Risk Forecasting with Vine Copulas

Mathematics-Economics

Master's Thesis, 4.111a

Nikolaj Haslev Mortensen

Viktor Boel Borchersen

AALBORG UNIVERSITY

STUDENT REPORT

Department of Mathematical Sciences
Mathematics-Economics
Skjernvej 4A
9220 Aalborg East
http://math.aau.dk

Title:

Risk Forecasting with Vine Copulas

Project:

Master's Thesis

Period:

Spring Semester 2024

Group:

4.111a

Participants:

Nikolaj Haslev Mortensen
Viktor Boel Borchersen

Supervisor:

Orimar Sauri

Date of Completion:

03/06/2024

Pagecount:

105 + Frontpage

Abstract:

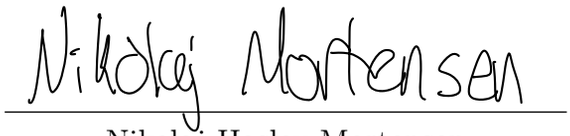
This project seeks to investigate the capabilities of vine copulas, where the marginals are modeled with ARMA-GARCH models, to forecast the risk measures (Conditional) Value at Risk($(C)VaR$). To do so, theory about copulas, vine copulas, and estimation and selection is presented. Afterwards, theory about $(C)VaR$ and backtesting of these measures is shown. Data from three sectors from the S&P-500 index is chosen namely financial, healthcare, and technology. The analysis starts with in-sample estimation of both the marginals and the vine copula, as well as goodness of fit testing on the data. The fits of the marginals were tested with an Anderson-Darling test where the null hypothesis was not rejected for any of the marginals. The R-vine copula outperformed the Gaussian- and Student's t-copula. Therefore, the analysis proceeded to an out-of-sample forecast of $(C)VaR$. The backtests of the forecasts of $(C)VaR$ indicated that the model was not able to forecast $(C)VaR$ well. The inadequate forecasts of the $(C)VaR$, were due to the high dimensional and the changing tendencies in the data over time. However, a mean-$CVaR$ portfolio allocation problem was performed for each sector anyway. The mean-$CVaR$ strategies all outperformed equally-weighted portfolios in terms of return but also had higher levels of $(C)VaR$.

Preface

This Master's Thesis is written during the spring semester of 2024 as a part of the Mathematics-Economics program at the Department of Mathematical Sciences at Aalborg University.

This report is composed of numbered chapters with corresponding sections and subsections. Citations and external references are made using the Vancouver method. References are made at the start of a chapter and further references are made in each (sub)section if applicable. Should the reference apply only to the preceding line of text, the reference is placed before the period. Should the reference apply to the previous section of text, the reference is placed after the period. Dates are denoted as day/month/year. The `R`-code used for the project and the supporting material for the application is attached in a separate zip-file and can be received from the authors by request. Lastly, we would like to extend our appreciation to our supervisor Orimar Sauri.

Signatures



Nikolaj Haslev Mortensen
nmort19@student.aau.dk

Viktor Boel Borchersen
vborch19@student.aau.dk

Nomenclature

Functions

$[x]^+$	Gives zero if $x \leq 0$ and x if $x > 0$, where $x \in \mathbb{R}$.
$\overset{a}{\sim}$	Asymptotically distributed.
$\overset{d}{\rightarrow}$	Convergence in distribution.
$Dom(\cdot)$	Domain of $(\cdot)$
$Ran(\cdot)$	Range of $(\cdot)$

Sets

$\bar{\mathbb{R}}$	The extended real line, $[-\infty, \infty]$.
$\bar{\mathbb{R}}^d$	The extended d-dimensional plane, $\bar{\mathbb{R}} \times \bar{\mathbb{R}} \times \cdots \times \bar{\mathbb{R}}$.
$A \triangle B$	The symmetric difference of two sets, i.e., $(A \backslash B) \cup (B \backslash A)$

Contents

1 Introduction

How does one measure financial risk? Throughout the years, a lot of different measures have been used to try and describe the uncertainty of the financial markets. Modeling spread using variance or standard deviation is a very popular choice, and has been around for a long time, as Karl Pearson first used the term *standard deviation* back in 1894 [1]. Another way of measuring risk, is by looking at potential losses, which has been done by the *Value at Risk* (VaR) since the late 1980s. It was further popularized when the Basel II Accord was published in 2004, as VaR was the preferred approach for measuring market risk [2]. The VaR at a given level, $\alpha \in (0, 1)$, tells one that there is an α probability of losing VaR or more on an investment. The drawback of VaR is that it does not give any information on potential losses beyond the threshold α. Therefore, the so-called Conditional Value at Risk ($CVaR$) was later introduced. The idea behind $CVaR$ is to model the *expected shortfall* of an investment, which is achieved by averaging the losses above the VaR level. Besides yielding more information than VaR, $CVaR$ is also what is called a coherent risk measure. A coherent risk has the property that diversification lowers risk.

There are three main ways of estimating VaR, and hence also $CVaR$: The historical method, the variance-covariance method, and the Monte Carlo method. The historical method relies on estimating the VaR based on historical observations. The variance-covariance method relies on fitting a parametric distribution to the return data, and then forecasting using time series. The Monte Carlo method relies on simulating from a multivariate distribution, and then estimating the VaR based on the quantiles of the Monte Carlo simulations. This method relies on finding a good multivariate distribution for the returns, as the method is then a valid approach.

There are multiple ways to estimate a multivariate distribution for returns. One way is to try and fit a multivariate copula to the data. However, as shown by studies of Sahamkhadam and Stephan, so-called *vine copula*-based models outperform simple multivariate copulas [3]. Vines are sequences of trees where each tree is constructed using bivariate copulas as building blocks. Compared to multivariate copulas, vine copulas yield the opportunity of choosing different dependence structure between assets, and not only one single copula family. To be able to estimate a vine-copula model, it is necessary to specify the marginal distribution of each asset return. There are many ways to model the marginal distributions. One way is to use a time-series model, like an ARMA-GARCH model.

1.1 Problem Statement

How can vine copulas be used to model the dependence structure of a high dimensional portfolio? Furthermore, how can a vine copula with ARMA-GARCH marginals be used to forecast Value at Risk and Conditional Value at Risk?

2 | Vine Copulas

This chapter is based on [4] and [5].

Modeling the dependence structure of a portfolio consisting of multiple assets requires a flexible multivariate distribution. Frequently used multivariate distributions such as multivariate Gaussian- or Student's t-distributions are restrictive in the sense that they require that every asset is modeled using the same distribution. A more flexible class of multivariate distributions is copulas. With copulas, the marginal distribution of each asset can be selected separately from the joint distribution, thereby giving a more flexible approach. However, as the number of assets in the portfolio increases, the dependence structure also becomes more complicated to model. When this is the case, multivariate copulas, such as elliptical- or Archimedean copulas, are typically too restrictive. Another approach is to use independent bivariate copulas, called pair copulas, as building blocks to create a multivariate distribution called an R-vine distribution. The flexibility of this approach is that each pair can be modeled using a different bivariate copula, which in theory should capture the dependence structure better compared to using a multivariate copula.

2.1 Copulas

This section is based on [6] and [7].

Copulas are multivariate distribution functions with uniform marginals. Copulas are used to model the dependence structure separately from the marginals. By separating the modeling problem of the dependence structure and marginals, copulas make it possible to construct a joint distribution from any set of univariate marginals.

Definition 2.1. **d-Dimensional Copula**
A function C is a d-dimensional copula if it has the following properties:

1. $Dom(C) = [0,1]^d$.
2. C is grounded, i.e., for every $u \in [0,1]^d$

$$C(u) = 0, \tag{2.1}$$

   if at least one coordinate of u is 0.

3. C is d-increasing, i.e., for every $a, b \in [0,1]^d$ such that $a \leq b$

$$V_C([a,b]) \geq 0, \tag{2.2}$$

where $V_C([a,b])$ is the C-volume of the d-box $B = [a,b]$ given by

$$V_C(B) = \sum sign(c)C(c), \tag{2.3}$$

where c is the vertices of the d-box B given by $c = (c_1, c_2, \ldots, c_d)$ and c_k is equal to either a_k or b_k. The sum is taken over every vertex of B and $sign(c)$ is given by

$$sign(c) := \begin{cases} 1, & \text{if } c \text{ has an even number of lower bounds,} \\ -1, & \text{if } c \text{ has an odd number of lower bounds.} \end{cases} \tag{2.4}$$

4. C has one-dimensional marginals C_k such that $C_k(u) = C(1, \ldots, 1, u, 1, \ldots, 1) = u$, for all $u \in [0,1]$.

Due to the properties of a d-dimensional copula being grounded and d-increasing, Lemma A.1 shows that it is also nondecreasing. Therefore, it follows from Definition A.2 that a d-dimensional copula is a d-dimensional distribution function on $[0,1]^d$ with uniform marginals.

Another important property of d-dimensional copulas is that they can be used to construct a multivariate distribution from any set of univariate marginals, which was first stated by Sklar.

Theorem 2.2. Sklar's Theorem
Let F be a d-dimensional distribution function with univariate marginal distribution functions $F_1, F_2, \ldots, F_d$. Then, for all $(x_1, x_2, \ldots, x_d) \in \bar{\mathbb{R}}^d$, there exists a d-dimensional copula C such that

$$F(x_1, x_2, \ldots, x_d) = C\left(F_1(x_1), F_2(x_2), \ldots, F_d(x_d)\right). \tag{2.5}$$

The copula C is unique if $F_1, F_2, \ldots, F_d$ are continuous. Otherwise, C is unique on $Ran(F_1) \times Ran(F_2) \times \cdots \times Ran(F_d)$. Conversely, if C is a d-dimensional copula and $F_1, F_2, \ldots, F_d$ are univariate marginal distribution functions, then F is a d-dimensional distribution function with univariate marginal distributions $F_1, F_2, \ldots, F_d$.

The proof is omitted. However, a proof can be found in [8].

Theorem 2.2 states that the information contained by the multivariate distribution function can be extracted by modeling the dependence structure between the variables and the marginals separately. A copula should therefore be chosen such that it models the observed dependence structure between the random variables. A way to construct a copula is to

take the quasi-inverse[1] of the marginals in (2.5) such that

$$C(u_1, \ldots, u_d) = F\left(F_1^{(-1)}(u_1), \ldots, F_d^{(-1)}(u_d)\right), \tag{2.6}$$

for all $(u_1, \ldots, u_d) \in [0,1]^d$. To construct a copula using (2.6), it is required that the joint distribution is known. In general, this is not the case when working with data. Another approach is therefore needed. This approach is to use already-known families of copulas and select a specific copula based on an information criterion such as the Akaike Information Criterion(AIC). In this project, two families of parametric copulas will be utilized namely elliptical copulas and Archimedean copulas. These two families of parametric copulas are explored in Section A.2. However, the characteristics of each copula are summarized in the following table.

Table 2.1: Characteristics of Elliptical Copulas and Archimedean Copulas

	Tail Structure	Tail Dependence
Elliptical		
Gaussian	Symmetric	None
Student's t	Symmetric	Both
Archimedean		
Gumbel	Asymmetric	Upper
Survival Gumbel	Asymmetric	Lower
Rotated Gumbel (90 and 270 degrees)	Asymmetric	None
Frank	Asymmetric	None

2.1.1 A Pair Copula Decomposition of a Multivariate Distribution

This subsection is based on [9].

In this subsection, it will be shown how a multivariate distribution can be constructed using bivariate distributions as building blocks. It is assumed that all joint-, marginal-, and conditional distributions are absolutely continuous with corresponding densities such that the copulas are unique. Given a vector of random variables $X = (X_1, \ldots, X_d)$, with joint density function $f(x_1, \ldots, x_d)$, the joint density can be factorized as

$$f(x_1, \ldots, x_d) = f_d(x_d) \cdot f(x_{d-1}|x_d) \cdot f(x_{d-2}|x_{d-1}, x_d) \cdots f(x_1|x_2, \ldots, x_d). \tag{2.7}$$

The joint density function can also be derived by using the chain-rule on (2.5)

$$f(x_1, \ldots, x_d) = c_{1,\ldots,d}\left(F_1(x_1), \ldots, F_d(x_d)\right) f_1(x_1) \cdots f_d(x_d), \tag{2.8}$$

[1]The definition of a quasi-inverse is found in Definition A.3.

where the density of a copula is given by

$$c_{1,\dots,d}\left(F_1(x_1),\dots,F_d(x_d)\right) = \frac{\partial C\left(F_1(x_1),\dots,F_d(x_d)\right)}{\partial F_1(x_1)\partial F_2(x_2)\cdots\partial F_d(x_d)}. \tag{2.9}$$

Remark that the sub-indices of the copulas, copula densities, and marginals are noted with the same sub-indices that their input has instead of their actual input, to simplify notation. For example, $c_{x_1,\dots,x_k}$ is denoted as $c_{1,\dots,k}$. The conditional density in the two-dimensional case is therefore given by

$$f(x_1|x_2) = c_{1,2}\left(F_1(x_1), F_2(x_2)\right) f_1(x_1). \tag{2.10}$$

The decomposition of the joint density is not unique. The non-uniqueness of the decompositions will be shown in the three-dimensional case. For $f(x_1|x_2,x_3)$ there exists the following decompositions

$$\begin{aligned} f(x_1|x_2,x_3) &= c_{1,2|3}\left(F(x_1|x_3), F(x_2|x_3)\right)\cdot f(x_1|x_3) \\ &= c_{1,2|3}\left(F(x_1|x_3), F(x_2|x_3)\right)\cdot c_{1,3}\left(F_1(x_1), F_3(x_3)\right)\cdot f_1(x_1), \qquad (2.11) \\ f(x_1|x_2,x_3) &= c_{1,3|2}\left(F(x_1|x_2), F(x_3|x_2)\right)\cdot f(x_1|x_2) \\ &= c_{1,3|2}\left(F(x_1|x_2), F(x_3|x_2)\right)\cdot c_{1,2}\left(F_1(x_1), F_2(x_2)\right)\cdot f_1(x_1). \qquad (2.12) \end{aligned}$$

where the conditional density of the copula in (2.11) is calculated as

$$c_{1,2|3}\left(F(x_1|x_3), F(x_2|x_3)\right) = \frac{\partial}{\partial x_1 \partial x_2} C_{1,2|3}\left(F(x_1|x_3), F(x_2|x_3)\right), \tag{2.13}$$

and the conditional density of the copula in (2.12) is calculated the same way. In general, each term on the right side of (2.7) can be decomposed into a bivariate copula times the conditional marginal

$$f(x_i|x_v) = c_{i,v_j|v_{-j}}\left(F(x_i|x_{v_{-j}}), F(x_{v_j}|x_{v_{-j}})\right)\cdot f(x_i|x_{v_{-j}}), \tag{2.14}$$

where the conditional density of a copula is given by

$$c_{i,v_j|v_{-j}}\left(F(x_i|x_{v_{-j}}), F(x_{v_j}|x_{v_{-j}})\right) = \frac{\partial C_{i,v_j|v_{-j}}\left(F(x_i|x_{v_{-j}}), F(x_{v_j}|x_{v_{-j}})\right)}{\partial x_i \partial x_{v_j}}, \tag{2.15}$$

and $i \in \{1,\dots,d\}$, $v \subseteq \{1,\dots,d\} \setminus i$, v_j is an arbitrary element of v and v_{-j} is v without the $j'th$ element. Furthermore, the conditional distribution function is given by

$$F(x_i|x_v) = \frac{\partial}{\partial F(x_{v_j}|x_{v_{-j}})} C_{i,v_j|v_{-j}}\left(F(x_i|x_{v_{-j}}), F(x_{v_j}|x_{v_{-j}})\right), \tag{2.16}$$

where $C_{i,v_j|v_{-j}}$ is a bivariate copula function. If v is univariate, then

$$F(x_i|x_v) = \frac{\partial}{\partial F(x_v)} C_{i,v}(F(x_i), F(x_v)). \tag{2.17}$$

To obtain $F(x_i|x_{v_{-j}})$ and $F(x_{v_j}|x_{v_{-j}})$ a recursive method is used. This method will be introduced later in this project. Every valid factorization of the joint distribution can be represented in a graphical model. This model is called a regular vine tree structure.

2.2 Vines

Vines are graphical models used to store the different pair copula decompositions of a multivariate distribution. To store the different dependence structures, vines are a sequence of connected trees where each edge in the prior tree becomes a node in the next, thereby characterizing the decomposition of a multivariate distribution.

Definition 2.3. *R*-Vine Tree Sequence
The set of trees $\mathcal{V} = (T_1, \dots, T_{d-1})$ is a regular (R-) vine tree sequence on d elements if

1. T_1 is a tree with nodes $N_1 = \{1, \dots, d\}$ and edges E_1.
2. For $j \geq 2$, T_j is a tree with nodes $N_j = E_{j-1}$ and edges E_j.
3. For $j \geq 2$ and $\{a, b\} \in E_j$ it must hold that $\#(a \cap b) = 1$.

The last condition in Definition 2.3, is called the proximity condition, which ensures that if an edge is connecting a and b in $T_j, j \geq 2$, then a and b must have a common node in T_{j-1}. The number of different R-vine tree structures is large, as there exists $(d!/2) \cdot 2^{\binom{d-2}{2}}$ different possibilities in the d-dimensional case. As a consequence of the many different ways an R-vine tree structure can be constructed, a way to store each unique sequence of trees is needed. To feasibly store an R-vine tree sequence, making inference possible, a $d \times d$-dimensional upper triangular matrix is constructed.

Definition 2.4. *R*-Vine Matrix
Let M be an upper triangular matrix with entries $m_{i,j}$ for $i \leq j$, where $1 \leq m_{i,j} \leq d$. Such a matrix is called an R-vine matrix, if

1. $\{m_{1,i}, \dots, m_{i,i}\} \subset \{m_{1,j}, \dots, m_{j,j}\}$ for $1 \leq i < j \leq d$.
2. $m_{i,i} \notin \{m_{1,i-1}, \dots, m_{i-1,i-1}\}$.
3. For $i = 3, \dots, d$ and $k = 1, \dots, i-1$ there exist (j, ℓ) with $j < i$ and $\ell < j$ so it holds that

$$\left\{m_{k,i}, \{m_{1,i}, \dots, m_{k-1,i}\}\right\} = \left\{m_{j,j}, \{m_{1,j}, \dots, m_{\ell,j}\}\right\} \text{ or} \tag{2.18}$$

$$\left\{m_{k,i}, \{m_{1,i}, \dots, m_{k-1,i}\}\right\} = \left\{m_{\ell,j}, \{m_{1,j}, \dots, m_{\ell-1,j}, m_{j,j}\}\right\}. \tag{2.19}$$

The last condition in Definition 2.4 corresponds to the proximity condition in Definition 2.3, and hence there is a bijection between R-vine tree structures and R-vine matrices. It is therefore possible to store the R-vine tree structures in R-vine matrices. An R-vine matrix makes it possible to store and visualize high dimensions of R-vine tree structures, which is effective in analysis of the dependence structure and the practical implementation of the

R-vines.

In order to further analyze the properties of different vine structures, some sets regarding the edges are of interest.

Definition 2.5. Complete Union, Conditioning and Conditioned Sets of an Edge [5, p.5]
The complete union of an edge, $e_i \in E_i$, is the set

$$U_{e_i} = \{n \in N_1 | \exists e_j \in E_j, j = 1, \dots, i-1, \text{ with } n \in e_1 \in e_2 \in \cdots \in e_{i-1} \in e_i\} \subset N_1. \tag{2.20}$$

For $e_i = \{a, b\} \in E_i$, where $a, b \in E_{i-1}$ for $i = 1, \dots, d-1$, the conditioning set of e_i is

$$D_{e_i} = U_a \cap U_b, \tag{2.21}$$

and the conditioned sets of e_i are

$$\mathcal{C}_{a_{e_i}} = U_a \backslash D_{e_i}, \quad \mathcal{C}_{b_{e_i}} = U_b \backslash D_{e_i}, \text{ and } \quad \mathcal{C}_{e_i} = \mathcal{C}_{a_{e_i}} \cup \mathcal{C}_{b_{e_i}} = U_a \triangle U_b. \tag{2.22}$$

From the sets in Definition 2.5 an edge can be written as $e = (\mathcal{C}_{a_e}, \mathcal{C}_{b_e} | D_e)$, where the conditioned set is prior to "|" and the conditioning set is shown after "|".

Example 2.6.
The purpose of this example is to illustrate an R-vine tree structure, R-vine matrices and the sets defined in Definition 2.5. An example of an R-vine tree sequence for $d = 6$, is illustrated in the following graph.

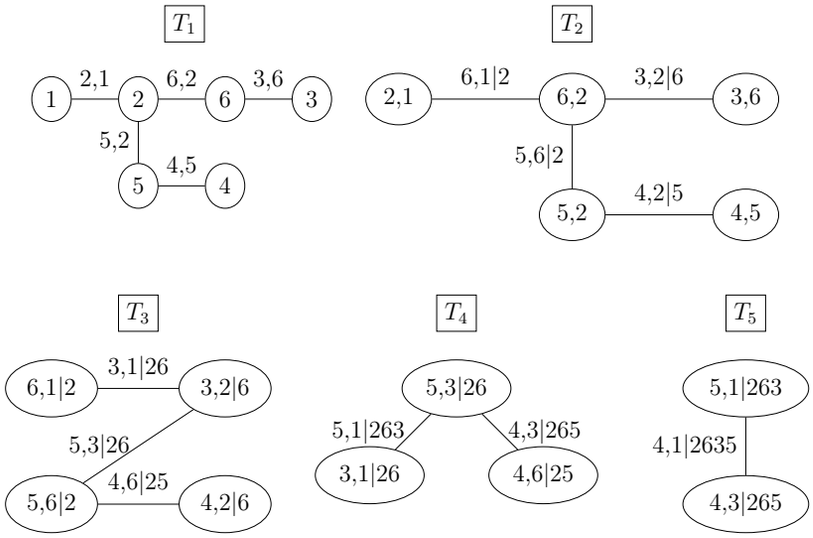


Figure 2.1: Example of an R-vine tree sequence for $d = 6$.

In Figure 2.1, the complete union for the edge $\{(2,1),(6,2)\}$ from T_2 is given as $\{1,2,6\}$. The conditioning set for the edge $\{(6,1|2),(3,2|6)\}$ from T_3 is given as $\{2,6\}$. The conditioned set for the edge $\{(6,1|2),(3,2|6)\}$ from T_3 is given as $\{1,3\}$.

The R-vine tree structure in Figure 2.1 is not unique in the way that, given T_1, it would be possible to construct other T_2 trees, by changing the conditioned and conditioning sets. For example, combining 5 and 1 by 2 in T_1, would give the edge $5,1|2$ in T_2, which is not the case in this example, and hence would have specified another R-vine tree structure.

When constructing an R-vine matrix corresponding to an R-vine tree sequence, it is possible to choose between two elements of the conditioned set. Therefore, the construction of an R-vine matrix is not unique. Examples of associated R-vine matrices for Figure 2.1 are given by

$$M_1 = \begin{bmatrix} 1 & 1 & 2 & 6 & 2 & 5 \\ & 2 & 1 & 2 & 6 & 2 \\ & & 6 & 1 & 3 & 6 \\ & & & 3 & 1 & 3 \\ & & & & 5 & 1 \\ & & & & & 4 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 2 & 2 & 5 & 2 & 6 & 2 \\ & 5 & 2 & 5 & 2 & 6 \\ & & 4 & 4 & 5 & 3 \\ & & & 6 & 4 & 5 \\ & & & & 3 & 4 \\ & & & & & 1 \end{bmatrix}. \tag{2.23}$$

2.3 *R*-Vine Copulas

It is possible to utilize the graphical structure of R-vines in the specification of multivariate distributions. This is achieved by specifying a bivariate copula, C_e, for each edge in the R-vine tree sequence, and hence adding a stochastic component to the structure.

Definition 2.7. *R*-vine Copula Specification
Let $\mathbf{F} = (F_1, \ldots, F_d)$ be a vector of continuous marginal invertible distribution functions, $\mathcal{V}$ be a d-dimensional R-vine tree structure, and $\mathcal{B} = \{C_e | e \in E_i \text{ for } i = 1, \ldots, d-1\}$ be a set of pair copulas. Then the triplet $(\mathbf{F}, \mathcal{V}, \mathcal{B})$ is called an R-vine copula specification.

Given an R-vine copula specification, it is possible to construct an R-vine distribution. Let F be a joint distribution function of a random vector $X = (X_1, \ldots, X_d)$. F is then an R-vine distribution if $C_e = C_{\mathcal{C}_{a_e}, \mathcal{C}_{b_e}|D_e}$ is the pair copula of $X_{\mathcal{C}_{a_e}}$ and $X_{\mathcal{C}_{b_e}}$ given $X_{D_e} = \{X_i | i \in D_e\}$ for all $e = \{a, b\} \in E_i$ for $i = 1, \ldots, d-1$. In this construction, it is assumed that the conditional copula, C_e, is independent of the conditioning variable X_{D_e}, which is called the simplifying assumption. The conditioning values still affect the R-vine distribution through

the dependence of the conditioning variables on the conditional distribution functions. It is also possible to construct a non-simplified version where the pair copulas would depend on the random vector X [4, p.207]. However, in this project, the focus will be on the simplified version.

In order to specify the density for an R-vine distribution, let X_j have marginal distribution F_j and let $c_{\mathcal{C}_{a_e},\mathcal{C}_{b_e}|D_e}$ denote the copula density of C_e.

Theorem 2.8. *R*-Vine Copula Density
Let $(\mathbf{F}, \mathcal{V}, \mathcal{B})$ be a d-dimensional R-vine copula specification. Then the density for the unique distribution F, which realizes the R-vine copula specification, is given as

$$f(x) = \prod_{j=1}^{d} f_j(x_j) \prod_{i=1}^{d-1} \prod_{e \in E_i} c_{\mathcal{C}_{a_e},\mathcal{C}_{b_e}|D_e} \left(F_{\mathcal{C}_{a_e}|D_e}(x_{\mathcal{C}_{a_e}}|x_{D_e}), F_{\mathcal{C}_{b_e}|D_e}(x_{\mathcal{C}_{b_e}}|x_{D_e}) \right), \tag{2.24}$$

where $x = (x_1, \ldots, x_d)$, $e = \{a, b\}$, $x_{D_e} = \{x_i | i \in D_e\}$, and f_j denotes the density of F_j for $j = 1, \ldots, d$.

The proof is omitted. However, a proof can be found in [10, p. 259-260].

An example of the use of the factorization (2.24), which corresponds with Example 2.6, can be found in Example B.1.

From Theorem 2.8 the distribution function of $X_{\mathcal{C}_{a_e}}$ and $X_{\mathcal{C}_{b_e}}$ given $X_{D_e} = x_{D_e}$, is given by

$$F_{\mathcal{C}_{a_e},\mathcal{C}_{b_e}|D_e}\left(x_{\mathcal{C}_{a_e}}, x_{\mathcal{C}_{b_e}}|x_{D_e}\right) = C_e\left(F_{\mathcal{C}_{a_e}}(x_{\mathcal{C}_{a_e}}|x_{D_e}), F_{\mathcal{C}_{b_e}}(x_{\mathcal{C}_{b_e}}|x_{D_e})\right). \tag{2.25}$$

Therefore, given a triplet $(\mathbf{F}, \mathcal{V}, \mathcal{B})$, Theorem 2.8 states that it is possible to specify the associated R-vine distribution and density, and hence also the associated R-vine copula.

Definition 2.9. *R*-Vine Copula
An R-vine copula is an R-vine distribution, with uniform marginal distributions on $[0, 1]$.

The conditional distributions $F_{\mathcal{C}_{a_e}}(x_{\mathcal{C}_{a_e}}|x_{D_e})$ and $F_{\mathcal{C}_{b_e}}(x_{\mathcal{C}_{b_e}}|x_{D_e})$ in (2.25) are found recursively using (2.16) . Firstly, (2.16) is rewritten using the notation presented in Definition 2.5. Therefore, let $e = \{a, b\} = (\mathcal{C}_{a_e}, \mathcal{C}_{b_e}|D_e) \in E_i$ and $e' \in E_{i-1}$ where $\mathcal{C}_{a_e} = \mathcal{C}_{a_{e'}}$, $\mathcal{C}_{b_{e'}} \in D_e$ and $D_{e'} = D_e \backslash \mathcal{C}_{b_{e'}}$. Then the conditional distribution is given as

$$\begin{aligned} F_{\mathcal{C}_{a_e}|D_e}(x_{\mathcal{C}_{a_e}}|x_{D_e}) &= \frac{\partial C_{\mathcal{C}_{e'}|D_{e'}}\left(F_{\mathcal{C}_{a_{e'}}|D_{e'}}(x_{\mathcal{C}_{a_{e'}}}|x_{D_{e'}}), F_{\mathcal{C}_{b_{e'}}|D_{e'}}(x_{\mathcal{C}_{b_{e'}}}|x_{D_{e'}})\right)}{\partial F_{\mathcal{C}_{b_{e'}}|D_{e'}}(x_{\mathcal{C}_{b_{e'}}}|x_{D_{e'}})} \\ &=: h\left(F_{\mathcal{C}_{a_{e'}}|D_{e'}}(x_{\mathcal{C}_{a_{e'}}}|x_{D_{e'}}), F_{\mathcal{C}_{b_{e'}}|D_{e'}}(x_{\mathcal{C}_{b_{e'}}}|x_{D_{e'}})\right). \end{aligned} \tag{2.26}$$

$F_{\mathcal{C}_{b_e}|D_e}(x_{\mathcal{C}_{b_e}}|x_{D_e})$ is constructed in a similar way. To construct a method to calculate (2.26) recursively, the conditional distribution is rewritten using the notation of an R-vine matrix:

$$\mathcal{C}_{a_{e'}} = \mathcal{C}_{a_e} = \{m_{k,k}\}, \tag{2.27}$$
$$\mathcal{C}_{b_{e'}} = \{m_{i-1,k}\}, \tag{2.28}$$
$$\mathcal{C}_{e'} = \mathcal{C}_{a_{e'}} \cup \mathcal{C}_{b_{e'}} = \{m_{k,k}, m_{i-1,k}\}, \tag{2.29}$$
$$D_e = \{m_{1,k}, \dots, m_{i-1,k}\}, \tag{2.30}$$

for $k = 2, \dots, d$ and $i = 1, \dots, k-1$. In the case $i = 1$ the conditioning set is empty. Additionally, the pair copula families and corresponding parameters are stored in the matrices $B = (b_{i,k})_{k=2,\dots,d,i=1,\dots,k-1}$ and $\Theta = (\theta_{i,k})_{k=2,\dots,d,i=1,\dots,k-1}$, respectively. Furthermore, the entry labels of the R-vine matrix are relabeled, such that the diagonal labels are in ascending order from 1 to d, to simplify notation. The diagonal is therefore given by $m_{k,k} = k$ for $k = 1, \dots, d$. M_1 in Example 2.6 would be relabeled as

$$M_1 = \begin{bmatrix} 1 & 1 & 2 & 3 & 2 & 5 \\ & 2 & 1 & 2 & 3 & 2 \\ & & 3 & 1 & 4 & 3 \\ & & & 4 & 1 & 4 \\ & & & & 5 & 1 \\ & & & & & 6 \end{bmatrix}. \tag{2.31}$$

Let the maximum matrix of M be denoted $\mathbb{M} = (\mathbf{m}_{i,k})_{k=1,\dots,d,i=1,\dots,k}$, where $\mathbf{m}_{i,k} = \max\{m_{1,k}, \dots, m_{i,k}\}$ for all $k = 1, \dots, d$ and $i = 1, \dots, k$. The element $\mathbf{m}_{i,k}$ is, therefore, the maximum of the k'th column of M from the top down to the i'th element. For $i = k$ the maximum $\mathbf{m}_{k,k}$ is equal to the element in diagonal, i.e., $\mathbf{m}_{k,k} = m_{k,k} = k$ for $k = 1, \dots, d$.

The conditional distributions in (2.26) is rewritten using the notation for an R-vine copula, and are given by

$$g_{i,k}^{(1)} = F_{m_{k,k}|\{m_{1,k},\dots,m_{i-1,k}\}}(x_{m_{k,k}}|x_{m_{1,k}}, \dots, x_{m_{i-1,k}}), \tag{2.32}$$
$$g_{i,k}^{(2)} = F_{m_{i,k}|\{m_{1,k},\dots,m_{i-1,k}\}}(x_{m_{1,k}}|x_{m_{1,k}}, \dots, x_{m_{i-1,k}}), \tag{2.33}$$

for $k = 2, \dots, d$ and $i = 1, \dots, k-1$. The conditional distributions are therefore iterating over the d columns of M and from the top down to the element over the diagonal in each column. This way of iterating ensures that every edge of the R-vine, with the corresponding copula type and parameters, is visited once. Both $g_{i,k}^{(1)}$ and $g_{i,k}^{(2)}$ can be given as either the first or second argument of the function $h(\cdot, \cdot; b_{i,k}, \theta_{i,k})$. Therefore, to store the positioning of the conditional distributions, two matrices V^{direct} and $V^{indirect}$ are used. The order $g_{i,k}^{(1)}, g_{i,k}^{(2)}$ is stored in V^{direct}, which is selected when $\mathbf{m}_{i,k} = m_{i,k}$. The order $g_{i,k}^{(2)}, g_{i,k}^{(1)}$ is stored in $V^{indirect}$, which is selected when $\mathbf{m}_{i,k} > m_{i,k}$. It can be shown by induction that it is the correct order of $g_{i,k}^{(1)}$ and $g_{i,k}^{(2)}$ that is chosen, when using V^{direct} and $V^{indirect}$. The proof will not be done in this project, but can be found in [5, p.12-14].

To initialize the algorithm the first row of V^{direct} is set to $(v_{1,1}^{direct}, v_{1,2}^{direct}, \dots, v_{1,d}^{direct}) = \big(F_1(x_1), F_2(x_2), \dots, F_d(x_d)\big)$. The first row of $V^{indirect}$, $(v_{1,1}^{indirect}, v_{1,2}^{indirect}, \dots, v_{1,d}^{indirect})$, does not need to be initialized since V^{direct} is always chosen for $i = 1$ since $\mathbf{m}_{1,k} = m_{1,k}$ for all $k = 1, \dots, d$.

Algorithm 1 Conditional Distributions for R-Vine Density

Input R-vine copula specification matrices M, B, Θ, where $m_{k,k} = k, k = 1, \dots, d$.
Output Density of the R-vine distribution at $x = (x_1, \dots, x_d)$, for an R-vine copula specification.

Set F = 1.
Let $V^{direct} = (v_{i,k}^{direct} | k = 1, \dots, d, i = 1, \dots, k)$.
Let $V^{indirect} = (v_{i,k}^{indirect} | k = 1, \dots, d, i = 1, \dots, k)$.
Set $(v_{1,1}^{direct}, v_{1,2}^{direct}, \dots, v_{1,d}^{direct}) = \big(F_1(x_1), F_2(x_2), \dots, F_d(x_d)\big)$.
Let $\mathbb{M} = (\mathbf{m}_{i,k} | k = 1, \dots, d, i = 1, \dots, k)$ where $\mathbf{m}_{i,k} = \max\{m_{i,k}, \dots, m_{d,k}\}, \forall k = 1, \dots, d$ and $i = 1, \dots, k$.
for $k = 2, \dots, d$ **do** (Iterating over the columns of M, except the first column)
    for $i = 1, \dots, k-1$ **do** (Iterating over the rows of M, above the diagonal)
        Set $g_{i,k}^{(1)} = v_{i,k}^{direct}$.
        if $\mathbf{m}_{i,k} = m_{i,k}$ **then**
            Set $g_{i,k}^{(2)} = v_{i,\mathbf{m}_{i,k}}^{direct}$.
        else
            Set $g_{i,k}^{(2)} = v_{i,\mathbf{m}_{i,k}}^{indirect}$.
        end if
        Set $F = F \cdot c\left(g_{i,k}^{(1)}, g_{i,k}^{(2)} | b_{i,k}, \theta_{i,k}\right)$.
        Set $v_{i+1,k}^{direct} = h\left(g_{i,k}^{(1)}, g_{i,k}^{(2)} | b_{i,k}, \theta_{i,k}\right)$ and $v_{i+1,k}^{indirect} = h\left(g_{i,k}^{(2)}, g_{i,k}^{(1)} | b_{i,k}, \theta_{i,k}\right)$.
    end for
end for
return F

2.3.1 Sampling From an *R*-Vine Copula Specification

This subsection is based on [11] and [12].

To sample from an R-vine copula specification, Algorithm 1 can be modified using the inverse probability integral transformation.

Theorem 2.10. The Inverse Transform Method
Let $x \in \mathbb{R}$, $y \in [0, 1]$, $F(x)$ be a CDF, and $F^{-1}(y)$ be the inverse function given by

$$F^{-1}(y) = \min\{x : F(x) \geq y\}. \tag{2.34}$$

> By defining $X = F^{-1}(U)$, where U has a continuous uniform distribution on $[0, 1]$, then X follows the distribution F.

The proof can be found in [11, p. 1].

Therefore, to sample from an R-vine copula specification given a sample, $u_1, \dots, u_d$, of i.i.d. $unif(0, 1)$ variables, let

$$x_1 = u_1, \tag{2.35}$$
$$x_2 = F^{-1}_{2|1}(u_2|x_1), \tag{2.36}$$
$$x_3 = F^{-1}_{3|1,2}(u_3|x_1, x_2), \tag{2.37}$$
$$\vdots$$
$$x_d = F^{-1}_{d|1,2,\dots,d-1}(u_d|x_1, \dots, x_{d-1}). \tag{2.38}$$

To calculate each x_j for $j = 1, \dots, d$, the inverse of the h-function in (2.26) is used. Firstly, the inverse conditional distributions are rewritten using the notation of an R-vine matrix

$$x_j = F^{-1}_{j|1,\dots,j-1}(u_j|x_1, \dots, x_{j-1}) = F^{-1}_{m_{k,k}|\{m_{1,k},\dots,m_{k-1,k}\}}(x_{m_{k,k}}|x_{m_{1,k}}, \dots, x_{m_{k-1,k}}). \tag{2.39}$$

To calculate x_j, it can be noted that

$$F_{m_{k,k}|\{m_{1,k},\dots,m_{k-1,k}\}}(x_{m_{k,k}}|x_{m_{1,k}}, \dots, x_{m_{k-1,k}}) \tag{2.40}$$
$$= h\left(F_{m_{k,k}|m_{1,k},\dots,m_{k-2,k}}(x_{m_{k,k}}|x_{m_{1,k}}, \dots, x_{m_{k-2,k}}), g^2_{k-1,k}\right) \tag{2.41}$$
$$= h\left(h\left(F_{m_{k,k}|\{m_{1,k},\dots,m_{k-3,k}\}}(x_{m_{k,k}}|x_{m_{1,k}}, \dots, x_{m_{k-3,k}}), g^2_{k-2,k}\right), g^2_{k-1,k}\right) \tag{2.42}$$
$$\vdots$$
$$= h\left(\dots h\left(x_{m_{k,k}}, g^2_{1,k}\right)\dots, g^2_{k-1,k}\right). \tag{2.43}$$

By taking the inverse of the h-functions, x_j is given by

$$x_j = h^{-1}\left(\dots h^{-1}\left(u_j, g^2_{k-1}\right)\dots, g^2_{1,k}\right). \tag{2.44}$$

To incorporate this into the modified algorithm, two for-loops over the rows are used. The first one updates x_j using v^{direct} and $v^{indirect}$, which is selected the same way as in Algorithm 1. In the second loop, v^{direct} and $v^{indirect}$ are calculated based on the prior column.

Algorithm 2 Simulation From an R-Vine Copula Specification

Input R-vine copula specification matrices M, B, Θ, where $m_{k,k} = k, k = 1, \ldots, d$.
Output Random observations $x = (x_1, \ldots, x_d)$ from the R-vine copula specification.

Let $u_1, \ldots, u_d$ be independent uniform samples.
Let $V^{direct} = (v_{i,k}^{direct} | k = 1, \ldots, d, i = 1, \ldots, k)$.
Let $V^{indirect} = (v_{i,k}^{indirect} | k = 1, \ldots, d, i = 1, \ldots, k)$.
Set $(v_{1,1}^{direct}, v_{1,2}^{direct}, \ldots, v_{1,d}^{direct}) = (u_1, u_2, \ldots, u_d)$.
Let $\mathbb{M} = (\mathbf{m}_{i,k} | k = 1, \ldots, d, i = 1, \ldots, k)$ where $\mathbf{m}_{i,k} = \max\{m_{i,k}, \ldots, m_{d,k}\}, \forall k = 1, \ldots, d$ and $i = 1, \ldots, k$.
Let $x_1 = v_{1,d}^{direct}$.
for $k = 2, \ldots, d$ **do** (Iterating over the columns of M except the first column)
    for $i = k-1, \ldots, 1$ **do** (Iterating over the rows of M above the diagonal)
        if $\mathbf{m}_{i,k} = m_{i,k}$ **then**
            Set $g_{i,k}^{(2)} = v_{i,\mathbf{m}_{i,k}}^{direct}$.
        else
            Set $g_{i,k}^{(2)} = v_{i,\mathbf{m}_{i,k}}^{indirect}$.
        end if
        Set $v_{1,k}^{direct} = h^{-1}\left(v_{1,k}^{direct}, g_{i,k}^{(2)} | b_{i,k}, \theta_{i,k}\right)$
    end for
    Let $x_k = v_{1,k}^{direct}$.
    for $i = 1, \ldots, k-1$ **do** (Iterating over the rows of M above the diagonal.)
        Set $g_{i,k}^{(1)} = v_{i,k}^{direct}$
        Set $v_{i+1,k}^{direct} = h\left(g_{i,k}^{(1)}, g_{i,k}^{(2)} | b_{i,k}, \theta_{i,k}\right)$ and $v_{i+1,k}^{indirect} = h\left(g_{i,k}^{(2)}, g_{i,k}^{(1)} | b_{i,k}, \theta_{i,k}\right)$.
    end for
end for
return $(x_1, \ldots, x_d)$

2.4 Parameter Estimation of Simplified *R*-Vine Copulas

Assume that there exists an R-vine copula specification, with an associated density as given in (2.24), and that the simplifying assumption holds. Given some i.i.d. d-dimensional data of sample size n, collected in $n \times d$ data matrix

$$u := (u_1^\top, \ldots, u_n^\top), \text{ with } u_j := (u_{j,1}, \ldots, u_{j,d})^\top \text{ for } j = 1, \ldots, n, \tag{2.45}$$

it is possible to formulate the likelihood for a simplified R-vine copula. Let $\theta = \{\theta_e, e \in E\}$, then the likelihood for a simplified R-vine copula is given as

$$L(\theta; u) = \prod_{j=1}^{n} \prod_{i=1}^{d-1} \prod_{e \in E_i} c_{\mathcal{C}_{a_e}, \mathcal{C}_{b_e} | D_e} \left(F_{\mathcal{C}_{a_e} | D_e}(u_{j,\mathcal{C}_{a_e}} | u_{j,D_e}), F_{\mathcal{C}_{b_e} | D_e}(u_{j,\mathcal{C}_{b_e}} | u_{j,D_e}) \right), \tag{2.46}$$

where $c_{\mathcal{C}_{a_e},\mathcal{C}_{b_e}|D_e}$ depends on the parameters $\theta_{\mathcal{C}_{a_e},\mathcal{C}_{b_e}|D_e}$. Furthermore, $F_{\mathcal{C}_{a_e}|D_e}$ and $F_{\mathcal{C}_{b_e}|D_e}$ are recursively calculated the same way as in Algorithm 1 and therefore they depend on the parameters used to calculate them.

To calculate the log-likelihood for a simplified R-vine copula, one can use a modified version of Algorithm 1. To calculate the log-likelihood, Line 1 is changed to $L = 0$ and Line 14 is changed to $L = L + \log\left(c\left(g_{i,k}^{(1)}, g_{i,k}^{(2)}; b_{i,k}, \theta_{i,k}\right)\right)$. Given multiple observations, $v_{i+1,k}^{direct}$ and $v_{i+1,k}^{indirect}$ must be calculated for every observation, which is achieved by taking the sum over every observation in Line 14. The algorithm for calculating the log-likelihood can be found in Section B.2.

When wanting to estimate the parameter vector using maximum likelihood for a simplified R-vine copula, where each pair copula only has one parameter, the number of parameters is $d(d-1)/2$. Therefore, it is desirable to find a good starting value, which can be found using sequential parameter estimation. The sequential estimation method maximizes the parameters tree-by-tree, and is, therefore, simpler than doing joint maximum likelihood, where all the parameters are estimated at once.

Definition 2.11. Sequential Estimation in Simplified *R*-Vine Copulas
Let θ_e be the copula parameters corresponding to $e = (\mathcal{C}_{a_e}, \mathcal{C}_{b_e}|D_e)$ in tree T_i for $i = 1, \ldots, d-1$. Furthermore, let $\theta(T_i)$ denote the copula parameters in tree T_i, with estimates denoted as $\hat{\theta}(T_i)$. If all copula parameters up to T_{i-1} are estimated, let $\hat{\theta}(T_{1,\ldots,i-1})$ denote the set of those parameters. Then, the sequential estimate of θ_e in T_i is based on pseudo-observations

$$u_{j,\mathcal{C}_{a_e}|D_e;\hat{\theta}(T_{1,\ldots,i-1})} := F_{\mathcal{C}_{a_e}|D_e}\left(u_{j,\mathcal{C}_{a_e}}|u_{j,D_e}; \hat{\theta}(T_{1,\ldots,i-1})\right), \tag{2.47}$$

$$u_{j,\mathcal{C}_{b_e}|D_e;\hat{\theta}(T_{1,\ldots,i-1})} := F_{\mathcal{C}_{b_e}|D_e}\left(u_{j,\mathcal{C}_{b_e}}|u_{j,D_e}; \hat{\theta}(T_{1,\ldots,i-1})\right), \tag{2.48}$$

for $j = 1, \ldots, n$. To find the estimate $\hat{\theta}_e$, one must maximize

$$\prod_{j=1}^{n} c_{\mathcal{C}_{a_e},\mathcal{C}_{b_e}|D_e}\left(u_{j,\mathcal{C}_{a_e}|D_e;\hat{\theta}(T_{1,\ldots,i-1})}, u_{j,\mathcal{C}_{b_e}|D_e;\hat{\theta}(T_{1,\ldots,i-1})}; \theta_e\right). \tag{2.49}$$

The concept of sequential estimation, as presented in Definition 2.11, is often used as a starting point for maximum likelihood estimation as stated before. However, it is also possible to use the result from the sequential estimation as the final estimate for the parameters.

2.5 Selecting *R*-Vine Copulas

The process of estimating and selecting an R-vine copula, can be summed up in three steps:

1. Selecting the R-vine tree structure in terms of (un)conditioned pairs.
2. Choosing bivariate copulas for all the pairs selected in step 1.
3. Estimating the copula parameter(s) for each pair copula chosen in step 2.

For a small number of dimensions, it would be possible to estimate the parameters for all the different structures. However, when the number of dimensions increases, the number of parameters increases quadratically, as stated in Section 2.4, and hence it becomes unfeasible to estimate all parameters for all outcomes. Therefore, consider a triplet $(\mathcal{V}, \mathcal{B}(\mathcal{V}), \Theta(\mathcal{B}(\mathcal{V})))$, where $\mathcal{V}$ denotes an R-vine tree structure, $\mathcal{B}(\mathcal{V})$ denotes the set of all $d(d-1)/2$ pair copulas for the edges in $\mathcal{V}$, and $\Theta(\mathcal{B}(\mathcal{V}))$ denotes the parameters estimated for all pair copulas in $\mathcal{B}(\mathcal{V})$. Given some data of size $n \times d$ and the specification of an adequate R-vine tree structure, the next step is to choose copulas $C_{\mathcal{C}_{a_e},\mathcal{C}_{b_e}|D_e}$ and estimate parameters. To select copulas for each edge in all trees, the AIC can be used.

Therefore, consider the set of bivariate copulas and associated parameters for $T_1 \in \mathcal{V}$. For an edge $e = \{a, b\} = (\mathcal{C}_{a_e}, \mathcal{C}_{b_e}|D_e) \in T_1$, the copula data is given as $u_{j,\mathcal{C}_{a_e}} := F_{\mathcal{C}_{a_e}}(x_{j,\mathcal{C}_{a_e}})$ and $u_{j,\mathcal{C}_{b_e}} := F_{\mathcal{C}_{b_e}}(x_{j,\mathcal{C}_{b_e}})$ for $j = 1, \ldots, n$. Let $\mathcal{B}_e$ be the set of possible parametric bivariate copulas for edge e, and C^B be an element in $\mathcal{B}_e$, with density c^B. Then, for each $C^B \in \mathcal{B}_e$ calculate the MLE of θ^B, using the copula data $u_e := \{u_{j,\mathcal{C}_{a_e}}, u_{j,\mathcal{C}_{b_e}}, \text{ for } j = 1, \ldots, n\}$. When each copula C^B is fit with parameters $\hat{\theta}^B$, calculate the AIC for each element as follows

$$AIC(C^B, \hat{\theta}^B; u_e) := -2 \sum_{j=1}^{n} \ln(c^B(u_{j,\mathcal{C}_{a_e}}, u_{j,\mathcal{C}_{b_e}}; \hat{\theta}^B) + 2j^B, \tag{2.50}$$

where j^B is the dimension of θ^B. The choice of bivariate copula for edge e, is the one which yields the lowest AIC value.

For T_i, where $i \geq 2$, the sequential estimation from Section 2.4 is utilized. For $e = \{a, b\} = (\mathcal{C}_{a_e}, \mathcal{C}_{b_e}|D_e) \in T_i$, pseudo-observations from the bivariate copula distribution $(U_{\mathcal{C}_{a_e}}, U_{\mathcal{C}_{b_e}})$ given U_{D_e} is available, and they will be denoted as

$$\hat{u}_{j,\mathcal{C}_{a_e}|D_e} := u_{j,\mathcal{C}_{a_e}|D_e;\hat{\theta}(T_1,\ldots,i-1)} \text{ and } \hat{u}_{j,\mathcal{C}_{b_e}|D_e} := u_{j,\mathcal{C}_{b_e}|D_e;\hat{\theta}(T_1,\ldots,i-1)}, \tag{2.51}$$

where $j = 1, \ldots, n$ and $u_{j,\mathcal{C}_{a_e}|D_e;\hat{\theta}(T_1,\ldots,i-1)}$ and $u_{j,\mathcal{C}_{b_e}|D_e;\hat{\theta}(T_1,\ldots,i-1)}$ are given as (2.47) and (2.48), respectively. Then, for each $e \in T_i$, where $i \geq 2$, follow the same procedure as for T_1. Given the choices of pair copulas and associated parameters for all edges in T_i, the pseudo-observations is defined and used for the selection of copulas for T_{i+1}. The procedure is executed until all pair copulas for every edge in all trees are specified.

As the number of R-vine tree sequences for d variables is $\frac{d!}{2} \cdot 2^{\binom{d-2}{2}}$, it is unfeasible to fit all possibilities of $(\mathcal{V}, \mathcal{B}(\mathcal{V}), \Theta(\mathcal{B}(\mathcal{V})))$ when d is large. Therefore, a greedy algorithm, called *Dißmann's Algorithm*, with top-down selection, is chosen for selecting an R-vine copula. The top-down selection means that one starts with the T_1, and sequentially selects all trees down to T_{d-1}. The first tree can be chosen as an arbitrary spanning tree, however, given a tree T_i where $i \in \{1, \dots, d-2\}$, the tree T_{i+1} has to obey the proximity condition, as defined in Definition 2.3. Let w be a weight for the edges, which in this project will either be the AIC as presented in (2.50), or the absolute empirical Kendall's tau given in (A.10).

Algorithm 3 Dißmann's Algorithm

Calculate $w_{i,k}$ for all pairs $\{i, k\}$ where $1 \leq i < k \leq d$.
Select the maximum spanning tree:

$$T_1 = \underset{T=(N,E)\,\text{spanning tree}}{\arg\max} \sum_{e=(\mathcal{C}_{a_e}, \mathcal{C}_{b_e}) \in E} w_{\mathcal{C}_{a_e}, \mathcal{C}_{b_e}}.$$

for each edge $e \in E_1$ **do**
    Select a copula, $C_{\mathcal{C}_{a_e}, \mathcal{C}_{b_e}}$ with parameter estimates $\hat{\theta}_{\mathcal{C}_{a_e}, \mathcal{C}_{b_e}}$.
    for $j = 1, \dots, n$ **do**

$$\begin{aligned} \hat{u}_{j,\mathcal{C}_{a_e}} &= F_{\mathcal{C}_{a_e}|\mathcal{C}_{b_e}}(u_{j,\mathcal{C}_{a_e}} | u_{j,\mathcal{C}_{b_e}}; \hat{\theta}_{\mathcal{C}_{a_e}, \mathcal{C}_{b_e}}), \\ \hat{u}_{j,\mathcal{C}_{b_e}} &= F_{\mathcal{C}_{b_e}|\mathcal{C}_{a_e}}(u_{j,\mathcal{C}_{b_e}} | u_{j,\mathcal{C}_{a_e}}; \hat{\theta}_{\mathcal{C}_{a_e}, \mathcal{C}_{b_e}}). \end{aligned}$$

    end for
end for
for $i = 2, \dots, d-1$ **do**
    Determine $w_{\mathcal{C}_{a_e}, \mathcal{C}_{b_e}|D_e}$ for all possible edges, $e = (\mathcal{C}_{a_e}, \mathcal{C}_{b_e}|D_e)$, in T_i.
    Let $E_{P,i}$ denote the set of edges which satisfy the proximity condition for T_i.
    Select the maximum spanning tree for the edges in $E_{P,i}$:

$$T_i = \underset{T=(N,E)\,\text{spanning tree with } E \subset E_{P,i}}{\arg\max} \sum_{e=(\mathcal{C}_{a_e}, \mathcal{C}_{b_e}|D_e) \in E} w_{\mathcal{C}_{a_e}, \mathcal{C}_{b_e}|D_e}$$

    for each edge $e \in E_i$ **do**
        Select a pair copula $C_{\mathcal{C}_{a_e}, \mathcal{C}_{b_e}|D_e}$ with parameter estimates $\hat{\theta}_{\mathcal{C}_{a_e}, \mathcal{C}_{b_e}|D_e}$.
        for $j = 1, \dots, n$ **do**

$$\begin{aligned} \hat{u}_{j,\mathcal{C}_{a_e}|D_e} &= F_{\mathcal{C}_{a_e}|\mathcal{C}_{b_e} \cup D_e}(u_{j,\mathcal{C}_{a_e}} | u_{j,\mathcal{C}_{b_e}}, u_{j,D_e}; \hat{\theta}_{\mathcal{C}_{a_e}, \mathcal{C}_{b_e}|D_e}), \\ \hat{u}_{j,\mathcal{C}_{b_e}|D_e} &= F_{\mathcal{C}_{b_e}|\mathcal{C}_{a_e} \cup D_e}(u_{j,\mathcal{C}_{b_e}} | u_{j,\mathcal{C}_{a_e}}, u_{j,D_e}; \hat{\theta}_{\mathcal{C}_{a_e}, \mathcal{C}_{b_e}|D_e}). \end{aligned}$$

        end for
    end for
end for
return the sequential model estimate $(\hat{\mathcal{V}}, \hat{\mathcal{B}}, \hat{\mathbf{\Theta}})$.

As Dißmann's algorithm is a greedy algorithm, it only makes locally optimal selections of trees. However, as mentioned in [4, p. 162], the performance of the algorithm is in general good. As for the concept of maximum spanning trees mentioned in Line 2 and Line 10 in Algorithm 3, it is, as the optimization problems show, a spanning tree where the weights are maximized. To solve those maximum spanning tree problems, a modification of Prim's algorithm is used [13, p.634-635]. As all pair copulas are tested for each edge in each tree, it can be beneficial to choose a restricted set of copulas, to reduce the complexity of the estimation. A restricted set of bivariate copulas, presented in Table 2.1, will therefore be chosen for the application part of the project.

2.6 Goodness-of-Fit

This section is based on [14] and [15].

The purpose of this section is to verify the choice of model using a goodness-of-fit(GOF) test. The test is based on the Bartlett identity, which is a relation between the Hessian matrix, $\mathbb{H}$, and the expected outer product of the corresponding score function, $\mathbb{C}$. Assume that the copula density is three times continuously differentiable and that the related expectations exist. Then

$$\begin{aligned}\mathbb{H}(\theta) &:= \mathbb{E}\left[\nabla_\theta^2 \ln\left(c_\theta\left(F_1(x_1),\dots,F_d(x_d)\right)\right)\right] \text{ and,}\\ \mathbb{C}(\theta) &:= \mathbb{E}\left[\nabla_\theta \ln\left(c_\theta\left(F_1(x_1),\dots,F_d(x_d)\right)\right)\left(\nabla_\theta \ln\left(c_\theta\left(F_1(x_1),\dots,F_d(x_d)\right)\right)\right)^\top\right],\end{aligned} \tag{2.52}$$

where ∇_θ is the gradient with respect to θ and the expectations are with respect to the joint distribution F. Under correct model specification, i.e., $\theta = \theta_0$, where θ_0 denotes the true parameter vector, the Bartlett identity is given as $-\mathbb{H}(\theta_0) = \mathbb{C}(\theta_0)$. From the Bartlett identity, the null hypothesis for the misspecification test is given as

$$H_0 : \mathbb{H}(\theta_0) + \mathbb{C}(\theta_0) = 0 \text{ against } H_1 : \mathbb{H}(\theta_0) + \mathbb{C}(\theta_0) \neq 0. \tag{2.53}$$

To test the hypothesis, estimation of $\mathbb{H}$ and $\mathbb{C}$ is required. Therefore, assume that $u := (u_1^\top,\dots,u_n^\top)$, with $u_j := (u_{j,1},\dots,u_{j,d})^\top$ for $j = 1,\dots,n$, is n i.i.d. d-dimensional samples of copula data, $\hat{\theta}_n = \hat{\theta}(u_1,\dots,u_n)$ is the estimated parameter vector, and U follows an R-vine copula distribution. Then, define

$$\mathbb{H}(\theta|U) := \nabla_\theta^2 l(\theta|U) \text{ and } \mathbb{C}(\theta|U) := \nabla_\theta l(\theta|U)\left(\nabla_\theta l(\theta|U)\right)^\top, \tag{2.54}$$

and the sample counterparts

$$\hat{\mathbb{H}}_j(\hat{\theta}_n) := \mathbb{H}(\hat{\theta}_n|u_j) \text{ and } \hat{\mathbb{C}}_j(\hat{\theta}_n) := \mathbb{C}(\hat{\theta}_n|u_j), \tag{2.55}$$

where $l(\theta|U)$ is the log-likelihood and $\hat{\mathbb{H}}_j(\hat{\theta}_n), \hat{\mathbb{C}}_j(\hat{\theta}_n) \in \mathbb{R}^{p\times p}$, where p is the length of the parameter vector. Then, the sample estimates of $\mathbb{H}(\theta)$ and $\mathbb{C}(\theta)$ is given as

$$\bar{\mathbb{H}}(\hat{\theta}_n) := \frac{1}{n}\sum_{j=1}^{n} \hat{\mathbb{H}}_j(\hat{\theta}_n) \text{ and } \bar{\mathbb{C}}(\hat{\theta}_n) := \frac{1}{n}\sum_{j=1}^{n} \hat{\mathbb{C}}_j(\hat{\theta}_n), \tag{2.56}$$

respectively. If all pair copulas are one-parametric, the dimensions of (2.55) are $d(d-1)/2 \times d(d-1)/2$, and hence $p = d(d-1)/2$. The dimension increases if higher parametric copulas are used and decreases if the independence copula is used. To construct the test statistic for the null hypothesis given in (2.53), the sum of the lower triangular matrices $\mathbb{H}$ and $\mathbb{C}$ are vectorized, i.e., the columns of the lower triangular matrices are stacked on top of each other, and is given as

$$d(\theta|U) := vec\left(\mathbb{H}(\theta|U) + \mathbb{C}(\theta|U)\right). \tag{2.57}$$

The empirical equivalents are given as

$$\hat{d}_j(\hat{\theta}_n) := d(\hat{\theta}_n|u_j), \tag{2.58}$$

$$\bar{d}(\hat{\theta}_n) := \frac{1}{n}\sum_{j=1}^{n} \hat{d}_j(\hat{\theta}_n), \tag{2.59}$$

which is both of dimension $p(p+1)/2$. Prior to (2.52), it was assumed that the copula density is three times continuously differentiable, which makes it possible to define the expected gradient matrix of $d(\theta|U)$ and corresponding estimate, as

$$\nabla D_\theta := \mathbb{E}\left[\nabla_{\theta_k} d_l(\theta|U)\right]_{l=1,\dots,p(p+1)/2,\ k=1,\dots,p} \text{ and} \tag{2.60}$$

$$\widehat{\nabla D_\theta} := \frac{1}{n}\sum_{j=1}^{n}\left[\nabla_{\theta_k}\hat{d}_j(\hat{\theta}_n|u_t)\right]_{l=1,\dots,p(p+1)/2,\ k=1,\dots,p}, \tag{2.61}$$

respectively. Under correct specification it holds that $\nabla D_{\theta_0} = 0$ which means that the asymptotic covariance matrix given as

$$\begin{aligned} V_{\theta_0} = \mathbb{E}\Big[&\big(d(\theta_0|U) - \nabla D_\theta \mathbb{H}(\theta_0)^{-1}\nabla_\theta l(\theta_0|U)\big) \\ &\cdot \big(d(\theta_0|U) - \nabla D_\theta \mathbb{H}(\theta_0)^{-1}\nabla_\theta l(\theta_0|U)\big)^\top\Big], \end{aligned} \tag{2.62}$$

is also equal to zero under the correct specification. Therefore, given the asymptotic covariance matrix and the fact that $\sqrt{n}\bar{d}(\hat{\theta}_n) \xrightarrow{d} N(0, V_{\theta_0})$ as $n \to \infty$, it is possible to formulate the GOF test statistic.

Proposition 2.12.
Under the correct copula specification and suitable regularity conditions[2], the test statistic is given as

$$\mathcal{T}_n := n \left(\bar{d}(\hat{\theta}_n) \right)^\top \hat{V}_{\hat{\theta}_n}^{-1} \bar{d}(\hat{\theta}_n), \tag{2.63}$$

where $\hat{V}_{\hat{\theta}_n}^{-1}$ is a consistent estimate of $V_{\theta_0}^{-1}$, and as a result $\mathcal{T}_n$ is asymptotically $\chi^2_{p(p+1)/2}$ distributed.

The proof is omitted. However, a proof for the general case of the misspecification test for a multivariate distribution can be found in [16]. Since the MLE for a vine copula is normally distributed;

$$\sqrt{n}\mathcal{I}(\theta_0)^{1/2} \left(\hat{\theta}_n - \theta_0 \right) \xrightarrow{d} N(0, I_p) \text{ as } n \to \infty, \tag{2.64}$$

this also holds for the case of vine copulas. In (2.64), $\mathcal{I}(\theta_0)$ denotes the Fisher Information Matrix.

Given the test statistic in Proposition 2.12, it is possible to construct an α-level test.

Corollary 2.13.
Let $\alpha \in (0, 1)$ and $\mathcal{T}_n$ be defined as in (2.63). Then reject $H_0 : \mathbb{H}(\theta_0) + \mathbb{C}(\theta_0) = 0$ if

$$\mathcal{T} > \left(\chi^2_{p(p+1)/2} \right)^{-1} (1 - \alpha), \tag{2.65}$$

which is an asymptotic α-level test, where $(\chi^2_{df})^{-1}(\beta)$ denotes the β-quantile of a χ^2_{df} distribution.

The χ^2-distribution in (2.65) depends solely on p, which is known from the set of copula families, and not the true parameter. Furthermore, the test is a so-called "blanket test", as it does not require any strategic choice in terms of for example weight- or kernel-functions and is therefore applicable for all copula families [17].

2.7 Modeling of Marginals

This section is based on [18] and [19].

In this section, it will be shown how the marginals of the log-returns will be modeled using an AutoRegressive Moving-Average(ARMA) and Generalized AutoRegressive Conditional

[2]The regularity conditions needed is assumptions $A.1$ - $A.10$ proposed in [16].

Heteroskedasticity (GARCH) model. The mean equation of a univariate time series describing the log-returns is given by

$$r_t = \mathbb{E}[r_t|\mathscr{F}_{t-1}] + \varepsilon_t, \tag{2.66}$$

where $\mathscr{F}_{t-1}$ is the information set at time $t-1$ and ε_t is the innovations of the time series. The conditional mean is modeled using an ARMA(m,n) process, and the conditional variance of the log-return series is modeled using a GARCH(p,q) model. A GARCH model is utilized to capture the characteristics of log-returns such as heavy tails, clustering of volatilities and the leverage effect.

The ARMA(m,n) mean equation is given by

$$r_t = \mu + \sum_{i=1}^{m} \phi_i r_{t-i} + \sum_{j=1}^{n} \theta_j \varepsilon_{t-j} + \varepsilon_t. \tag{2.67}$$

The GARCH(p,q) variance equation is given by

$$\varepsilon_t = z_t \sigma_t, \tag{2.68}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \tag{2.69}$$

where $\omega > 0$, $\alpha_i \geq 0$ for $i = 1, \ldots, p$, $\beta_j \geq 0$ for $j = 1, \ldots, q$, and $z_t \sim i.i.d(0,1)$. The distribution of the innovations is selected based on the data. In this project, the distribution of the innovations are allowed to follow either a standardized Gaussian-, standardized Student's t-, or a standardized Normal-Inverse Gaussian(NIG)-distribution. The density function for an NIG-distribution is given by

$$f(x;\mu,\alpha,\beta,\delta) = \frac{\alpha\delta}{\pi} \exp\left(\delta\sqrt{\alpha^2-\beta^2} + \beta(x-\mu)\right) \frac{K_1\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\sqrt{\delta^2+(x-\mu)^2}}, \tag{2.70}$$

where $K_\lambda(x) = \frac{1}{2}\int_0^\infty t^{\lambda-1}\exp\left(\frac{1}{2}x(t+t^{-1})\right)dt$ is a modified Bessel function of the third kind. An NIG-distribution's advantage over the Gaussian distribution is the ability to model heavy tails, kurtosis, and jumps, which can occur when working with a price process of financial assets [20, p. 224]. To select the distribution of the innovations the AIC is used. The order of the ARMA(m,n)-GARCH(p,q) model is also chosen with AIC. The fitted model's residuals, $\hat{z}_t$, are then converted into uniform marginals by using the CDF corresponding to the selected distribution of the innovations

$$\hat{u}_{t,} = F_j(\hat{z}_{t,j}), \tag{2.71}$$

where $\hat{u}_{t,j} \sim U[0,1]^d$ and $j = 1, \ldots, d$.

3 | (Conditional) Value at Risk

This section is based on [21] and [22].

Two popular risk measures in modern portfolio theory are Value at Risk (VaR) and the closely related term Conditional Value at Risk ($CVaR$), which is also called expected shortfall. Contrary to standard deviation, which models spread, the concern when modeling VaR and $CVaR$ is potential losses. Therefore, let $V(t)$ denote the value of a portfolio at time t, and

$$L_{[t,t+s]} := -(V(t+s) - V(t)), \tag{3.1}$$

denote the loss of the portfolio in the time period from t to $t+s$. As the value $V(t+s)$ is not observable at time t, it is a random variable. Therefore, $L_{[t,t+s]}$ is also a random variable, which distribution is called the loss distribution. In order to model $V(t)$, let $f : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ be a measurable function given as

$$f(t, P_t) = V(t), \tag{3.2}$$

which is a function of time, t, and the random variable $P_t = (P_{t,1}, \ldots, P_{t,d})$, which is the log-prices of d different assets at time t. Furthermore, let $r_t := P_t - P_{t-1}$ denote the log-returns of the d assets from $t-1$ to t. Then, it is possible to write the one-step-ahead loss function as

$$L_{[t,t+1]} = L := -\left(f(t+1, P_t + r_{t+1}) - f(t, P_t)\right). \tag{3.3}$$

Since the price, P_t, is known at time t, the distribution of the loss function only depends on the distribution of the return r_{t+1}. Hence, the conditional loss distribution is defined as

$$F_{L|\mathscr{F}_t}(l) := \mathbb{P}\left(\ell(r_{t+1}) \leq l | \mathscr{F}_t\right) = \mathbb{P}\left(L \leq l | \mathscr{F}_t\right), \tag{3.4}$$

where $\mathscr{F}_t$ is the information up to time t, $l \in \mathbb{R}$, and ℓ is the one-step-ahead loss operator defined as

$$\ell(x) := -\left(f(t+1, P_t + x) - f(t, P_t)\right), \tag{3.5}$$

where $x \in \mathbb{R}^d$. It is possible to use the conditional loss distribution, (3.4), to measure risk in terms of potential losses.

Definition 3.1. Value at Risk
Let L be a loss function with loss distribution function F_L and $\alpha \in (0,1)$. Then the VaR of a portfolio, at a confidence level α, is given as

$$VaR_\alpha(L) := \inf\{\ell \in \mathbb{R} : \mathbb{P}(L > \ell) \leq 1 - \alpha\} = \inf\{\ell \in \mathbb{R} : F_L(\ell) \geq \alpha\}. \tag{3.6}$$

The VaR of a portfolio measures the probability of obtaining a loss greater than or equal to ℓ, given a specific confidence level α. Hence, the VaR of a portfolio can be written as $VaR_\alpha(L) = F_L^{-1}(\alpha)$, which is the α-quantile of the loss distribution function. However, calculating the VaR for a portfolio, doesn't give any information of the size of the loss, as VaR only calculates the quantiles, but doesn't model the tail. Therefore, it could be beneficial to use a risk measure which models the tail, and hence try to estimate the expected loss.

Definition 3.2. Conditional Value at Risk
Let L be a loss function with $\mathbb{E}[|L|] < \infty$ and loss distribution function F_L and $\alpha \in (0, 1)$. Then the $CVaR$ of a portfolio, at a confidence level α, is given as

$$CVaR_\alpha(L) := \mathbb{E}[L|L \geq VaR_\alpha(L)] = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(L)du. \tag{3.7}$$

Given the possible losses L, the $CVaR$ calculates the average of the VaR for $u \geq \alpha$, which is why $CVaR$ is also called expected shortfall. The second equation in (3.7) is due to the fact that

$$\mathbb{E}[L|L \geq VaR_\alpha(L)] = \frac{1}{1-\alpha}\int_{F_L^{-1}(\alpha)}^\infty l\, dF_L(l) \tag{3.8}$$

$$= \frac{1}{1-\alpha}\int_\alpha^1 F_L^{-1}(u)du \tag{3.9}$$

$$= \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(L)du. \tag{3.10}$$

The main reason for choosing $CVaR$ as the risk measure, instead of VaR, is the fact that $CVaR$ is a coherent risk measure, which VaR is not. A coherent risk measure is a functional, which satisfies the properties: Monotonicity, sub-additivity, positive homogeneity, and translation invariance. The properties are explained in Appendix C. The main motivation for working with coherent risk measures is that diversification reduces the risk.

3.1 Estimation of *(C)VaR*

As the possible loss L is a latent variable, VaR and $CVaR$ need to be estimated. To estimate the loss distribution for $L = \ell(r_{t+1})$, given only the information, $\mathscr{F}_t$, three methods can be utilized: Variance-covariance estimation, historical simulation and estimation and Monte Carlo simulation and estimation. In this project, the Monte Carlo simulation method is utilized to estimate VaR and $CVaR$. This method is based, as the name suggests, on Monte Carlo simulations from an estimated model. First step is therefore to make a model for the returns, given the data $r_{t-n+1}, \ldots, r_t$. Then one samples M independent realizations, $r_{t+1,1}, \ldots, r_{t+1,M}$, and uses the realizations to calculate

$\{\tilde{L}_m = \ell(r_{t+1,m}) : m = 1, \ldots, M\}$. The empirical α-quantile is then found to calculate VaR, and the $CVaR$ is calculated as the average of the losses above the α-quantile. In general, the main problem with this method is to find a good multivariate model for the returns, which in this project is attempted with an R-vine copula.

3.2 Backtesting of *VaR* and *CVaR*

This section is based on [23], [24], and [25, p. 442-444].

Backtesting is used to measure the accuracy and thereby the effectiveness of $(C)VaR$ on historical data. To measure the accuracy of $(C)VaR$, forecasted values of $(C)VaR$ are compared to the actual losses.

3.2.1 Backtesting *VaR*

To determine the validity of the VaR forecasts, it is examined how many times the value of the loss function L exceeds the VaR forecast, which is called an αVaR violation. Let $I_t(\alpha)$ be a binary variable indicating whether an αVaR violation has occurred

$$I_t(\alpha) = \begin{cases} 1 & \text{if } L_t > VaR_\alpha(L_t) \\ 0 & \text{else} \end{cases}. \tag{3.11}$$

A VaR forecast is said to be valid if and only if the violation process $\big(I_t(\alpha)\big)_{T+1 \leq t \leq T+O}$, where $O \in \mathbb{N}$ is number of out-of-sample VaR forecasts, satisfies:

1. The probability of a loss function exceeding the VaR forecast must be equal to the confidence level α,

$$\mathbb{P}[I_t(\alpha) = 1] = \mathbb{E}[I_t(\alpha)] = \alpha. \tag{3.12}$$

2. VaR violations are distributed independently over time.

If this is the case, then the VaR violation process is a martingale difference

$$\mathbb{E}[I_t(\alpha)|\mathscr{F}_{t-1}] = \alpha. \tag{3.13}$$

Therefore, if the violation process $\big(I_t(\alpha)\big)_{T+1 \leq t \leq T+O}$ is said to be valid, then each $I_t(\alpha)$ should follow a Bernoulli distribution independently of each other, which will be tested with a Dynamic Quantile(DQ) test. A DQ test compares the number of violations of VaR with the expected number of violations. Furthermore, a DQ test tests for correlation between the violations of the VaR forecasts. To construct a test statistic for the DQ test,

the following linear regression model is utilized

$$Hit_t(\alpha) = \beta_0 + \sum_{i=1}^{p} \beta_l Hit_{t-i}(\alpha) + \sum_{j=1}^{q} u_j g\left(Hit_{t-j}(\alpha), Hit_{t-j-1}(\alpha), \dots, z_{t-j}, z_{t-j-1}, \dots\right) + \varepsilon_t, \tag{3.14}$$

where $t \in \{T+p, \dots, T+O\}$, $p \geq q$,

$$Hit_t(\alpha) = I_t(\alpha) - \alpha = \begin{cases} 1-\alpha, & \text{if } L_t > VaR_\alpha(L_t) \\ -\alpha, & \text{else} \end{cases}, \tag{3.15}$$

$g(\cdot)$ is a function of past violations and the variables z_{t-j} from the available information set $\mathscr{F}_{t-1}$, and ε_t is a discrete i.i.d. process with mean zero. $g(\cdot)$ can for example be the past returns, r_{t-j}, or the forecasted VaR, $\widehat{VaR}_\alpha(L_{t-j|t-j-1})$. In this project $g(\cdot) = \widehat{VaR}_\alpha(L_{t-j|t-j-1})$ and $q = 1$, since there will only be done one-day-ahead forecasts. To test whether the present violations of the VaR is uncorrelated with the past violations, i.e. $\beta_1 = \cdots = \beta_p = u_1 = 0$ and that the expected value of violations is equal to α, i.e., $\beta_0 = 0$, the null hypothesis is given by

$$H_0 : \beta_0 = \beta_1 = \cdots = \beta_p = u_1 = 0. \tag{3.16}$$

The linear regression model in (3.14) can be rewritten in matrix form as

$$Hit_t(\alpha) = \Psi X + \varepsilon_t, \quad t \in \{T+p, \dots, T+O\},\ p \geq 1, \tag{3.17}$$

where $\Psi = \left(\beta_0, \beta_1, \dots, \beta_p, u_1\right)$ is a vector of $p+2$ parameters and X is a matrix containing the explanatory variables. The DQ test statistic is then given as

$$DQ_\alpha = \frac{\hat{\Psi}_{OLS}^\top X^\top X \hat{\Psi}_{OLS}}{\alpha(1-\alpha)} \overset{a}{\sim} \chi^2_{(p+2)}, \tag{3.18}$$

where $\hat{\Psi}_{OLS}$ is the OLS estimates of Ψ given by

$$\hat{\Psi}_{OLS} = (X^\top X)^{-1} X^\top \widehat{Hit}_\alpha \overset{a}{\sim} N\left(0, \alpha(1-\alpha)(X^\top X)^{-1}\right), [3] \tag{3.19}$$

where $\widehat{Hit}_\alpha = \left(Hit_{T+1}(\alpha), Hit_{T+2}(\alpha), \dots, Hit_{T+O}(\alpha)\right)^\top$.

[3] This result can be found in [26, p. 12]

3.2.2 Backtesting *CVaR*

$CVaR$ is backtested with a $CVaR$ regression test. Let $\widehat{CVaR}_\alpha = \big(CVaR_\alpha(L_t)\big)_{T+1\leq t\leq T+O}$ be a process of the $CVaR$ forecasts with confidence level $\alpha \in (0,1)$. Furthermore, let $r = (r_t)_{T+1\leq t\leq T+O}$ be the actual portfolio returns. The $CVaR$ regression is then given by

$$r_t = \gamma_1 + \gamma_2 \widehat{CVaR}_{\alpha,t} + u_t^{CVaR}, \quad t \in \{T+1,\dots,T+O\}, \tag{3.20}$$

where $\gamma_1, \gamma_2 \in \mathbb{R}$, $\widehat{CVaR}_{\alpha,t}$ is the forecasted $CVaR$ at time t at confidence level α, and $CVaR_\alpha(u_t^{CVaR}|\mathscr{F}_{t-1}) = 0$ almost surely. As a consequence of the forecasts, $\widehat{CVaR}_{\alpha,t}$, being generated by the information set $\mathscr{F}_{t-1}$ and the condition on the error term, u_t^{CVaR}, the conditional $CVaR$ of r_t is given by

$$CVaR_\alpha(r_t|\mathscr{F}_{t-1}) = \gamma_1 + \gamma_2 \widehat{CVaR}_{\alpha,t}, \quad t \in \{T+1,\dots,T+O\}. \tag{3.21}$$

If the $CVaR$ is correctly specified then it must hold that $\widehat{CVaR}_{\alpha,t} = CVaR_\alpha(r_t|\mathscr{F}_{t-1})$ almost surely. The null- and alternative hypothesis are therefore given by

$$H_0 : (\gamma_1,\gamma_2) = (0,1) \quad \text{against} \quad H_1 : (\gamma_1,\gamma_2) \neq (0,1). \tag{3.22}$$

It is however noted by [25, p. 443] that estimating γ_1 and γ_2 in (3.20) using generalized method of moments or maximum likelihood estimation is infeasible. To overcome this problem, it can be utilized that (3.20) is an example of a $CVaR$ regression on the form $r_t = W_t^\top \gamma + u_t^{CVaR}$, where W_t is a vector containing the covariates. The estimation of the parameters is then done using a joint regression technique on

$$r_t = V_t^\top \beta + u_t^{VaR} \quad \text{and} \quad r_t = W_t^\top \gamma + u_t^{CVaR}, \quad t \in \{T+1,\dots,T+O\}, \tag{3.23}$$

where V_t and W_t are covariate vectors, $VaR_\alpha(u_t^{VaR}|\mathscr{F}_{t-1}) = 0$ and $CVaR_\alpha(u_t^{CVaR}|\mathscr{F}_{t-1}) = 0$ almost surely. By using a joint regression model, it is feasible to estimate the parameters (β,γ). The backtest for $CVaR$ used in this project is the Strict $CVaR$ Backtest where $V_t = W_t = (1, \widehat{CVaR}_{\alpha,t})$ and the regression system is given by

$$r_t = \beta_1 + \beta_2 \widehat{CVaR}_{\alpha,t} + u_t^{VaR} \quad \text{and} \quad r_t = \gamma_1 + \gamma_2 \widehat{CVaR}_{\alpha,t} + u_t^{CVaR}, \tag{3.24}$$

for $t \in \{T+1,\dots,T+O\}$. The null- and alternative hypothesis are given as

$$H_0 : (\gamma_1,\gamma_2) = (0,1) \quad \text{against} \quad H_1 : (\gamma_1,\gamma_2) \neq (0,1). \tag{3.25}$$

The null hypothesis is tested with a Wald-type test statistic

$$W_{CVaR} = O(\hat{\gamma} - (0,1))\hat{\Omega}_\gamma^{-1}(\hat{\gamma} - (0,1))^\top, \tag{3.26}$$

where $\hat{\Omega}_\gamma$ is the covariance of γ.

4 | Application

The purpose of this chapter is to evaluate the performance of the $(C)VaR$ forecasts, produced by an R-vine copula model, on three different portfolios consisting of assets from three different sectors within the S&P 500. Before evaluating the performance of the $(C)VaR$ forecasts, an analysis of the dependence structure in each sector will be conducted. The accuracy of the $(C)VaR$ forecasts will be evaluated using the theory presented in Section 3.2 on an equally-weighted portfolio. Lastly, the $CVaR$ models will be used to create a portfolio allocation strategy.

4.1 Data Introduction

The data used to construct and evaluate $(C)VaR$ risk forecasts, utilizing the methods presented in Chapter 2 and Chapter 3, will in this section be introduced. To test different dependence structures and compare risk levels in different market sectors, the analysis will be made on three different sectors. The three sectors are finance, healthcare, and technology. Additionally, the companies in each sector must be traded in the S&P 500 to be considered. The `R`-package used to obtain sector information and the asset prices is `Quantmod` [27].

Sector:	Finance	Healthcare	Technology
Number of companies:	61	60	71

The asset price for each asset in each sector in a one-day sampling interval from 03/01/2007−14/02/2024 is obtained using the data API available at Polygon [28]. A data filtration procedure is applied to each sector separately. In each sector, the timestamps that correspond to a weekend or a holiday are filtered out. Next, the latest publicly traded company is found in each sector, and the first observation in each sector is set to the day this company went public. The data for each sector is then split into training- and test data. The length of the training data is chosen to be 75 percent of the filtered data.

Table 4.1: The partition of the filtered data into training- and test data for each sector.

	Finance	Healthcare	Technology
Training data	07/07/2015-15/12/2021	01/08/2014-23/09/2021	06/07/2016-17/03/2022
Test data	16/12/2021-13/02/2024	24/09/2021-13/02/2024	18/03/2022-13/02/2024

To study the price development of each sector, an equally-weighted portfolio is constructed.

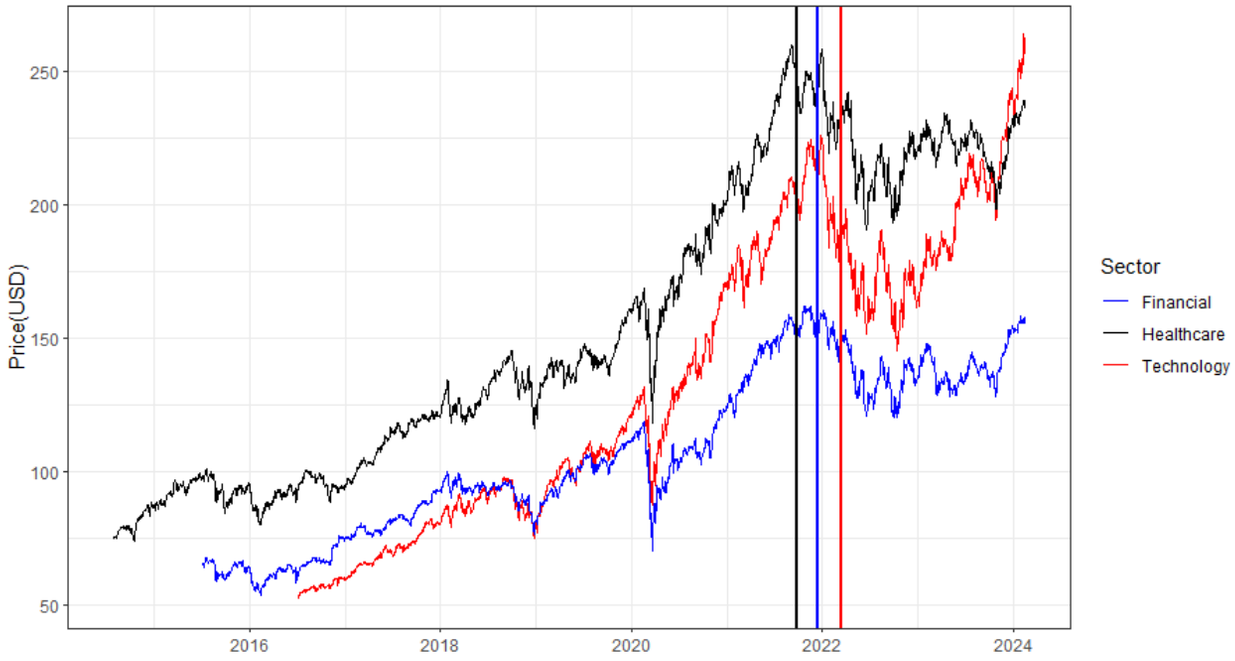


Figure 4.1: The price of an equally-weighted portfolio for each sector. The vertical lines shows where the training period ends for each sector.

From Figure 4.1 it can be inferred that each sector in the observed period in general has experienced a price hike. It can also be observed that each sector is influenced by the same market factors as they have the same price fluctuation tendencies and follow the same trends. It can also be observed that the behavior in each sector changes when the training period ends, which may be due to the conflict between Russia and Ukraine. To describe the behavior of each sector the joint distribution of each sector is constructed using an R-vine copula. However, to use an R-vine copula, a tailored marginal distribution for each asset in each sector must be constructed. The characteristics of each asset are therefore analyzed. In this section, the characteristics will only be shown for the first five listed assets in each sector, however, the characteristics for every asset in each sector can be found in the attached zip-file.

Table 4.2: Statistics for log-returns in the financial sector. For the three test-statistics, Augmented Dickey–Fuller(ADF) test, Jarque-Bera(JB) test, and Ljung-box(LB) test, where the lag is set to 3, it is the p-values that are given.

	Mean	Median	Min	Max	SD	Skew	Kurt	ADF	JB	LB(3)
JPM	0.000448	0.000302	-0.162	0.166	0.0177	-0.0467	13.3	0.01	0	$1.93 \cdot 10^{-8}$
V	0.000638	0.00145	-0.146	0.13	0.0159	-0.0757	9.57	0.01	0	$4.69 \cdot 10^{-7}$
MA	0.000725	0.00137	-0.136	0.154	0.0176	-0.00186	8.86	0.01	0	0.00802
BAC	0.000317	0.000305	-0.167	0.164	0.0203	-0.0453	9.4	0.01	0	0.00365
WFC	$-6.63 \cdot 10^{-5}$	0	-0.173	0.136	0.0204	-0.308	8.77	0.01	0	$8.76 \cdot 10^{-7}$

Table 4.3: Statistics for log-returns in the healthcare sector. For the three test-statistics, Augmented Dickey–Fuller(ADF) test, Jarque-Bera(JB) test, and Ljung-box(LB) test, where the lag is set to 3, it is the p-values that are given.

	Mean	Median	Min	Max	SD	Skew	Kurt	ADF	JB	LB(3)
LLY	0.00104	0.000963	-0.111	0.146	0.017	0.749	10.9	0.01	0	0.000761
UNH	0.000775	0.000954	-0.19	0.12	0.0162	-0.403	14.5	0.01	0	$1.97 \cdot 10^{-12}$
JNJ	0.000187	0.000275	-0.106	0.0769	0.0115	-0.389	9.72	0.01	0	0.000203
ABBV	0.000349	0.000396	-0.104	0.099	0.0136	-0.0696	6.91	0.01	0	0.00102
MRK	$5 \cdot 10^{-4}$	0.00113	-0.177	0.129	0.0171	-0.947	12.7	0.01	0	0.799

Table 4.4: Statistics for log-returns in the technology sector. For the three test-statistics, Augmented Dickey–Fuller(ADF) test, Jarque-Bera(JB) test, and Ljung-box(LB) test, where the lag is set to 3, it is the p-values that are given.

	Mean	Median	Min	Max	SD	Skew	Kurt	ADF	JB	LB(3)
MSFT	0.00108	0.00104	-0.159	0.133	0.0175	-0.214	8.02	0.01	0	$4.51 \cdot 10^{-13}$
AAPL	0.00108	0.000942	-0.138	0.113	0.0185	-0.216	5.84	0.01	0	0.00291
NVDA	0.00214	0.00282	-0.208	0.261	0.0313	0.122	6.98	0.01	0	0.0149
GOOGL	0.000732	0.00109	-0.124	0.0919	0.0182	-0.318	4.73	0.01	0	0.0256
GOOG	0.000742	0.00106	-0.118	0.0994	0.0181	-0.296	4.81	0.01	0	0.0303

An Augmented Dickey–Fuller(ADF) test is carried out to test whether the log-returns are stationary. From Table 4.2, Table 4.3, and Table 4.4 it can be inferred that it can't be rejected that the log-returns for each asset is stationary. The log-returns will therefore not be transformed in this project. To test for autocorrelation in the log-returns a Ljung-box(LB) test is used. The number of lags being tested is set to three, since it will be the maximum order for both the ARMA and the GARCH part. In the tables above, there is one asset, MRK, where it can be rejected that there is autocorrelation in the log-returns when the number of lags being tested is three. However, most of the assets' log-returns exhibit serial correlation. It will therefore later be examined whether the ARMA-GARCH models are able to eliminate the presence of serial correlation.

It is also tested whether the log-returns for each asset have the same skewness and kurtosis as the Gaussian distribution, thus indicating whether the log-returns follow a Gaussian distribution. The test is a Jarque-Bera(JB) test where the null hypothesis is that the skewness and kurtosis are equal to the Gaussian distribution. The null hypothesis is rejected for every asset thus indicating that the log-returns are not normally distributed. It will therefore be tested whether a standardized Student's t-distribution or a standardized NIG-distribution of the innovations of the ARMA-GARCH models gives a lower AIC value, compared to the standardized Gaussian distribution.

4.2 Fitting Marginal Distributions

As presented in Section 2.7 the modeling of the marginal distribution of each asset is done with an ARMA(m,n)-GARCH(p,q) model. This is achieved using the function `ugarchfit()` from the `R`-package `rugarch` [29]. The order of both the ARMA and GARCH parts are restricted to $m,n,p,q \in \{0,1,2,3\}$. To select the order, the AIC is utilized. Normally, one would choose the model with the lowest AIC. However, as the observed AIC values lie very close to each other, an additional selection criterion is implemented.

1. Calculate the AIC for each combination of orders for the ARMA- and GARCH part.
2. Find the smallest AIC value.
3. Check if there are other models with an AIC value within a one percent difference of the smallest AIC value. If this is not the case choose this model.
4. Among the constructed models within one percent of the smallest AIC value, select the model based on the following criteria
   (a) For each model find the highest order among the ARMA and GARCH orders. Only keep the models which have the lowest highest order among the models, i.e., between three models ARMA$(1,1)$-GARCH$(1,1)$, ARMA$(0,0)$-GARCH$(0,1)$, ARMA$(1,2)$-GARCH$(1,1)$, the last model is discarded since the highest order for that model is two and the two other models have the same highest order of one.
   (b) Lastly, take the sum of the orders of each model, i.e., for an ARMA$(1,1)$-GARCH$(1,1)$ the sum would be four. Select the model with the lowest sum of the orders.

The procedure is applied to every asset and yields the following result for the first five listed assets of the financial sector.

Table 4.5: Selected marginal distribution and parameters for the first five listed assets in the financial sector.

					Parameters Value (Standard Error)					
Stock	Model	Distribution	L	AIC	μ	AR	MA	ω	α	β
JPM	ARMA$(0,0)$ GARCH$(1,1)$	NIG	-1920	2.37		-	-		0.166 (0.0365)	0.786 (0.0453)
					0.0148 (0.018)	- - -	- - -	$1.4 \cdot 10^{-5}$ $(4.6 \cdot 10^{-6})$	- - -	- - -

Table 4.5: Selected marginal distribution and parameters for the first five listed assets in the financial sector *(continued)*.

					Parameters Value (Standard Error)					
Stock	Model	Distribution	L	AIC	μ	AR	MA	ω	α	β
V	ARMA$(0,0)$ GARCH$(1,1)$	NIG	-1940	2.39	0.0198 (0.0172)	- - -	- - -	$7.37 \cdot 10^{-6}$ $(2.6 \cdot 10^{-6})$	0.154 (0.0323) - -	0.821 (0.0357) - -
MA	ARMA$(0,0)$ GARCH$(1,1)$	NIG	-1940	2.39	0.0247 (0.017)	- - -	- - -	$8.35 \cdot 10^{-6}$ $(2.73 \cdot 10^{-6})$	0.167 (0.0318) - -	0.814 (0.0329) - -
BAC	ARMA$(0,0)$ GARCH$(1,1)$	STD	-2000	2.47	0.0336 (0.0176)	- - -	- - -	$1.64 \cdot 10^{-5}$ $(5.54 \cdot 10^{-6})$	0.124 (0.0268) - -	0.836 (0.0348) - -
WFC	ARMA$(0,0)$ GARCH$(1,1)$	STD	-1900	2.35	0.013 (0.0158)	- - -	- - -	$8.78 \cdot 10^{-6}$ $(3.18 \cdot 10^{-6})$	0.135 (0.0289) - -	0.851 (0.0305) - -

The results for every asset in every sector can be found in the attached zip-file. For the rest of the dependence structure analysis, the results will only be shown for the financial sector and the results for the other two sectors can be found in Appendix D and the attached zip-file.

From Table 4.5 it can be inferred that the selected order for each asset is $m = n = 0$ and $p = q = 1$. This is the case for every asset in the financial sector. The only parameter that is estimated using an ARMA model is therefore the intercept μ. The mean of log-returns is therefore fixed. This may be the reason behind the relatively high standard error of the μ estimates. The distribution that is selected the most for the innovations is the standardized NIG distribution. The standardized Gaussian distribution is not chosen for any of the assets, which corresponds with the results from the Jarque-Bera tests.

The fitted models' residuals are converted into uniform marginals by applying (2.71). It can be tested whether the residuals are a random sample from the chosen distribution of the innovations, with the parameters estimated for the ARMA-GARCH model, by performing an Anderson-Darling test.

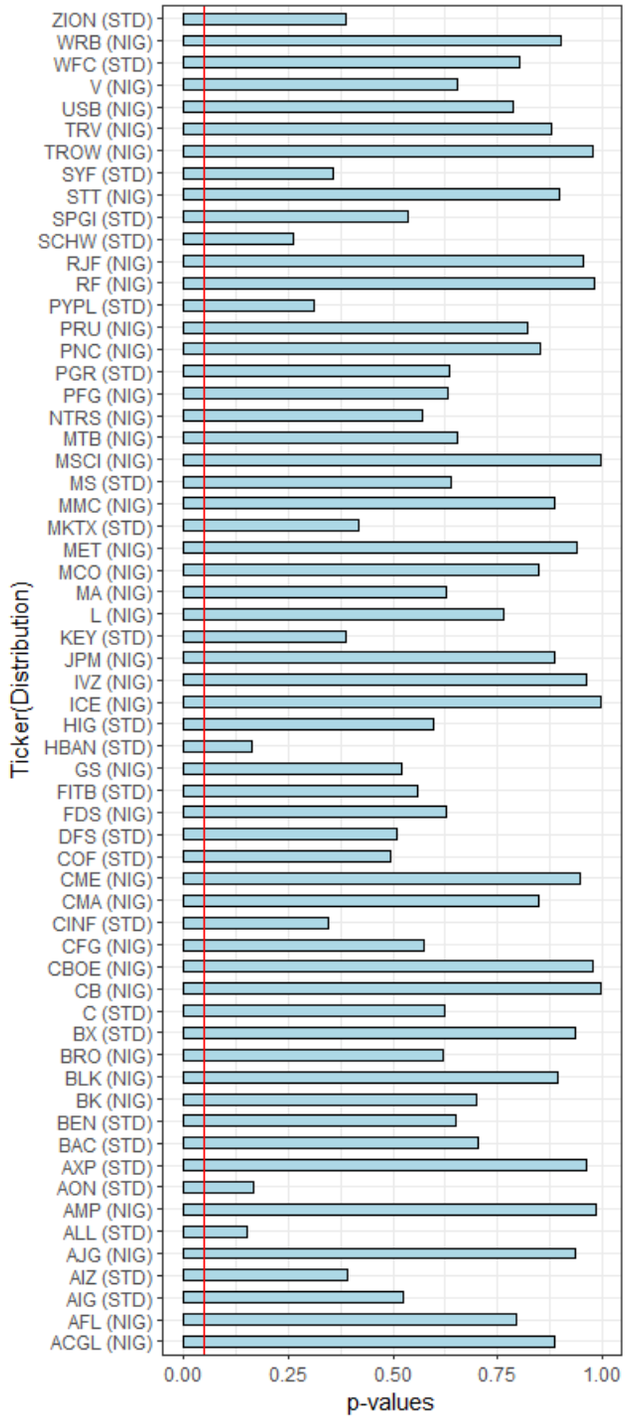


Figure 4.2: The p-values from an Anderson-Darling test performed on the residuals of the ARMA-GARCH models in the financial sector, with the respective distribution of the innovations shown in the parentheses.

The Anderson-Darling test indicates that it cannot be rejected that each asset's residuals

are sampled from their respective chosen distribution. To further test the choice of distribution of the residuals, a QQ-plot is constructed for the two stocks with the lowest p-value and the two stocks with the highest p-value for each selected distribution.

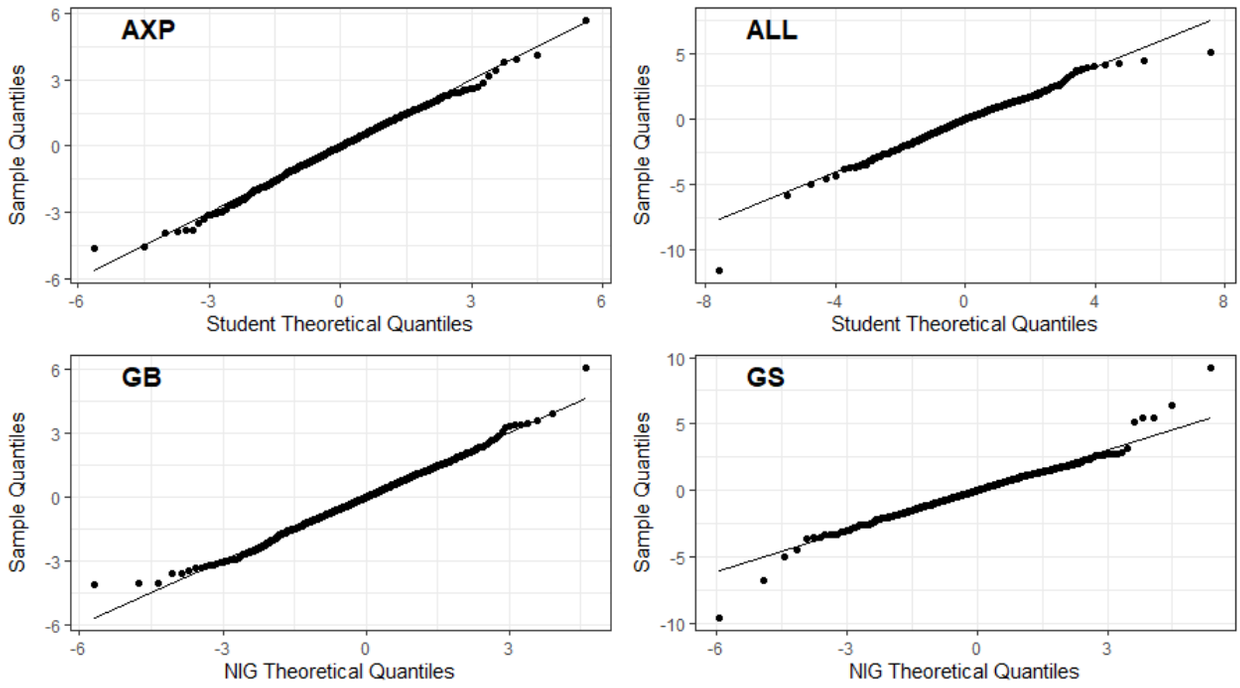


Figure 4.3: QQ-plots for the assets with tickers AXP, ALL, GB, and GS. The first column contains the assets with the highest p-value and the second column contains the assets with the lowest p-value for each distribution.

From Figure 4.3 it can be inferred that both the standardized Student's t-distribution and the standardized NIG distribution are able to capture the characteristics of the residuals. For the two assets with the lowest p-value from the Anderson-Darling test, the QQ-plots imply that the selected distribution is not able to capture the heavy tails of the residuals as well as for the two assets with the highest p-values.

Lastly, it will be investigated whether the fitted models are able to eliminate the serial correlation observed in the log-returns.

Table 4.6: Ljung-Box test on residuals(R) and squared residuals(R^2) from the fitted models from the financial sector, for lag $= 1, 2, 3$.

		Ljung-Box Test, R			Ljung-Box Test, R^2		
Stock	Lag	1	2	3	1	2	3
JPM	Statistic	0.0352	2.68	3.2	0.576	0.576	0.951
	p-value	0.851	0.262	0.362	0.448	0.75	0.813
V	Statistic	4.36	4.64	6.59	0.0158	0.0581	0.373
	p-value	0.0369	0.0983	0.086	0.9	0.971	0.946
MA	Statistic	2.61	2.63	3.44	0.0171	0.564	0.589
	p-value	0.106	0.269	0.328	0.896	0.754	0.899
BAC	Statistic	1.42	3.99	4	0.858	1.09	1.14
	p-value	0.234	0.136	0.261	0.354	0.58	0.768
WFC	Statistic	1.36	2.87	4.64	3.46	3.62	3.72
	p-value	0.243	0.238	0.2	0.0627	0.163	0.294

From Table 4.6 it can be inferred that the null hypothesis of the Ljung-Box test is not rejected for lag $= 1, 2, 3$ for the standardized residuals except for V with lag $= 1$, which implies that there is no autocorrelation between the residuals. For the squared standardized residuals, the null hypothesis is not rejected for every asset, which implies that there is no presence of autoregressive conditional heteroskedasticity in the residuals.

4.3 Fitting Vine Copulas

To fit an R-vine copula, the fitted models' residuals are converted to uniform marginals using (2.71). The uniform marginals for the first five listed assets in the financial sector are transformed to pseudo copula data to show the pairwise dependence structure between the five assets. The pseudo copula data is created by setting

$$u_j = (u_{j,1}, \dots, u_{j,5}) = \left(\hat{F}_1(x_{j,1}), \dots, \hat{F}_5(x_{j,5})\right), \quad j = 1, \dots, n, \tag{4.1}$$

where $n \in \mathbb{N}$ is the number of observations.

Figure 4.4: Pseudo copula data using marginals from the fitted models. The lower triangle contains empirical normalized contour plots. The upper triangle contains pair scatter plot of copula data and empirical Kendall's tau. Lastly, the diagonal contains marginal histograms of copula data.

The lower triangle of Figure 4.4 indicates that there is tail dependence between the assets. From the plots, it is hard to conclude whether the dependence structure is asymmetric or symmetric. The diagonal of Figure 4.4 shows that the marginals are approximately uniform. In the upper triangle of Figure 4.4 the empirical Kendall's tau is given. To select an appropriate R-vine copula one of the selection criteria that will be utilized is the empirical Kendall's tau.

To fit an R-vine copula to the uniform marginals, the function `RVineStructureSelect` from the `R`-package `VineCopula` is used [30]. Additionally, to simulate samples from the fitted R-vine copula the function `RVineSim` from the same `R`-package `VineCopula` is used.

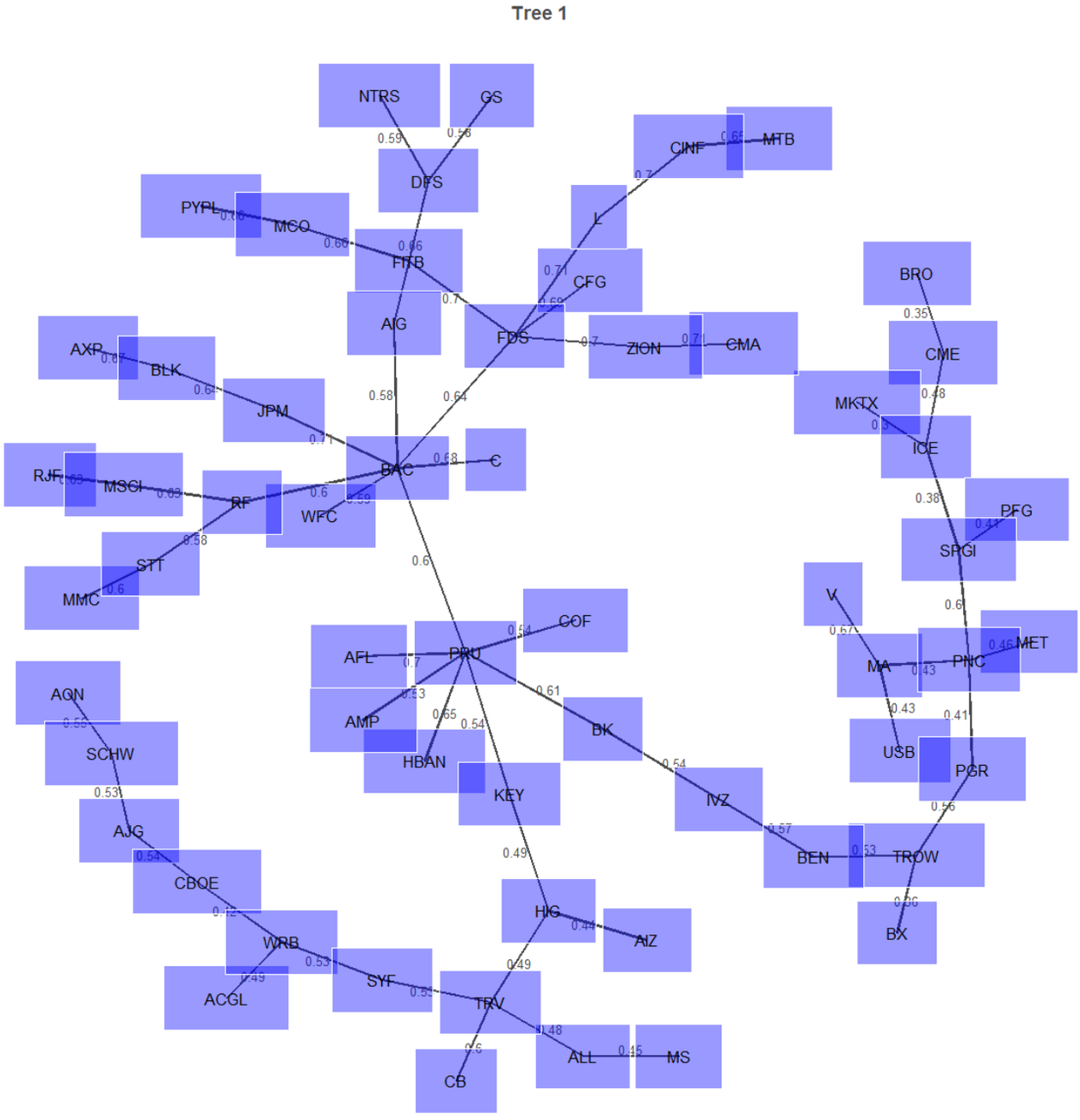


Figure 4.5: The first R-vine tree fitted on the marginals from the financial sector. The values on the edges are the empirical Kendall's tau.

The first tree of the R-vine tree sequence is plotted in Figure 4.5. The weight for each edge is in this case chosen to be the empirical Kendall's tau. It can be inferred that the selected pairs are also the expected pairs from the empirical Kendall's tau in Figure 4.4.

Table 4.7: The number of times each bivariate copula is selected in each tree, where Π denotes the bivariate independence copula.

Tree	Π	N	t	G	F	G 180	G 90	G 270
1	0	0	59	0	0	1	0	0
2	0	0	40	0	19	0	0	0
3	0	1	20	1	34	2	0	0
4	2	0	18	0	35	2	0	0
5	3	3	22	2	24	2	0	0
6	9	2	18	2	14	9	1	0
7	11	3	10	3	26	0	0	1
8	18	2	12	6	11	1	0	3
9	17	6	9	3	15	0	2	0
10	15	3	11	2	14	2	2	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
46	10	2	1	0	1	0	1	0
47	7	1	1	0	3	1	1	0
48	8	1	0	0	3	1	0	0
49	7	0	0	0	2	1	2	0
50	8	0	0	1	1	0	0	1
51	4	1	2	0	1	2	0	0
52	6	0	1	0	0	1	0	1
53	4	0	1	1	1	1	0	0
54	6	0	0	0	1	0	0	0
55	4	1	0	0	0	0	1	0
56	4	0	0	1	0	0	0	0
57	4	0	0	0	0	0	0	0
58	3	0	0	0	0	0	0	0
59	1	1	0	0	0	0	0	0
60	0	0	0	0	1	0	0	0

From Table 4.7 it can be inferred that in the first tree, the only two selected bivariate copulas are the bivariate Student's t-copula and the rotated 180 degrees Gumbel copula. The characteristics of the bivariate Student's t-copula is a symmetric heavy tail structure and the characteristics of the bivariate rotated 180 Gumbel is an asymmetric tail structure with a lower tail dependence. From tree three to tree six the bivariate Frank copula and the bivariate Student's t-copula are the most used pair copulas. The bivariate Frank copula's characteristics are tail asymmetry and no tail dependence. For higher-order trees, the dependence strength in the bivariate copulas decreases, which can explain the transition

from the bivariate Student's t-copula to copulas with no tail dependence, i.e., the bivariate independence copula, the bivariate Gaussian copula, the bivariate Frank copula, and the bivariate rotated 90 and 270 Gumbel copula.

4.3.1 Model Comparison

The R-vine copulas that will be compared in this section, are R-vine copulas constructed with sequential estimation and R-vine copulas estimated with maximum log-likelihood using the sequential estimates as starting values. The maximum number of iteration steps in the maximum log-likelihood is set to 20 to make it computationally feasible. It is therefore not given, that the maximum log-likelihood converges. Furthermore, two different weights in Algorithm 3 will be tested, namely AIC and Kendall's tau. The different R-vine copula models are compared to a multivariate Gaussian copula and a multivariate Student's t-copula. To construct these two multivariate copulas, their bivariate counterparts are used as building blocks in an R-vine copula. The R-vine copulas constructed this way, is equivalent to a multivariate Gaussian copula and a multivariate Student's t-copula, respectively [31, p. 5-6].

The different R-vine copula models are compared on the number of parameters(par), log-likelihood(loglike), AIC, BIC, and the goodness-of-fit test(GOF test) presented in Section 2.6. To apply the GOF test, the vector $d(\theta|U)$ must be estimated. The vector $d(\theta|U)$ has dimension $p(p+1)/2$, where $p = d(d-1)/2$, making the GOF test computationally infeasible for a portfolio consisting of 61, 60, and 71 assets, respectively. The GOF test is therefore used on portfolios consisting of the first 5, 10, and 20 listed assets in the financial sector to test how the GOF test reacts to an increase in the dimension of a portfolio.

Table 4.8: Fitted models on the first 5 listed assets in the financial sector, number of parameters, estimated log-likelihood, AIC, BIC, and GOF test.

Model	par	loglike	AIC	BIC	GOF test (p-value)
R-vine AIC seq	16	3742.065	-7452.129	-7365.847	0.986
R-vine AIC seq mle	16	3742.268	-7452.536	-7366.254	0.988
R-vine tau seq	16	3739.560	-7447.119	-7360.837	0.975
R-vine tau seq mle	16	3742.228	-7452.457	-7366.174	0.989
Gaussian seq	10	3428.217	-6836.434	-6782.508	0.000
Gaussian seq mle	10	3428.217	-6836.435	-6782.508	0.000
Student's t seq	20	3736.766	-7433.532	-7325.679	1.000
Student's t seq mle	20	3737.018	-7434.035	-7326.182	1.000

Table 4.9: Fitted models on the first 10 listed assets in the financial sector, number of parameters, estimated log-likelihood, AIC, BIC, and GOF test.

Model	par	loglike	AIC	BIC	GOF test (p-value)
R-vine AIC seq	65	8201.244	-16272.49	-15921.97	1.000
R-vine AIC seq mle	65	8202.495	-16274.99	-15924.47	1.000
R-vine tau seq	58	8184.734	-16253.47	-15940.69	1.000
R-vine tau seq mle	58	8187.357	-16258.71	-15945.94	1.000
Gaussian seq	45	7538.926	-14987.85	-14745.18	0.574
Gaussian seq mle	45	7538.928	-14987.86	-14745.19	0.576
Student's t seq	90	8172.539	-16165.08	-15679.74	1.000
Student's t seq mle	90	8175.472	-16170.94	-15685.61	1.000

Table 4.10: Fitted models on the first 20 listed assets in the financial sector, number of parameters, estimated log-likelihood, AIC, BIC, and GOF test.

Model	par	loglike	AIC	BIC	GOF test (p-value)
R-vine AIC seq	224	15381.02	-30314.04	-29106.08	1.000
R-vine AIC seq mle	224	15386.80	-30325.60	-29117.65	1.000
R-vine tau seq	196	15309.97	-30227.93	-29170.97	1.000
R-vine tau seq mle	196	15320.87	-30249.75	-29192.79	1.000
Gaussian seq	190	14073.92	-27767.85	-26743.24	1.000
Gaussian seq mle	190	14073.95	-27767.89	-26743.29	1.000
Student's t seq	380	15224.86	-29689.73	-27640.52	Too many parameters to be estimated
Student's t seq mle	380	15242.68	-29725.36	-27676.16	Too many parameters to be estimated

The worst-performing model is the multivariate Gaussian copula. The null hypothesis for the GOF test for the multivariate Gaussian copula is rejected for 5 assets, accepted for 10 assets, and gives the same p-value for 20 assets as the other models. However, the relative difference in both AIC and BIC between the multivariate Gaussian copula and the other models is consistent in the three scenarios, indicating that the multivariate Gaussian copula fits the data worse compared to the other models. The multivariate Student's t-copula gives either a higher or the same p-value in the GOF test compared to the other models. However, the multivariate Student's t-copulas don't achieve the lowest AIC- or BIC values or the highest log-likelihood. Comparing Kendall's tau to the AIC approach, the R-vine copulas using AIC gives a higher log-likelihood and a lower AIC. However, the

R-vine copulas using Kendall's tau give a lower BIC value. Using sequential estimation together with maximum log-likelihood improves every model on every benchmark. The difference between using only sequential estimation and maximum log-likelihood is however insignificant, especially considering the computational time it takes to estimate the R-vine copula using maximum log-likelihood.

Additionally, since it is not possible to perform the GOF test on the entire financial sector portfolio, the GOF test is applied to every bivariate copula in the R-vine copula using the AIC as selection criteria. Firstly, to establish a baseline, the GOF test is performed on the bivariate copulas in the R-vine copula AIC seq, constructed for the first 20 listed assets. The results for first 5 and 10 listed assets in the financial sector can be found in Section D.2.

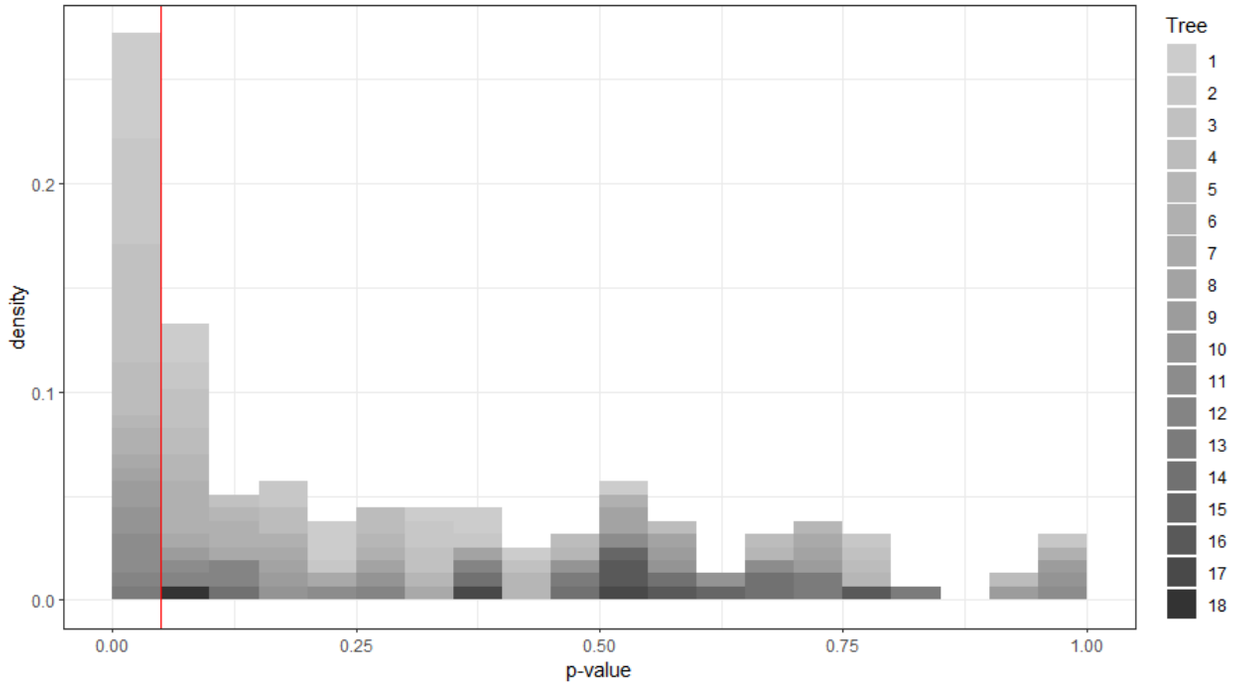


Figure 4.6: GOF test on the bivariate copulas in the R-vine copula AIC seq for the first 20 listed assets in the financial sector.

Even though the p-value for the GOF test for the R-vine AIC seq is 1 in Table 4.10, it is not the case that the null hypothesis is not rejected for every bivariate copula. It also must be noted that for the independence copula, the GOF test is not able to give a p-value. For example, the bivariate copula in the 19th tree in Figure 4.6 is the bivariate independence copula and is therefore not present in the plot.

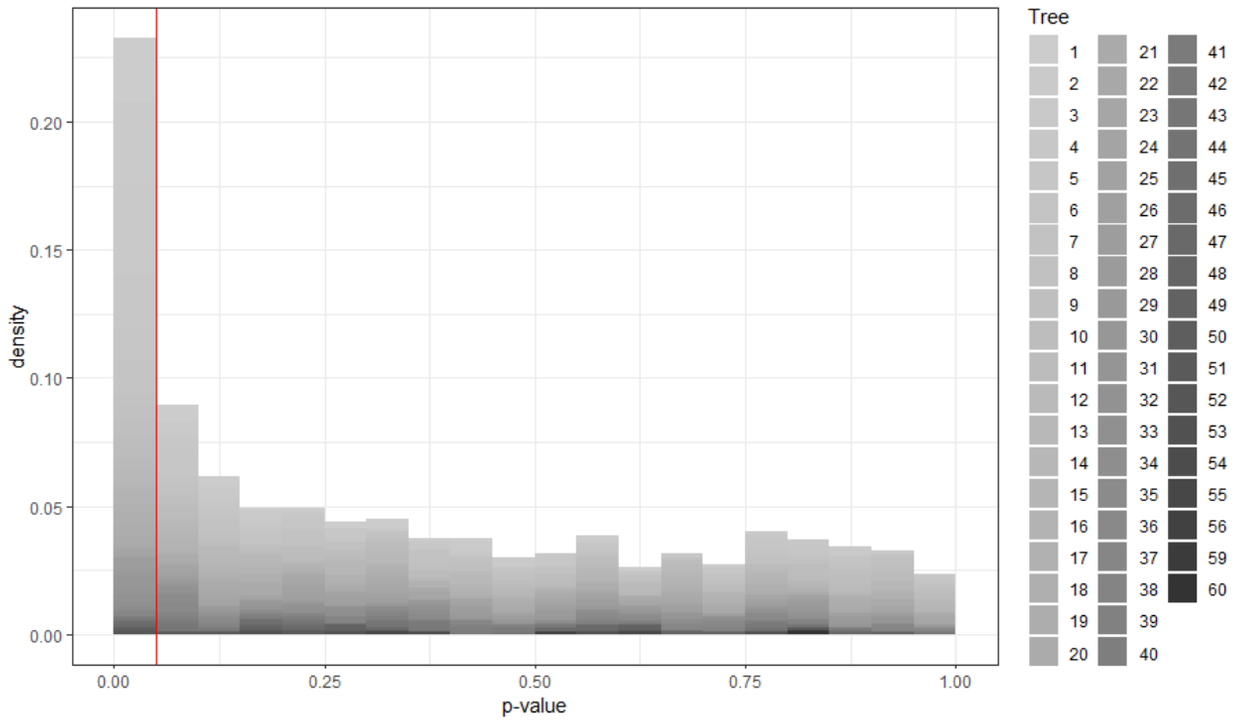


Figure 4.7: GOF test on the bivariate copulas in the R-vine copula AIC seq for every asset in the financial sector.

From Figure 4.7 it can be inferred that the null hypothesis of the GOF test is rejected for approximately 23 percent of the non-independence bivariate copulas. The null hypothesis is not only rejected for the bivariate copulas in the higher order trees where the dependence between the conditional marginals is diminishing, but also for the bivariate copulas in the lower order trees. Comparing Figure 4.7 with Figure 4.6 the overall distributions of p-values are similar, indicating that a GOF test on the R-vine copula AIC seq would not have rejected the null hypothesis. The GOF test is also performed on the bivariate copulas used to create the multivariate Gaussian copula. The results can be found in Section D.2. In Table 4.8, the null hypothesis of the GOF test is rejected for the first 5 assets in the financial sector for the multivariate Gaussian copula. The null hypothesis is rejected for 80% of the bivariate copulas in the multivariate Gaussian copula for the first 5 assets, confer Figure D.3. Comparing the multivariate Gaussian copula with the R-vine copula AIC, the rejection rate is significantly higher for the bivariate copulas in the multivariate Gaussian copula compared to bivariate copulas in the R-vine copula AIC which is consistent with the results in Table 4.8, Table 4.9, and Table 4.10.

Based on the results from the log-likelihood, AIC, BIC, and GOF test, the R-vine copula AIC seq will be used in the rest of the application part of this project.

4.4 Out-of-Sample Risk Forecasting

The performance of one-day-ahead $(C)VaR$ forecasts using the R-vine copulas fitted in Section 4.3, will in this section be evaluated using the tests presented in Section 3.2. To obtain one-day-ahead $(C)VaR$ forecasts the following procedure is used.

1. Sample data and calculate standardized daily returns.
2. Fit ARMA-GARCH models for the residuals. The ARMA-GARCH models are refitted every timestep using a one-step rolling window procedure. The order is however kept fixed. The one-step rolling window procedure removes the first observation in the dataset and adds the observation observed at t, thus keeping the length of the dataset the same throughout the procedure.
3. Obtain uniform marginals from the cumulative marginal distribution function i.e. $\hat{u}_{t,j} = F_{t,j}(\hat{z}_{t,j})$ for $t = T+1, \ldots, T+O$ and $j = 1, \ldots, d$, where O is the number of out-of-sample forecasts.
4. Fit an R-vine copula model using Dißmans algorithm. The R-vine copula is refitted every 21 trading day.
5. Simulate M uniform random samples from the R-vine copula fitted in step 4.
6. Transform the M samples using the inverse marginal distribution for each asset and get the standardized residuals.
7. Use the standardized residuals to compute the return forecasts.

$$\hat{r}_{t,j} = \hat{\mu}_{t,j} + \hat{\sigma}_{t,j} F_{t,j}^{-1}(\text{residual}_{t,j}), \quad t = T+1, \ldots, T+O, \ j = 1, \ldots d, \tag{4.2}$$

   where $\hat{\mu}$ and $\hat{\sigma}$ is the forecasted mean and volatility from the ARMA-GARCH model, respectively.
8. Based on the return forecasts, calculate equally-weighted portfolios.

$$\hat{r}_{t,m} = \frac{1}{d} \sum_{j=1}^{d} \hat{r}_{t,j,m}, \quad t = T+1, \ldots, T+O, \ m = 1, \ldots, M. \tag{4.3}$$

9. Estimate the α-quantile for the daily portfolio $(C)VaR$ based on the daily distributions of the simulated portfolio returns

$$VaR_\alpha(-\hat{r}_t) = F_{-\hat{r}_t}^{-1}(\alpha), \tag{4.4}$$

$$CVaR_\alpha(-\hat{r}_t) = \text{mean}\left(-\hat{r}_{t,m} > VaR_\alpha(-\hat{r}_t)\right). \tag{4.5}$$

The number of Monte Carlo simulations is set to $M = 10.000$. The procedure has also been tested for $M = 5.000$ and $M = 20.000$. However, when considering the trade-off

between precision and computational time, $M = 10.000$ was chosen. The VaR forecast procedure is applied to every sector for $\alpha = 0.025, 0.05,$ and 0.10. First, the VaR forecasts are compared between the sectors for $\alpha = 0.05$, and afterward, the performance of the VaR forecasts for each sector will be evaluated.

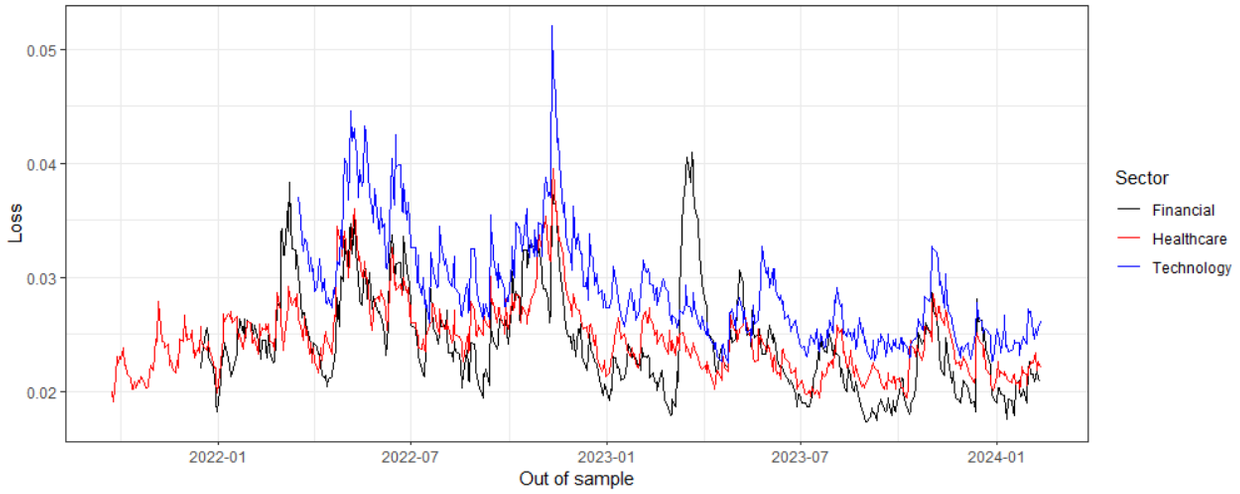


Figure 4.8: Comparison between the 95% quantile VaR forecasts for each sector.

From Figure 4.8 it can be inferred that the VaR forecasts for the technology sector is generally higher than the two other sectors. Furthermore, comparing the VaR forecasts for the financial and the healthcare sectors, the financial sector has greater fluctuations but also has the most periods with the lowest VaR.

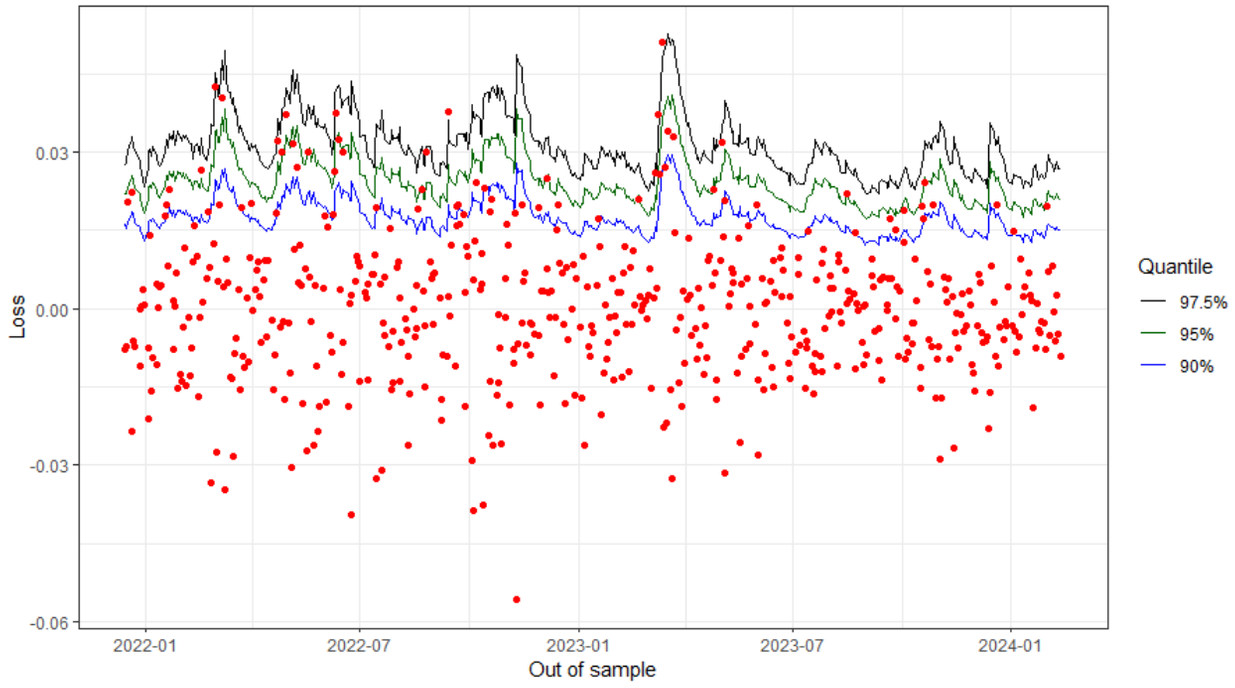


Figure 4.9: VaR forecasts for an equally-weighted portfolio consisting of the assets in the financial sector. The red dots are the actual losses for the equally-weighted portfolio.

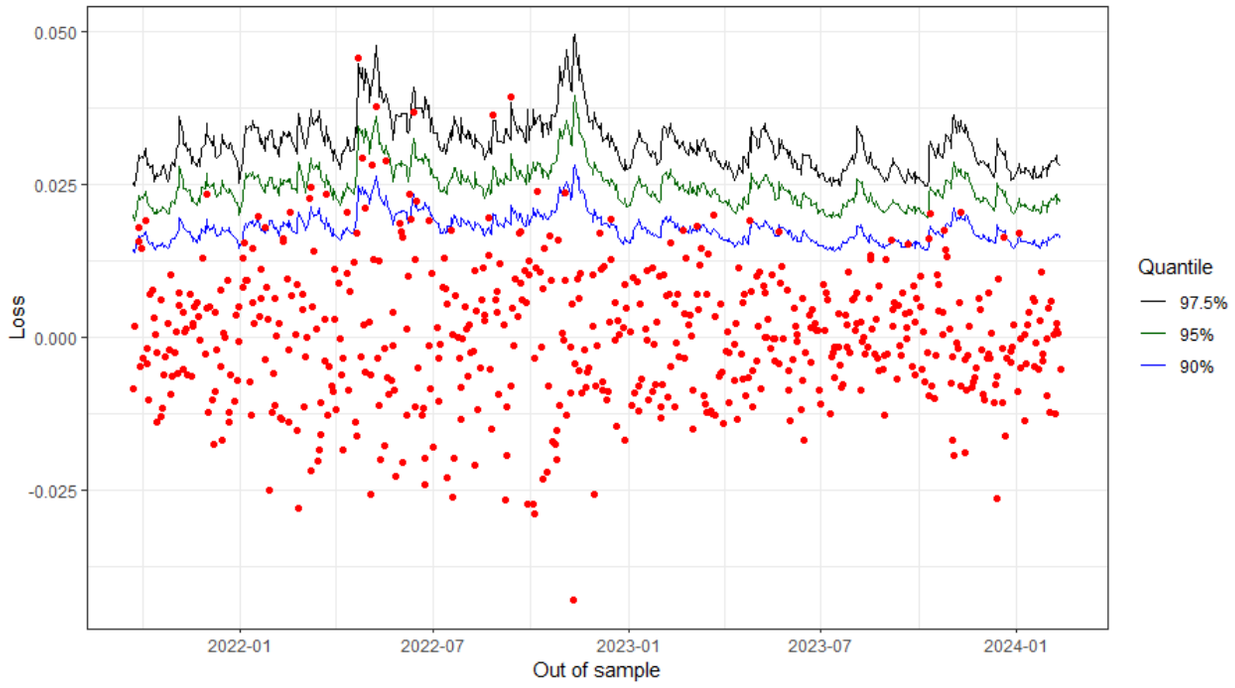


Figure 4.10: VaR forecasts for an equally-weighted portfolio consisting of the assets in the healthcare sector. The red dots are the actual losses for the equally-weighted portfolio.

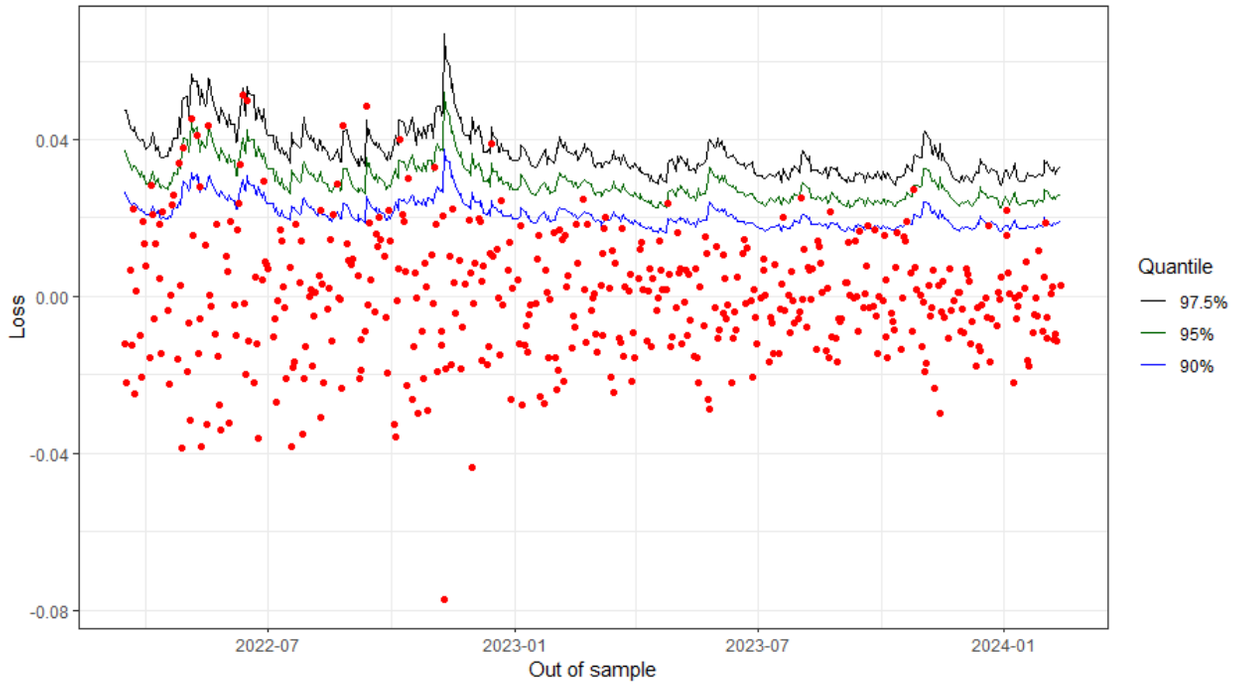


Figure 4.11: VaR forecasts for an equally-weighted portfolio consisting of the assets in the technology sector. The red dots are the actual losses for the equally-weighted portfolio.

In theory, the loss should exceed the VaR forecasts 2.5%, 5%, and 10% of the total observations in the out-of-sample period. From Figure 4.9, Figure 4.10, and Figure 4.11 it can be difficult to conclude whether this is the case for all three α-levels. However, it can for all three sectors be noted that many observations are on the edge of being counted as an αVaR violation.

Table 4.11: Observed αVaR violations compared to the expected number of αVaR violations, and the ratio between observed αVaR violations and the total number of observations for each tested α-level in each sector.

Sector	α-level	Observed violations	Expected violations	Mean violations
Financial	0.025	3	13.550	0.006
	0.05	13	27.100	0.024
	0.10	55	54.200	0.101
Healthcare	0.025	3	15.000	0.005
	0.05	5	30.000	0.008
	0.10	31	60.000	0.052
Technology	0.025	2	11.975	0.004
	0.05	8	23.950	0.017
	0.10	33	47.900	0.069

From Table 4.11 it can be noted that the only case where the number of observed VaR violations is equal to the expected number of VaR violations is in the financial sector for $\alpha = 0.10$. In every other case, the VaR forecasts are too conservative. To validate the performance of the VaR forecasts, the DQ test presented in Section 3.2 is used.

Table 4.12: DQ test performed on the three sectors for $\alpha = 0.025, 0.05,$ and 0.10 and lag $p = 4$ in each sector.

Sector	DQ 2.5% VaR p-value	DQ 5% VaR p-value	DQ 10% VaR p-value
Financial	0.00110	0.00226	0.000736
Healthcare	0.0954	0.000316	0.0000256
Technology	0.260	0.000729	0.00228

For the financial sector, the null hypothesis of the DQ test is rejected for every α-level. Even though the observed $0.10VaR$ violations are equal to the expected $0.10VaR$ violations the null hypothesis is rejected, which may imply that the VaR violations are correlated for lag $p = 4$. For the financial sector it is therefore tested what impact the choice of lag in the DQ test has on the null hypothesis.

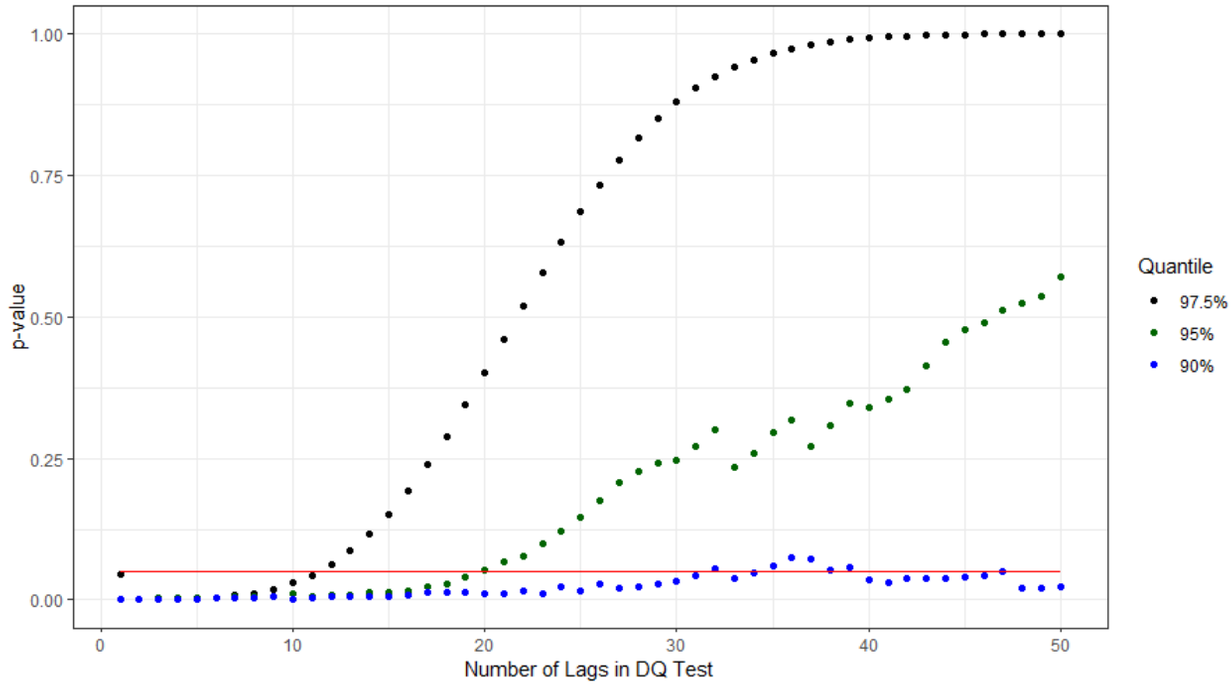


Figure 4.12: The DQ test performed on the VaR forecasts for the financial sector for different lags.

The null hypothesis for the DQ test is not rejected for lag greater than 11 for $\alpha = 0.025$. The consequence of a higher choice of lag is that as the number of parameters in the linear regression model, (3.14), increases, the parameter values get closer to zero making it more likely that the null hypothesis is not rejected. Regardless, the null hypothesis for $\alpha = 0.10$ is only not rejected for lag $p = 32, 35, 36, 37, 38, 39$, and rejected again after $p = 40$ implying that the $\widehat{VaR}_{0.10}$ is not a valid risk measure.

For the healthcare and technology sectors, the null hypothesis is not rejected for $\alpha = 0.025$. However, this is not the case for $\alpha = 0.05$ and 0.10. The overall performance of the VaR forecasts is therefore not good. As mentioned earlier many observations are close to being an αVaR violation. It will therefore be tested whether modeling the marginals with an ARMA$(1,1)$-GARCH$(1,1)$ model yields a better result.

Table 4.13: Observed αVaR violations compared to the expected number of αVaR violations, and the ratio between observed αVaR violations and the total number of observations for each tested α-level in each sector, where the marginals are modeled using an ARMA$(1,1)$-GARCH$(1,1)$.

Sector	α-level	Observed violations	Expected violations	Mean violations
Financial	0.025	6	13.550	0.011
	0.05	19	27.100	0.035
	0.10	60	54.200	0.111
Healthcare	0.025	3	15.000	0.005
	0.05	5	30.000	0.008
	0.10	35	60.000	0.058
Technology	0.025	2	11.975	0.004
	0.05	13	23.950	0.027
	0.10	38	47.900	0.079

Modeling every marginal with an ARMA$(1,1)$-GARCH$(1,1)$ has increased the overall number of αVaR violations. This is especially the case for $\alpha = 0.10$ where many observations are on the edge of being a $0.10VaR$ violation in Figure 4.9, Figure 4.10, and Figure 4.11. However, it must be noted that the two different approaches to selecting the marginals give very similar VaR forecasts. To illustrate how close the VaR-forecasts for the two approaches are to each other, the mean square error(MSE) given by,

$$MSE_\alpha = \frac{1}{O} \sum_{t=T+1}^{T+O} \left(VaR_\alpha \left(-\hat{r}_t^{\text{ approach 1}} \right) - VaR_\alpha \left(-\hat{r}_t^{\text{ approach 2}} \right) \right)^2, \tag{4.6}$$

is calculated.

Table 4.14: The MSE between the VaR forecasts using the approach introduced at the beginning of this section and VaR forecasts where every marginal is modeled using ARMA$(1,1)$-GARCH$(1,1)$ for each sector.

Sector	2.5%	5%	10%
Financial	$1.12 \cdot 10^{-6}$	$7.28 \cdot 10^{-7}$	$5.21 \cdot 10^{-7}$
Healthcare	$1.23 \cdot 10^{-6}$	$6.23 \cdot 10^{-7}$	$3.84 \cdot 10^{-7}$
Technology	$1.62 \cdot 10^{-6}$	$1.11 \cdot 10^{-6}$	$7.65 \cdot 10^{-7}$

From Table 4.14 it can be inferred that the difference between the two approaches is almost negligible. However, the small difference between the two approaches results in increasements in the p-values for the DQ test for every α-level across the three sectors.

Table 4.15: DQ test, where the marginals are modeled using an ARMA$(1,1)$-GARCH$(1,1)$ model, performed on the three sectors for $\alpha = 0.025, 0.05,$ and 0.10 and lag $p = 4$.

Sector	DQ 2.5% VaR p-value	DQ 5% VaR p-value	DQ 10% VaR p-value
Financial	0.0000551	0.00258	0.0959
Healthcare	0.109	0.000609	0.00368
Technology	0.288	0.0167	0.437

The null hypothesis of the DQ test is no longer rejected for $\alpha = 0.10$ for both the financial and technology sectors. Only modeling the marginals with an ARMA$(1,1)$-GARCH$(1,1)$ model therefore gives a better result in the examined time period.

The validity of the $CVaR$ forecasts is also examined using the test presented in Subsection 3.2.2. The VaR forecasts for each sector is too conservative in their estimates of VaR. The $CVaR$ forecasts are also too conservative.

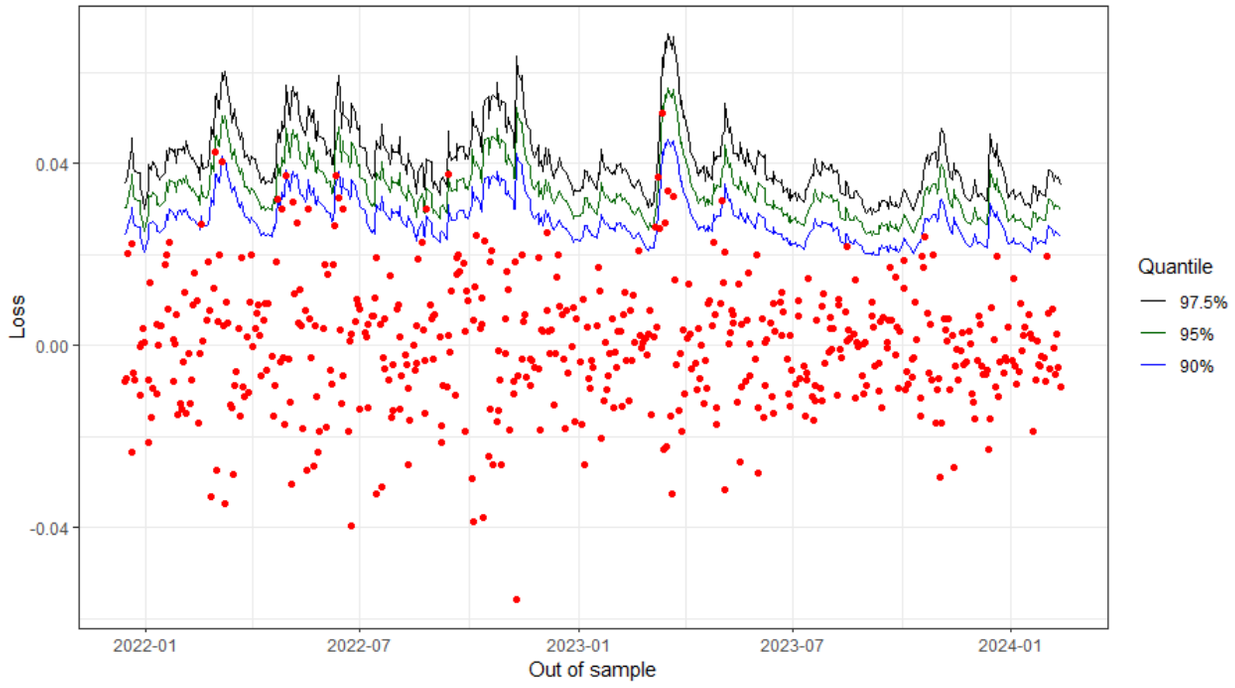


Figure 4.13: $CVaR$ forecasts for an equally-weighted portfolio consisting of the assets in the financial sector. The red dots are the actual losses for the equally-weighted portfolio.

Table 4.16: Strict $CVaR$ backtest performed on the three sectors for $\alpha = 0.025, 0.05,$ and 0.10.

Sector	2.5% CVaR p-value	5% CVaR p-value	10% CVaR p-value
Financial	0	0	$1 \cdot 10^{-7}$
Healthcare	0	0	0
Technology	0	0	0

From Table 4.16, it can be concluded that the null hypothesis of the strict $CVaR$ backtest is rejected in every sector and for any α-level. The ARMA-GARCH R-vine copula approach is therefore not able to produce a valid $CVaR$ forecast. This is also the case when the marginals are modeled using only ARMA$(1,1)$-GARCH$(1,1)$, confer Table 4.17. For the two approaches, the VaR forecasts are similar, and hence, it is to be expected that their $CVaR$ forecasts perform similar.

Table 4.17: Strict $CVaR$ backtest, where the marginals are modeled using an ARMA$(1,1)$-GARCH$(1,1)$ model, performed on the three sectors for $\alpha = 0.025, 0.05,$ and 0.10.

Sector	2.5% CVaR p-value	5% CVaR p-value	10% CVaR p-value
Financial	0	0	0
Healthcare	0	0	0
Technology	0	0	0

As inferred from Figure 4.1 the price tendencies change over time, which might influence the model's performance. Therefore, it is tested in Appendix E for the financial sector, whether fitting and testing the model on a time period with the same tendencies yields a better result. The results from the ratio between observed- and expected violations and the backtests strongly imply that the model performs better in the shorted period.

4.5 Mean-*CVaR* Portfolio Allocation

The purpose of this section is to formulate a mean-$CVaR$ portfolio allocation problem and execute the strategy on the presented data. Let w_t be the weight vector at time t, and consider the auxiliary variable $\nu_{t,m} = \left[-w_t^\top r_{t,m} - VaR_\alpha\right]^+$, which are the losses greater than VaR_α, for a given α-level. Then, as proposed in [32], it is possible to formulate a convex linear programming problem to solve the mean-$CVaR$ minimization problem:

$$\min_{w,VaR,\nu} \quad VaR_\alpha + \frac{1}{M(1-\alpha)} \sum_{m=1}^{M} \nu_{t,m}, \tag{4.7}$$

$$\text{subject to} \quad w_t^\top r_{t,m} + VaR_\alpha + \nu_{t,m} \geq 0, \quad \forall m \in \{1, \ldots, M\}, \tag{4.8}$$

$$w_t^\top 1 = 1, \tag{4.9}$$

$$0 \leq w_{t,i} \leq 1, \text{ for } i = 1, \ldots, d, \tag{4.10}$$

$$w_t^\top r_t \geq \mu, \tag{4.11}$$

where d is the number of assets in the portfolio and μ is a predetermined return level. If the goal is to simply minimize the $CVaR$ for a portfolio, the last constraint, (4.11), is omitted.

The mean-$CVaR$ problem will be executed for each of the sectors, which are then compared. The problem will be solved using the `minCVaR`-function from the `NMOF` package in `R` [33]. The analysis will be done for an α-level of 0.05, and the predetermined return level will be set at 0.001. The accumulated return of the portfolio will be compared to the accumulated return for an equally-weighted portfolio, and the amount of assets which outperforms the strategy in terms of return will be investigated as well. Furthermore, the VaR and $CVaR$

for the portfolios will be calculated based on the historical quantiles, which again will be compared to an equally-weighted portfolio. The analysis is done using the forecasts from Section 4.4.

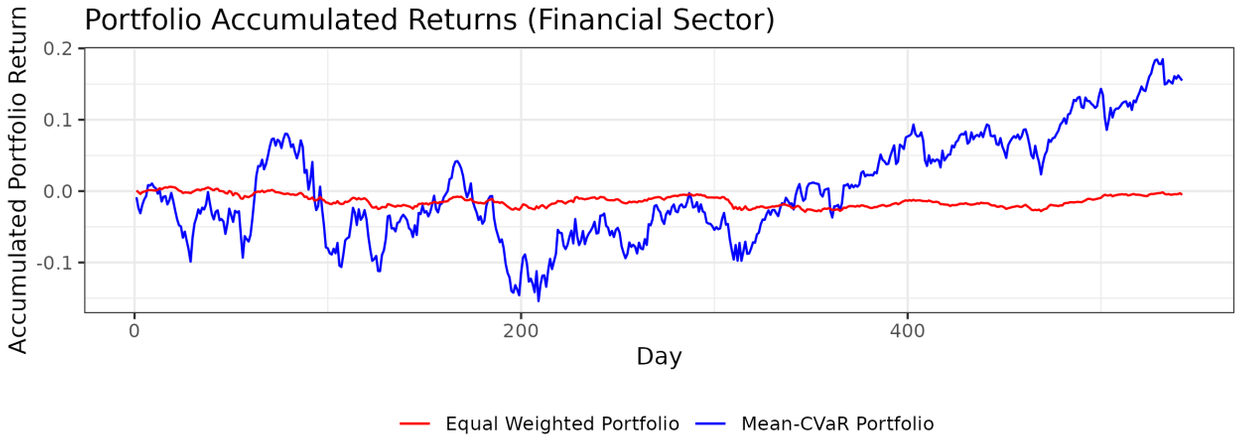


Figure 4.14: Accumulated portfolio returns for the mean-$CVaR$ portfolio (blue) and equally-weighted portfolio (red) for the financial sector.

From Figure 4.14 it can be concluded, that the mean-$CVaR$ portfolio outperforms the equally-weighted portfolio in terms of return for the financial sector. The return for the mean-$CVaR$ portfolio strategy is 0.155 and the return for the equally-weighted portfolio is -0.00525.

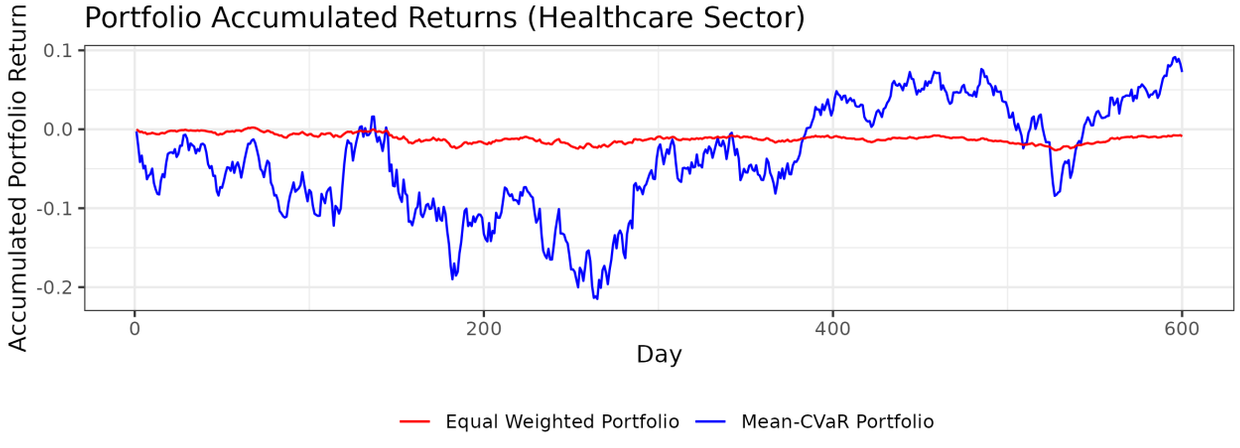


Figure 4.15: Accumulated portfolio returns for the mean-$CVaR$ portfolio (blue) and equally-weighted portfolio (red) for the healthcare sector.

From Figure 4.15 it can be concluded, that the mean-$CVaR$ portfolio outperforms the equally-weighted portfolio in terms of return for the healthcare sector. The return for the

mean-$CVaR$ portfolio strategy is 0.0725 and the return for the equally-weighted portfolio is -0.00900.

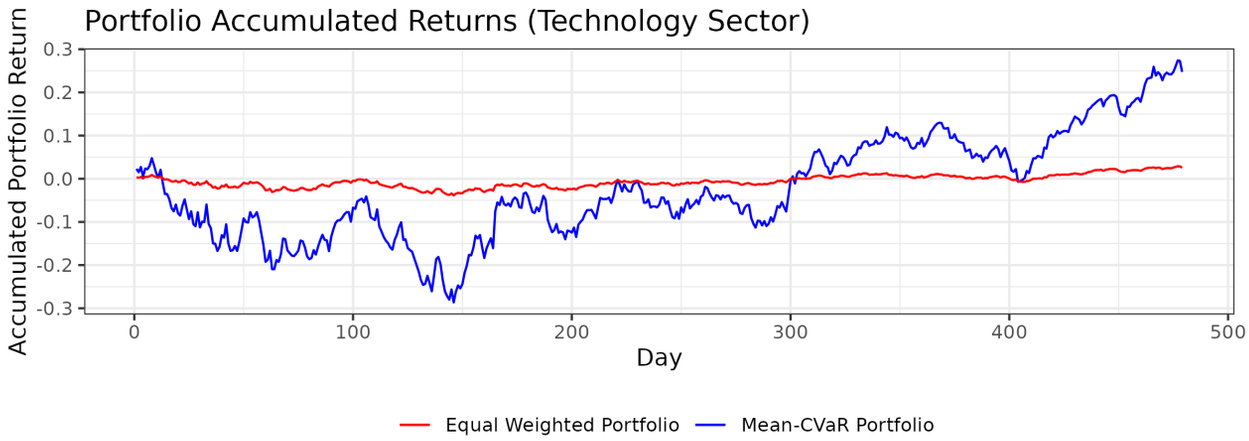


Figure 4.16: Accumulated portfolio returns for the mean-$CVaR$ portfolio (blue) and equally-weighted portfolio (red) for the technology sector.

From Figure 4.15 it can be concluded, that the mean-$CVaR$ portfolio outperforms the equally-weighted portfolio in terms of return for the technology sector. The return for the mean-$CVaR$ portfolio strategy is 0.247 and the return for the equally-weighted portfolio is 0.0257. For all three sectors, the mean-$CVaR$ strategy performs better in the last half of the trading period, compared to the first half. These performances fit well with the price-plots in Figure 4.1, as the prices for all three sectors decline in the first period of the test dataset.

The returns for each of the three portfolios are also compared to buy-and-hold strategies in each individual asset. The outperforming assets (in terms of return) in each sector and their respective return are given in Table 4.18.

Table 4.18: Tickers and returns for the outperforming assets in each sector.

Financial		Healthcare		Technology	
Ticker	Return	Ticker	Return	Ticker	Return
V	0.262	LLY	1.163	MSFT	0.319
MA	0.285	UNH	0.236	NVDA	1.069
AXP	0.257	MRK	0.541	AVGO	0.731
PGR	0.581	ABBV	0.479	AMD	0.429
CB	0.241	AMGN	0.301	CRM	0.290
AJG	0.344	ISRG	0.0842	ADBE	0.303
AIG	0.254	SYK	0.207	ORCL	0.344
TRV	0.310	VRTX	0.810	INTU	0.298
AFL	0.308	REGN	0.375	IBM	0.362
AMP	0.268	BSX	0.367	AMAT	0.312
ALL	0.334	CI	0.511	NOW	0.312
ACGL	0.649	MCK	0.913	LRCX	0.528
HIG	0.287	HCA	0.176	KLAC	0.598
RJF	0.162	CNC	0.203	PANW	0.672
CBOE	0.355	CAH	0.709	SNPS	0.590
BRO	0.173	AXON	0.403	CDNS	0.652
WRB	0.397	MOH	0.333	ANET	0.745
L	0.246	UHS	0.106	APH	0.306
				MSI	0.324
				IT	0.431
				ON	0.257
				MPWR	0.456
				CDW	0.297
				FTV	0.302
				FICO	0.972
				PTC	0.474
				HUBB	0.628
				JBL	0.803
				FSLR	0.667

The number of outperforming assets is 18, 18, and 29 for financial, healthcare, and technology, respectively. As the number of total assets in each sector is 61, 60, and 71 for financial, healthcare, and technology, respectively, the number of outperforming assets in each sector is less than a third for financial and healthcare, and less than half for technology.

Besides comparing the strategies on return alone, the VaR and $CVaR$ are calculated for both mean-$CVaR$ strategies and equally-weighted portfolios. This is done by finding quantiles of the observed returns from the forecasted period.

Table 4.19: VaR and $CVaR$ for the mean-$CVaR$ strategy for the three sectors, calculated on historical observations with $\alpha = 0.05$, compared to equally-weighted portfolios for the three sectors.

	Financial	Healthcare	Technology
VaR	0.0187	0.0187	0.0252
CVaR	0.0282	0.0252	0.0333
VaR Equal Weight	0.00257	0.00182	0.00349
CVaR Equal Weight	0.00352	0.00248	0.00516

The results in Table 4.19 show that the VaR and $CVaR$ for the mean-$CVaR$ portfolio strategy are around 10 times higher than the equally-weighted portfolio for financial and healthcare, and around seven times higher for technology. This corresponds with the plots of the accumulated portfolios in Figure 4.14, Figure 4.15, and Figure 4.16, as the values throughout the period for the mean-$CVaR$ portfolios fluctuates more than the equally-weighted portfolios. Furthermore, the result of a higher $(C)VaR$ also aligns well with the fact that our $(C)VaR$ forecasts are not good, and hence, the minimization problem uses bad estimates.

The strategy has also been tested with different levels of μ, which is shown in Figure F.1. The strategy is executed on the assets in the financial sector, with three different levels of return: $\mu = 0, 0.001$ and 0.005. The strategy performs best, in terms of return, for $\mu = 0.001$, which is also the one chosen in this project. When μ gets too high, the strategy is not able to reach the given level of return, which results in a worse performance - both in terms of return and $CVaR$.

5 | Discussion

In the construction of the R-vine copula the simplifying assumption was made, to make inference possible for high dimensions. In [34], it is tested whether the assumption is useful or might be too simplistic. It is concluded that even though the actual model is far away from the simplified model, it still yields a rather good approximation.

The choices of selection criteria for fitting the R-vine copula using Dißmann's algorithm have in this project been the AIC and empirical Kendall's tau. The AIC is chosen based on the result mentioned in [4, p. 157], where Czado refers to the AIC selection approach being superior to three other approaches based on a large-scale Monte Carlo study. The other three approaches are: Choosing the highest p-value based on the Cramér-von-Mises test, the smallest distance between Kendall's tau and tail dependence, and the best bivariate copula based on a Vuong test, which is explained further in [4, chap. 9]. The empirical Kendall's tau approach was also tested since it is more computationally feasible compared to the other approaches [35, p. 19]. Another choice of selection criterion could have yielded better results.

The choice of estimation method is, as mentioned above, a frequentist approach in terms of Dißmann's algorithm. Another approach is a Bayesian model selection approach, which is discussed in [36]. In this paper, Czado and Gruber tests a Bayesian model selection against Dißmann's algorithm. They conclude that the Bayesian selection approach is superior in model selection, and that the results in terms of risk forecasting and investment decisions are superior as well. Using the Bayesian approach in this project, could therefore have yielded better results, based on the conclusion from Czado and Gruber.

In this project the marginals are modeled using an ARMA-GARCH model. The innovations have the possibility to follow either a standardized Gaussian-, standardized Student's t-, or standardized NIG distribution. Compared to only using the standardized Gaussian distribution, the standardized Student's t- and the standardized NIG distribution makes it possible to model heavy tails and skewness, which are the characteristics of asset returns. Based on the Anderson-Darling tests made in Section 4.2, the fits for the marginals yield good results, as the fitted distributions are accepted for all assets. Including more distributions for the innovations, could have resulted in better fits. It was tested in Section 4.4 whether an ARMA$(1,1)$-GARCH$(1,1)$ model performed better than the one chosen by the selection approach in Section 4.2. The ARMA$(1,1)$-GARCH$(1,1)$ yielded a better result in the DQ test. Therefore, it would be reasonable to try other methods to model the marginals, as it was possible to achieve better results by simply changing the order of the ARMA-GARCH model.

To determine whether the VaR forecasts in this project were valid, the DQ test was used. In Figure 4.12 it can be observed that the choice of lag has a big impact on whether the null hypothesis was rejected or not rejected. In [23] the power of the DQ test is examined. The paper found out that the larger the number of forecasts, based upon parameters with

estimation error, the more the backtest will wrongly reject the null hypothesis. Additionally, the paper concluded that the DQ test exhibits low power in their application. This is in line with the results in this project where a different model for marginals changes the results from the DQ test even though the difference in the VaR forecasts was negligible.

The amount of assets chosen in this project is 60, 61, and 71 for the sectors, respectively. The results for those amounts of assets were not satisfying. A GOF test was performed on a portfolio consisting of 5, 10, and 20 assets in the financial sector. The null hypothesis of the GOF test was not rejected for these amounts of assets, indicating that the R-vine copula is capable of modeling the joint distribution for a portfolio consisting of less than 20 assets. Applying the $(C)VaR$ procedure in a portfolio consisting of fewer assets could therefore have led to better results. For example, a similar study has achieved good results with a portfolio consisting of 10 assets [37].

From Figure 4.1, it can be inferred that the data does not follow the same trend throughout the whole period. When COVID-19 arrived at the start of 2020, there was a big downward spike in all sectors, followed by a steep upward trend, especially for healthcare and technology. Then the conflict between Russia and Ukraine started in 2022, which resulted in a drop in all sectors. The change in trends might have influenced the model. Therefore, it was also tested in Appendix E whether a period with similar tendencies in both the training and testing data would perform better in the financial sector. From Appendix E it can be inferred that this is the case, since the results from the ratio between observed- and expected violations and the backtests strongly imply that the model performs better in the shorted period.

6 | Conclusion

The purpose of this project was to explore the capabilities of an R-vine copula, where the marginals were modeled using an ARMA-GARCH model, to forecast the risk of a portfolio. The risk measures chosen to model the risk of a portfolio were VaR and $CVaR$. The $(C)VaR$ were forecasted and validated on three portfolios, consisting of data from different sectors from the S&P 500.

In Chapter 2, theory regarding copulas, including Sklar's Theorem and different copula families were introduced. Then the pair copula construction of a multivariate distribution was proposed, which gave the opportunity to construct a multivariate distribution based on bivariate copulas, so-called pair copulas. To graphically construct the multivariate distribution using pair copulas, vines were introduced. Vines are sequences of trees, which, with the pair copulas as edges, constructs R-vine copulas. The estimation of an R-vine copula can be achieved using sequential estimation, and an algorithm for selecting an R-vine copula was proposed. Dißmann's algorithm is a greedy algorithm, which sequentially chooses the best tree in the R-vine tree sequence. The selection is based on a weight, which in this project was either the AIC or the absolute empirical Kendall's tau. To test the fit of the model, a GOF test, which is based on the Bartlett identity, was shown. Lastly, theory regarding modeling of the marginals was presented. The ARMA-GARCH model for modeling time series data was used, and three different possibilities for distributions of the innovations were chosen: the standardized Gaussian-, the standardized Student's t-, and the standardized NIG distribution.

In Chapter 3, theory regarding risk measures was introduced. VaR and $CVaR$ were defined, and properties were discussed. Estimation using the Monte Carlo approach was chosen for the risk measures, and backtests of VaR and $CVaR$ were introduced. For VaR, the Dynamic Quantile test was chosen and for $CVaR$, a strict $CVaR$ regression test was chosen.

In Chapter 4, the application part of this project was presented. Assets from S&P 500 were spread into 3 sectors; finance with 61 assets, healthcare with 60 assets, and technology with 71 assets. The log-returns for each asset were characterized and examined to fit an ARMA-GARCH model. The fits of the marginals were tested using the Anderson-Darling and Ljung-Box tests. The tests implied that all marginal models fitted the data well. The analysis therefore proceeded to fit an R-vine copulas for each sector. The estimation and selection were done with Dißman's algorithm, where the AIC and the absolute empirical Kendall's tau were chosen as the weights, and eight different bivariate copulas were chosen for possible pair copulas. The last part of the in-sample analysis was a goodness of fit test and model comparison. The comparison showed that the R-vine copulas - both using the AIC and absolute empirical Kendall's tau - yielded better AIC and BIC values compared to multivariate Gaussian- and Student's t-copulas.

Lastly, the performance of the models was validated with out-of-sample risk forecasting.

The *(C)VaR* forecasts were investigated for three different α-levels, namely: 0.025, 0.05, and 0.10. In every sector, the *VaR* forecasts were all conservative, as the number of observed violations did not reach the number of expected violations. The Dynamic Quantile test was only not rejected for $\alpha = 0.025$ for technology and healthcare. To test whether the choice of lags of the ARMA-GARCH models made a difference, an ARMA$(1,1)$-GARCH$(1,1)$ model was fitted for all marginals, which improved the results of the Dynamic Quantile test. The *CVaR* backtest yielded p-values of zero for all α-levels for all sectors, which inferred that the models were not good at forecasting $CVaR$. A mean-$CVaR$ portfolio strategy was constructed for all three sectors and compared to equally-weighted portfolios. For all sectors, the mean-$CVaR$ portfolio performed better in terms of return, but gave larger VaR and $CVaR$ values, than the equally-weighted portfolios.

7 | Bibliography

[1] Pearson K. III. Contributions to the Mathematical Theory of Evolution. University College, London. 1894;.

[2] Fees E, Hege U. The Basel II Accord: Internal ratings and bank defferentiation. Center for Financial Studies (CFS); 2004. 2004/25. Available at `https://ideas.repec.org/p/zbw/cfswop/200425.html`.

[3] Sahamkhadam M, Stephan A. Portfolio optimization based on forecasting models using vine copulas: An empirical assessment for global financial crises. Journal of Forecasting. 2023;42:2139–2166.

[4] Czado C. *Analyzing Dependent Data with Vine Copulas*. Springer; 2019.

[5] Dißmann J, Brechmann EC, Czado C, Kurowicka D. Selecting and estimating regular vine copulae and application to financial returns. Computational Statistics and Data Analysis. 2012;59:52–69.

[6] Nielsen RB. *An Introduction to Copulas*. Springer; 2005.

[7] Borchersen VB, Kellmer AT, Jensen RK, Mortensen NH. Pairs Trading: A Copula and Lévy Process Approach. Aalborg University; 2023.

[8] Durante F, Fernández-Sánchez J, Sempi C. *A topological proof of Sklar's theorem*. Applied Mathematics Letters; 2013.

[9] Aas K, Czado C, Frigessi A, Bakken H. Pair-copula constructions of multiple dependence. ScienceDirect. 2009;44:182–198.

[10] Bedford T, Cooke RM. Probability density decomposition for conditionally dependent random variables modeled by vines. Annals of Mathematics and Artificial Intelligence. 2001;32:245–268.

[11] Sigman K. *Inverse Transform Method*. Columbia University; 2010. Available at `https://www.columbia.edu/~ks20/4404-Sigman/4404-Notes-ITM.pdf`.

[12] Dißmann JF. *Statistical Inference for Regular Vines and Application* [Paper]. Technische Universität München; 2010 [last visited 25/03/2024]. Available at `https://mediatum.ub.tum.de/doc/1079308/document.pdf`.

[13] Cormen TH, Leiserson CE, Rivest RL, Stein C. *Introduction to Algorithms*. The MIT Press; 2009.

[14] Huang W, Prokhorov A. A Goodness-of-fit Test for Copulas. Econometric Reviews. 2014;33(7):751–771.

[15] Schepsmeier U. A goodness-of-fit test for regular vine copula models. Econometric Reviews. 2013 06;38.

[16] White H. Maximum Likelihood Estimation of Misspecified Models. Econometrica. 1982 01;50(1).

[17] Genest C, Rémillard B, Beaudoin D. Goodness-of-fit tests for copulas: A review and a power study. Insurance: Mathematics and Economics. 2009;44:199–213.

[18] Würtz D, Chalabi Y, Luksan L. Parameter Estimation of ARMA Models with GARCH/APARCH Errors An R and SPlus Software Implementation. Journal of Statistical Software. 2021;vol. 49,11:2964–2980. Available at `https://www.math.pku.edu.cn/teachers/heyb/TimeSeries/lectures/garch.pdf`.

[19] van de Wiel DP. *Valuation of insurance products using a Normal Inverse Gaussian distribution* [Paper]. Tilburg University; 2005 [last visited 18/04/2024]. Available at `https://arno.uvt.nl/show.cgi?fid=146193`.

[20] Cont R. Empirical properties of asset returns: stylized facts and statistical issues. QUANTITATIVE FINANCE. 2001;1:223–236.

[21] McNeil AJ, Frey R, Embrechts P. Quantative Risk Management: Concepts, Techniques and Tools. Princeton University Press; 2005.

[22] Petters AO, Dong X. *An Introduction to Mathematical Finance with Applications*. Springer; 2016.

[23] Dumitrescu EI, Hurlin C, Pham V. *Backtesting Value-at-Risk: From Dynamic Quantile to Dynamic Binary Tests* [Paper]. halshs-00671658; 2012 [last visited 02/04/2024]. Available at `https://shs.hal.science/halshs-00671658/document`.

[24] Solomonov I. *VaR Techniques* [Paper]. Uppsala Universitet; 2021 [last visited 02/04/2024]. Available at `https://uu.diva-portal.org/smash/get/diva2:1574101/FULLTEXT01.pdf`.

[25] Bayer S, Dimitriadis T. Regression-Based Expected Shortfall Backtesting*. Journal of Financial Econometrics. 2020 09;20(3):437–471. Available at `https://doi.org/10.1093/jjfinec/nbaa013`.

[26] Engle RF, Manganelli S. *CAViaR: CONDITIONAL VALUE AT RISK BY QUANTILE REGRESSION* [Paper]. NATIONAL BUREAU OF ECONOMIC RESEARCH; 1999 [last visited 04/04/2024]. Available at `https://www.nber.org/system/files/working_papers/w7341/w7341.pdf`.

[27] Ryan JA, Ulrich JM. quantmod: Quantitative Financial Modelling Framework; 2024. R package version 0.4.26. Available at `https://CRAN.R-project.org/package=quantmod`.

[28] Dashboard [Website]; 2024 [last visited 15/04/2024]. Available at https://polygon.io/dashboard.

[29] Galanos A. rugarch: Univariate GARCH models.; 2023. R package version 1.5-1.

[30] Nagler T, Schepsmeier U, Stoeber J, Brechmann EC, Graeler B, Erhardt T. VineCopula: Statistical Inference of Vine Copulas; 2023. R package version 2.5.0. Available at https://CRAN.R-project.org/package=VineCopula.

[31] Czado C, Min A, Baumann T, Dakovic R. Pair-copula constructions for modeling exchange rate dependence; 2009. p. 5–6. Available at https://api.semanticscholar.org/CorpusID:756658.

[32] Rockafellar RT, Uryasev S. Optimization of conditional value-at risk. Journal of Risk. 2000;3:21–41. Available at https://api.semanticscholar.org/CorpusID:854622.

[33] Schumann E. Numerical Methods and Optimization in Finance (NMOF) Manual. Package version 2.8-0); 2011–2023. Available at http://enricoschumann.net/NMOF/.

[34] Hoff IH, Aas K, Frigessi A. On the simplified pair-copula construction — Simply useful or too simplistic? Journal of Multivariate Analysis. 2010;101:1296 – 1310.

[35] Czado C, Nagler T. Vine copula based modeling. Annual Review of Statistics and Its Applications;.

[36] Gruber L, Czado C. Sequential Bayesian Model Selection of Regular Vine Copulas. Bayesian Analysis. 2015;10(4):937 – 963. Available at https://doi.org/10.1214/14-BA930.

[37] Zhang B, Wei Y, Yu J, Lai X, Peng Z. Forecasting VaR and ES of stock index portfolio: A Vine copula method. Physica A: Statistical Mechanics and its Applications. 2014;416:112–124. Available at https://www.sciencedirect.com/science/article/pii/S0378437114007286.

[38] *Copulas for Finance A Reading Guide and Some Applications* [Paper]; 2000 [last visited 15/03/2023]. Available at http://www.thierry-roncalli.com/download/copula-survey.pdf.

[39] Lindskog F, Mcneil A, Schmock U. Kendall's tau for elliptical distributions. Contributions to Economics. 2003;p. 149–156.

A | Supporting Theory for Copulas

This chapter is based on [6] and [7].

The following lemma and definition are used to infer that a multivariate copula is a multivariate distribution function.

Lemma A.1.
Let $S_1, S_2, \ldots, S_d$ be nonempty subsets of $\bar{\mathbb{R}}$ and let H be a grounded and d-increasing function with domain $S_1 \times S_2 \times \cdots \times S_d$. Then H is non-decreasing in each argument, meaning that if $(t_1, \ldots, t_{k-1}, x, t_{k+1}, \ldots, t_d), (t_1, \ldots, t_{k-1}, y, t_{k+1}, \ldots, t_d) \in dom(H)$ and $x < y$ then

$$H(t_1, \ldots, t_{k-1}, x, t_{k+1}, \ldots, t_d) \leq H(t_1, \ldots, t_{k-1}, y, t_{k+1}, \ldots, t_d). \tag{A.1}$$

Definition A.2. *d*-Dimensional Distribution Function
A function H with $dom(H) = \bar{\mathbb{R}}^d$ is a d-dimensional distribution function if

1. H is d-increasing.
2. $H(t) = 0$ for all $t \in \bar{\mathbb{R}}^d$ where $t_k = -\infty$ for at least one k and $H(\infty, \infty, \ldots, \infty) = 1$.

Definition A.3. Quasi-Inverse[6, p. 21]
Let F be a distribution function. The quasi-inverse of a distribution function F is any function $F^{(-1)}$ with domain $[0, 1]$ such that

1. If $t \in ran(F)$ then

$$F(F^{(-1)}(t)) = t, \quad \forall t \in ran(F). \tag{A.2}$$

2. If $t \notin ran(F)$ then

$$F^{(-1)}(t) = \inf\{x | F(x) \geq t\} = \sup\{x | F(x) \leq t\}. \tag{A.3}$$

If F is strictly increasing, then quasi-inverse is the same as the normal inverse.

A.1 Dependence Measure

This section is based on [4, chap. 2] and [6, chap. 5].

Two different measures of dependence are defined to investigate how copulas can be used to model the dependence structure between random variables.

A.1.1 Kendall's Tau

Kendall's tau measures the dependence between random variables by utilizing the principle of concordance. A pair of random variables are concordant if large values of one of the variables are associated with large values of the other variable and small values of one of the variables are associated with small values of the other variable. The opposite is discordant where small values of one of the variables are associated with large values of the other variable. Kendall's tau is defined as the probability of concordance minus the probability of discordance between two random variables X and Y.

Definition A.4. Kendall's Tau
Let X and Y be continuous random variables. Kendall's tau between X and Y is then defined as

$$\tau(X,Y) := \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_2) > 0\right) - \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_2) < 0\right), \tag{A.4}$$

where (X_1, Y_1) and (X_2, Y_2) are i.i.d. copies of (X, Y).

The range of Kendall's tau is $[-1, 1]$ since it is defined as the difference between two probabilities. Additionally, Kendall's tau can be expressed only in terms of a copula. The measure does therefore not depend on the marginal distributions.

Theorem A.5.
Let (X_1, Y_1) and (X_2, Y_2) be continuous i.i.d. variables. Kendall's tau is then given by

$$\tau = 4 \int_0^1 \int_0^1 C(u,v) \frac{\partial^2}{\partial u \partial v} C(u,v) du dv - 1. \tag{A.5}$$

The proof is omitted. However, the proof can be found in [4, p.32].

To select an R-vine copula specification a sequential method based on Kendall's tau is utilized. The sequential method is based on data. An empirical version of Kendall's tau is therefore introduced. Kendall's tau given in (A.4) can be rewritten as

$$\begin{aligned}
\tau &= \mathbb{P}((X_1 - X_2)(Y_1 - Y_2) > 0) - \mathbb{P}((X_1 - X_2)(Y_1 - Y_2) < 0) && \text{(A.6)} \\
&= \mathbb{E}\left[\mathbb{1}_{(X_1 - X_2)(Y_1 - Y_2) > 0} - \mathbb{1}_{(X_1 - X_2)(Y_1 - Y_2) < 0}\right] && \text{(A.7)} \\
&= \mathbb{E}\left[sign\left((X_1 - X_2)(Y_1 - Y_2)\right)\right], && \text{(A.8)}
\end{aligned}$$

where

$$sign(x) = \begin{cases} 1 & \text{if } x > 0, \\ -1 & \text{if } x < 0, \\ 0 & \text{else.} \end{cases} \tag{A.9}$$

Given independent samples $(X_1, Y_1), (X_2, Y_2), \ldots, (X_d, Y_d)$, the empirical Kendall's tau is given by

$$\begin{aligned} \hat{\tau}_n &:= \frac{\#\text{ of concordant pairs} - \#\text{ of discordant pairs}}{\#\text{ of all pairs}} \\ &= \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} sign\left((X_i - X_j)(Y_i - Y_j)\right). \end{aligned} \tag{A.10}$$

A.1.2 Tail Dependence

This subsection is based on [6, s.214].

Another measure of dependence is tail dependence. Tail dependence is given by the probability of joint occurrence of extremely small or large values.

Definition A.6. Tail dependence
Let X and Y be continuous random variables with marginal distribution functions F and G, respectively. Furthermore, let C be the copula of X and Y. The upper- and lower tail dependence coefficients are then defined as

$$\lambda_U := \lim_{q \uparrow 1} \mathbb{P}\left(Y > G^{-1}(q) \middle| X > F^{-1}(q)\right) = \lim_{q \uparrow 1} \frac{1 - 2q + C(q,q)}{1 - q} \quad \text{and} \tag{A.11}$$

$$\lambda_L := \lim_{q \downarrow 0} \mathbb{P}\left(Y \le G^{-1}(q) \middle| X \le F^{-1}(q)\right) = \lim_{q \downarrow 0} \frac{C(q,q)}{q}, \tag{A.12}$$

respectively, if the limits exists.

If $\lambda_L \in (0, 1]$, then C has lower tail dependence and conversely if $\lambda_L = 0$, then C has no tail dependence. This is also the case for λ_U and upper tail dependence.

A.2 Copula Families

This section is based on [6] and [38].

There exist many different families of copulas. In this project copulas that model elliptical distributions and Archimedean copulas will be used. Both families, given that the data is perfectly positive correlated or perfectly negative correlated, are equal to the Fréchet-Hoeffding lower- and upper bound, respectively.

Theorem A.7. Fréchet-Hoeffding Bounds
Let C be a d-dimensional copula. Then for every $u \in dom(C)$ it holds that

$$W^d(u) \leq C(u) \leq M^d, \tag{A.13}$$

where Fréchet-Hoeffding lower- and upper bound is given by

$$W^d(u) = \max(u_1 + u_2 + \cdots + u_d - d + 1, 0), \tag{A.14}$$
$$M^d(u) = min(u_1, u_2, \ldots, u_d). \tag{A.15}$$

$W^d(u)$ is only a copula for $d = 2$ and M^d is a copula for all $d \geq 2$.

Although $W^d(u)$ is only a copula for $d = 2$, it is possible for $d \geq 3$ to find a d-dimensional copula C, such that $C(u) = W^d(u)$ for any $u \in [0, 1]^d$.

Another case of extreme dependence is when the individual variables in the data are independent. If this is the case, the data can then be modeled using the independence copula.

Definition A.8. Independence Copula
Let $\Pi : [0, 1]^d \to [0, 1]$ be a d-dimensional copula. Π is then the independence copula if

$$\Pi(u_1, u_2, \ldots, u_d) = u_1 \cdot u_2 \cdots u_d \tag{A.16}$$

A.2.1 Elliptical Copulas

In this project, two elliptical copulas will be explored, namely the Gaussian copula and the Student's t-copula.

Definition A.9. Multivariate Gaussian Copula
Let the matrix $P \in \mathbb{R}^{d \times d}$ be symmetric and positive definite with $diag(P) = 1$ and Φ_P be a standard d-dimensional Gaussian distribution with correlation matrix P. The d-dimensional Gaussian copula is then defined as

$$C(u_1, u_2, \ldots, u_d; P) = \Phi_P \left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_d) \right), \tag{A.17}$$

and the density is given by

$$c(u_1, u_2, \ldots, u_d; P) = \frac{1}{|P|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} m^\top (P^{-1} - I_d) m \right) \tag{A.18}$$

where the n'th entry of the vector m is $m_n = \Phi^{-1}(u_n)$.

The Gaussian copula is equal to the Fréchet-Hoeffding upper bound if every entry in the correlation matrix P converges towards 1. Fréchet-Hoeffding lower bound is only a copula for $d = 2$. In the bivariate case, the Gaussian copula converges towards the Fréchet-Hoeffding lower bound if the pair-wise correlation converges towards -1. Lastly, if the correlation matrix is equal to the identity matrix, then the Gaussian copula is equal to the independence copula.

In the bivariate case let ρ be the off-diagonal of the correlation matrix P. The upper- and lower-tail dependence of bivariate Gaussian copula is then given by

$$\lambda_L = \lambda_U = \begin{cases} 1 & \text{if } \rho = 1, \\ 0 & \text{if } \rho \in (-1, 1) \end{cases}. \tag{A.19}$$

Kendall's tau can also be derived for the bivariate Gaussian copula and is given by

$$\tau_\rho^{Ga} = \frac{2}{\pi} \arcsin(\rho). \tag{A.20}$$

Definition A.10. Multivariate Student's *t*-Copula
Let the matrix $P \in \mathbb{R}^{d \times d}$ be symmetric and positive definite with $diag(P) = 1$ and $T_{P,\nu}$ be the standardized multivariate Student's t-distribution with ν degrees of freedom and correlation matrix P. The d-dimensional Student's t-copula is then defined as

$$C(u_1, u_2, \ldots, u_d) = T_{P,\nu}\left(T_\nu^{-1}(u_1), T_\nu^{-1}(u_2), \ldots, T_\nu^{-1}(u_d)\right), \tag{A.21}$$

where T_ν^{-1} denotes the inverse of the univariate Student's t-distribution. The density of (A.21) is given as

$$c(u_1, u_2, \ldots, u_d) = |P|^{-\frac{1}{2}} \frac{\Gamma(\frac{\nu+d}{2}) \left(\Gamma(\frac{\nu}{2})\right)^d \left(1 + \frac{1}{\nu} m^\top P^{-1} m\right)^{-\frac{\nu+d}{2}}}{\left(\Gamma(\frac{\nu+d}{2})\right)^d \Gamma(\frac{\nu}{2}) \prod_{n=1}^d \left(1 + \frac{m_n^2}{\nu}\right)^{-\frac{\nu+1}{2}}}, \tag{A.22}$$

where the n'th entry of the vector m is $m_n = T_\nu^{-1}(u_n)$.

In the bivariate case let ρ be the off-diagonal of the correlation matrix P. The upper- and lower-tail dependence of bivariate Student's t-copula is then given by

$$\lambda_U = \lambda_L = 2T_{\nu+1}\left(-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\right), \tag{A.23}$$

where $T_{\nu+1}$ is the cumulative distribution function of the Student's t-distribution with $\nu + 1$ degrees of freedom. For both the Student's t-copula and the Gaussian copula, the upper- and lower tail dependence coincide, which is the case for all elliptical and symmetric

distributions. Furthermore, if $\nu \to \infty$ then $\lambda_U = \lambda_L \to 0$ which is due to the Student's t-distribution converging towards the Gaussian distribution for $\nu \to \infty$.

Kendall's tau is the same for all elliptical distributions [39]. Kendall's tau is, therefore, for the bivariate Student's t-copula, given as

$$\tau_P^t = \frac{2}{\pi} \arcsin(\rho). \tag{A.24}$$

A.2.2 Archimedean Copulas

Definition A.11. Archimedean copula
Let $\phi : [0,1] \to [0,\infty]$ be a continuous, convex and strictly decreasing function such that $\phi(1) = 0$ and $\phi(0) = \infty$. The Archimedean copula is then defined as

$$C(u_1, u_2, \ldots, u_d) = \phi^{-1}\left(\phi(u_1) + \phi(u_2) + \cdots + \phi(u_d)\right). \tag{A.25}$$

Archemidean copulas are by construction symmetric since any permutation of $u_1, \ldots, u_d$ results in the same value of the copula. Archemidean copulas are also associative meaning that $C(C(u,v),w) = C(u,C(v,w))$ for all $u, v, w \in [0,1]$. Lastly, for any constant $c > 0$, $c\phi$ is also a generator of C.

There exist many different types of Archimedean copulas. The Archimedean copulas that are used in the construction of the R-vine copulas are Gumbel, Gumbel rotated 180 degrees (survival Gumbel), Gumbel rotated 270 or 90 degrees, and Frank Copula. A rotated copula is created in the following way for the bivariate case

$$u_1 = F_1(x_1), \quad 1 - u_2 = F_2(x_2), \quad \text{rotated 90 degrees} \tag{A.26}$$
$$1 - u_1 = F_1(x_1), \quad u_2 = F_2(x_2), \quad \text{rotated 270 degrees} \tag{A.27}$$
$$1 - u_1 = F_1(x_1), \quad 1 - u_2 = F_2(x_2), \quad \text{rotated 180 degrees.} \tag{A.28}$$

Table A.1: Archimedean Copulas in the Bivariate Case

Copula	$\varphi(u\,\vert\,\theta)$	$C(u,v;\theta)$	$\theta \in$
Gumbel	$(-\ln(u))^\theta$	$-\exp\left(-\left((-\ln(u)^\theta + (-\ln(v)^\theta\right)^{\frac{1}{\theta}}\right)$	$[1,\infty)$
Survival Gumbel	$\ln(1-\theta\ln(u))$	$uv\exp(-\theta\ln(u)\ln(v))$	$(0,1]$
Frank	$-\ln\left(\frac{e^{-\theta u}-1}{e^{-\theta}-1}\right)$	$-\theta^{-1}\ln\left(1+\frac{(e^{-u-\theta}-1)(e^{-v-\theta}-1)}{e^{-\theta}-1}\right)$	$(-\infty,\infty)\backslash\{0\}$

The tail dependence and Kendall's tau for the chosen Archimedean copulas are given in the following table.

Table A.2: Tail Dependence and Kendall's Tau for the Presented Archimedean Copulas.

Copula	λ_L	λ_U	τ
Gumbel	0	$2-2^{1/\theta}$	$1-\frac{1}{\theta}$
Survival Gumbel	$2-2^{1/\theta}$	0	$1-\frac{1}{\theta}$
Frank	0	0	$1+\frac{4}{\theta}\left(\int_0^\theta \frac{u}{e^u-1}du-1\right)$
Gumbel(90 degrees)	0	0	$-1-\frac{1}{\theta}$
Gumbel(270 degrees)	0	0	$-1-\frac{1}{\theta}$

B | Vine Copula

This section is based on [6, s.46].

B.1 Factorization of a Multivariate Distribution

Elaborating on Example 2.6, a decomposition of a multivariate distribution for $d = 6$ will be shown using the methods presented in this section and Subsection 2.1.1.

Example B.1.
Assume that a six-dimensional density $f(x_1, x_2, \dots, x_6)$ is decomposed in the following way using the factorization given in (2.7)

$$\begin{aligned} f(x_1, x_2, \dots, x_6) =& f_1(x_1) f_{2|1}(x_2|x_1) f_{6|1,2}(x_6|x_1, x_2) f_{3|1,2,6}(x_3|x_1, x_2, x_3) \\ & \cdot f_{5|1,2,3,6,3}(x_5|x_1, x_2, x_6, x_3) f_{4|1,2,6,3,5}(x_4|x_1, x_2, x_6, x_3, x_5). \end{aligned} \tag{B.1}$$

The conditional densities can be rewritten using (2.14)

$$f_{2|1}(x_2|x_1) = c_{2,1}(F_1(x_1), F_2(x_2)) f_2(x_2), \tag{B.2}$$

$$f_{6|1,2}(x_6|x_1, x_2) = c_{6,1|2}(F_{6|2}(x_6|x_2), F_{1|2}(x_1|x_2)) f_{6|2}(x_6|x_2), \tag{B.3}$$

$$f_{3|1,2,6}(x_3|x_1, x_2, x_6) = c_{3,1|2,6}(F_{3|2,6}(x_3|x_2, x_6), F_{1|2,3}(x_1|x_2, x_3)) f_{3|2,6}(x_3|x_2, x_6), \tag{B.4}$$

$$\begin{aligned} f_{5|1,2,6,3}(x_5|x_1, x_2, x_6, x_3) =& c_{5,1|2,6,3}(F_{5|2,6,3}(x_5|x_2, x_6, x_3), F_{1|2,6,3}(x_1|x_2, x_6, x_3)) \\ & \cdot f_{5|2,6,3}(x_5|x_2, x_6, x_3), \end{aligned} \tag{B.5}$$

$$\begin{aligned} f_{4|1,2,6,3,5}(x_4|x_1, x_2, x_6, x_3, x_5) =& c_{4,1|2,6,3,5}(F_{4|2,6,3,5}(x_4|x_2, x_6, x_3, x_5), F_{1|2,6,3,5}(x_1|x_2, x_6, x_3, x_5)) \\ & \cdot f_{4|2,6,3,5}(x_4|x_2, x_6, x_3, x_5). \end{aligned} \tag{B.6}$$

The conditional densities above are also rewritten using (2.14)

$$f_{6|2}(x_6, x_2) = c_{6,2}(F_6(x_6), F_2(x_2)) f_6(x_6), \tag{B.7}$$

$$f_{3|2,6}(x_3|x_2, x_6) = c_{3,2|6}(F_{3|6}(x_3|x_6), F_{2|6}(x_2|x_6)) f_{3|6}(x_3|x_6), \tag{B.8}$$

$$f_{5|2,6,3}(x_5|x_2, x_6, x_3) = c_{5,3|2,6}(F_{5|2,6}(x_5|x_2, x_6), F_{3|2,6}(x_3|x_2, x_6)) f_{5|2,6}(x_5|x_2, x_6), \tag{B.9}$$

$$\begin{aligned} f_{4|2,6,3,5}(x_4|x_2, x_6, x_3, x_5) =& c_{4,3|2,6,5}(F_{4|2,5}(x_4|x_2, x_6, x_5), F_{3|2,6,5}(x_3|x_2, x_6, x_5)) \\ & \cdot f_{4|2,6,5}(x_4|x_2, x_6, x_5). \end{aligned} \tag{B.10}$$

Again can the conditional densities be rewritten

$$f_{3|6}(x_3|x_6) = c_{3,6}(F_3(x_3), F_6(x_6)) f_3(x_3), \tag{B.11}$$

$$f_{5|2,6}(x_5|x_2, x_6) = c_{5,6|2}(F_{5|2}(x_5|x_2), F_{6|2}(x_6|x_2)) f_{5|2}(x_5|x_2), \tag{B.12}$$

$$f_{4|2,6,5}(x_4|x_2, x_6, x_5) = c_{4,6|2,5}(F_{4|2,5}(x_4|x_2, x_5), F_{6|2,5}(x_6|x_2, x_5)) f_{4|2,5}(x_4|x_2, x_5), \tag{B.13}$$

and the conditional densities here can be rewritten as

$$f_{5|2}(x_5, x_2) = c_{5,2}(F_5(x_5), F_2(x_2))f_5(x_5), \tag{B.14}$$
$$f_{4|2,5}(x_4|x_2, x_5) = c_{4,2|5}(F_{4|5}(x_4|x_5), F_{2|5}(x_2|x_5))f_{4|5}(x_4|x_5). \tag{B.15}$$

The last conditional density is rewritten as

$$f_{4|5}(x_4|x_5) = c_{4,5}(F_4(x_4), F_5(x_5))f_4(x_4), \tag{B.16}$$

whereby the joint density is given as

$$\begin{aligned} f(x_1, x_2, \ldots, x_6) = &\prod_{j=1}^{6} f_j(x_j) \\ &\cdot c_{2,1}(F_1(x_1), F_2(x_2)) \cdot c_{6,2}(F_6(x_6), F_2(x_2)) \\ &\cdot c_{3,6}(F_3(x_3), F_6(x_6)) \cdot c_{5,2}(F_5(x_5), F_2(x_2)) \\ &\cdot c_{4,5}(F_4(x_4), F_5(x_5)) \\ &\cdot c_{6,1|2}(F_{6|2}(x_6|x_2), F_{1|2}(x_1|x_2)) \\ &\cdot c_{3,2|6}(F_{3|6}(x_3|x_6), F_{2|6}(x_2|x_6)) \\ &\cdot c_{5,6|2}(F_{5|2}(x_5|x_2), F_{6|2}(x_6|x_2)) \\ &\cdot c_{4,2|5}(F_{4|5}(x_4|x_5), F_{2|5}(x_2, x_5)) \\ &\cdot c_{3,1|2,6}(F_{3|2,6}(x_3|x_2, x_6), F_{1|2,3}(x_1|x_2, x_3)) \\ &\cdot c_{5,3|2,6}(F_{5|2,6}(x_5|x_2, x_6), F_{3|2,6}(x_3|x_2, x_6)) \\ &\cdot c_{4,6|2,5}(F_{4|2,5}(x_4|x_2, x_5), F_{6|2,5}(x_6|x_2, x_5)) \\ &\cdot c_{5,1|2,6,3}(F_{5|2,6,3}(x_5|x_2, x_6, x_3), F_{1|2,6,3}(x_1|x_2, x_6, x_3)) \\ &\cdot c_{4,3|2,6,5}(F_{4|2,5}(x_4|x_2, x_6, x_5), F_{3|2,6,5}(x_3|x_2, x_6, x_5)) \\ &\cdot c_{4,1|2,6,3,5}(F_{4|2,6,3,5}(x_4|x_2, x_6, x_3, x_5), F_{1|2,6,3,5}(x_1|x_2, x_6, x_3, x_5)), \end{aligned} \tag{B.17}$$

which corresponds with (2.24). Furthermore, it is also possible to specify the matrices

$\mathcal{B}$ and Θ, which follows the structure of (2.31):

$$\mathcal{B} = \begin{bmatrix} - & C_{2,m_{1,2}} & C_{3,m_{1,3}} & C_{4,m_{1,4}} & C_{5,m_{1,5}} & C_{6,m_{1,6}} \\ & - & C_{3,m_{2,3}|m_{1,3}} & C_{4,m_{2,4}|m_{1,4}} & C_{5,m_{2,5}|m_{1,5}} & C_{6,m_{2,6}|m_{1,6}} \\ & & - & C_{4,m_{3,4}|m_{2,4},m_{1,4}} & C_{5,m_{3,5}|m_{2,5},m_{1,5}} & C_{6,m_{3,6}|m_{2,6},m_{1,6}} \\ & & & - & C_{5,m_{4,5}|m_{3,5},m_{2,5},m_{1,5}} & C_{6,m_{4,6}|m_{3,6},m_{2,6},m_{1,6}} \\ & & & & - & C_{6,m_{5,6}|m_{4,6},m_{3,6},m_{2,6},m_{1,6}} \\ & & & & & - \end{bmatrix} \tag{B.18}$$

$$= \begin{bmatrix} - & C_{2,1} & C_{3,2} & C_{4,3} & C_{5,2} & C_{6,5} \\ & - & C_{3,1|2} & C_{4,2|3} & C_{5,3|2} & C_{6,2|5} \\ & & - & C_{4,1|2,3} & C_{5,4|3,2} & C_{6,3|2,5} \\ & & & - & C_{5,1|4,3,2} & C_{6,4|3,2,5} \\ & & & & - & C_{6,1|4,3,2,5} \\ & & & & & - \end{bmatrix}, \tag{B.19}$$

and

$$\Theta = \begin{bmatrix} - & \theta_{2,m_{1,2}} & \theta_{3,m_{1,3}} & \theta_{4,m_{1,4}} & \theta_{5,m_{1,5}} & \theta_{6,m_{1,6}} \\ & - & \theta_{3,m_{2,3}|m_{1,3}} & \theta_{4,m_{2,4}|m_{1,4}} & \theta_{5,m_{2,5}|m_{1,5}} & \theta_{6,m_{2,6}|m_{1,6}} \\ & & - & \theta_{4,m_{3,4}|m_{2,4},m_{1,4}} & \theta_{5,m_{3,5}|m_{2,5},m_{1,5}} & \theta_{6,m_{3,6}|m_{2,6},m_{1,6}} \\ & & & - & \theta_{5,m_{4,5}|m_{3,5},m_{2,5},m_{1,5}} & \theta_{6,m_{4,6}|m_{3,6},m_{2,6},m_{1,6}} \\ & & & & - & \theta_{6,m_{5,6}|m_{4,6},m_{3,6},m_{2,6},m_{1,6}} \\ & & & & & - \end{bmatrix} \tag{B.20}$$

$$= \begin{bmatrix} - & \theta_{2,1} & \theta_{3,2} & \theta_{4,3} & \theta_{5,2} & \theta_{6,5} \\ & - & \theta_{3,1|2} & \theta_{4,2|3} & \theta_{5,3|2} & \theta_{6,2|5} \\ & & - & \theta_{4,1|2,3} & \theta_{5,4|3,2} & \theta_{6,3|2,5} \\ & & & - & \theta_{5,1|4,3,2} & \theta_{6,4|3,2,5} \\ & & & & - & \theta_{6,1|4,3,2,5} \\ & & & & & - \end{bmatrix}. \tag{B.21}$$

B.2 Log-Likelihood Algorithm

To calculate the log-likelihood of an R-vine copula the following algorithm is used.

Algorithm 4 Log-Likelihood for an R-Vine

Input R-vine copula specification matrices M, B, Θ, where $m_{k,k} = k, k = 1, \ldots, d$.
Output Density of the R-vine distribution at $x = (x_1, \ldots, x_d)$, for an R-vine copula specification.

Set L = 0.
Let $V^{direct} = (v_{i,k}^{direct} | k = 1, \ldots, d, i = 1, \ldots, k)$.
Let $V^{indirect} = (v_{i,k}^{indirect} | k = 1, \ldots, d, i = 1, \ldots, k)$.
Set $(v_{1,1}^{direct}, v_{1,2}^{direct}, \ldots, v_{1,d}^{direct}) = (F_1(x_1), F_2(x_2), \ldots, F_d(x_d))$.
Let $\mathbb{M} = (\mathbf{m}_{i,k} | k = 1, \ldots, d, i = 1, \ldots, k)$ where $\mathbf{m}_{i,k} = \max\{m_{i,k}, \ldots, m_{d,k}\}, \forall k = 1, \ldots, d$ and $i = 1, \ldots, k$.
for $k = 2, \ldots, d$ **do** (Iterating over the columns of M except the first column)
    for $i = 1, \ldots, k-1$ **do** (Iterating over the rows of M above the diagonal)
        Set $g_{i,k}^{(1)} = v_{i,k}^{direct}$.
        if $\mathbf{m}_{i,k} = m_{i,k}$ **then**
            Set $g_{i,k}^{(2)} = v_{i,\mathbf{m}_{i,k}}^{direct}$.
        else
            Set $g_{i,k}^{(2)} = v_{i,\mathbf{m}_{i,k}}^{indirect}$.
        end if
        Set $L = L + \log\left(c\left(g_{i,k}^{(1)}, g_{i,k}^{(2)}; b_{i,k}, \theta_{i,k}\right)\right)$
        Set $v_{i+1,k}^{direct} = h\left(g_{i,k}^{(1)}, g_{i,k}^{(2)} | b_{i,k}, \theta_{i,k}\right)$ and $v_{i+1,k}^{indirect} = h\left(g_{i,k}^{(2)}, g_{i,k}^{(1)} | b_{i,k}, \theta_{i,k}\right)$.
    end for
end for
return L

C | Coherent Risk Measures

This section is based on [21, p. 239-240].

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, Δ a time horizon, and $L^0(\Omega, \mathscr{F}, \mathbb{P})$ the set of all random variables on $(\Omega, \mathscr{F})$ which are almost surely finite. Financial risks, which in this case is portfolio losses over the time horizon Δ, are then represented as a set of random variables $\mathcal{M} \subset L^0(\Omega, \mathscr{F}, \mathbb{P})$, which is a convex cone, i.e., $L_1, L_2 \in \mathcal{M} \implies L_1 + L_2 \in \mathcal{M}$ and $\lambda L_1 \in \mathcal{M}, \forall \lambda > 0$. A risk measure $\rho : \mathcal{M} \to \mathbb{R}$, is called a coherent risk measure, if it satisfies the following axioms:

1. **Translation invariance**; For all $L \in \mathcal{M}$ and every $l \in \mathbb{R}$ it holds that $\rho(L + l) = \rho(L) + l$.
2. **Subadditivity**; For all $L_1, L_2 \in \mathcal{M}$ it holds that $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$.
3. **Positive homogeneity**; For all $L \in \mathcal{M}$ and every $\lambda > 0$ it holds that $\rho(\lambda L) = \lambda \rho(L)$.
4. **Monotonicity**; For $L_1, L_2 \in \mathcal{M}$ such that almost surely $L_1 \leq L_2$ it holds that $\rho(L_1) \leq \rho(L_2)$.

D | Results Supporting the Application Part of the Project

In this appendix results supporting the application part of this project will be shown. However, the characteristics of the log-returns and the fit of the marginals for every asset are not included and can be found in the attached zip file. The results are shown separately for each sector.

D.1 Financial Sector

D.1.1 Fitting Vine Copulas

Table D.1: The number of times each bivariate copula is selected in each tree, where Π notes the bivariate independence copula.

Tree	Π	N	t	G	F	G 180	G 90	G 270
1	0	0	59	0	0	1	0	0
2	0	0	40	0	19	0	0	0
3	0	1	20	1	34	2	0	0
4	2	0	18	0	35	2	0	0
5	3	3	22	2	24	2	0	0
6	9	2	18	2	14	9	1	0
7	11	3	10	3	26	0	0	1
8	18	2	12	6	11	1	0	3
9	17	6	9	3	15	0	2	0
10	15	3	11	2	14	2	2	2
11	16	3	9	3	15	2	2	0
12	21	7	10	2	3	2	4	0
13	17	3	11	3	10	1	1	2
14	18	3	9	2	8	2	3	2
15	17	3	5	6	11	3	0	1
16	17	7	3	3	7	4	2	2
17	17	5	3	2	13	4	0	0
18	21	2	3	2	9	2	2	2
19	21	3	4	3	6	1	2	2
20	15	7	4	1	11	2	0	1
21	20	4	2	3	3	4	3	1
22	22	4	2	1	5	1	2	2

Table D.1: The number of times each bivariate copula is selected in each tree, where Π notes the bivariate independence copula *(continued).*

Tree	Π	N	t	G	F	G 180	G 90	G 270
23	17	0	3	5	9	3	1	0
24	17	3	0	1	12	0	1	3
25	18	1	3	1	7	2	0	4
26	18	3	3	3	7	0	1	0
27	16	3	4	3	3	1	1	3
28	19	2	3	2	5	0	1	1
29	16	2	3	1	6	1	1	2
30	18	3	3	2	4	0	1	0
31	14	2	2	1	4	1	4	2
32	16	1	0	1	5	2	2	2
33	18	1	3	2	3	0	1	0
34	11	2	1	0	8	4	1	0
35	13	4	0	0	7	2	0	0
36	10	1	1	1	6	4	1	1
37	12	2	0	0	5	4	0	1
38	9	4	5	0	5	0	0	0
39	13	1	2	1	4	0	1	0
40	12	1	1	1	4	1	1	0
41	11	4	2	0	1	1	1	0
42	10	0	0	3	2	2	0	2
43	9	1	2	1	2	0	3	0
44	11	0	1	0	3	2	0	0
45	11	2	0	0	1	1	0	1
46	10	2	1	0	1	0	1	0
47	7	1	1	0	3	1	1	0
48	8	1	0	0	3	1	0	0
49	7	0	0	0	2	1	2	0
50	8	0	0	1	1	0	0	1
51	4	1	2	0	1	2	0	0
52	6	0	1	0	0	1	0	1
53	4	0	1	1	1	1	0	0
54	6	0	0	0	1	0	0	0
55	4	1	0	0	0	0	1	0
56	4	0	0	1	0	0	0	0
57	4	0	0	0	0	0	0	0

Table D.1: The number of times each bivariate copula is selected in each tree, where Π notes the bivariate independence copula *(continued)*.

Tree	Π	N	t	G	F	G 180	G 90	G 270
58	3	0	0	0	0	0	0	0
59	1	1	0	0	0	0	0	0
60	0	0	0	0	1	0	0	0

D.2 Goodness-of-Fit

The GOF test is carried out on the financial sector for the first 5 and 10 listed assets. Additionally, the GOF test is also made on each bivariate copula in the multivariate Gaussian copula to compare with the results from the *R*-vine AIC seq.

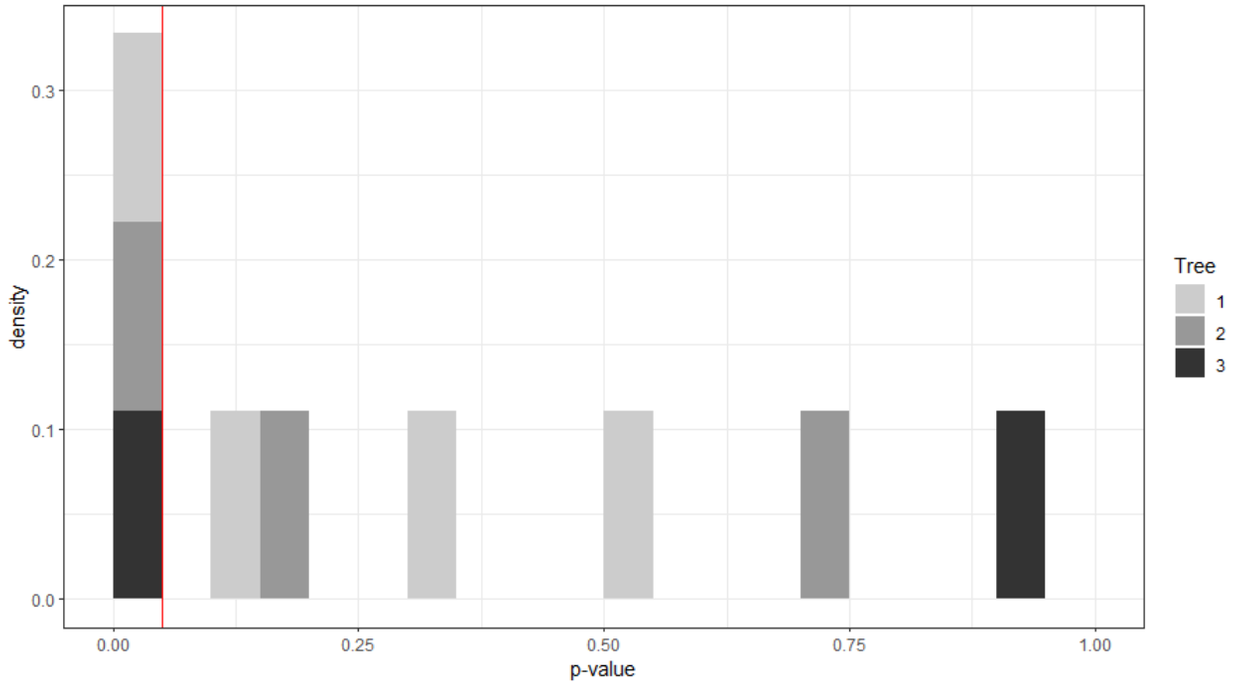


Figure D.1: GOF test on the bivariate copulas in the *R*-vine copula AIC seq for the first 5 listed assets in the financial sector.

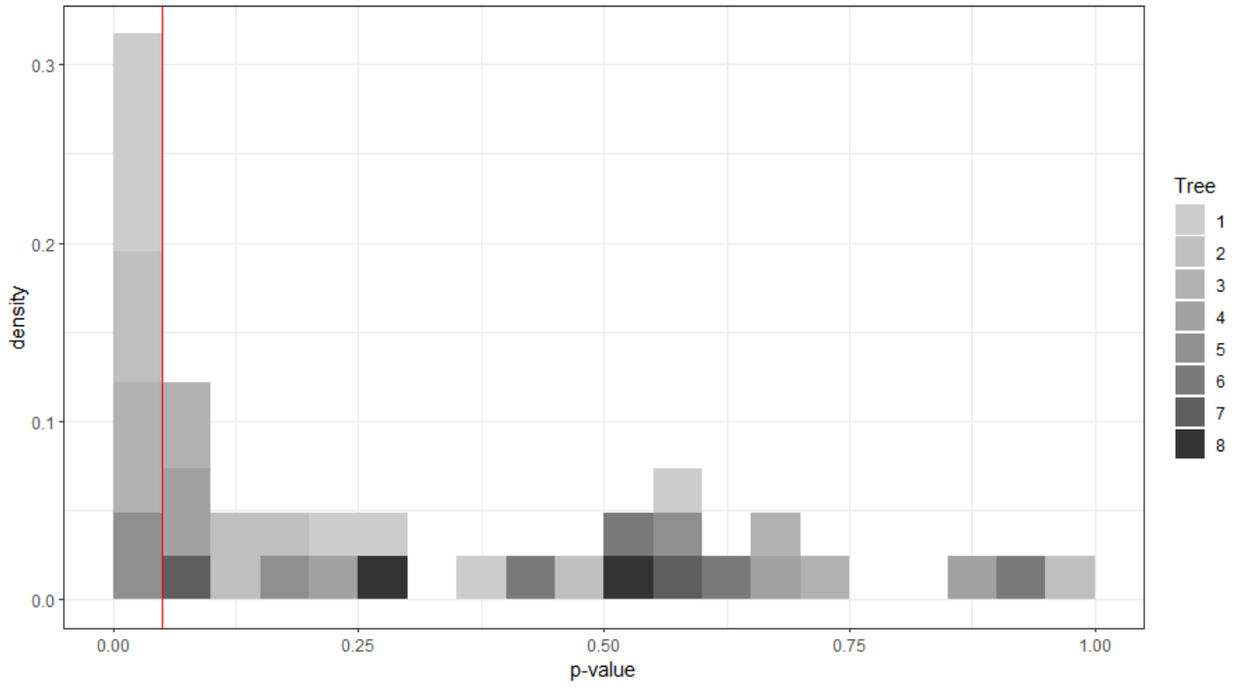


Figure D.2: GOF test on the bivariate copulas in the R-vine copula AIC seq for the first 10 listed assets in the financial sector.

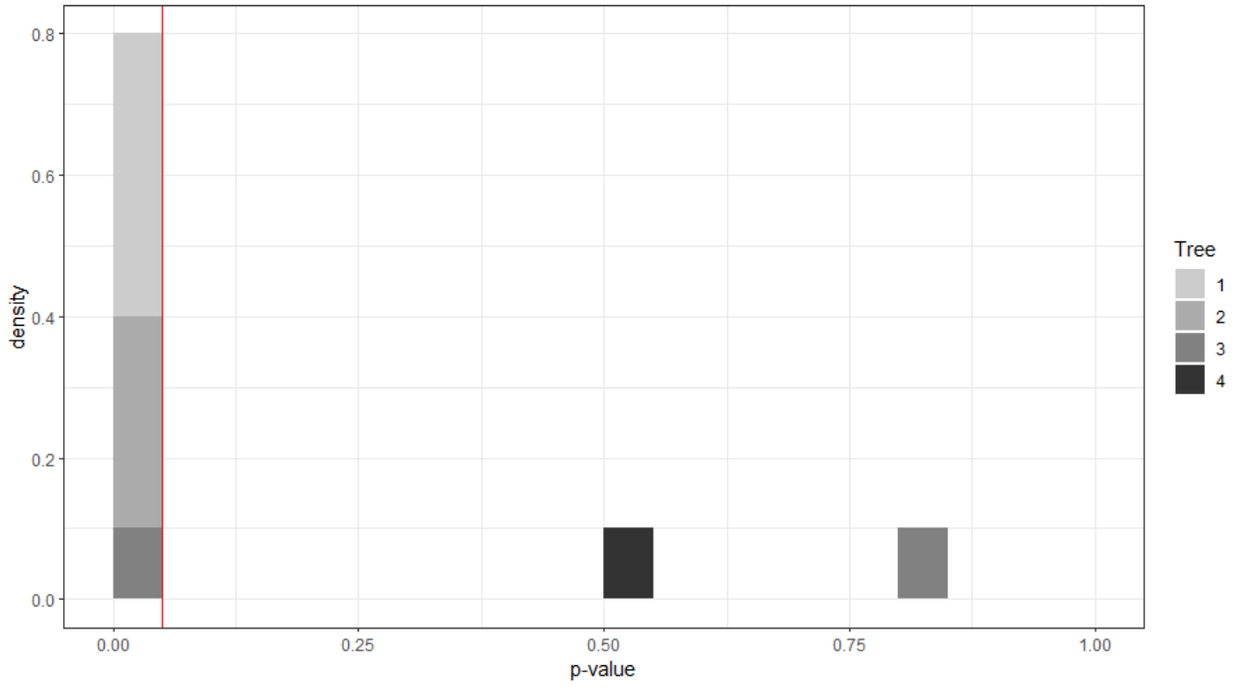


Figure D.3: GOF test on the bivariate copulas in the Gaussian seq for the first 5 listed assets in the financial sector.

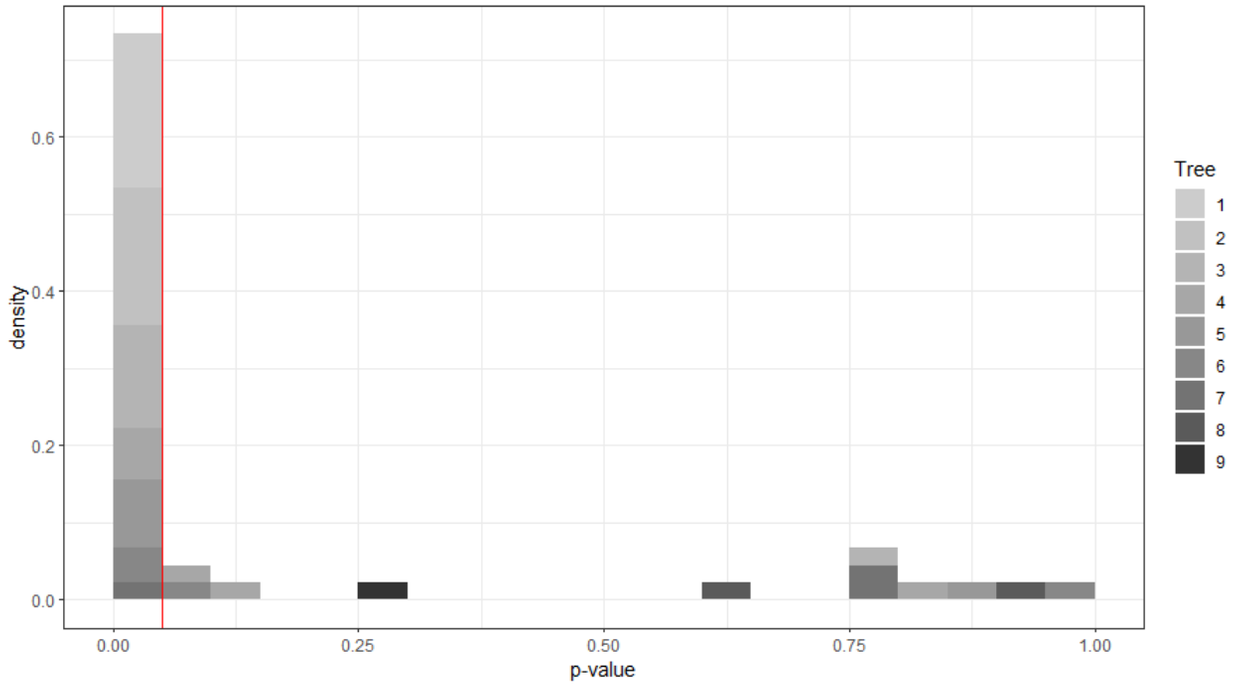


Figure D.4: GOF test on the bivariate copulas in the Gaussian seq for the first 10 listed assets in the financial sector.

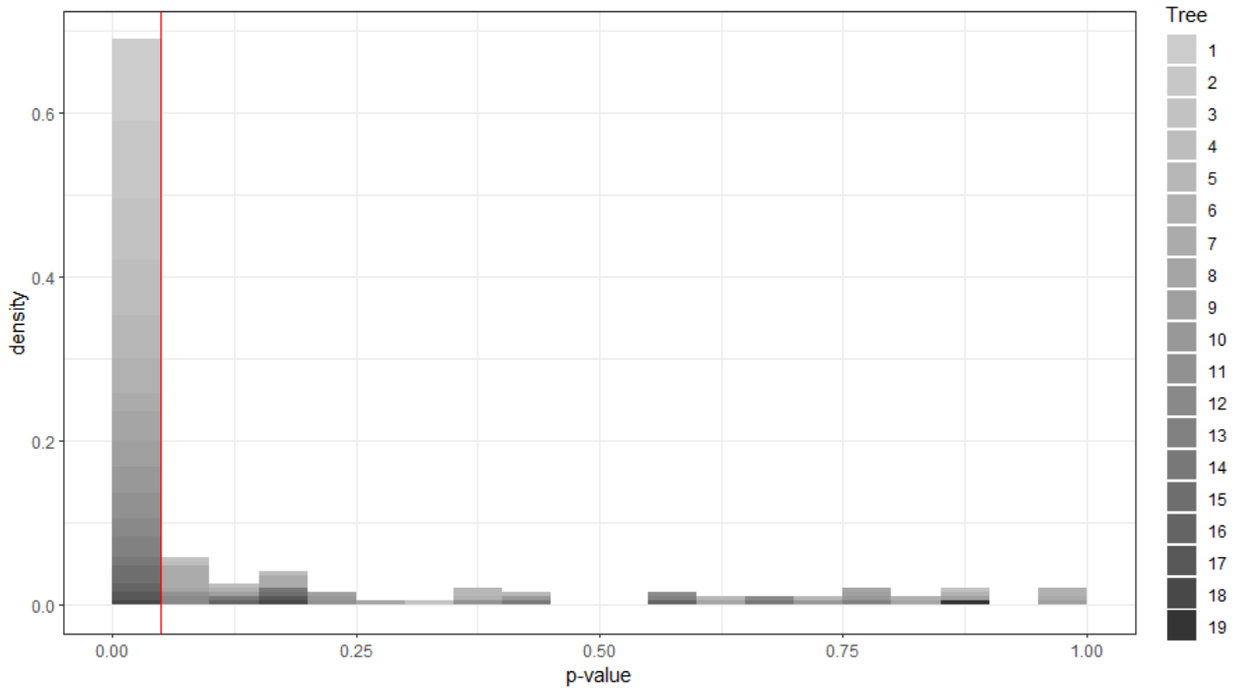


Figure D.5: GOF test on the bivariate copulas in the Gaussian seq for the first 20 listed assets in the financial sector.

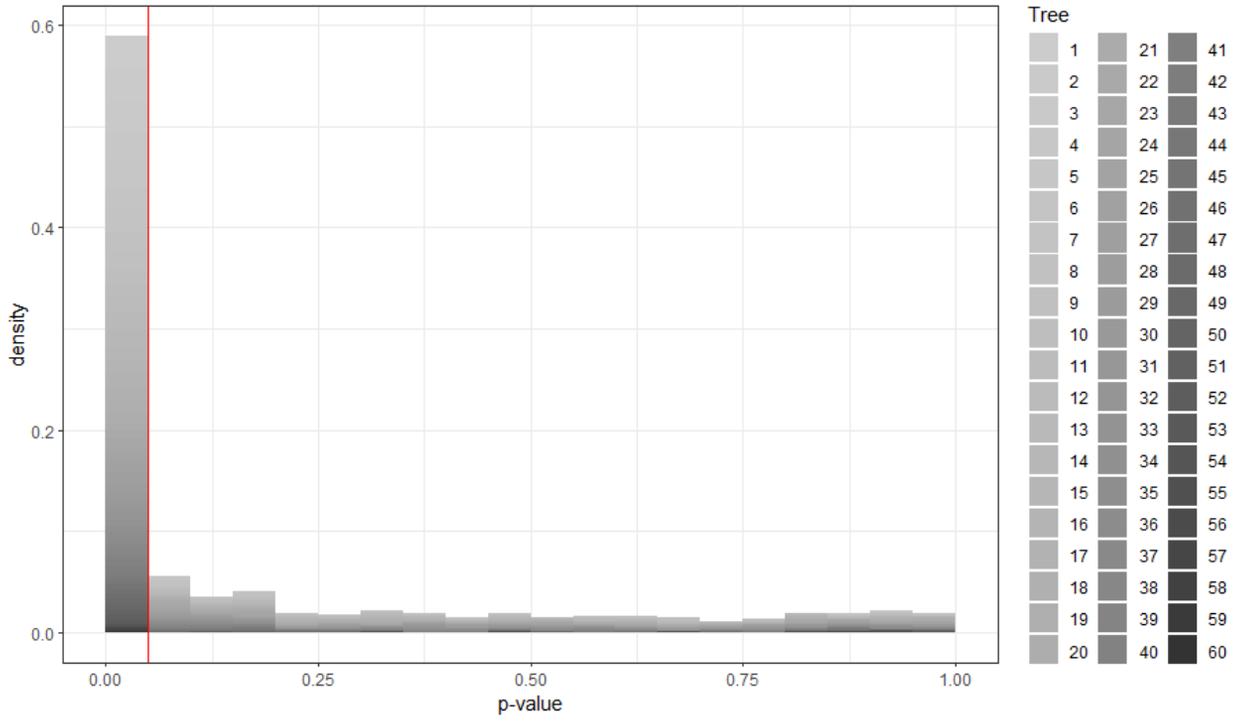


Figure D.6: GOF test on the bivariate copulas in the Gaussian seq for every asset in the financial sector.

D.3 Healthcare Sector

This section contains supporting results for the healthcare sector.

D.3.1 Fitting Marginal Distribution

Table D.2: Selected marginal distribution and parameters for each asset in the healthcare sector.

					Parameters Value (Standard Error)					
Stock	Model	Distribution	L	AIC	μ	AR	MA	ω	α	β
LLY	ARMA(0, 0) GARCH(1, 1)	STD	-2150	2.4	-0.00757 (0.0156)	- - -	- - -	$1.35 \cdot 10^{-5}$ ($4.6 \cdot 10^{-6}$)	0.113 (0.027) - -	0.85 (0.0334) - -

Table D.2: Selected marginal distribution and parameters for each asset in the healthcare sector *(continued).*

					Parameters Value (Standard Error)					
Stock	Model	Distribution	L	AIC	μ	AR	MA	ω	α	β
UNH	ARMA(0, 0) GARCH(1, 1)	STD	-2200	2.45	0.0116 (0.016)	- - -	- - -	$5.18 \cdot 10^{-6}$ ($1.97 \cdot 10^{-6}$)	0.0933 (0.0186) - -	0.892 (0.0205) - -
JNJ	ARMA(0, 0) GARCH(1, 1)	STD	-2220	2.47	0.0398 (0.0168)	- - -	- - -	$7.06 \cdot 10^{-6}$ ($1.89 \cdot 10^{-6}$)	0.118 (0.0226) - -	0.827 (0.0296) - -
ABBV	ARMA(0, 0) GARCH(1, 1)	STD	-2290	2.55	0.00311 (0.0176)	- - -	- - -	1.34e-05 ($4.19 \cdot 10^{-6}$)	0.132 (0.031) - -	0.798 (0.0448) - -
MRK	ARMA(0, 0) GARCH(1, 1)	STD	-2320	2.59	0.0231 (0.0171)	- - -	- - -	$1.76 \cdot 10^{-5}$ ($6.2 \cdot 10^{-6}$)	0.15 (0.0358) - -	0.81 (0.0431) - -

It is tested whether the residuals are a random sample from the chosen distribution of the innovations, with the parameters estimated for the ARMA-GARCH model, by performing an Anderson-Darling test.

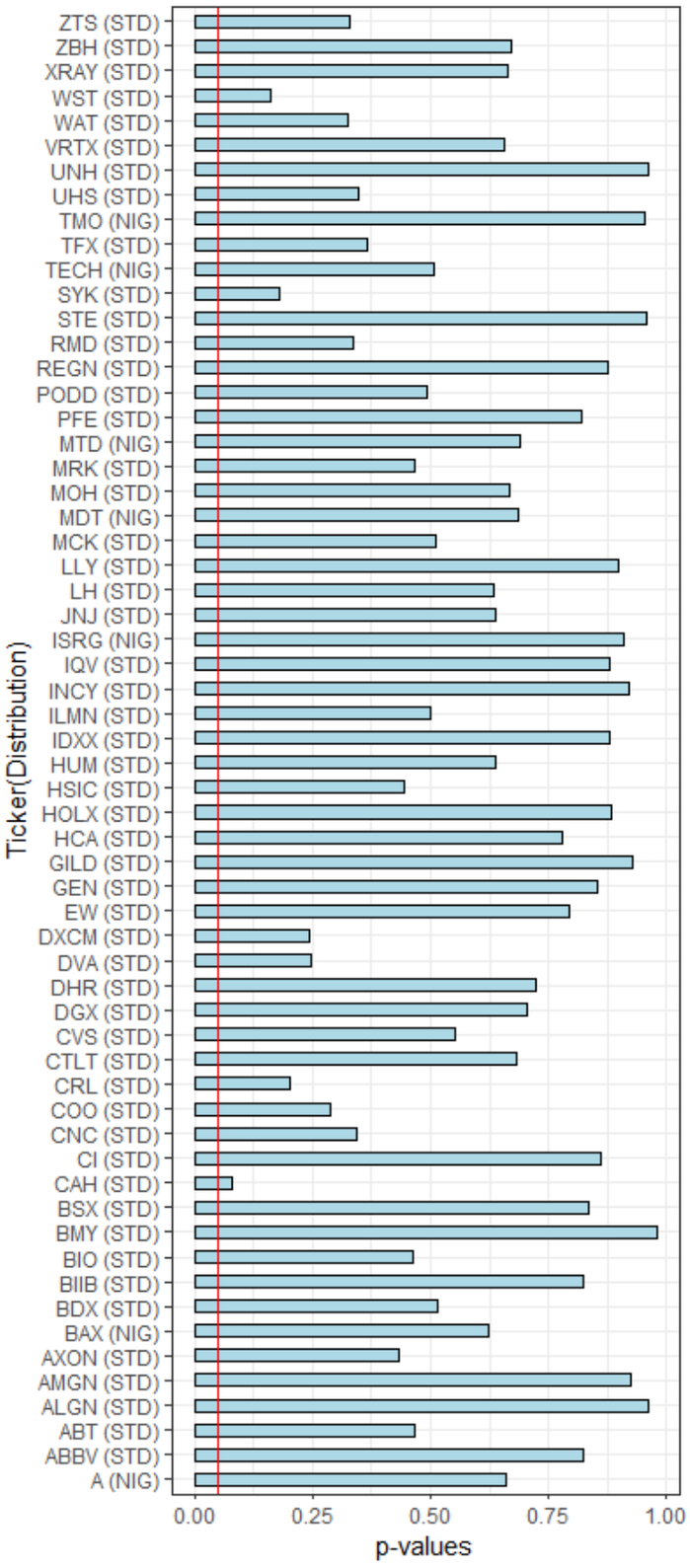


Figure D.7: The p-values from an Anderson-Darling test performed on the residuals of an ARMA(m, n)-GARCH(p, q)-model in the healthcare sector, with the respective distribution of the innovations shown in the parentheses.

As for the financial sector, based on the Anderson-Darling test, all marginals can't be rejected for the healthcare sector, confer Figure D.7.

Table D.3: Ljung-Box test on residuals(R) and squared residuals(R^2) from fitted models on healthcare data, for lag $= 1, 2, 3$.

		Ljung-Box Test, R			Ljung-Box Test, R^2		
Stock	Lag	1	2	3	1	2	3
LLY	Statistic	2.57	3.39	4.01	0.314	0.516	0.829
	p-value	0.109	0.184	0.26	0.575	0.773	0.842
UNH	Statistic	0.238	0.325	1.62	0.475	1.26	1.9
	p-value	0.625	0.85	0.654	0.49	0.533	0.593
JNJ	Statistic	1.58	1.81	1.96	0.226	0.434	0.437
	p-value	0.209	0.405	0.58	0.634	0.805	0.932
ABBV	Statistic	0.187	0.662	0.775	0.000124	0.00713	0.201
	p-value	0.665	0.718	0.855	0.991	0.996	0.977
MRK	Statistic	0.379	0.588	0.654	0.0928	0.167	0.209
	p-value	0.538	0.745	0.884	0.761	0.92	0.976

D.3.2 Fitting Vine Copulas

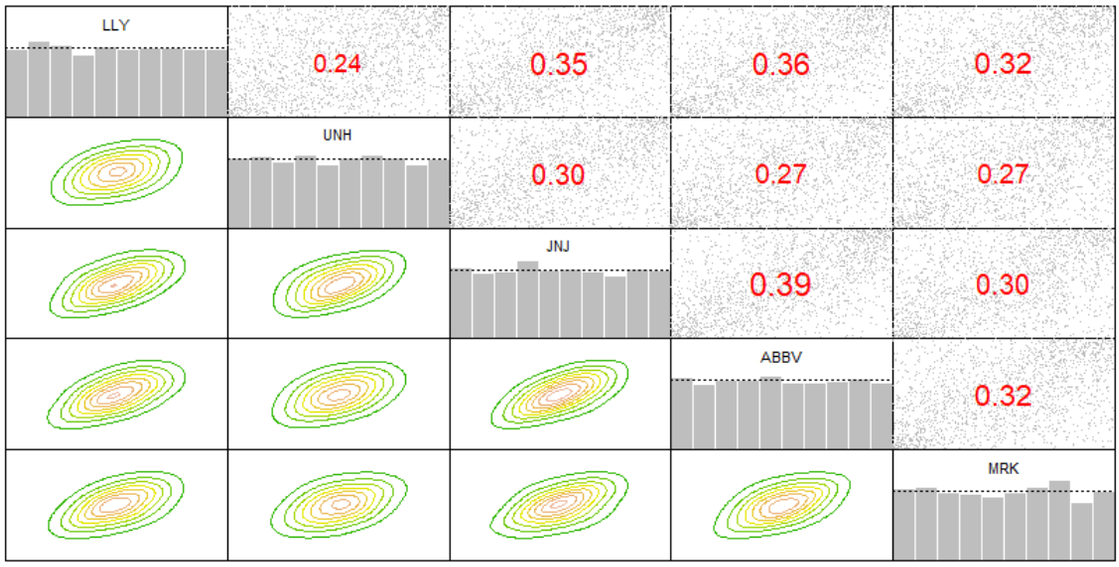


Figure D.8: Pseudo copula data using marginals from the fitted models. The lower triangle: empirical normalized contour plots, upper triangle: pair scatter plot of copula data and empirical Kendall's tau and diagonal: marginal histograms of copula data.

Table D.4: The number of times each bivariate copula is selected in each tree, where Π notes the bivariate independence copula.

Tree	Π	N	t	G	F	G 180	G 90	G 270
1	0	0	52	0	4	3	0	0
2	0	0	20	0	38	0	0	0
3	0	0	11	0	44	2	0	0
4	1	1	14	0	38	2	0	0
5	1	1	16	2	32	3	0	0
6	5	1	12	2	29	5	0	0
7	9	1	12	7	20	4	0	0
8	9	4	10	2	23	2	2	0
9	19	5	8	2	14	2	1	0
10	17	2	6	4	16	5	0	0
11	16	1	6	3	15	5	2	1
12	10	4	6	4	18	2	1	3

Table D.4: The number of times each bivariate copula is selected in each tree, where Π notes the bivariate independence copula *(continued).*

Tree	Π	N	t	G	F	G 180	G 90	G 270
13	15	7	5	5	9	3	0	3
14	13	7	13	0	11	2	0	0
15	13	5	10	3	8	2	2	2
16	17	10	3	1	9	1	2	1
17	18	4	4	2	13	1	1	0
18	19	1	5	1	10	3	2	1
19	13	3	5	3	9	5	2	1
20	15	5	4	1	11	1	1	2
21	19	3	7	2	3	2	2	1
22	16	5	3	2	7	2	2	1
23	15	6	2	3	6	2	2	1
24	13	5	4	2	7	3	1	1
25	17	0	0	5	9	4	0	0
26	13	3	3	4	6	3	1	1
27	13	0	3	4	7	3	3	0
28	15	2	4	2	7	2	0	0
29	12	3	2	1	7	2	2	2
30	14	2	1	1	7	1	2	2
31	11	1	0	3	9	2	1	2
32	13	4	3	0	5	0	1	2
33	14	2	2	2	4	1	1	1
34	13	3	3	0	2	3	2	0
35	19	1	2	1	1	1	0	0
36	7	5	3	1	4	1	3	0
37	8	3	3	2	6	0	0	1
38	13	0	5	0	2	2	0	0
39	12	1	2	2	3	1	0	0
40	9	2	2	1	4	1	1	0
41	14	1	0	0	4	0	0	0
42	5	3	2	0	4	0	2	2
43	8	2	0	1	1	3	1	1
44	7	2	1	2	0	2	2	0
45	10	0	2	2	1	0	0	0
46	10	0	0	0	3	0	1	0
47	8	2	1	0	2	0	0	0

Table D.4: The number of times each bivariate copula is selected in each tree, where Π notes the bivariate independence copula *(continued).*

Tree	Π	N	t	G	F	G 180	G 90	G 270
48	5	0	3	0	3	0	0	1
49	6	1	0	0	2	1	0	1
50	3	2	0	0	4	0	1	0
51	2	1	0	1	5	0	0	0
52	1	1	1	1	2	0	1	1
53	3	0	1	1	1	1	0	0
54	3	2	0	0	1	0	0	0
55	1	0	0	1	1	1	0	1
56	0	1	0	0	2	0	1	0
57	1	1	0	1	0	0	0	0
58	2	0	0	0	0	0	0	0
59	0	0	0	0	0	0	1	0

D.3.3 Goodness-of-Fit

Table D.5: Fitted models on Fitted models on the first 5 assets in the healthcare sector, number of parameters, Estimated log-likelihood, AIC and BIC for the fitted models.

Model	par	loglike	AIC	BIC	White test (p-value)
R-vine AIC seq	17	1384.233	-2734.466	-2641.051	0.119
R-vine AIC seq mle	17	1385.408	-2736.817	-2643.402	0.173
R-vine tau seq	16	1380.696	-2729.392	-2641.472	0.0217
R-vine tau seq mle	16	1383.344	-2734.688	-2646.768	0.0300
Gaussian seq	10	1235.433	-2450.865	-2395.916	0.0000795
Gaussian seq mle	10	1235.433	-2450.866	-2395.916	0.0000802
Student's t seq	20	1377.817	-2715.633	-2605.733	0.642
Student's t seq mle	20	1379.554	-2719.108	-2609.208	0.741

Table D.6: Fitted models on Fitted models on the first 10 assets in the healthcare sector, number of parameters, Estimated log-likelihood, AIC and BIC for the fitted models.

Model	par	loglike	AIC	BIC	White test (p-value)
R-vine AIC seq	61	4634.725	-9147.450	-8812.256	1.000
R-vine AIC seq mle	61	4637.003	-9152.006	-8816.812	1.000
R-vine tau seq	54	4615.871	-9123.742	-8827.013	1.000
R-vine tau seq mle	54	4620.237	-9132.473	-8835.744	1.000
Gaussian seq	45	4189.245	-8288.491	-8041.217	0.961
Gaussian seq mle	45	4189.246	-8288.492	-8041.218	0.962
Student's t seq	90	4594.452	-9008.904	-8514.355	1.000
Student's t seq mle	90	4597.936	-9015.872	-8521.323	1.000

Table D.7: Fitted models on Fitted models on the first 20 assets in the healthcare sector, number of parameters, Estimated log-likelihood, AIC and BIC for the fitted models.

Model	par	loglike	AIC	BIC	White test (p-value)
R-vine AIC seq	211	10567.058	-20712.12	-19552.67	1
R-vine AIC seq mle	211	10572.724	-20723.45	-19564.01	1
R-vine tau seq	186	10507.087	-20642.17	-19620.11	1
R-vine tau seq mle	186	10519.046	-20666.09	-19644.02	1
Gaussian seq	190	9487.053	-18594.11	-17550.06	1
Gaussian seq mle	190	9487.056	-18594.11	-17550.07	1
Student's t seq	380	10391.356	-20022.71	-17934.62	1
Student's t seq mle	380	10412.944	-20065.89	-17977.79	1

D.3.4 Out-of-Sample *CVaR* Forecast

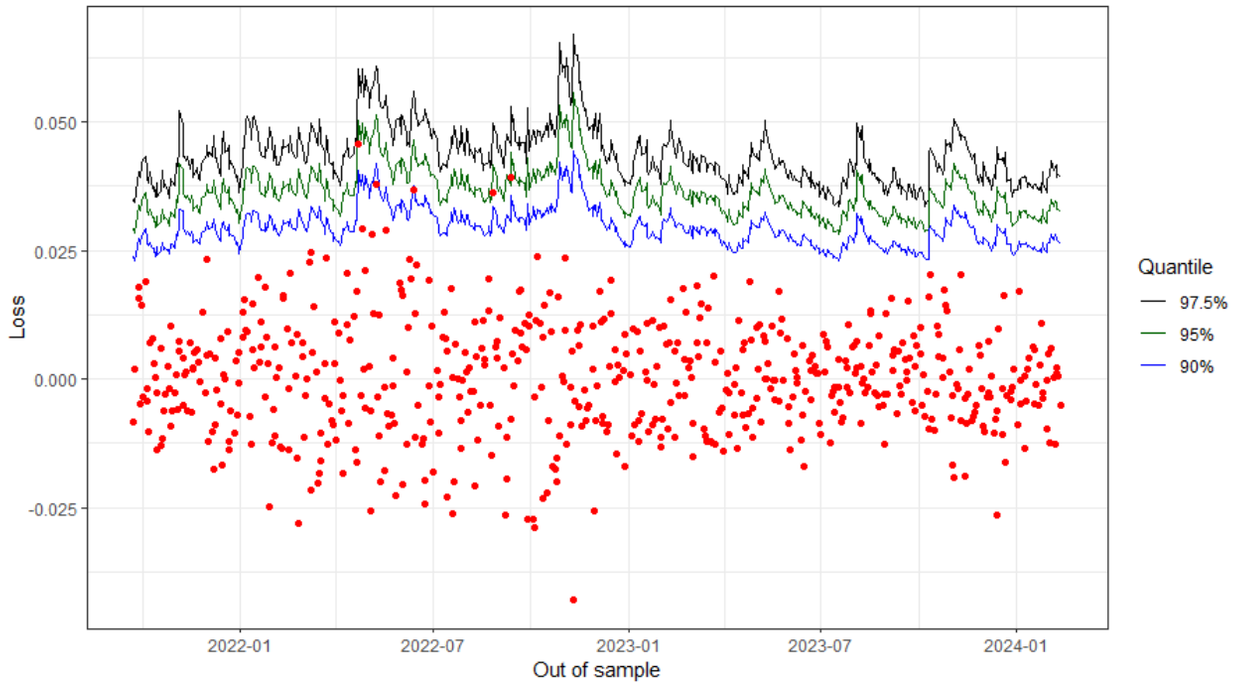


Figure D.9: Out-of-sample CVaR forecast for an equally-weighted portfolio consisting of the assets in the healthcare sector.

D.4 Technology Sector

This section contains supporting results for the technology sector.

D.4.1 Fitting Marginal Distribution

Table D.8: Selected marginal distribution and parameters for each asset in the technology sector.

					Parameters Value (Standard Error)					
Stock	Model	Distribution	L	AIC	μ	AR	MA	ω	α	β
MSFT	ARMA$(0,1)$ GARCH$(1,1)$	NIG	-1660	2.32	0.0147 (0.0158)	-	-0.121 (0.0275)	$7.7 \cdot 10^{-6}$ $(2.63 \cdot 10^{-6})$	0.155 (0.031)	0.824 (0.0326)
						-	-		-	-
						-	-		-	-
						-	-		-	-

Table D.8: Selected marginal distribution and parameters for each asset in the technology sector *(continued).*

					Parameters Value (Standard Error)					
Stock	Model	Distribution	L	AIC	μ	AR	MA	ω	α	β
AAPL	ARMA(0, 0) GARCH(1, 1)	STD	-1780	2.48	0.033 (0.0179)	- - -	- - -	$1.2 \cdot 10^{-5}$ $(4.09 \cdot 10^{-6})$	0.163 (0.0337) - -	0.823 (0.0323) - -
NVDA	ARMA(0, 0) GARCH(1, 1)	STD	-1780	2.49	0.0216 (0.0186)	- - -	- - -	$6.35 \cdot 10^{-5}$ $(2.23 \cdot 10^{-5})$	0.161 (0.0389) - -	0.783 (0.0508) - -
GOOGL	ARMA(0, 0) GARCH(1, 1)	NIG	-1670	2.34	0.019 (0.0192)	- - -	- - -	$5.79 \cdot 10^{-6}$ $(2.52 \cdot 10^{-6})$	0.0949 (0.0209) - -	0.891 (0.0235) - -
GOOG	ARMA(0, 0) GARCH(1, 1)	NIG	-1660	2.32	0.0203 (0.019)	- - -	- - -	$6.04 \cdot 10^{-6}$ $(2.52 \cdot 10^{-6})$	0.1 (0.0217) - -	0.886 (0.0239) - -

Table D.9: Ljung-Box test on residuals(R) and squared residuals(R^2) from the fitted models from the technology data for lag $= 1, 2, 3$.

		Ljung-Box Test, R			Ljung-Box Test, R^2		
Stock	Lag	1	2	3	1	2	3
MSFT	Statistic	0.425	0.817	0.848	0.0497	0.264	0.499
	p-value	0.514	0.665	0.838	0.824	0.876	0.919
AAPL	Statistic	3.5e-05	0.937	3.39	1.3	1.51	2.1
	p-value	0.995	0.626	0.335	0.254	0.471	0.553
NVDA	Statistic	2.36	2.37	2.48	0.11	0.439	0.898
	p-value	0.124	0.306	0.479	0.74	0.803	0.826
GOOGL	Statistic	0.000651	0.307	1.3	2.87e-09	0.429	0.429
	p-value	0.98	0.858	0.729	1	0.807	0.934
GOOG	Statistic	0.0163	0.228	1.68	0.0248	0.412	0.452
	p-value	0.898	0.892	0.642	0.875	0.814	0.929

It is tested whether the residuals are a random sample from the chosen distribution of the innovations, with the parameters estimated for the ARMA-GARCH model, by performing an Anderson-Darling test.

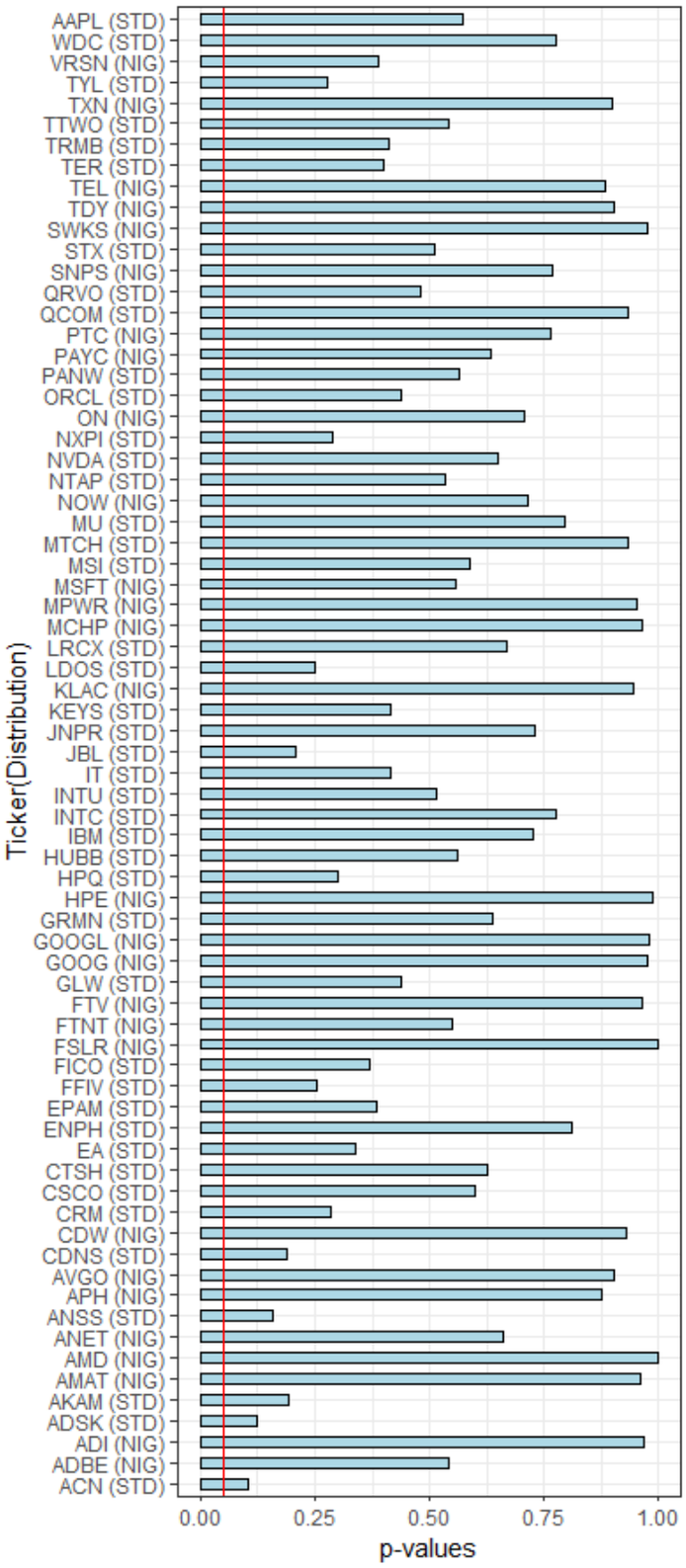


Figure D.10: The p-values from an Anderson-Darling test performed on the residuals of an ARMA(m, n)-GARCH(p, q)-model in the technology sector, with the respective distribution of the innovations shown in the parentheses.

The Anderson-Darling test indicates that it cannot be rejected that each asset's residuals are sampled from their respective chosen distribution.

D.4.2 Fitting Vine Copulas

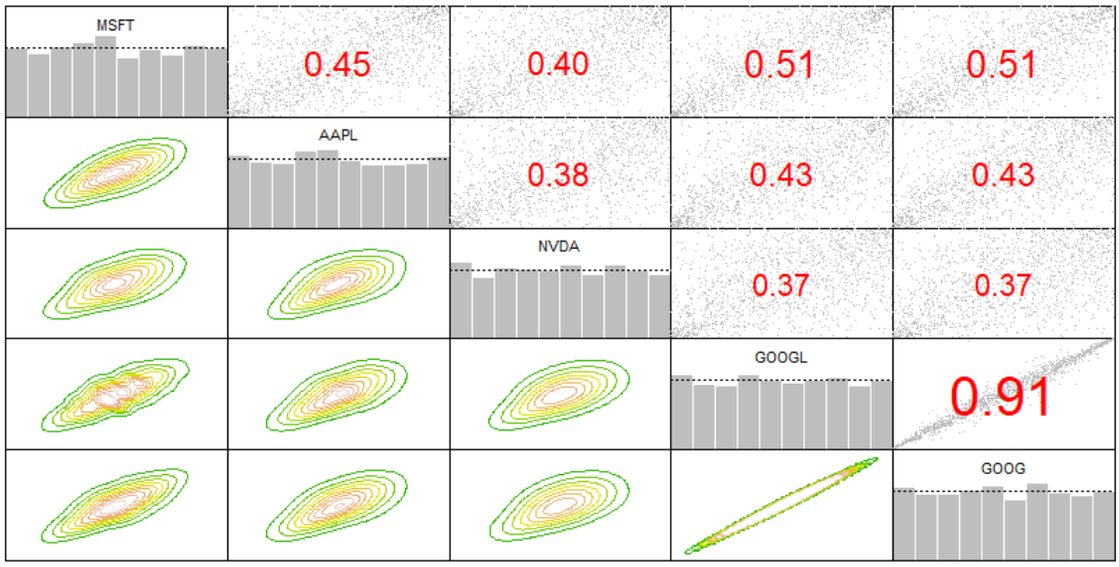


Figure D.11: Pseudo copula data using marginals from the fitted models. The lower triangle: empirical normalized contour plots, upper triangle: pair scatter plot of copula data and empirical Kendall's tau and diagonal: marginal histograms of copula data.

Table D.10: The number of times each bivariate copula is selected in each tree, where Π notes the bivariate independence copula.

Tree	Π	N	t	G	F	G 180	G 90	G 270
1	0	0	55	0	2	13	0	0
2	0	1	23	1	42	2	0	0
3	0	1	9	2	51	5	0	0
4	2	2	9	5	45	4	0	0
5	2	4	13	1	42	3	1	0
6	11	3	11	0	32	5	1	2
7	11	4	12	3	30	4	0	0
8	13	6	10	3	29	2	0	0
9	19	7	5	3	18	6	0	4

Table D.10: The number of times each bivariate copula is selected in each tree, where Π notes the bivariate independence copula *(continued)*.

Tree	Π	N	t	G	F	G 180	G 90	G 270
10	16	11	8	4	15	5	1	1
11	21	6	7	4	17	4	1	0
12	19	2	12	5	15	1	4	1
13	25	7	8	3	9	2	2	2
14	18	9	7	1	18	2	0	2
15	20	4	6	3	18	3	1	1
16	22	5	8	1	12	3	4	0
17	26	6	6	3	9	1	3	0
18	27	3	3	3	9	3	2	3
19	19	10	9	3	4	2	3	2
20	19	8	4	4	13	1	1	1
21	25	3	5	0	13	1	2	1
22	26	4	5	1	11	1	0	1
23	18	7	1	3	15	2	1	1
24	17	4	6	5	11	1	0	3
25	15	7	3	3	14	3	1	0
26	21	4	4	1	9	3	3	0
27	24	2	1	1	10	2	2	2
28	21	1	4	2	9	3	2	1
29	20	3	7	5	4	2	1	0
30	20	3	2	1	11	4	0	0
31	17	5	3	2	12	0	0	1
32	17	6	3	0	6	6	0	1
33	18	5	2	3	7	2	0	1
34	16	3	4	2	11	1	0	0
35	17	6	1	2	4	3	3	0
36	19	3	5	2	4	0	0	2
37	19	2	2	4	5	0	1	1
38	14	2	0	3	8	1	2	3
39	15	5	1	2	7	0	0	2
40	12	5	2	1	6	3	1	1
41	13	1	4	2	9	1	0	0
42	14	2	4	3	3	1	2	0
43	16	0	1	2	4	3	0	2
44	16	2	0	1	5	2	0	1

Table D.10: The number of times each bivariate copula is selected in each tree, where Π notes the bivariate independence copula *(continued)*.

Tree	Π	N	t	G	F	G 180	G 90	G 270
45	12	2	2	2	8	0	0	0
46	8	2	2	3	7	2	0	1
47	12	1	1	2	6	1	0	1
48	12	2	2	2	5	0	0	0
49	13	3	1	0	3	0	0	2
50	9	4	1	1	4	0	2	0
51	11	4	0	1	3	0	0	1
52	10	0	2	1	5	1	0	0
53	8	4	0	1	5	0	0	0
54	9	0	2	0	3	2	0	1
55	8	2	1	0	2	0	3	0
56	6	2	1	0	4	0	1	1
57	10	2	0	1	1	0	0	0
58	8	1	1	1	2	0	0	0
59	2	1	3	1	4	0	0	1
60	4	2	3	0	1	1	0	0
61	7	0	1	0	1	0	1	0
62	5	1	0	1	1	0	0	1
63	5	0	1	0	0	1	0	1
64	5	0	1	0	1	0	0	0
65	4	0	1	0	0	0	1	0
66	2	1	0	0	1	0	1	0
67	2	1	1	0	0	0	0	0
68	2	1	0	0	0	0	0	0
69	2	0	0	0	0	0	0	0
70	1	0	0	0	0	0	0	0

D.4.3 Goodness-of-Fit

Table D.11: Fitted models on he first 5 assets in the technology sector, number of parameters, Estimated log-likelihood, AIC and BIC for the fitted models and GOF test

Model	par	loglike	AIC	BIC	White test (p-value)
R-vine AIC seq	13	4432.921	-8839.841	-8771.345	0.943
R-vine AIC seq mle	13	4431.709	-8837.418	-8768.922	0.887
R-vine tau seq	13	4430.336	-8834.673	-8766.177	0.913
R-vine tau seq mle	13	4431.665	-8837.329	-8768.833	0.884
Gaussian seq	10	3976.458	-7932.916	-7880.227	1.000
Gaussian seq mle	10	3976.459	-7932.917	-7880.228	1.000
Student's t seq	20	4404.905	-8769.809	-8664.431	1.000
Student's t seq mle	20	4404.031	-8768.062	-8662.683	1.000

Table D.12: Fitted models on he first 10 assets in the technology sector, number of parameters, Estimated log-likelihood, AIC and BIC for the fitted models and GOF test

Model	par	loglike	AIC	BIC	White test (p-value)
R-vine AIC seq	50	7017.151	-13934.30	-13670.86	1
R-vine AIC seq mle	50	7016.548	-13933.10	-13669.65	1
R-vine tau seq	46	6996.956	-13901.91	-13659.54	1
R-vine tau seq mle	46	7001.269	-13910.54	-13668.17	1
Gaussian seq	45	6349.087	-12608.17	-12371.07	1
Gaussian seq mle	45	6349.088	-12608.18	-12371.07	1
Student's t seq	90	6946.603	-13713.21	-13239.00	1
Student's t seq mle	90	6951.437	-13722.87	-13248.67	1

Table D.13: Fitted models on he first 20 assets in the technology sector, number of parameters, Estimated log-likelihood, AIC and BIC for the fitted models and GOF test

Model	par	loglike	AIC	BIC	White test (p-value)
R-vine AIC seq	187	13232.27	-26090.54	-25105.25	1
R-vine AIC seq mle	187	13232.99	-26091.98	-25106.69	1
R-vine tau seq	159	13158.11	-25998.22	-25160.46	1
R-vine tau seq mle	159	13166.07	-26014.13	-25176.37	1
Gaussian seq	190	11981.90	-23583.79	-22582.70	1
Gaussian seq mle	190	11981.90	-23583.79	-22582.70	1
Student's t seq	380	13054.47	-25348.93	-23346.74	1
Student's t seq mle	380	13073.83	-25387.65	-23385.46	1

D.4.4 Out-of-Sample *CVaR* Forecast

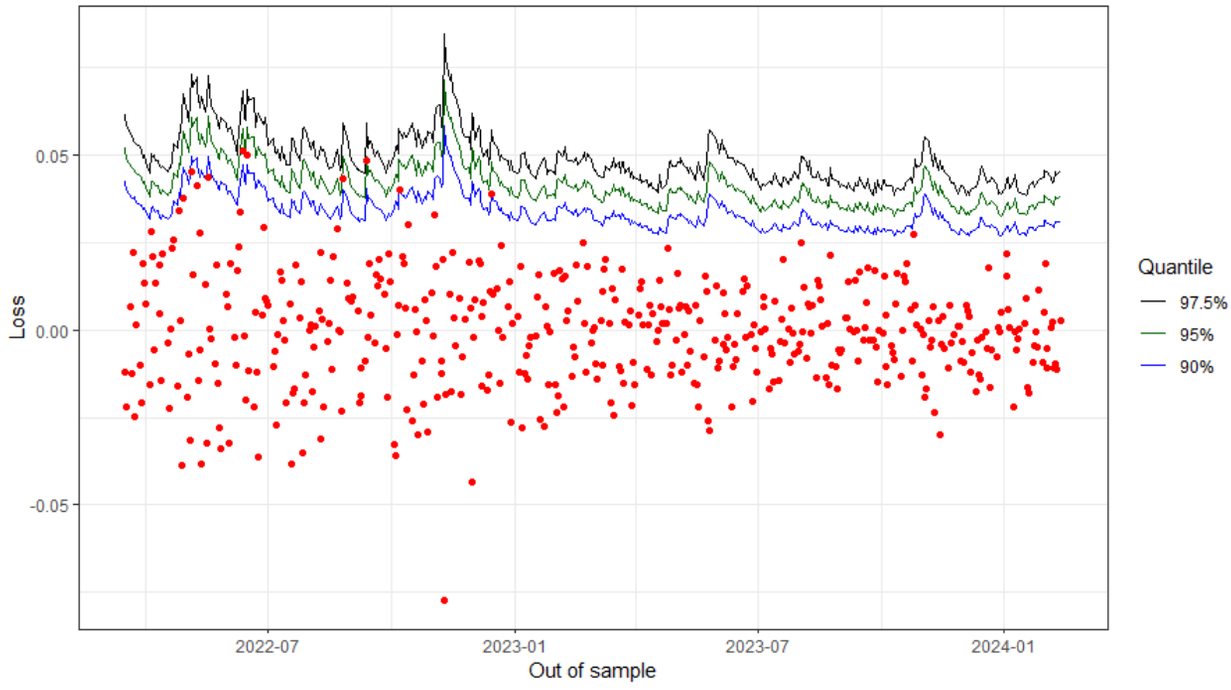


Figure D.12: Out-of-sample CVaR forecast for an equally-weighted portfolio consisting of the assets in the technology sector.

E | Shorted Period *(C)VaR* Forecasts

In this appendix, the performance of the $(C)VaR$ forecasts is evaluated on a shorter time period, for the financial sector, due to observed changes in tendencies in the price movements of in Figure 4.1. The new period goes from 07/07/2015 to 02/01/2020, the training/test split is 75/25, and the procedure from Section 4.4 is used on the data.

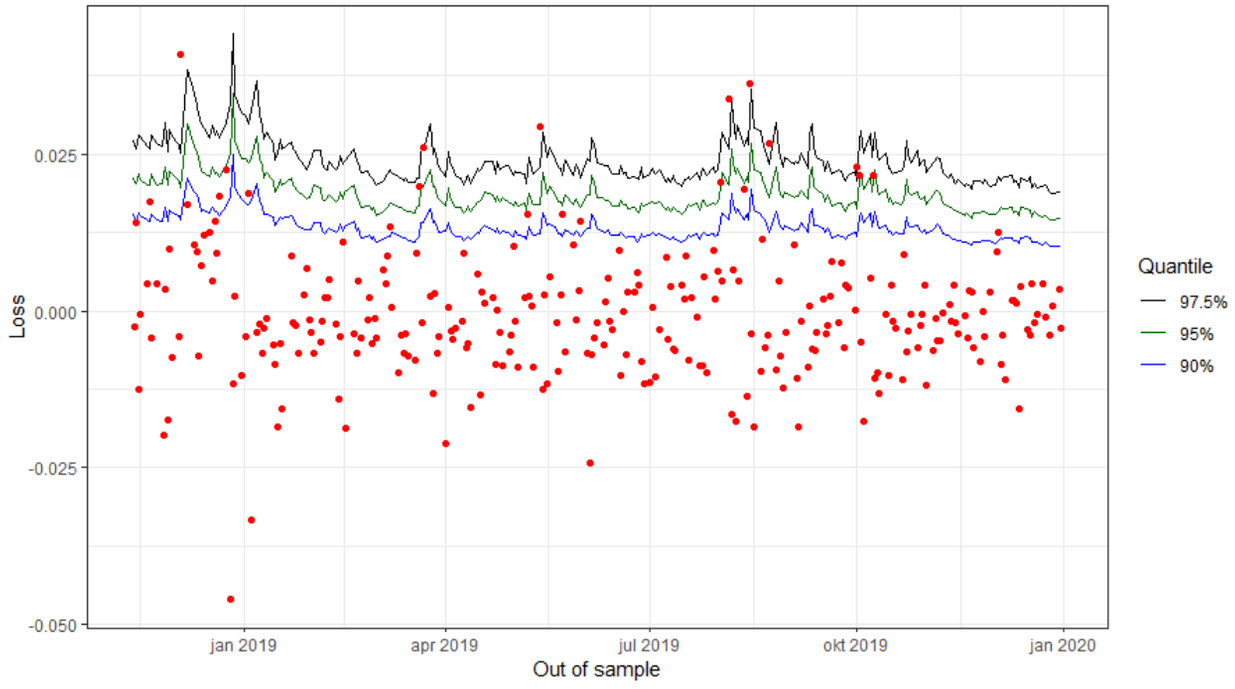


Figure E.1: VaR forecasts for an equally-weighted portfolio consisting of the assets in the financial sector from 14/11/2018 to 02/01/2020. The red dots are the actual losses for the equally-weighted portfolio.

Table E.1: Observed αVaR violations compared to the expected number of αVaR violations, and the ratio between observed αVaR violations and the total number of observations for each tested α-level in each sector from 14/11/2018 to 02/01/2020.

Sector	α-level	Observed violations	Expected violations	Mean violations
Financial	0.025	7	7.075	0.025
	0.05	11	14.150	0.039
	0.10	21	28.300	0.074

From Figure E.1 and Table E.1, it can be inferred that the model performs better in general in terms of observed and expected violations, than the model for the whole period. The

ratio for $\alpha = 0.10$ is a bit lower, however, for $\alpha = 0.025$ and 0.05 the ratio is significantly better than in Table 4.11. To further evaluate the performance on the shorted period, the DQ test is performed as well.

Table E.2: DQ test performed on the financial sector for the shorted period for $\alpha = 0.025, 0.05,$ and 0.10 and lag $p = 4$.

Sector	DQ 2.5% VaR p-value	DQ 5% VaR p-value	DQ 10% VaR p-value
Financial	0.916	0.363	0.496

From Table E.2, it can be inferred that the results from the DQ test on the shorted period are significantly better than for the whole period, Table 4.12. The p-values are all above 0.05, compared to all being below for the whole period.

The validity of the $CVaR$ forecasts on the shorted period, will be performed in the same manner as for the whole period.

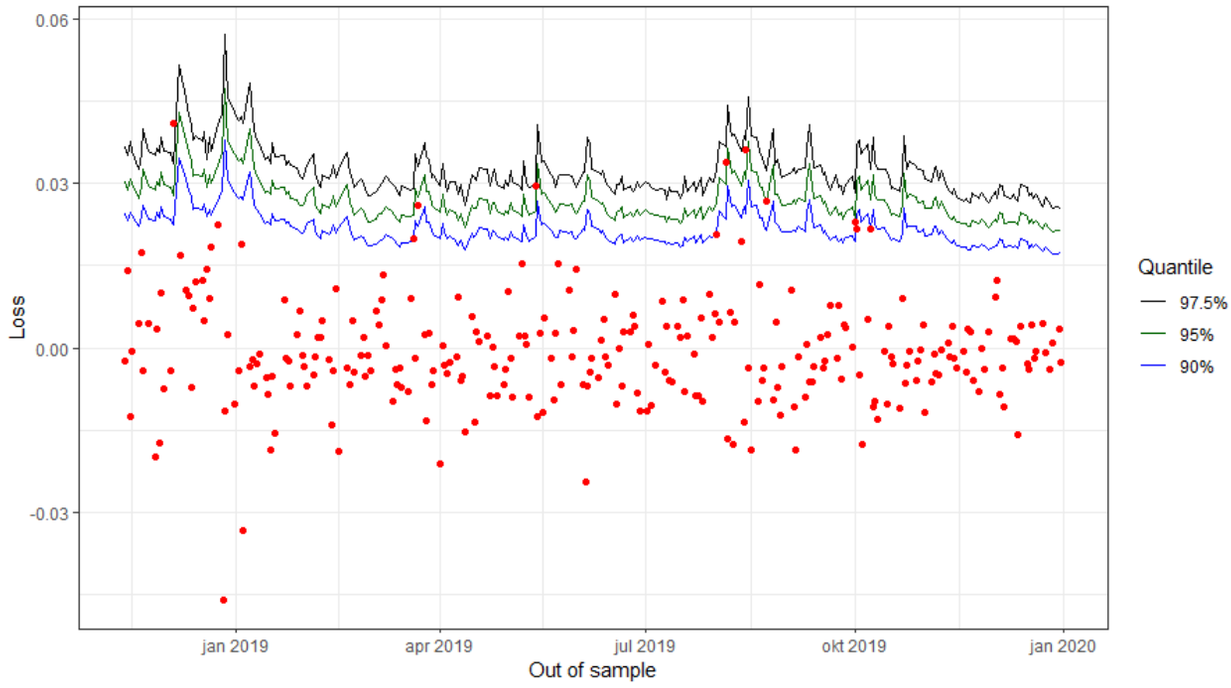


Figure E.2: $CVaR$ forecasts for an equally-weighted portfolio consisting of the assets in the financial sector from 14/11/2018 to 02/01/2020. The red dots are the actual losses for the equally-weighted portfolio.

Table E.3: Strict $CVaR$ backtest performed on the financial sector for $\alpha = 0.025, 0.05,$ and 0.10.

Sector	2.5% CVaR p-value	5% CVaR p-value	DQ 10% CVaR p-value
Financial	0.746	0.504	0.212

The strict $CVaR$ backtest on the shorted period infers that the $CVaR$ forecast are significantly better for the shorted period than for the whole period. The p-values are all above 0.05, whereas the p-values were all zero for the whole period. Therefore, Table E.3, further confirms that the model performs better on the shorted period, where the tendencies of the price movements are more similar.

F | Mean-$CVaR$ with Different Returns

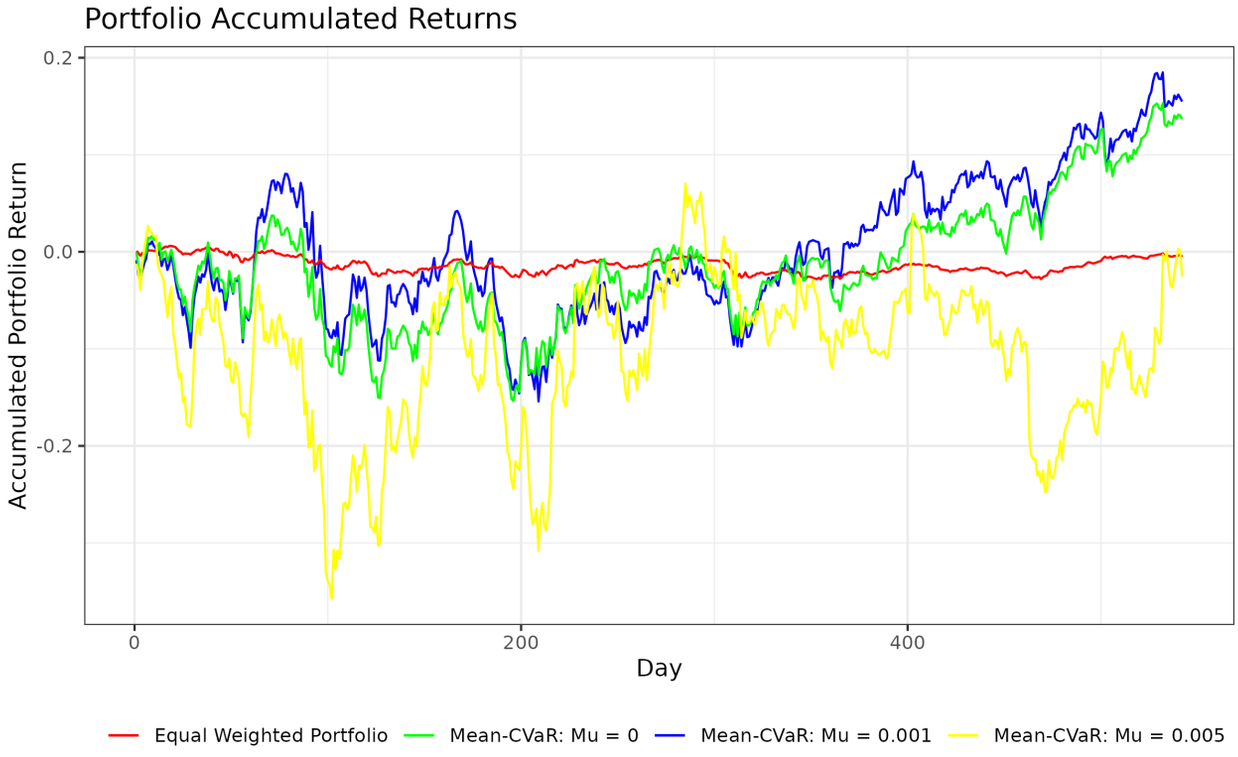


Figure F.1: Accumulated portfolio returns for the mean-$CVaR$ portfolio, with $\mu = 0.001$ (blue), $\mu = 0$ (green), and $\mu = 0.005$ (yellow), and equally-weighted portfolio (red) for the financial sector.