generalized Autoregressive Heteroscedasticity (GARCH) model (Bollerslev, 1986) is achieved:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_p \sigma_{t-p}^2 .$$

Detailed explanations of the parameters in GARCH model are the same as those in ARMA(p,q) model. In real applications, MLE is usually preferred to predict those parameters.

Set $\eta = (\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)^T$, with σ_t^2 as the conditional variance at time t and log-returns ε_t^q , then the conditional likelihood is:

$$l^b(\eta) = \sum_{t=q+1}^{n} l_t(\eta) = -\frac{n-1}{2}\log(2\pi) - \frac{1}{2}\sum_{t=2}^{n} log\sigma_t^2 - \frac{1}{2\sum_{t=q+1}^{n} \frac{\varepsilon_t^q + 1}{\sigma_t^2}} .$$

Under certain technical conditions, ML estimators will be achieved as consistent through twice derivatives. In our thesis research, functions in R packages *rugarch* and *rmgarch* take response of estimating the parameters. Together, for the univariate case, if GARCH (1,1) is considered, we have:

$$\widehat{\sigma_t^2} = \widehat{\alpha_0} + \hat{\alpha}\varepsilon_{t-1}^2 + \hat{\beta}\widehat{\sigma_{t-1}^2}$$
$$= \frac{\varepsilon_{t-1}^2 + \hat{\beta}\varepsilon_{t-2}^2 + \widehat{\beta^2}\varepsilon_{t-3}^2 + \cdots}{1 + \hat{\beta} + \widehat{\beta^2} + \cdots}.$$

Here, we substituted $\hat{\alpha}$ with 1-$\hat{\beta}$, since, from the R output, $\hat{\beta} + \hat{\alpha}$ equals to 0.995 approximately in most cases.

4.2.2 DCC-GARCH model and GO-GARCH model

The financial objects covered in the thesis are stock portfolios consisting of exchangeable shares. Thus, multivariate GARCH models are appropriate and they are built based on the created univariate GARCH specification objects.

Now suppose that we have a d dimensional time-varying process x_t, with the mean of each element $x_{i,t}$ equals to $\mu_{i,t}$. If information is available until time t-1, then

$$x_t|\mathcal{F}_{t-1} = \mu_t + \varepsilon_t.$$

Notice the error term ε_t is time-varying and depends on the historical residuals estimated by GARCH models. With the increase of portfolios' dimensions, the dependence structure generated by unequal error vectors are more and more complex, which makes the prediction less reliable. A straightforward idea to solve it is to reduce or amalgamate the d x d symmetric time-dependent covariance matrix H_t in to a d*(d+1)/2 dimensional vector. Currently most commonly used multivariate GARCH models are operated based on this idea to deal with the time-varying covariance matrix (Alexander, 2001). More specifically, take a closer look at a three-dimensional example with GARCH (1,1) applied. One GARCH model is established with Vec Specification, which is the basis of DCC and GO-GRACH specification. Basic ideas of Vec Specification will be defined below.

Set H_t as the conditional covariance matrix with elements $H_{ii,t}$; ϕ be the grand mean; α and β be the two corresponding parameters for σ_t and ε_t, respectively.

$$H_t = \begin{bmatrix} h_{11,t} & h_{12,t} & h_{13,t} \\ h_{21,t} & h_{22,t} & h_{23,t} \\ h_{31,t} & h_{32,t} & h_{33,t} \end{bmatrix}.$$

Next, define vech(.) as an operator that stacks the upper triangular part of a symmetric 3 x 3 matrix into a $d^* = 3(3+1)/2$ dimensional vector. This vector is thus denoted as $\boldsymbol{h_t} = \boldsymbol{vech}(H_t)$. Then we have: $\boldsymbol{h_t} = \boldsymbol{vech}\,(H_t) = (h_{11,t}, h_{12,t}, h_{13,t}, h_{22,t}, h_{23,t}, h_{33,t})^T$, this definitely gives us a simpler structure of the covariance matrix. The simplified model can thus be rewritten explicitly as:

$$\begin{pmatrix} h_{11,t} \\ h_{12,t} \\ h_{13,t} \\ h_{22,t} \\ h_{23,t} \\ h_{33,t} \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \\ \phi_5 \\ \phi_6 \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} & \alpha_{15} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} & \alpha_{25} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{34} & \alpha_{35} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44} & \alpha_{45} \\ \alpha_{51} & \alpha_{52} & \alpha_{53} & \alpha_{54} & \alpha_{55} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1}^2 \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{1,t-1}\varepsilon_{3,t-1} \\ \varepsilon_{2,t-1}^2 \\ \varepsilon_{2,t-1}\varepsilon_{3,t-1} \\ \varepsilon_{3,t-1}^2 \end{pmatrix}$$

$$+ \begin{pmatrix} \beta_{11} & \beta_{12} & \beta_{13} & \beta_{14} & \beta_{15} \\ \beta_{21} & \beta_{22} & \beta_{23} & \beta_{24} & \beta_{25} \\ \beta_{31} & \beta_{32} & \beta_{33} & \beta_{34} & \beta_{35} \\ \beta_{41} & \beta_{42} & \beta_{43} & \beta_{44} & \beta_{45} \\ \beta_{51} & \beta_{52} & \beta_{53} & \beta_{54} & \beta_{55} \end{pmatrix} \begin{pmatrix} h_{11,t-1} \\ h_{12,t-1} \\ h_{13,t-1} \\ h_{22,t-1} \\ h_{23,t-1} \\ h_{33,t-1} \end{pmatrix}.$$

On this basis, Engle (2002) proposed a model called Dynamic Conditional Correlation (DCC) model to solve time-varying correlations.

Suppose we have returns e_t from d-dimensional portfolios with mean zero and conditional covariance matrix H_t, $\delta_t \sim \cdot\, (0, I)$. Then the DCC-model is given by:

$$r_t = \mu_t + e_t$$

$$e_t = H_t^{\frac{1}{2}} \delta_t$$

$$H_t = D_t R_t D_t \ .$$

Notation:

r_t: log returns of the portfolio, d x 1 vector;

$H_t^{\frac{1}{2}}$: can be obtained through the DCC specification process;

D_t: d x d diagonal matrix of conditional standard deviations of e_t at time t;

R_t: d x d correlation matrix of e_t at time t. It should be positive definite. All the elements should not be larger than 1.

Next, the main task of the DCC-GARCH model is to estimate the error term δ_t. Currently, there is still no unanimous of which multivariate distribution fit the parameters in GARCH models the best. In our thesis, the multivariate Gaussian distribution is selected to extract δ_t. And, maximum likelihood method is used for estimation.

Although DCC-GARCH model is designed to overcome the restrictions that general GARCH models are executed on the hypothesis of constant correlation (Weide, 2002), the conditional variances of DCC specification are still estimated from its univariate correspondence. This should be an issue especially for large covariance matrices. In view of this, Alexander (1990's) designed new approaches called Orthogonal GARCH (O-GARCH) and Generalized Orthogonal GARCH (GO-GARCH) model successively from 1994 to 2001. These models assume that there exists at least one orthogonal matrix to link the observed data linearly to a set of independent components, which can be explained as econometric factors that will influence the financial activities. Since the transformed components are assumed to be independent, it is logical to get the conditional variances from a cluster of univariate GARCH models. Nevertheless, just like a concept called Inverse, no one can guarantee there must exists the linkage (an inverse). And, this "map & inverse map" idea is challenged with certain identification concerns. Suppose the diagonal elements of H are not all distinct, then the transformed orthogonal matrix, say P, will no longer be coincide with with the linkage Z. And the estimation of the matrix solely hinges on the sample data set, instead of the historical information. On the contrast, GO-GARCH has prominent advantages over its ancestor. It allows Z to be given by any possible invertible matrix. Estimation of this matrix requires the historical information to solve the identification problems and thus gives rise to introducing higher dimensional portfolios in our investigation at an affordable money and time costs. In our modeling process, GO-GARCH will be preferred in higher dimensions while DCC-GARCH is used in lower ones.

The key assumption of GO-GARCH is that the observed economic process $\{x_t\}$ is governed by a linear combination of independent economic components $\{y_t\}$:

$$x_t = \mathrm{Z}y_t ,$$

with Z as the linear map that links original data with independent components, $y_t \sim N(0, H_t)$. The conditional covariance of $\{x_t\}$ are given by:

$$V_t = ZH_tZ^T.$$

After applying the linkage, similar to the definition of DCC-GARCH model, we have $R_t = D_t^{-1}V_tD_t^{-1}$. Considering the bivariate case, we set the linkage map Z_θ as:

$$Z_\theta = \begin{bmatrix} 1 & 0 \\ cos\theta & sin\theta \end{bmatrix}.$$

In this way, θ measures the extent to which the observed data are mapped to the independent components. For example, if $\theta = 0$, the map is not invertible and the model is failed to be established; whereas, if $\theta = \pi/2$, the original data are already uncorrelated and we do not have to transform them. The conditional covariance matrix is given by:

$$\sigma_{1t}^2 = h_{1t}$$
$$\sigma_{2t}^2 = h_{1t}cos^2\theta + h_{2t}sin^2\theta .$$

Restricted to the available reference, Z is chosen to be constant over time for the prediction of covariance in this thesis. Since the direction of Z does not ever change during the process, this constant assumption is valid for use.

4.2.3 Summarization of the modeling procedure

As discussed, modeling process will move forward to the Log-Return equation.

$$X_{i,t} = \mu_{i,t} + \sigma_{i,t}\epsilon_{i,t}\ , i = 1,2, \dots , n$$

Step 1: Build Univariate GARCH model to specify the conditional variance.

Step 2: Use the results from Step 1 to fit DCC-GARCH model.

Step 2^{*}:Model the time series data with GO-GARCH model.

Step 3: Extract residuals, expectations and covariance matrix from either Step 2 or Step 2^{*}.

Step 4: Simulate residuals on known multivariate distribution to avoid calculating failures.

Step 5: Specify marginal distributions, fit copulas and obtain estimated parameters.

Step 6: Check if the copula-generated distributions fit well by contour plots.

Step 7: Estimate Log Returns for the forecast and predict VaR(α).

Step 8: Use backtesting to test the goodness of selected copula estimations at the very dimension and try to fix the problems.

Following these detailed procedures, we can model the Log Returns reliably, but which copulas output better results is still open to doubt. More significantly, to what extent the copula methods can be effective need to be verified as well. In next chapter, we will address a solution through practical cases.

Chapter 5

Check the Goodness of Copula Methods

5.1 Basics of Backtesting

With the improvement of the predictive modeling methods in measuring risks, researchers have developed many ways to examine the reliability of the models. Since 1998, regulatory guidelines have required banks with substantial trades to launch tests to gauge the accuracy of their VaR models (Sean, 2005). Abide by the guidelines, scholars designed a variety of backtests to examine the adequacy of VaR measures. Basically, the backtesting procedures evaluate the quality of the forecast of a risk model by comparing the actual results with those generated by the VaR models. While these tests differ in details, most of them focus on particular transformations of the reported VaR and realized losses. Currently, backtests can be classified by whether they examine an unconditional coverage property or independence property of a VaR measure. Sometimes, we are also interested in checking the validity of VaR models at a series of quantiles, rather than a given single quantile. Furthermore, backtests in terms of pre-specified loss functions are also worth concern since sometimes the past data on profits or loss give us a nominal good fit but still need to be checked statistically.

Specifically, consider a event that the loss of a portfolio exceeds its reported VaR, $\mathrm{VaR}(\alpha)$ at quantile α. Denoting the profit or loss of the portfolio between time t to t+1 as $r_{t,t+1}$, the so called "hit" function $I_{t+1}(\alpha)$ can be defined as (Christoffersen,1998):

$$I_{t+1}(\alpha) = \begin{cases} 1 & if \;\; r_{t,t+1} \le -\mathrm{VaR}(\alpha) \\ 0 & if \;\; r_{t,t+1} > -\mathrm{VaR}(\alpha) \end{cases}.$$

Thereupon, the hit function sequence, e.g. (1,0,1,0,0,0,0,1), tells us the times that the loss of the portfolio has excess the labeled $\mathrm{VaR}(\alpha)$ during the observed period. In addition, from the sampling sequence, we can know the corresponding time pause intervals between the exceeded points. For instance, the vector (1,0,1,0,0,0,0,1) owns two time-pause intervals, 1 for the first interim (with one zero occurred between 1s) and 4 for the second interim (with four zero occurred between 1s).

5.2 Unconditional Coverage Tests

Traditionally, some proposed VaR tests solely keen on the property of unconditional coverage. In other words, those tests are interested in the times that reported VaRs have violated. If the violations are much different from $\alpha\times100\%$ of the total observing days, the tests are rejected.

Early unconditional coverage tests include Kupiec's (1995) proportion of failure (POF) test and the "traffic light" method as a surcharge factor.

For POF test, if the number of recorded violations differ considerably from $\alpha\times100\%$ of the sample size, it is reasonable to doubt the accuracy of the underlying model. For example, suppose a financial institution reported a 1% VaR of \$10,000,000 over a 1-day horizon, then the institution would be expected to realize a loss in excess of \$10,000,000 for 1% of the trading times. The current regulatory framework requires that financial institutions use their own internal risk models to calculate and report their 1% value-at-risk, VaR (0.01), over a 10-day horizon. In other words, if more than ten violations of the 1% VaR are recorded in a 250-day span, which accounts to 4% of the sample period, the VaR model is thus claimed inaccurate, and immediate steps are required to improve the risk management system. Given time period t, which means we have t observations, Kupiec's test statistics can be written as:

$$\mathrm{POF} = 2\log\left(\left(\frac{1-\hat{\alpha}}{1-\alpha}\right)^{t-I(\alpha)}\left(\frac{\hat{\alpha}}{\alpha}\right)\right)$$

$$\hat{\alpha} = \frac{1}{t}I(\alpha)$$

$$\mathrm{I}(\alpha) = \sum_{i=1}^{t} I_i(\alpha) .$$

Where $\hat{\alpha}$ is the estimated proportions of the exceedance during time period t, $I_i(\alpha)$ is the "hit" function at time i. Now, consider the test statistics POF, if $\hat{\alpha}\times 100\%$ is exactly equal to the criterion value $\alpha\times 100\%$, there is no evidence to say the VaR model is inadequate. But, if the estimator $\hat{\alpha}$ is biased, we will have sufficient evidence to say that the proposed VaR measure either overestimate or underestimate the risk of the portfolio.

Apart from POF statistic, we can also design another one by transforming the predicted percent of violation $\hat{\alpha}$ on given time period as a scaled version. Under the assumption that the VaR measure is accurate, then,

$$\mathrm{z} = \frac{\sqrt{t}(\hat{\alpha} - \alpha)}{\sqrt{\alpha(1-\alpha)}},$$

where z ~ N (0,1). And the hypothesis test can be conducted the same way as what POF statistic does.

5.3 Independence Tests

In light of the failure of unconditional coverage test to detect dependency property of VaR measure, we need to find the specialized independency tests to examine the independency property of VaR hit series, $I_t(\alpha)$.

Earlier, Christoffersen (1998) developed Markov test to examine whether or not the likelihood of a VaR violation depends on the occurrence of violations on the historical days. For instance, if more 1% VaR violations come out after a violation, this might manifest the likelihood of 1% VaR following a violation will be increased.

Recently, Christoffersen and Pelletier (2004) noticed that the interim between two VaR violations exhibits no duration dependency, that is, the chance that a violation occurs in the next ten-day period should never rely on whether a violation occurred in the past 10 to 100 days.

According to Christofersen's researches, likelihood ratio test for independence of exceedances is worthy of conducting, we firstly define q^* as the likelihood, then:

$$q_0^* = P_r(I_t = 0|i_{t-1} = 0)$$
$$q_1^* = P_r(I_t = 0|i_{t-1} = 1)\ .$$

These two probabilities give us the conditional coverage test we will talk about in the next part. The actual probabilities of experiencing and not experiencing exceedance are differentiated by the exceedance at time t-1 (0 means no exceedance while 1 means exceedance occurred). The null hypothesis thus is: H_0: $q_0^* = q_1^* = q^*$.

Hence, if a VaR measure is observed over n+1 periods, there will be n consecutive observations and the relationship is: $n_{00} + n_{01} + n_{10} + n_{11} = n$. And, the null hypothesis can thus be transformed as: $\frac{n_{00}}{n_{00}+n_{01}} = \frac{n_{10}}{n_{10+n_{11}}}$.

Our likelihood ratio is:

$$\Lambda = \frac{(1-\widehat{q^*})^{n_{01}+n_{11}}\ (\widehat{q^*})^{n_{00}+n_{10}}}{(1-\widehat{q_0^*})^{n_{01}}\ (\widehat{q_0^*})^{n_{00}}\ \ (1-\widehat{q_1^*})^{n_{11}}\ (\widehat{q_1^*})^{n_{10}}}\ .$$

With $-2\log(\Lambda) \sim \chi^2(1,0)$. We want to use this statistic to check if the independence property of the portfolio is fulfilled.

5.4 Joint Tests of Unconditional Coverage and Independence

This joint test sometimes is also called conditional coverage test and is used to be compared with the testing results of previously described POF method and its transformed tests. We denote the Conditional Coverage (CC) test by the following formula:

$$L_{cc} = L_{pof} + L_{idp}\ .$$

Notice that both of the two terms on the right side can be used separately, because they represent the POF test and Independency test, respectively. And their joint influence lead to the so called CC test.

In all, a backtesting process is the test to check whether or not the fitted VaR exceedance is equal to the so called empirical threshold at a given confidence value α. If failed to reject, we can conclude the given model fits well, otherwise, the model needs to be modified or abandoned.

In R, the function 'VARTest' from package 'rugarch' helps do the backtesting test.

Chapter 6

Reliability Verification of Copula Methods

6.1 Data Summary

In our study, different dimensional cases are selected to build portfolios. The purpose of our work contains two tips: 1. The influence of dimensions on the goodness of copula methods. 2. The accuracy of different copulas on estimation.

To achieve our purpose, time point price data of stocks on NASDAQ and NYSE are extracted online by R functions. The time period we used is from Jan. 1st, 2005 to Jan. 13th 2016. After collected, the stocks are clustered and assigned into different portfolios. As mentioned before, the raw datasets are transformed to log forms for time series analysis. Techniques treated with the data (BAC as an example) are shown below to get the log-returns of each stock.

```
##BAC
bac<-BAC[,4]
bac[,1]<-diff(log(bac[,1]))
bac<-bac[-1,]
bac<-as.xts(bac)
```

Next, take portfolio {BAC, DAL, DG} as an example to see the details of the transformed dataset. The following output shows the first 6 rows of this portfolio's historical log-returns.

	boa.logRet	delta.logRet	dollargeneral.logRet
2009-11-16	-0.0069074058	0.024067026	0.016146980
2009-11-17	-0.0063211336	-0.011328000	-0.002600738
2009-11-18	0.0361184964	-0.017880425	-0.021496858

2009-11-19	-0.0166516336	-0.022150743	0.014091030
2009-11-20	0.0006216973	0.003944778	0.014757125
2009-11-23	0.0123534616	0.014332493	0.008152845

Now, we examine the data included in our analysis to get a brief overview of them. Table 2 summarizes the portfolios chosen as the sample. The first column is the dimensions of the samples in our research, which cover 3, 5 and 9 dimensional cases.

Portfolio Dimensions	Samples	Sample size (rw size)	Proportions	Markets
3	Amazon, Ford, Shell	1000 (250)	1:1:1	NYSE, NASDAQ
	BA, SPY, GS			
	DDD,SPY,GS			
	BAC, DAL, DG			
	APPL, GOOG, MSFT			
5	APPLE, NFLX,DDD,BA,TUC	1750 (500)	1:1:1:1:1	NYSE, NASDAQ
	AMZN, FORD, GS, SHELL, SPY			
	DDD,TUC,FORD,BAC,MSFT			
	WMT, C, WMAR, MS, ACET			
	SHELL,CITI,BAC,MS,ACUR			
	APPLE, DAL, ENZN, SPY, NFLX			
9	APPLE, WMT, NFLX, SPY,C,F,AMZN,MSFT,GS	2550 (250)	1:1:1:1:1:1:1:1:1	NYSE
	ACET, C, ACUR, BAC, MS, DDD,TUC, SH, SPY			
	WSTL, C, ACUR, BA, MS, DDD, TUC, WMAR, ENZN			

Table 2 Description of the sampling portfolios

The second column lists all the portfolios. We have 5 samples for 3 dimensional portfolios, 6 for 5 dimensional case, and 3 for the 9 dimensions. Although the raw data are collected for more than 2800 days, we only extract 1000, 1750 and 2550 of them for 3, 5 and 9 dimensional cases, respectively. What needs to be specified here is that the title "rw size" stands for Rolling Window sizes. The so called Rolling Window analysis is an algorithm on moving time intervals that helps us estimate time series data forward stably. In addition, all of the assets in each portfolio are equally proportional, which are equity securities of listed companies on NYSE and NASDAQ.

In order to ensure random sampling and typical, we selected elements of the portfolios by random sampling methods. For instance, portfolio {Amazon, Ford, Shell} is built with three representative stocks from Standard & Poor's 500 Index, and the stocks are selected randomly by industries. A portfolio like {AAPL, GOOG, MSFT} consists of three technical corporations and it is designed to see if the functions of copulas will be affected by high related assets. For higher dimensional cases, we firstly opted a "stock pool" from Standard & Poor's 500 Index with different industries (stratifications) and then randomly group assets together to form portfolios for our analysis.

Although nominally some portfolios have highly related equity securities while others seem not, we cannot simply conclude that the trends of the assets are mutually influenced. For example, referring to the financial statements of Ford and Amazon, Vanguard Group, Inc. is the biggest shareholder of Amazon and the third biggest shareholder of Ford by the first season of 2016. Thus, it is believed that the two stocks are more or less correlated with each other and the relationship would be manifested on the second market. More details can be found in the output for each case.

6.2 Accuracy test for Copula Estimation on trivariate cases

In this section, we put our models into practice and check if they are reliable for real usage. The modeling process of portfolio {BAC, DAL, DG} with Gumbel Copula will be introduced in detail and backtesting results with further analysis are given with more attention. This portfolio contains tickers from three different sectors. A 750-day rolling analysis is conducted with 250-day rolling window, 1:1:1 weights and a 5% significant level. With the residual vectors estimated by GARCH model, we can then fit the residual distributions prepared for copula simulation. Figures 4, 5, and 6 give us a visualized view of how to choose the multivariate distributions for the simulated residuals. As for the curves in these plots, the blue ones represent the log-return residuals estimated by GO-GARCH

(1,1) model, while the yellow curves stand for normal distribution with parameters extracted from GO-GARCH (1,1). The green and black ones are for skewed-t distribution and skewed normal distribution, respectively. From these plots, it is obvious that skewed-t distribution fits the estimated residuals the best, not only for the two tails, but also at the peak point. However, since simulating residuals on skewed-t distribution is time wasting and expensive, t distribution is preferred in most cases. (Franke, Hardle and Hafner, 2008).

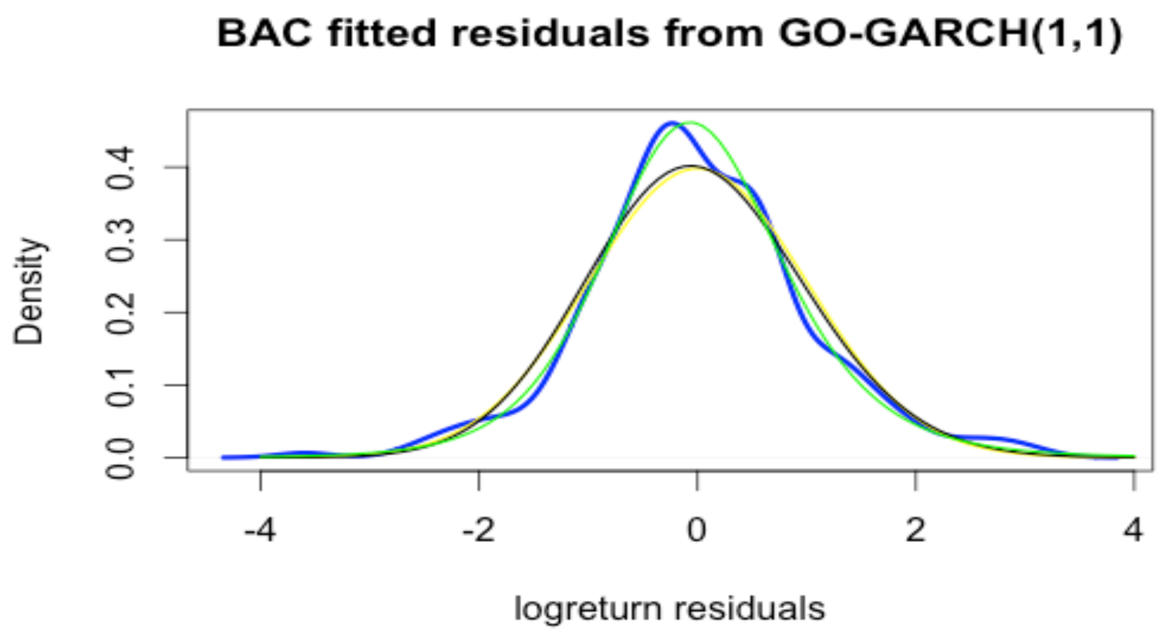

Figure 4 Residual plot for BAC fitted by GO-GARCH(1,1)

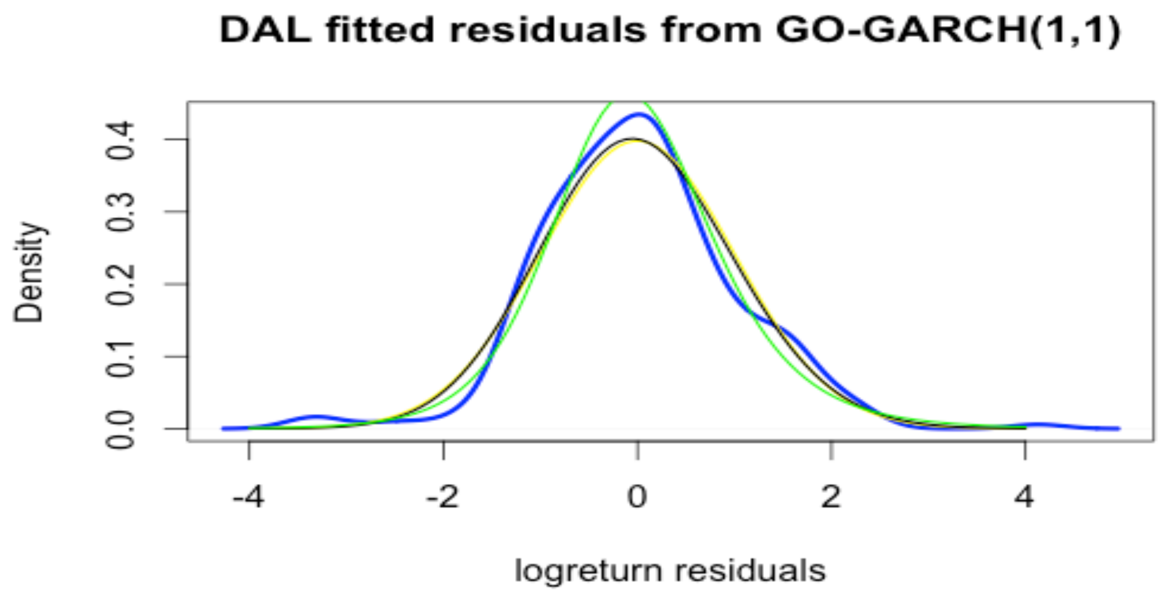

Figure 5 Residual plot for DAL fitted by GO-GARCH(1,1)

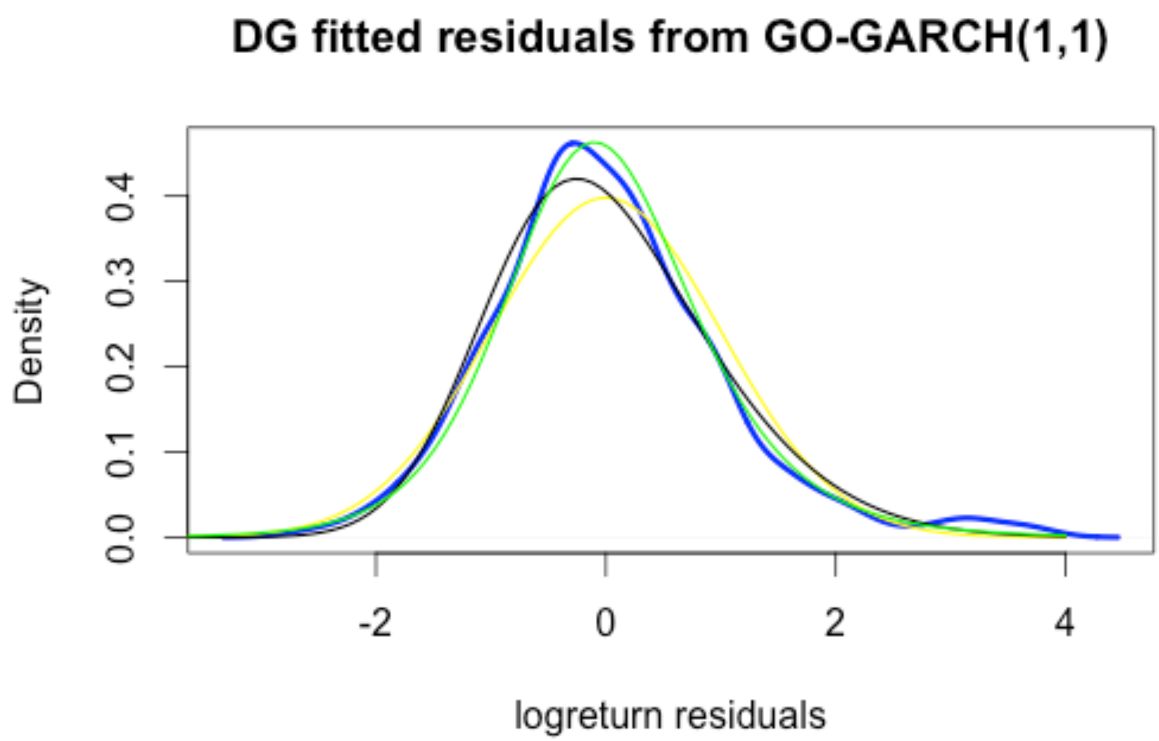

Figure 6 Residual plot for DG fitted by GO-GARCH(1,1)

For other trivariate cases, t distribution is also the best choice for log-return residuals. Due to the limit of the space, we do not report all the plots here but only display portfolio {BAC, DAL, DG}'s as an illustration.

After simulating the residuals, we are ready to fit copulas with them and get aimed parameters for the final VaR exceedance tests. Prior to applying backtesting, plotting parameters and variables is an easy way to check the goodness of the copula models as well. Figure 7 displays 750 residuals for the portfolio {BAC, DAL, DG} estimated from four commonly used Elliptical Copulas and Archimedean Copulas. In Figure 7, the first panel is for Gumbel, the second reports normal copula, the third one displays t copula, and the last plot is for Clayton Copula. It can be seen that normal and t copulas produce a roughly elliptical results, while the results of Gumbel copula are less partial to the outward.

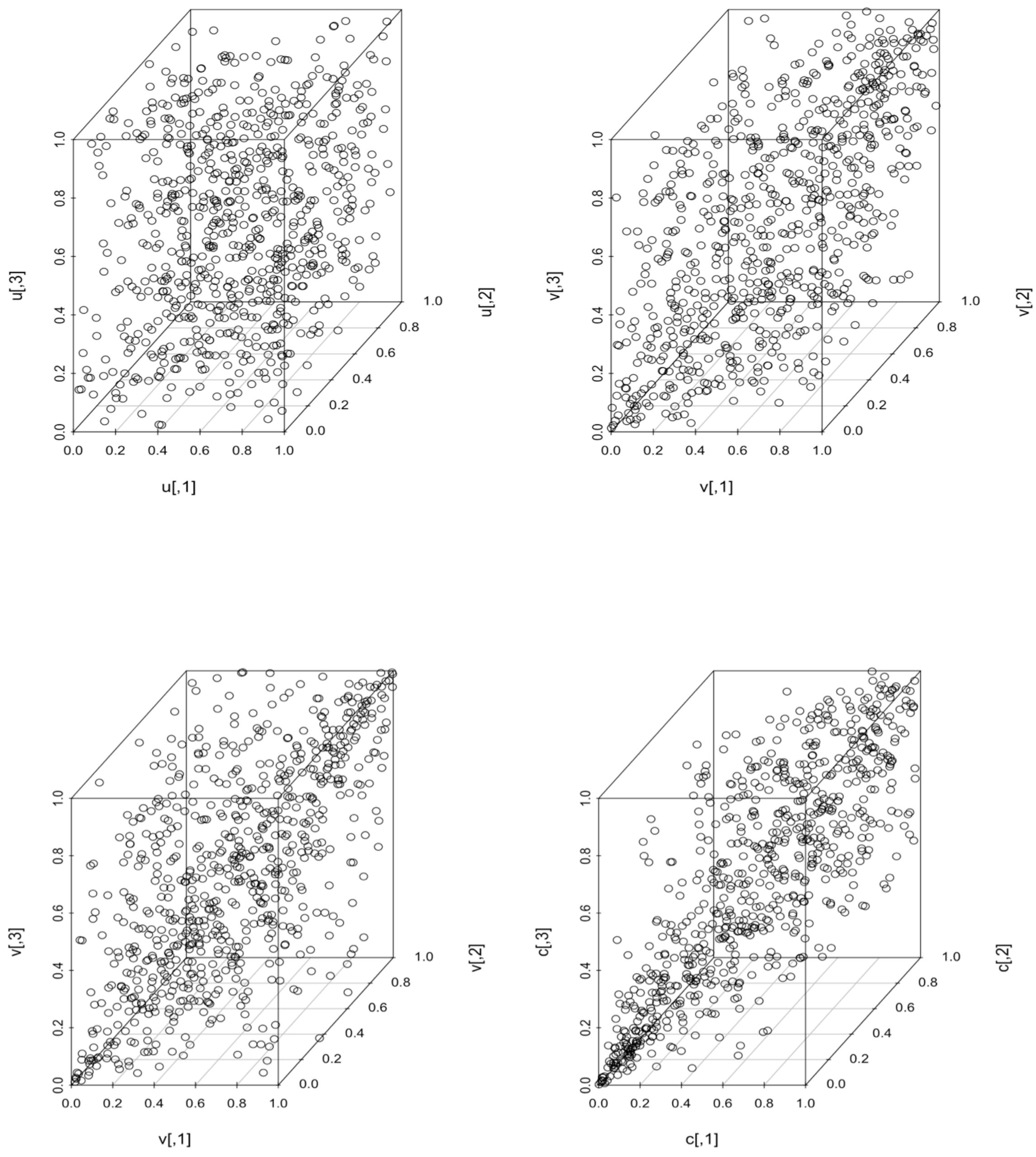

Figure 7 scatter plot of Gumbel copula fitted residuals (1st panel), of normal copula (2nd panel), t copula (3rd), and clayton (4th)

As discussed at the start of Chapter 6, the goodness of fit is evaluated through backtesting for the VaR results. Both Unconditional Coverage Test and Conditional Test are conducted and compared for analysis. Charts in this section contain the results of backtesting methods for each applicable copula with expected coverage at 5%. To be more convincing, we made use of 3, 5 and 9 dimensional portfolios one by one as the cases for the tests. According to the definition, AMH Copula is only available at bivariate case (Franke, 2007), so it would not be involved in our analysis. Figure 8 visualizes the VaR prediction made by Gumbel Copula, although a few points exceed the criterial threshold, most of them remain acceptable. The rolling analysis is based on 750 days. Again, due to the limit of space, we only plot the VaR prediction with Gumbel Copula at here, VaR predictions for other copulas are of the similar format. Note the red line is the estimated $\widehat{VaR(\alpha)}$.

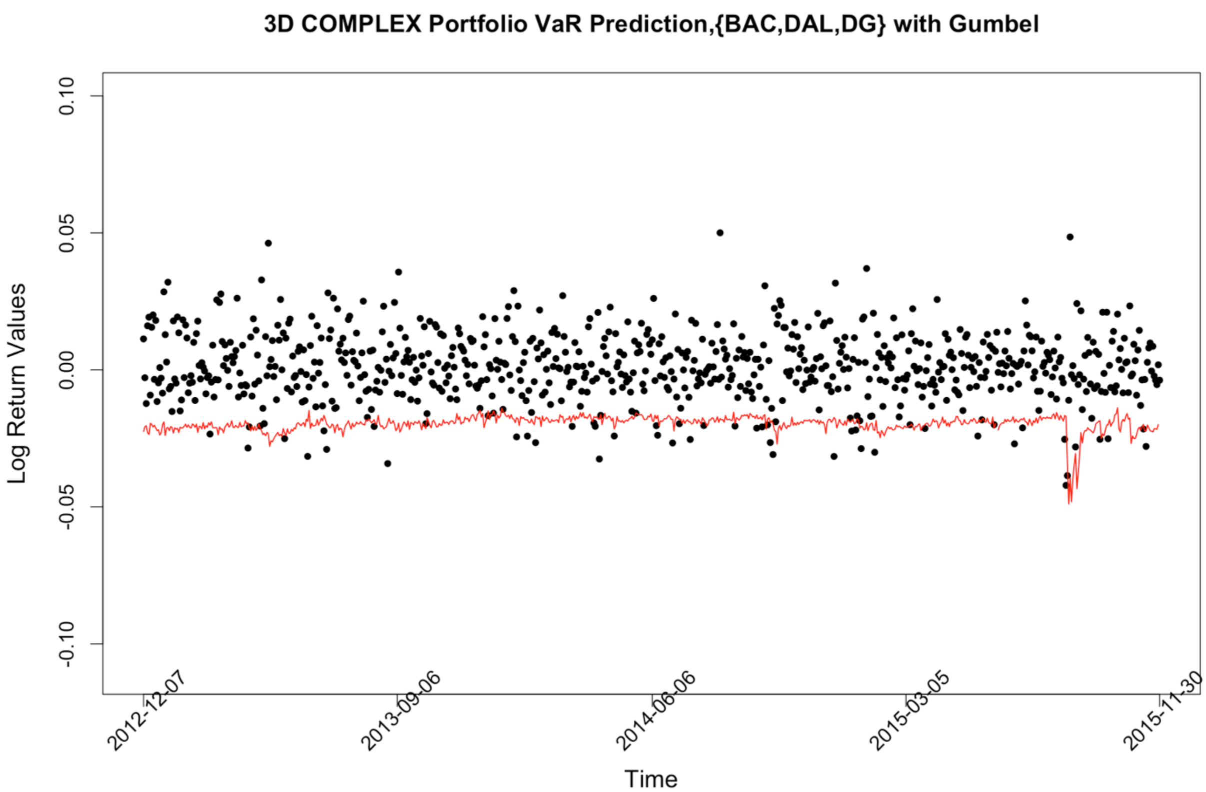

Figure 8 VaR 5% prediction for trivariate portfolio {BAC, DAL, DG} with Gumbel Copula

The summary of the backtesing results for this portfolio is given in Table 3:

Portfolio: {BAC, DAL, DG}			
copula	**unconditional**	**conditional**	**Exceedance**
Gumbel	0.5480806	0.7367212	5.6% (42 out of 750)
Clayton	NA	NA	NA
Gaussian	0.4590989	0.4405354	5.6% (42 out of 750)
t copula	0.06521229	0.1360852	4.53% (34 out of 750)
NAC	0.5632115	0.5253472	5.5% (41 out of 750)

Table 3 Bactesting results of portfolio {BAC, DAL, DG}, with 5 copula methods applied on 5% significant level

For this portfolio, 5 copulas are applied to estimate the log-returns. After simulating all the parameters by DCC-GARCH (1,1) model, Maximum Likelihood Estimation was introduced to estimate the Copula parameters, with skewed t distribution defined as the margins. The parameters to be estimated consist of marginal parameters β and copula parameter α. As suggested from Table 3, all the tests, except for Clayton Copula, fail to reject the null hypothesis that we expect exactly 5% of the 750 days (approximately 37 days) of exceedance. It is worth noting that the Nested Archimedean Copula in this case is iterated with a Gumbel generator. From the results, NAC with Gumbel generator gives us the best testing result, with 41 days or 5.6% out of the 750 days, while t Copula produces results with the most difference to our null hypothesis. As for Gaussian and Gumbel Copulas, they actually pass the test with the same exceedance result, that is, 5.6% or 42 days out of 750 observed days. More importantly, Clayton Copula generates NA results since the copula parameter is not permitted to be negative unless the dimension of portfolios is less than 3. This error appears to be caused by the data itself, yet it also uncovers the restriction of Clayton Copula.

The next portfolio we design contains three “cornerstones” from NASDAQ, and they are Amazon, Ford and Shell. Referring to their financial statement of the first season of 2016, the components of their top ten shareholders are almost the same, which means investment banks or other personal investors believed these three stocks are not correlated too much and investing a portfolio with the stocks might be a wise choice.

Portfolio: {AMZN, Ford, SHELL}			
copula	**unconditional**	**conditional**	**Exceedance**
Gumbel	0.3544971	0.7822533	4.53% (34 out of 750)
Clayton	0.8027796	0.4134458	5.2% (39 out of 750)
Gaussian	0.9333781	0.3770437	5.06% (38 out of 750)
t copula	0.5515784	0.1781517	4.53% (34 out of 750)
NAC	0.9333781	0.9950386	5.06% (38 out of 750)

Table 4 Bactesting results of portfolio {AMZN, F, SH}, with 5 copula methods applied on 5% significant level

From Table 4, all the tests fail to reject the null hypothesis, which means all of the copulas fit this model quite well. Compared with the last case, the p-values and exceedance numbers here look better and more reliable. More specifically, NAC with Clayton generator and Gaussian Copula perform the best results, while Gumbel and t copula provides the weakest fitness, say, 4.53% or 34 days out of 750 observed days. Figure 9 shows the VaR prediction at 5% significant level, where the rolling window is still 250 days. Since the information manifested by the plot is highly similar for all the cases in 3 dimensions, we will not plot another one for the following 3 dimensional cases in this section any more.

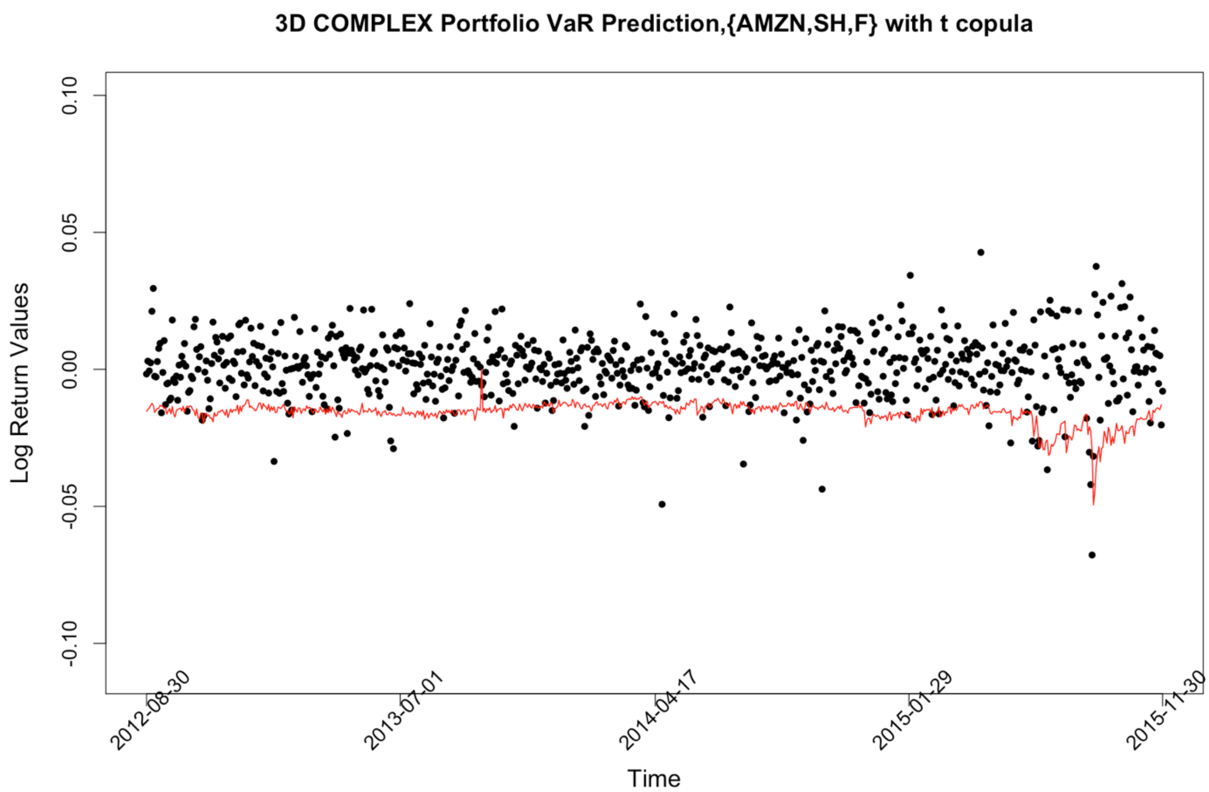

Figure 9 Estimated VaR 5% (solid line), P&L and Exceedance

Portfolio: {BA, SPY, GS}			
copula	**unconditional**	**conditional**	**Exceedance**
Gumbel	0.0652123	0.1644217	6.5% (49 out of 750)
Clayton	0.0037568	0.0058276	7.5% (56 out of 750)
Gaussian	0.3673953	0.4207369	5.7% (43 out of 750)
t copula	0.0315854	0.07015144	6.8% (51 out of 750)
NAC	0.0213867	0.06917631	6.9% (52 out of 750)

Table 5 Backtesting results of portfolio {BA, SPY, GS}, with 5 copula methods applied on 5% significant level

Table 5 displays backtesting results for portfolio {BA, SPY, GS}. At this time, the three stocks are from financial sectors with difference in focused businesses. We combine the three stocks together with 1:1:1 weights and 1000 log-return observations. From the table, Gaussian Copula fits the best, and Gumbel Copula ranks the second. As the above two cases, NAC iterated with Clayton as a generator performs better results than copulas like Gumbel or Clayton. In addition, as for Gumbel, t and NAC with Clayton, Conditional Coverage (CC) Tests fails to reject the null hypothesis, while Unconditional Coverage (UC) tests reject it. Based on the theoretical knowledge mentioned before, CC test has better performance over its counterpoints. Both CC and UC tests reject the null hypothesis under the model built with Clayton Copula, which convince us the restriction of Clayton Copula.

Portfolio: {DDD, SPY, GS}			
copula	**unconditional**	**conditional**	**Exceedance**
Gumbel	0.5515784	0.7822533	4.5% (34 out of 750)
Clayton	NA	NA	NA
Gaussian	0.1386943	0.3316343	3.87% (29 out of 750)
t copula	0.1386943	0.3316343	3.87% (29 out of 750)
NAC	0.4418941	0.6722767	4.4% (33 out of 750)

Table 6 Backtesting results of portfolio {DDD, SPY, GS}, with 5 copula methods applied on 5% significant level

In this portfolio, Clayton Copula fail to model the VaR's once more because its parameters become negative again during the modeling process. Gumbel Copula and NAC fit better than the other two copulas.

In this section, VaR exceedance tests are taken on trivariate cases. Modeling processes are put into effect with fitted residuals and margins. Given the tables and plots, we can generate the following brief conclusions:

- Although skewed t distribution fits the residuals estimated from time series models the best, in real cases, t distribution is usually preferred because of costs.
- From the scatterplots of copula fitted residuals, we know that Gumbel Copula and Gaussian Copula produce more equally distributed residuals, while residuals estimated from t and Clayton copula assemble more on the tails.
- For trivariate cases, NAC, Gumbel, elliptical, and Clayton copulas pass the Exceedance tests, which means they are workable for portfolios at this dimension level.
- Clayton Copula has some calculating restrictions which prevent it from being effective for particular data.
- NAC performs better than generator copulas themselves on backtesting, which means it can be a stable estimating method at this dimension level.

6.3 Accuracy tests for Copula estimation on 5 dimensional cases

With the backtesting results and short conclusions on trivariate cases, we extend the Copula methods to higher dimensions to see if the methods are still reliable. In this section, 5 dimensional portfolios are established and analyzed. The weight of each asset in each portfolio is 1:1:1: 1:1. Since the increase of dimension makes the covariance matrix larger, we design a 1750 sample size with rolling window as 500 days. This will provide us more historical information compared with the 750-day experimental design. GO-GARCH (1,1) model is preferred at this time, and the residuals are then simulated on t distribution to provide data for copula estimation. The margins of copulas are still chosen as skewed t distribution, as discussed in Chapter 3.

Consider a portfolio {AAPL, NFLX, DDD, BA, TUC}, we temporarily named it as complex.5d portfolio since the assets are selected from different sectors with complex backgrounds. The raw data are collected from Jan. 2005 to Dec. 2015.

Figure 10 displays the contour plots of fitted copulas on the given complex.5d portfolio. From this figure, we can see that the copulas overlap most estimated residuals.

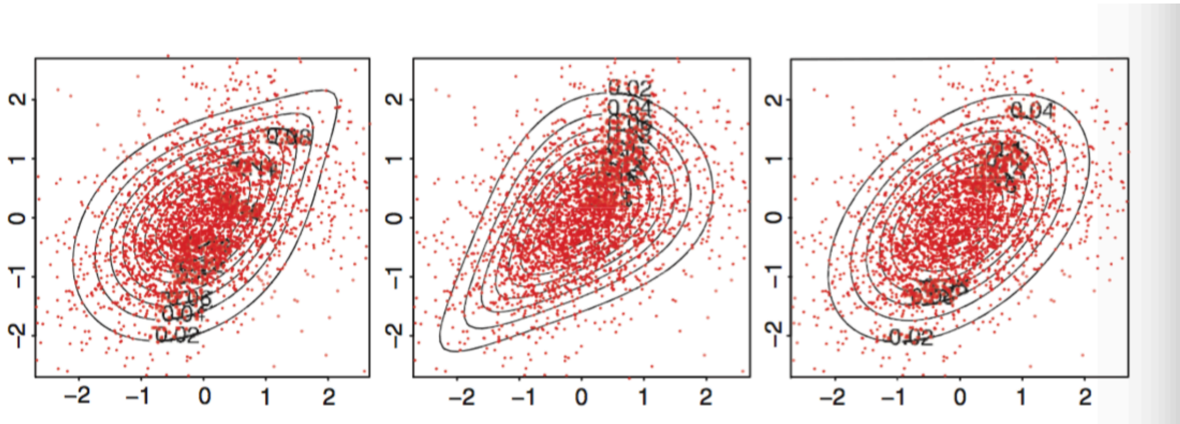

Figure 10 Estimated Residuals stretched with fitted Copulae: Gumbel copula, clayton and normal

Figure 11 plots the time-varying θ estimated by t copula. It can be seen that θ oscillates a lot over the time period. This phenomenon occurs because the price of Apple's equity security experienced huge increase during this observing period and thus made the whole portfolio face higher risk than the others. Figure 12 suggests the same point of view. P & L returns are witnessed a wide shake, with the corresponding $\widehat{VaR(\alpha)}$ changed as well.

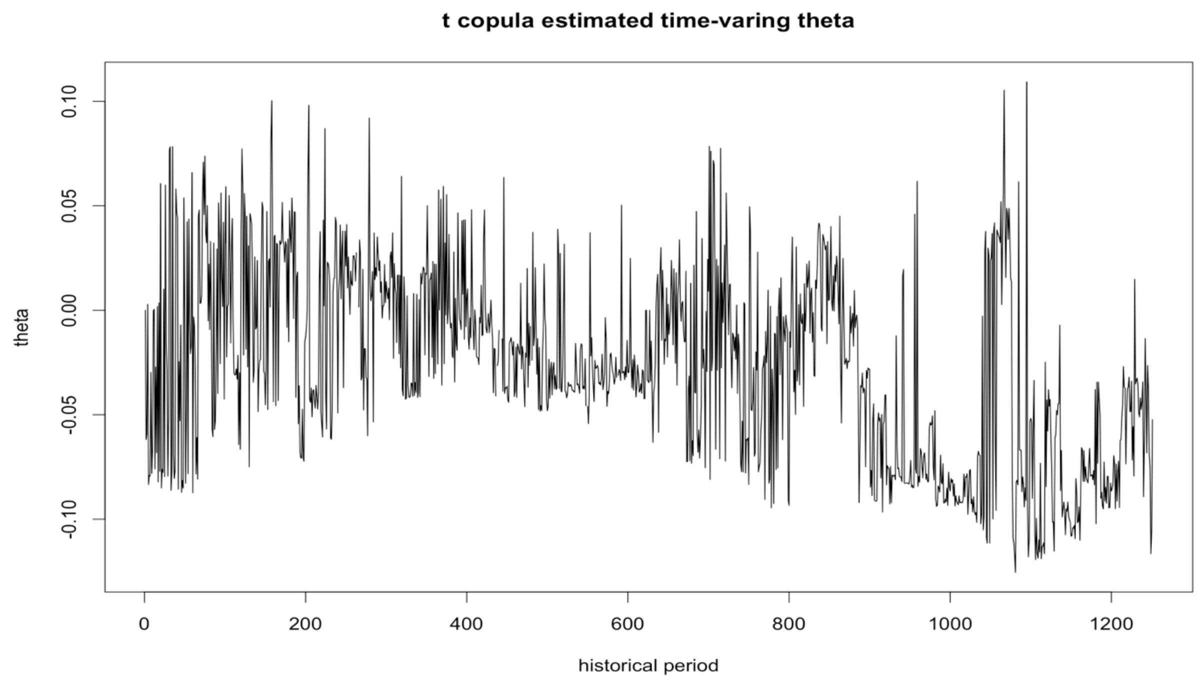

Figure 11 estimated t copula parameter theta

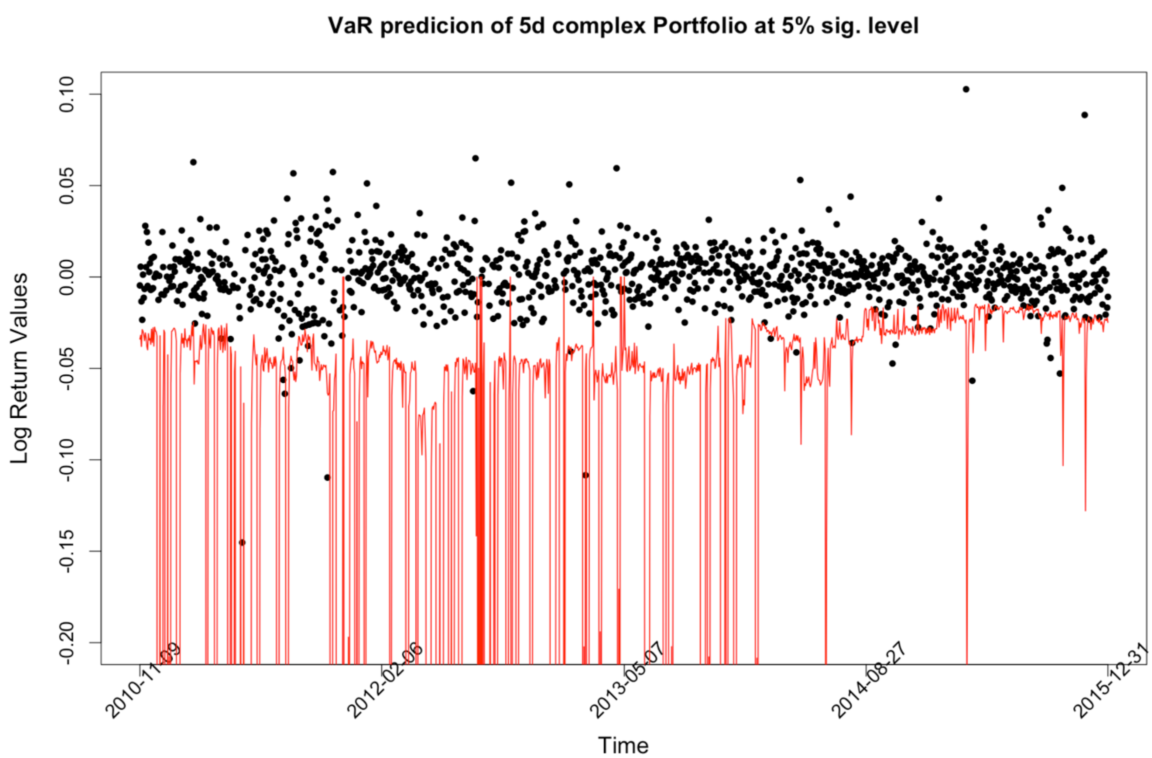

Figure 12 VaR prediction at 5% level (red line) and estimated Log-Returns

Although this portfolio has large covariance matrix and experiences a huge oscillation, 4 copulas models the P & L fairly satisfying according to Table 7, except that Clayton Copula fails to pass the Exceedance tests. The reason of the failure is that its marginal parameters turned to be negative again.

Portfolio: {AAPL, NFLX, DDD, BA, TUC}			
copula	**unconditional**	**conditional**	**Exceedance**
Gumbel	0.3208206	0.5854642	4.4% (55 out of 1250)
Clayton	NA	NA	NA
Gaussian	0.7440254	0.803617	4.8% (60 out of 1250)
t copula	0.9481972	0.7883348	4.96% (62 out of 1250)
NAC	0.7893542	0.8231094	4.4% (55 out of 1250)

Table 7 Backtesting results of portfolio {AAPL, NFLX}, with 5 copula methods applied on 5% significant level

In order to figure out if these copulae really work well on 5 dimension, another 4 portfolios are designed, with internal weights as 1:1:1:1:1. Assumptions and constraints of the model are as same as the definitions past sections. Because the modeling process were introduced repeatedly in complex.5d portfolio case and previous 3 dimensional examples, we will only print out the backtesting results and make further inference.

The backtesting results of the 4 designed portfolios are presented in Tables 8, 9, 10 and 11, with five different Copulas applied for each portfolio. At 5% significant level, all these copula methods predict Value-at-Risk pretty good. The decision regarding which copula fits the dependency structure the best can hardly be made for portfolio {AMZN, F, GS, SHELL, SPY} and {DDD, TUC, F, BAC, MSFT}, because either Archimedean Copula or Elliptical Copula fails to reject the null hypothesis.

Portfolio: {AMZN, FORD, GS, SHELL, SPY}			
copula	**unconditional**	**conditional**	**Exceedance**
Gumbel	0.4063832	0.6386581	5.5% (69 out of 1250)
Clayton	0.9481972	0.7883348	4.96% (62 out of 1250)
Gaussian	0.8450712	0.9801981	4.9% (61 out of 1250)
t copula	0.4063832	0.6386581	5.5% (69 out of 1250)
NAC	0.8964532	0.9354267	4.9% (61 out of 1250)

Table 8 Backtesting results of portfolio {AMZN, FORD, GS, SH, SPY}, with 5 copula methods applied on 5% significant level

Portfolio: {DDD, TUC, FORD, BAC, MSFT}			
copula	**unconditional**	**conditional**	**Exceedance**
Gumbel	0.115282	0.2105959	6% (75 out of 1250)
Clayton	0.8462265	0.7121321	5.1% (64 out of 1250)
Gaussian	0.1386943	0.3316343	5.6% (70 out of 1250)
t copula	0.1386943	0.3316343	5.7% (71 out of 1250)
NAC	0.4418941	0.6722767	6% (75 out of 1250)

Table 9 Backtesting results of portfolio {DDD, TUC, F, BAC, MSFT}, with 5 copula methods applied on 5% sig. level

Portfolio: {WMT, C, WMAR, MS, ACET }			
copula	**unconditional**	**conditional**	**Exceedance**
Gumbel	0.05237355	0.1324256	6.24% (78 out of 1250)
Clayton	NA	NA	NA
Gaussian	0.2592072	0.146309	4.3% (54 out of 750)
t copula	0.03932489	0.107305	6.3% (9 out of 750)
NAC	0.9481972	0.866156	4.96% (62 out of 750)

Table 10 Backtesting results of portfolio {WMT, C, WMAR, MS, ACET}, with 5 copula methods applied on 5% significant level

Portfolio: {SHELL, CITI, BAC, MS, ACUR }			
copula	**unconditional**	**conditional**	**Exceedance**
Gumbel	0.05237355	0.1520078	6.24% (78 out of 1250)
Clayton	0.3208206	0.073336	4.4% (55 out of 1250)
Gaussian	0.3208206	0.569273	4.4% (55 out of 1250)
t copula	0.9483278	0.0510023	5.04% (63 out of 1250)
NAC	0.4690168	0.04213229	4.56% (57 out of 750)

Table 11 Backtesting results of portfolio {SH, CITI, BAC, MS, ACUR}, with 5 copula methods applied on 5% significant level

The results of VaR Exceedance tests of portfolios {WMT, C, WMAR, MS, ACET} and {SH, CITI, BAC, MS, ACUR} can be found in Tables 10 and 11. Note that in Table 11, when predicting VaR with Gaussian Copula, DCC-GARCH (1,1) was firstly tried, yet the likelihood estimating process failed to converge, so we turned to GO-GARCH (1,1), which provides accurate parameters. Furthermore, for {SH, CITI, BAC, MS, ACUR}, NAC fails to pass the CC test but passes the UC test, this information convinces us that independency check is necessary in verifying copula fitting.

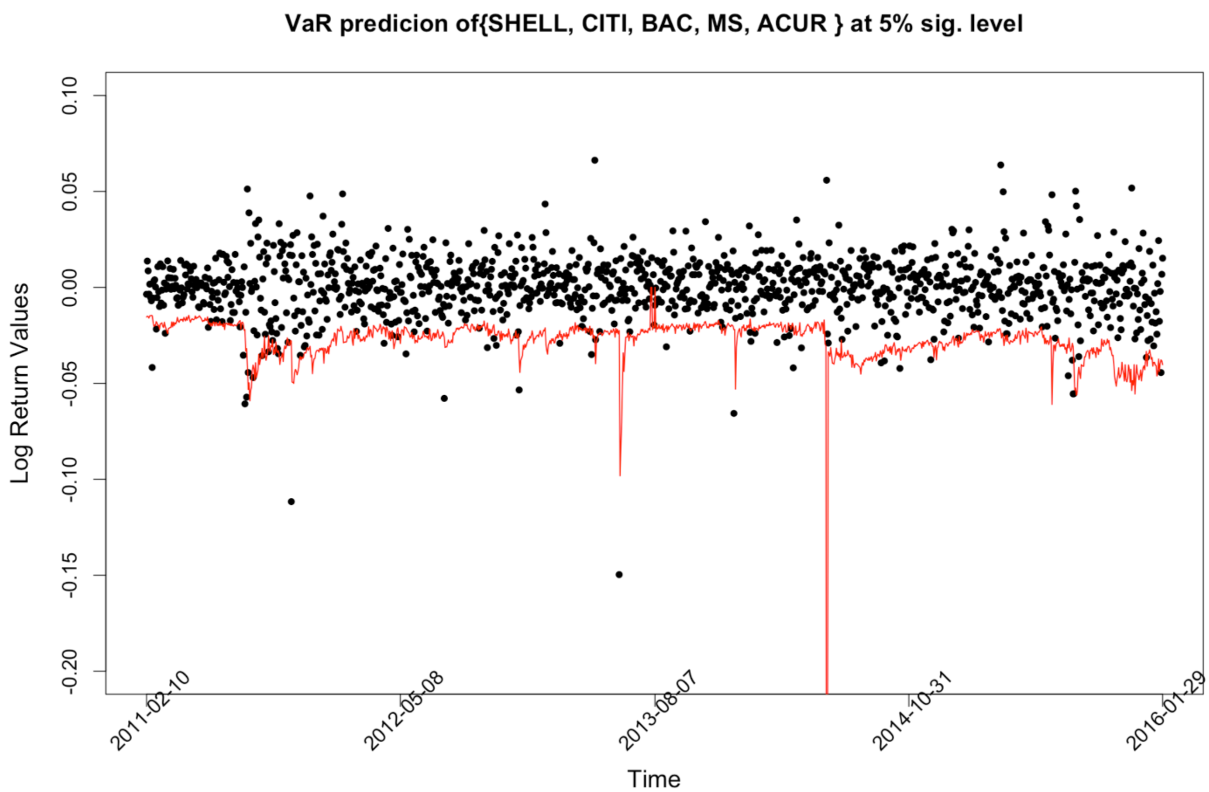

Figure 13 Estimated VaR 5% (solid line), P&L and Exceedance

Figure 13 visualizes the $\widehat{VaR(5\%)}$ (the solid red line) with the points stand for time-point P&L. The data are estimated from 1750-day log-returns with 500-day rolling window.

Portfolio: {AAPL, DAL, EMZN, SPY, NFLX }			
copula	**unconditional**	**conditional**	**Exceedance**
Gumbel	0.022374765	0.04202986	8% (100 out of 1250)
Clayton	NA	NA	NA
Gaussian	0.002615453	0.00519584	6.96% (87 out of 1250)
t copula	0.115282	0.2235687	6% (75 out of 1250)
NAC	0.6524971	0.8679374	5.3% (66 out of 1250)

Table 12 Backtesting results of portfolio {AAPL,DAL,EMZN, SPY, NFLX}, with 5 copula methods applied on 5% significant level

For Portfolio {AAPL, DAL, EMZN, SPY, NFLX}, Gumbel and Gaussian Copulas fail to pass the UC and CC tests. With no exception, the estimated parameters of Clayton Copula are negative again, which results in NA outputs. In this case, Nested Archimedean Copula with Gumbel generator passes the tests with 5.3% or 66 exceedance out of the 1250 days. This improves our belief on the goodness of NAC because it gives us more accurate estimations.

From the discussion in this section, we can summarize the following tips:

- Archimedean Copula, Elliptical Copula and NAC are reliable estimating methods on 5 dimensional VaR prediction, even some assets experience huge oscillations.
- The Copulas' passing rates of backtesting decrease a little compared with that on 3 dimension.
- Conditional Coverage test is more reliable since Independence properties cannot be checked through UC test and pure exceedance numbers and rates.
- Overall, NAC estimates better parameters for VaR prediction on 5 dimensional cases.
- Clayton Copula has calculating restrictions so that it cannot be widely used.

6.4 Extended accuracy checks for higher dimensional cases

In this section, higher dimensional portfolios are designed and our modeling process is executed on these cases. For dimensions larger than 5, Archimedean copulas are able to capture different kinds of tail dependencies, whereas their functional symmetry (also called exchangeability) exerts negative influence on the stability of calculations especially for higher dimensional portfolios. Thus we need to use NAC or HAC to avoid the drawback.

As described in Table 2, three portfolios are designed for 9 dimensional cases. Higher dimensions make the dependency structures more complex, so we can increase the sample size to a 2550-day period with 250-day rolling windows. Also, the margins of each copula are defined as following skewed t distribution. We select 'itau' as the method for copula estimations rather than ML method, because it ensures the stability of calculation at higher dimensions. The following tables represent the backtesting results of our copula models. Unfortunately, regarding the results of UC or CC test, none of them provide solid evidence to show any copula is effective at this dimension.

Portfolio: {AAPL, WMT,NFLX, SPY, C, F, AMZN, MSFT, GS }			
copula	**unconditional**	**conditional**	**Exceedance**
HAC Gumbel	0.2428564	0.4969938	4.5% (103 out of 2300)
Gaussian	0.00	0.00	1.7% (38 out of 2300)
NAC Clayton	NA	NA	NA
HAC Clayton	0.02281092	0.02024528	4% (92 out of 2300)
t Copula	0.001822061	0.003110557	1.8% (40 out of 2300)

Table 13 Backtesting results of 9d portfolio {AAPL,, GS }, with 5 copula methods applied on 5% significant level

Note that in Table 13, NAC with Clayton generator produce NAs because maximum likelihood method generates infinite elements in the covariance matrix. Only HAC Gumbel is reliable regarding UC and CC tests, and there are 4.5% or 103 exceedance out of 2300 P&L predictions.

Portfolio: {ACET, C, ACUR, BAC, MS, DDD, TUC, SH, SPY }			
copula	**unconditional**	**conditional**	**Exceedance**
HAC Gumbel	0.003	0.01	3.7% (86 out of 2300)
Gaussian	0.00	0.00	1.5% (34 out of 2300)
NAC Clayton	NA	NA	NA
HAC Clayton	0.0073	0.025	3.4% (80 out of 2300)
t Copula	0	0	1.5% (35 out of 2300)

Table 14 Backtesting results of 9d portfolio {ACET,..., SPY }, with 5 copula methods applied on 5% sig. level

Portfolio: {WSTL, C, ACUR, BA, MS, DDD, TUC, WMAR, ENZN }			
copula	**unconditional**	**conditional**	**Exceedance**
HAC Gumbel	0.02	0.005	3.7% (86 out of 2300)
Gaussian	0.00	0.00	1.7% (30 out of 2300)
NAC Clayton	NA	NA	NA
HAC Clayton	0.002	0.007	3.7% (85 out of 2300)
t Copula	0	0	1.8% (42 out of 2300)

Table 15 Backtesting results of 9d portfolio {WSTL,..., ENZN }, with 5 copula methods applied on 5% sig. level

For portfolios shown in Tables 14 and 15, none of the Copula method is effective, they all reject the null hypothesis at a level close to zero, let alone the failed execution of NAC with Clayton generator.

Hereby, our concern is to which dimension can Copula methods be efficient in describing the portfolios dependency structures. After repeating the test procedures, only NAC is effective on 6 dimension, while others are not stable. For portfolios with dimensions more than 6, like 7, 8, or 9, it is groundless to stick to use Copula methods for VaR prediction.

Chapter 7

Conclusion

Currently in financial investment and risk management areas, predicting VaR at a given significant level α accurately is always of main concern, especially when considering multi-dimensional portfolios. While it is known that Copulas can solve the dependency structure problem, no literature gives exact conclusion on the accuracy of each Copula method or their applicable dimensions. In this thesis, a modeling idea is presented and the verification of its accuracy is executed. Main conclusions of our thesis can be summarized as a list below:

- Elliptical Copulas and Archimedean Copulas can be good ways for dealing with dependency structures of portfolios reached to 5 dimension, while Hierarchical Archimedean Copulas is still efficient in VaR(α) prediction on 6 dimensional portfolios.
- With the increase of dimensions, the accuracy of Copulas decreases constantly, while NAC and HAC are preferred because of its nested structures, which ensures the stability and validity.
- Skewed t distribution is the most accurate margins for fitting Copulae.
- Clayton Copula is restricted to many calculating restrictions and thus has limitation in practical usage.
- GO-GARCH model is more efficient and stable in higher dimensional cases while other multivariate time series specifications are stable only at low dimensions and univariate cases.

As for the modeling process, we combine time series models with Copula estimation. The reason is that we need accurate data before decomposing the dependency structure into the combination of a copula with marginal functions of the provided residuals. To achieve this purpose, multivariate GARCH is proven to be effective ways to fit the raw residuals, covariance matrix, and estimated means we want. The whole modeling procedures are formulated based on the Value-at-Risk equation, or, Log-returns equation: $X_{i,t} = \mu_{i,t} + \sigma_{i,t}\epsilon_{i,t}$, $i = 1,2, \dots, n$.

In addition, graphics are important tools to help us explore the properties and make inference in our study. The selections of simulating multivariate distributions, of Copulas' margins and the goodness of Copula methods, for example, are all rely on plots.
In this study, the optimizing methods are discussed to be maximum likelihood method at first, whereas with the increase of dimensions, 'ml' are no longer reliable and we turn to 'itau', further studies can be focused on the impact of optimizing methods as well.

Bibliography

[1] Aas, K., & Berg, D. (2010). Modeling Dependence Between Financial Returns Using Pair-Copula Constructions. *Vine Copula Handbook Dependence Modeling,* 305-328.

[2] Balcilar, M., & Ozdemir, Z. A. (2012). The export-output growth nexus in Japan: A bootstrap rolling window approach. *Empir Econ Empirical Economics, 44*(2), 639-660.

[3] Bee, M. (2011). Adaptive Importance Sampling for simulating copula-based distributions. *Insurance: Mathematics and Economics, 48*(2), 237-245.

[4] Chi, X., & Goodwin, B. K. (2012, August). *A High-Dimensional, Multivariate Copula Approach to Modeling Multivariate Agricultural Price Relationships and Tail Dependencies* [Scholarly project].

[5] Christoffersen, P. F. (1998). Evaluating Interval Forecasts. *International Economic Review, 39*(4), 841.

[6] Christoffersen, P. F., & Pelletier, D. (2004). Backtesting Value-at-Risk: A Duration-Based Approach. *SSRN Electronic Journal SSRN Journal, 2*(1), 84-108.

[7] Danciulescu, C. (n.d.). Backtesting Value-at-Risk Models: A Multivariate Approach. *SSRN Electronic Journal SSRN Journal*.

[8] Engle, R. F. (n.d.). Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models. *SSRN Electronic Journal SSRN Journal*.

[9] Engle, R. F. (2002). Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models. *SSRN Electronic Journal SSRN Journal, 15*(4), 157-168.

[10] Franke, J., Härdle, W., & Hafner, C. (2008). *Statistics of financial markets: An introduction* (Third ed.). Springer.

[11] Hering, C., & Hofert, M. (2015). Goodness-of-fit Tests for Archimedean Copulas in High Dimensions. *Innovations in Quantitative Risk Management Springer Proceedings in Mathematics & Statistics,* 357-373.

[12] Hofert, M., Mächler, M., & Mcneil, A. J. (2012). Likelihood inference for Archimedean copulas in high dimensions under known margins. *Journal of Multivariate Analysis, 110*, 133-150.

[13] Hofert, M., Mächler, M., & Mcneil, A. J. (2012). Likelihood inference for Archimedean copulas in high dimensions under known margins. *Journal of Multivariate Analysis, 110*, 133-150.

[14] Hu, W., & Kercheval, A. N. (n.d.). The skewed t. *Econometrics and Risk Management Advances in Econometrics,* 55-83.

[15] Jäschke, S. (2014). Estimation of risk measures in energy portfolios using modern copula techniques. *Computational Statistics & Data Analysis, 76*, 359-376.

[16] Mendenhall, W., & Scheaffer, R. L. (1973). *Mathematical statistics with applications*. North Scituate, MA: Duxbury Press.

[17] Okhrin, O., & Ristig, A. (2014). Hierarchical Archimedean Copulae: The HAC Package. *Journal of Statistical Software J. Stat. Soft., 58*(4).

[18] Orskaug, E. (2009). *Multivariate DCC-GARCH Model -With Various Error Distributions* (Unpublished master's thesis). Norwegian University of Science and Technology, Department of Mathematical Sciences.

[19] Patton, A. (2013). Copula Methods for Forecasting Multivariate Time Series. *Handbook of Economic Forecasting, 2*, 899-960.

[20] Trivedi, P. K., & Zimmer, D. M. (2006). Copula Modeling: An Introduction for Practitioners. *Foundations and Trends® in Econometrics FNT in Econometrics, 1*(1), 1-111.

[21] Weide, R. V. (2002). GO-GARCH: A multivariate generalized orthogonal GARCH model. *J. Appl. Econ. Journal of Applied Econometrics, 17*(5), 549-564.

[22] Yamai, Y., & Yoshiba, T. (2005). Value-at-risk versus expected shortfall: A practical perspective. *Journal of Banking & Finance, 29*(4), 997-1015.

[23] Yan, J. (2007). Enjoy the Joy of Copulas: With a Package copula. *Journal of Statistical Software J. Stat. Soft., 21*(4).

[24] Yoshiba, T. (2015). Risk Aggregation with Copula for Banking Industry. *Mathematics for Industry Applications Practical Conceptualization Mathematics = Fruitful Innovation,* 247-259.