TUM

Technische Universität München

DEPARTMENT OF MATHEMATICS

FINANCIAL RISK MEASURES ESTIMATION USING VINE COPULA MODELS: APPLICATION TO BVK PORTFOLIO
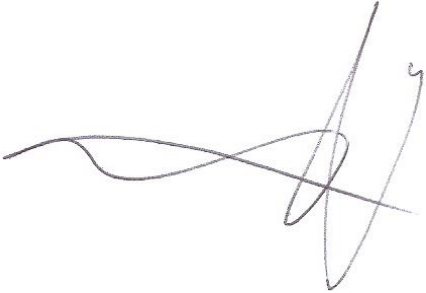
Javier Blasco Aguado

Supervisor:	Prof. Claudia Czado, Ph.D.
Advisor:	Prof. Claudia Czado, Ph.D.
	Karoline Bax, Ph.D.
	Stephan Zeisberger, M.Sc.

Master's Thesis

July 2023

I hereby declare that this thesis is entirely the result of my own work except where otherwise indicated. I have only used the resources given in the list of references.

Munich, 31/07/2023   Javier Blasco Aguado

Abstract

In finance, obtaining a precise estimation of the risk measures of the investment portfolio is crucial for achieving good performance.

This thesis compares three methods for estimating the value at risk (VaR) and expected shortfall (ES), the most important risk measures in portfolio analysis. Two of the methods capture cross-sectional and serial dependence, while the third one uses the first four moments of the portfolio assets and ignores serial and cross-sectional dependencies. The first model uses an ARMA-GARCH vine copula model utilising the software developed by Sommer (2022). In the first step it models the serial dependence with ARMA-GARCH models for each asset, separately. In a second step R-vine copulas are used to quantify cross-sectional dependence. The second method uses S-vine copulas proposed by Nagler et al. (2022) for modelling both dependencies at the same time based on translation invariance. For the last method Fleishman's transformation introduced in Fleishman (1978) is used for estimating the risk measures. For all the methods, the estimation of risk measures is performed using Monte Carlo on the portfolio's log returns. To conduct a time varying analysis, it is applied within a rolling window. The research includes several evaluation measures and backtests to compare the methods.

A case study is conducted representing the BVK portfolio, composed of private equity, equity, fixed income, real estate and hedge fund indices. Two markets with 15 and 7 assets each, and three portfolios (equal weighted, market capitalization, and BVK) are considered in the study over a time period of 2005 to 2022. This comprehensive analysis helps us discover the strengths and weaknesses of each method.

Contents

List of Figures

List of Tables

1 Introduction

The modern landscape of financial markets is characterised by increasing complexity and globalisation, which has amplified the need for effective risk management strategies. In this context, financial risk measures play a crucial role in assessing and quantifying the potential risks associated with various financial instruments, portfolios, and investment strategies. Accurate estimation of risk measures is essential for making informed investment decisions, optimising portfolios, and safeguarding against adverse market conditions.

These measures are typically calculated to assess the probability of incurring losses and, if they occur, quantifying them. Therefore, the most widely used risk measures in the world of finance are the Value at Risk (VaR) and the Expected Shortfall (ES). Historically, different methods have been developed and used to calculate these. In general, we can highlight estimation through observed historical returns or through the use of Monte Carlo simulation. However, these methods are based solely on past observations without taking into account the temporal dependency or serial dependence between the observations. Therefore, in this thesis, we present two methods for estimating these risk measures by allowing for serial dependence modelling and compare them with a third method that does not model it.

Thus, considering a portfolio composed of multiple assets, we aim to model the dependence that arises over time between the returns for each asset, known as serial dependence, as well as the dependence between the assets that make up the portfolio at each moment in time, known as cross-sectional dependence. The first approach in the study is based on Maarouf (2021). In a first step it models the serial dependence using ARMA-GARCH models for each asset. These models, introduced by Box and Jenkins (1976) and Engle (1982) respectively, capture the trend and volatility of each series through a linear regression based on previous observations. In a second step, cross sectional dependence is captured using vine copula models Czado (2019). These methods are based on modelling the joint dependence of multiple variables using the vine structure, which enables measuring pairwise dependence.

The second method is based on Nagler et al. (2022), a special class of vine copulas to simultaneously model serial and cross-sectional dependence. These models, known as stationary vine copulas, model the serial dependence by assuming that the time series are stationary.

As this work is conducted in collaboration with Bayerische Versorgungskammer, the method employed by the company is also introduced in the study. This method does not model serial or cross-sectional dependence, nor does it assume any distribution in the returns. Based on Fleishman (1978), this method utilises the first four moments of the assets forming the portfolio to estimate the risk measures.

The comparison of the methods is carried out in the case study. For this purpose, we consider a market consisting of 15 indices representing the BVK portfolio over 17 years, to include periods of high and low volatility. Due to the lack of convergence of the Fleishman's transformation method, a reduced study with 7 indices is also conducted. The results obtained in the study show that the ARMA-GARCH model combined with vine copulas provides the best estimation of the risk measures. Both stationary vine copulas and Fleishman's transformation show similar results. However, Fleishman's transformation convergence is not always ensured.

The thesis is divided into two parts: theoretical and practical. In the theoretical part, we introduce all the necessary concepts and models to implement the methods, covering Sections 2 to 10. We begin by introducing some basic concepts and notation used throughout the work. In Section 3, we briefly define the main characteristics of financial series. Risk measures are

1

explained in Section 4, where we establish the foundation of the theory since the objective of the work is to obtain the best estimation of these measures. Once all the background is introduced, we develop all the necessary models for the different estimation methods. In Section 5, we develop the ARMA-GARCH models to model univariate financial series. We continue in Section 6 by presenting vine copulas, and in Section 7, we extend them to stationary vine copulas. In Section 8, we introduce the Fleishman's transformation. Moving on to Section 9, we introduce measures and backtests to assess and compare the risk measures we want to estimate. Finally, in Section 10, we describe the methods employed using the described models and explain how to estimate the risk measures. In the practical part, we apply the methods developed earlier to a database introduced in Section 11. Three portfolios are considered, all composed of the same assets, which are the main components of BVK's investment portfolio. Due to the limitations of the Fleishman's transformation method, a reduced portfolio is also employed. In Section 12, we estimate the risk measures for the complete portfolio. Subsequently, in Section 13, we do the same for the reduced portfolio.

2 General notation and background

Definition 2.1. *Let $F_X : \mathbb{R} \to [0,1]$ be the distribution function of a random variable X, i.e. $F_X(x) := \mathbb{P}(X \leq x)$ for all $x \in \mathbb{R}$. Then*

$$F_X^{\leftarrow} : (0,1) \to \mathbb{R}$$
$$y \mapsto inf\{x \in \mathbb{R} \colon F_X(x) \geq y\}$$

*is the **generalised inverse** or **quantile function**.*

Definition 2.2. *Let $X_1, ..., X_n$ be i.i.d. random variables with distribution function F and $x_{1,n} \leq x_{2,n} \leq \ldots \leq x_{n,n}$ the corresponding order statistics. Then,*

- *the **empirical distribution function** is given by*

$$F_n(x) = \frac{1}{n} \sum_{k=1}^{n} 1_{\{x_k \leq x\}} \quad x \in \mathbb{R},$$

i.e. $F_n(x) = k/n$ for $x_{k,n} \leq x \leq x_{k+1,n}$.

- *the **empirical quantile function** is given by*

$$F_n^{\leftarrow}(y) = inf\{x \in \mathbb{R} : F_n(x) \geq y\} = x_{[yn],n} \quad y \in [0,1], \tag{2.1}$$

where $[z] = inf\{x \in \mathbb{Z} : z \leq x\}$.

Proposition 2.1. *Let X be a random variable with distribution function F_X and quantile function $F_X^{\leftarrow}$. Then:*

- *i) Let $U \sim \mathbf{U}(0,1)$, then $F_X^{\leftarrow}(U) = X$*

- *ii) $F_X = U \Leftrightarrow F$ is continuous.*

Proof. Proof is found in Angus (1994) $\qquad \qquad \square$

3 Financial time series

In this section we introduce the most important characteristics of financial time series. This term refers to the time series used in finance for individual assets prices and returns, but also for portfolio prices and return values.

For years, many studies have been done for forecasting financial prices. However it is very difficult to predict their movement, since it depends on many variables, including human factors. Also, the use of different currencies makes it difficult to compare assets. To solve these problems usually financial time series are compared at the return scale. Thus, given a time series p_t, which represent the price of any asset at time t, we define the **log returns** as follow:

$$r_t = ln\left(\frac{p_t}{p_{t-1}}\right) \tag{3.2}$$

considering that p_t takes only positive values. Since, financial prices changes constantly over time, this is not a big restriction. However, a discrete version of the log returns can also be used $r_t = \frac{p_t - p_{t-1}}{p_{t-1}}$.

Most of return time series show similar characteristics. Cont (2001) summarises the most important characteristics or stylised facts of financial asset returns introduced in Mandelbrot (1963):

- **Absence of autocorrelations.** Linear autocorrelations of asset returns are often small or insignificant. We look at it in more depth in Section 5.

- **Heavy tails.** Returns distributions usually show a non Gaussian behaviour with kurtosis greater then three, $K > 3$. However the precise forms of the tails is difficult to determine.

- **Gain/loss asymmetry.** Stock markets usually shows larger draw downs in stock prices than upward movements.

- **Volatility clustering.** The volatility of returns usually show clustering. This means, high volatility periods are follow by high volatility periods and the same for low volatility periods.

- **Autocorrelation on the squared returns.** While returns show near zero autocorrelation, absolute and squared returns show significant autocorrelations. Check Section 5.2.

- **Leverage effect.** Volatility is affected differently by positive and negative returns. Usually negative returns increases volatility over positive returns. Most time series show negative skewness , $sk < 0$.

Figure 3.1 show the returns time series of Microsoft as an example of a financial time series. In this case, we have the daily closing price from 23/04/2018 until 19/04/2023. Some of the stylized facts are observed directly, the volatility clustering is very clear since high and low volatility is observed along different periods. It is also observable that the mean of the log returns is around zero. Heavy tails and leverage effects are not so visible in Figure 3.1, however estimating kurtosis and skewness we obtain $\hat{k} = 7.1$ and $\hat{sk} = -0.28$ respectively.

Considering a market with N assets we can combine them linearly to obtain a portfolio. Thus, considering δ_i the amount of money invested in asset i with the corresponding price $p_{i,t}$ at time t, the value of the portfolio V_t^* at time t is:
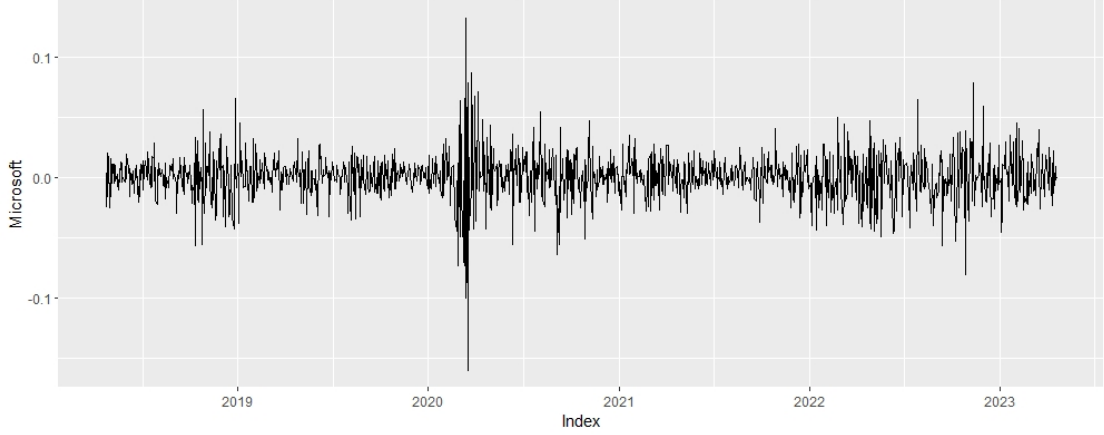
Figure 3.1: Microsoft returns time series between 23/04/2018 - 19/04/2023.

$$V_t^* = \sum_{i=1}^{N} \delta_i p_{i,t} \tag{3.3}$$

The value of the portfolio can also be written using the weight scale, defining the **weight of asset i in portfolio** P, $w_i^P = \frac{\delta_i}{\sum_{i=1}^{N} \delta_i}$, such that $\sum_{i=1}^{N} w_i^P = 1$. Equivalent to (3.3) we can write the value of the portfolio using the weights as follow:

$$V_t = \sum_{i=1}^{N} w_i p_t^i \tag{3.4}$$

Then, using (3.2) we obtain the returns of the portfolio $r_t^V = ln \frac{V_t}{V_{t-1}}$.

4 Risk measures

In this chapter we define the different risk measures and their characteristics. First we introduce some background definitions on financial theory.

Definition 4.1. *The triplet* $(\Omega, \mathcal{F}, \mathbb{P})$ *is called a* **probability space** *if:*

- *the* **sample space**, Ω, *is a non-empty set.*

- *a* $\boldsymbol{\sigma - algebra}$, $\mathcal{F}$, *on* Ω *is a system of subsets of* Ω *satisfying:*

  1. $\Omega \in \mathcal{F}$.

  2. $A \in \mathcal{F} \Rightarrow A^c := \Omega \setminus A \in \mathcal{F}$.

  3. For each sequence $\{A_n\}_{n\in\mathbb{N}}$, *with* $A_n \in \mathcal{F}$ *for all* $n \in \mathbb{N}$ *it holds that* $\left(\bigcup_{n\in\mathbb{N}} A_n\right) \in \mathcal{F}$.

- *a* **probability measure**, $\mathbb{P}$, *on the* **measurable space**, $(\Omega, \mathcal{F})$ *is a map* $\mathbb{P} : \mathcal{F} \to [0, \infty]$ *with* $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$ *and each sequence of* $\{A_n\}_{n\in\mathbb{N}}$ *of disjoint sets from* $\mathcal{F}$ *satisfies* $\mathbb{P}(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \mathbb{P}(A_n)$.

Definition 4.2. *A* **financial position** *is a mapping* $X : \Omega \to \mathbb{R}$, *being* Ω *a sample space.* $X(\omega)$ *is the* **discounted net worth** *of the financial position, for* $\omega \in \Omega$. *Moreover,* $X \in \mathcal{X}$, *the* **space of financial positions**.

In our case Ω represents the set of all possible scenarios, including market conditions, regulations, etc. Also, the discounted net worth is the profit and losses of the position at the end of the trading period.

Definition 4.3. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space and* $\mathcal{X}$ *be a non-empty set of* $\mathcal{F}$*-measurable real-valued random variables. Then any mapping* $\rho : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ *is called a* **risk measure**.

Definition 4.4. *Given a risk measure* $\rho : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ *we introduce some possible properties for* ρ:

1. **Monotonicity**. *For all* X *and* $Y \in \mathcal{X}$, $X \leq Y \Rightarrow \rho(X) \geq \rho(Y)$.

   Monotonicity means that for two different financial positions X *and* $Y \in \mathcal{X}$. *If the profits and losses of* $X(\omega) \leq Y(\omega)$ $\forall \omega \in \Omega$ *then the risk of financial position* X *is higher than the risk of financial position* Y.

2. **Translation invariance**. *For all* $X \in \mathcal{X}$ *and all real numbers* $m \in \mathbb{R}$, *we have* $\rho(X+m) = \rho(X) - m$.

   Translation invariance axiom means that adding a cash amount of m *to the initial position, simply decreases (resp. increases) the risk measure by* m.

3. **Convexity**. *For all* X *and* $Y \in \mathcal{X}$ *and* $\lambda \in [0,1]$, $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$.

   Convexity refers that diversification reduce risks.

4. **Positive homogeneity**. *For all* $X \in \mathcal{X}$ *and* $\lambda \geq 0$, $\rho(\lambda X) = \lambda\rho(X)$.

   Positive homogeneity refers to the size of the financial position. The size of the position does not affect on the risk.

5. **Subadditivity**. *For all X and $Y \in \mathcal{X}$, $\rho(X+Y) \leq \rho(X) + \rho(Y)$*

   Subadditivity stated that a merger does not create extra risk.

   Then, ρ is called a:

   - *monetary risk measure, if it fulfils monotonicity and translation invariance.*

   - *convex risk measure, if it fulfils monotonicity, translation invariance and convexity.*

   - *coherent risk measure, if it fulfils monotonicity, translation invariance, convexity and positive homogeneity.*

Definition 4.5. *Let ρ be a monetary risk measure. Then the **acceptance set of** ρ is defined as:*

$$\mathcal{A}_\rho := \{X \in \mathcal{X} : \rho(X) \leq 0\}.$$

4.1 Value at risk

Definition 4.6. *Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{X} = \mathcal{L}^0(\Omega, \mathcal{F}, \mathbb{P})$ the set of all random variables corresponding to the probability space, which are $\mathcal{P} - a.s.$ finite, i.e. $\mathbb{P}(X < \infty) = 1$. We define the **value at risk at level** α **on the financial position** $X \in \mathcal{X}$ $(VaR_\alpha(X))$ as follow:*

$$VaR_\alpha(X) := \inf\{m \in \mathbb{R} : \mathbb{P}(m + X < 0) \leq \alpha\}.$$

By definition $VaR_\alpha(X)$ is the smallest amount of capital m which, if added to X and invested into the risk free asset, keeps the probability of a negative outcome below the level α.

From Definition 4.6 and Definition 2.1 we obtain:

$$
\begin{aligned}
VaR_\alpha(X) :&= \inf\{m \in \mathbb{R} : \mathbb{P}(m + X < 0) \leq \alpha\} \\
&= \inf\{m \in \mathbb{R} : \mathbb{P}(X < -m) \leq \alpha\} \\
&= \inf\{m \in \mathbb{R} : \mathbb{P}(X \geq -m) \geq 1 - \alpha\} \\
&= \inf\{m \in \mathbb{R} : \mathbb{P}(-X \leq m) \geq 1 - \alpha\} = F^{\leftarrow}_{-X}(1 - \alpha)
\end{aligned}
\tag{4.5}
$$

Proposition 4.1. *Value at risk is a monetary risk measure on $\mathcal{X}$ and it is positively homogeneous.*

The VaR can be estimated using the theoretical distribution of X as shown in Equation (4.5). Also, a Monte Carlo approach is utilised using the empirical quantile function. Considering $x_1, x_2, \ldots, x_n$ sample obtained from the random variable $X \in \mathcal{X}$, we obtain the following empirical VaR estimation:

$$q_\alpha(X) := \widehat{VaR}_\alpha(X) = F^{\leftarrow}_{-n}(1 - \alpha) \tag{4.6}$$

where $F^{\leftarrow}_{-n}$ represents the empirical loss quantile function of $-x_{n,n} \leq -x_{n-1,n} \leq \ldots \leq -x_{1,n}$ defined in Equation (2.1).

Criticism on VaR

Value at risk is well known and used in the finances. However, it shows some limitations. The main limitations are :

- VaR only controls the probability of a loss, but it does not capture the size of such a loss if it occurs.

- The acceptance set of VaR is typically not convex, hence VaR is not a convex risk measure. As a consequence, VaR may penalise diversification as shown in Artzner et al. (1999)

For solving this limitations we also introduce another risk measure, the expected shortfall.

4.2 Expected shortfall

As we have seen, value at risk can not control the size of the loss in case of occurring. For measuring the loss we introduce now the concept of expected shortfall.

Definition 4.7. *We define **average value at risk at level** $\alpha \in (0,1)$ **of a financial position** $X \in \mathcal{X}$, $AVaR_\alpha(X)$, as follow*

$$AVaR_\alpha(X) := \frac{1}{\alpha} \int_0^\alpha VaR_\gamma(X) \, d\gamma$$

*The **expected shortfall at level** $\alpha \in (0,1)$ **on the financial position** $X \in \mathcal{X}$ is defined by*

$$ES_\alpha(X) := E[-X| - X \geq VaR_\alpha(X)]$$

Then, the expected shortfall measures the loss, being calculated as the expected loss when it happens.

Proposition 4.2. *If $X \in \mathcal{X}$ has a continuous distribution function, then for $\alpha \in (0,1)$:*

$$AVaR_\alpha(X) = ES_\alpha(X) := E[-X| - X \geq VaR_\alpha(X)] = \frac{1}{\alpha} \int_0^\alpha VaR_\gamma(X) \, d\gamma \geq VaR_\alpha(X)$$

Proof. First we should remark that being U a uniform variable, then $U|U \geq 1 - \alpha$ is a uniform distribution on $[1 - \alpha, 1]$. Thus it density is given by $f_{U|U \geq 1-\alpha}(u) = \frac{1}{1-(1-\alpha)} \mathbb{1}_{[1-\alpha,1]}(u) = \frac{1}{\alpha} \mathbb{1}_{[1-\alpha,1]}(u)$. Then, we have from the definition of expected shortfall and using Proposition 2.1 i):

$$\begin{aligned}
ES_\alpha(X) &= E[-X| - X \geq VaR_\alpha(X)] \\
&= E[F_{-X}^\leftarrow(U)|F_{-X}^\leftarrow(U) \geq F_{-X}^\leftarrow(1-\alpha)] \\
&= E[F_{-X}^\leftarrow(U)|U \geq 1 - \alpha] \\
&= \int_{1-\alpha}^1 F_{-X}^\leftarrow(x) \frac{1}{\alpha} \, dx \\
&= \frac{1}{\alpha} \int_0^\alpha F_{-X}^\leftarrow(1-x) \, dx \\
&= \frac{1}{\alpha} \int_0^\alpha VaR_\alpha(x) \, dx = AVaR_\alpha(X)
\end{aligned}$$

$\square$

Embrechts and Wang (2015) prove using several approaches that the expected shortfall is a coherent risk measure. Then, expected shortfall fulfils convexity and positive homogeneity, which implies that portfolio diversification reduce the expected shortfall.

Expected shortfall estimation is usually done with a Monte Carlo approach. Using VaR value, we estimate ES as the mean of the loss empirical distributions which are higher than VaR at level α. Then, given the observed sample $x_1, x_2 \ldots, xn$ from the random variable $X \in \mathcal{X}$, we estimate the expected shortfall:

$$\widehat{ES}_\alpha(X) = \frac{1}{\#\{x_t | -x_t \geq VaR_\alpha(X)\}} \sum_{t \in \{1, \ldots, T\} | -x_t \geq VaR_\alpha(X)} -x_t \tag{4.7}$$

5 ARMA-GARCH models

In this section we give a short introduction to ARMA-GARCH models. These are used to model univariate time series. ARMA models the mean part of the time series are introduced in Section 5.1, while GARCH models the variance are in Section 5.2. Before developing ARMA-GARCH models we introduce some notions related to the time series.

A time series $\{r_t, t \in T\}$ is considered as an specific realisation of the stochastic process $\{R_t, t \in T\}$, where the index t usually represents time. Since we only have one observation per time point, we introduce stationarity.

Definition 5.1. *A process $\{R_t, t \in T\}$ is **strictly stationary** if the joint distribution of all sizes does not change over time:*

$$F(R_{t_1}, R_{t_2}, ..., R_{t_i}) = F(R_{t_1+k}, R_{t_2+k}, ..., R_{t_i+k}) \quad \forall i, k$$

Due to the lack of information for checking the strictly stationarity in observed time series we define a weaker version of stationarity used throughout this section.

Definition 5.2. *A process $\{R_t, t \in T\}$ is **weakly stationary** if both the mean of R_t and the covariance between R_t and R_{t-k} are time invariant, where k is an arbitrary integer. More specifically we have:*

$$
\begin{aligned}
E[R_t] &= \mu \quad constant \\
Cov(R_t, R_{t-k}) &= \gamma_k
\end{aligned}
\tag{5.8}
$$

In Equation (5.8), we assume that the first two moments of R_t are finite. Also, the $\gamma_k = Cov(R_t, R_{t-k})$ is called **the lag-l autocovariance of** R_t. It has the properties that $\gamma_0 = Var(R_t)$ and $\gamma_{-k} = \gamma_k$. In the finance literature, it is common to assume that a return series, $\{r_t, t \in T\}$, is weakly stationary.

An important characteristic of stationary processes is that they are closed under linear combination, i.e. the linear combination of stationary processes is also stationary.

For modelling using ARMA-GARCH models we also need the concepts of white noise, autocorrelation function and partial autocorrelation function.

Definition 5.3. *A time series process A_t is called a **white noise** if $\{A_t, t \in T\}$ is a sequence of independent and identically distributed with finite mean and variance:*
1) $E[A_t] = 0 \quad \forall t.$
2) $Var(A_t) = \sigma_A^2 \quad \forall t.$
3) $Cov(A_t, A_{t+k}) = 0 \quad \forall t, k.$

Definition 5.4. *Consider a weakly stationary time series process $\{R_t, t \in T\}$. The **autocorrelation function (ACF)** is the correlation coefficients ρ_k between R_t and R_{t-k} for all k. For each k, the correlation is called **lag-k autocorrelation** of R_t and it is obtained as follow:*

$$\rho_k = \frac{Cov(R_t, R_{t-k})}{\sqrt{Var(R_t)Var(R_{t-l})}} = \frac{Cov(R_t, R_{t-k})}{Var(R_t)} = \frac{\gamma_k}{\gamma_0}$$

From the definition we obtain that ρ_k meets the following characteristics:

- $\rho_k = \rho_{-k}$

11

- $-1 \leq \rho_k \leq 1$

- $\rho_0 = 1$

Definition 5.5. *Given a weakly stationary time series process $\{R_t, t \in T\}$. The **partial correlation coefficient of order k**, ρ^p_{kk} is the correlation coefficient between R_t and R_{t-k}, after the effect of the intermediate variables $R_{t-1}, ..., R_{t-(k-1)}$ are accounted for:*

$$\rho^p_{kk} = Corr(R_t, R_{t-k}|R_{t-1}, ..., R_{t-(k-1)}),$$

Wei (2005) shows that the partial correlation between R_t and R_{t-k} can also be obtained as the regression coefficient associated with R_{t-k} when regressing R_t on its k lagged variables $R_{t-1}, ..., R_{t-k}$ as follow:

$$R_t = \alpha_{k1} R_{t-1} + \alpha_{k2} R_{t-2} + ... + \alpha_{kk} R_{t-k} + A_t$$

with $\rho^p_{kk} = \alpha_{kk}$ and A_t being a white noise.

*We define the **partial autocorrelation function (PACF) of order k** to the first k partial correlation coefficients.*

In Wei (2005) it is shown that PACF can be obtained as follow:

$$\rho^p_{11} = \rho_1 \tag{5.9}$$

$$\rho^p_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} \tag{5.10}$$

$$\vdots$$

$$\rho^p_{ll} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{l-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{l-3} & \rho_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{l-1} & \rho_{l-2} & \rho_{l-3} & \cdots & \rho_1 & \rho_l \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{l-2} & \rho_{l-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{l-3} & \rho_{l-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{l-1} & \rho_{l-2} & \rho_{l-3} & \cdots & \rho_1 & 1 \end{vmatrix}} \tag{5.11}$$

being ρ_i the lag-i autocorrelation of r_t.

5.1 ARMA models

ARMA models were introduced in Box and Jenkins (1976) for modelling the mean of univariate time series. These models are divided in two parts, MA given in Section 5.1.1, and AR in Section 5.1.2. In Section 5.1.3 we define ARMA models as a combination of MA and AR. Finally, in Sections 5.1.4 and 5.1.5 we give the estimation and forecast method for ARMA models.

5.1.1 Moving average models (MA)

Moving average (MA) models were introduced by Box and Jenkins (1976). These models explain the corresponding time series process R_t as a linear combination of the previous errors. The order of the model is given by the number of lagged error terms used in the model. Then we can define a **moving average model of order q MA(q)**:

$$R_t = \mu + A_t - \theta_1 A_{t-1} - ... - \theta_q A_{t-q} \tag{5.12}$$

being μ the mean of the time series R_t, θ_i parameters $\forall i$ and A_{t-j} white noise $\forall j$. From Equation (5.12) we have that MA processes are stationary, since they are a linear combination of stationary processes, A_{t-i} $i \in \{0, 1, ..., q\}$.

Using the **backshift operator** $B^j R_t = R_{t-j}$, the MA(q) model is rewritten:

$$R_t = \mu + (1 - \theta_1 B - ... - \theta_q B^q) A_t = \mu + \theta(B) A_t$$

with $\theta_q(B) = (1 - \theta_1 B - ... - \theta_q B^q)$.

We say that a moving average model, $R_t = \theta_q(B) A_t$ is invertible if R_t can be written as a linear combination of their past observations, $R_{t-1}, R_{t-2}, ...$ Box and Jenkins (1976) showed that a MA process is invertible if the roots of $\theta_q(B) = 0$ lie outside of the unit circle.

To get a better idea of moving average models we show the characteristics of MA(1) process with $\mu = 0$. Thus, from Equation (5.12) we have $R_t = A_t - \theta_1 A_{t-1} = (1 - \theta_1 B) A_t$, which is invertible for $|\theta_1| < 1$. The mean of the process is $E[R_t] = E[A_t] - \theta_1 E[A_{t-1}] = 0$.

The autocovariance function is given by:

$$
\begin{aligned}
\gamma_k = Cov(R_t, R_{t-k}) &= Cov(A_t - \theta_1 A_{t-1}, A_{t-k} - \theta_1 A_{t-(k+1)}) \\
&= E[(A_t - \theta_1 A_{t-1})(A_{t-k} - \theta_1 A_{t-(k+1)})] \\
&= E[A_t A_{t-k}] - \theta_1 E[A_{t-1} A_{t-k}] - \theta_1 E[A_t A_{t-(k+1)}] + \theta_1^2 E[A_{t-1} A_{t-(k+1)}]
\end{aligned}
$$

hence, the autocovariance of the process are:

$$
\gamma_k = \begin{cases} (1 + \theta_1^2) \sigma_A^2, & k = 0 \\ -\theta_1 \sigma_A^2, & k = 1 \\ 0, & k > 1 \end{cases}
$$

with autocorrelation functions:

$$
\rho_k = \begin{cases} \frac{-\theta_1}{1 + \theta_1^2}, & k = 1 \\ 0, & k > 1 \end{cases}
$$

In case of the PACF we obtain from (5.9)-(5.11):

$$\rho_{kk}^p = \frac{-\theta_1^k (1 - \theta_1^2)}{1 - \theta_1^{2(k+1)}} \neq 0, \quad k \geq 1$$

In general, a $MA(q)$ process has only the first $q - lag$ autocorrelations different from zero, and not null partial autocorrelations for all ρ_{kk}^p with $k \in \mathbb{N}$. Then, to identify the correct order of a moving average model we check the empirical ACF and PACF.

5.1.2   Autoregressive models (AR)

The autoregressive model (AR) were also introduced in Box and Jenkins (1976), they explain each time point of the time series process, R_t as a linear combination of the last p time points of the time series, $R_{t-1}, ..., R_{t-p}$; being p the order of the model. Then, we write an **autoregressive model of order p, AR(p)**:

$$R_t = c + \phi_1 R_{t-1} + ... + \phi_p R_{t-p} + A_t \tag{5.13}$$

with A_t being a white noise and c a constant. Also, the model can be rewritten using the backshift operator:

$$\phi_p(B)R_t = c + A_t$$

where $\phi_p(B) = (1 - \phi_1 B - ... - \phi_p B^p)$. Considering a weak stationary autoregressive time series process, the mean is given by:

$$\mu = E[R_t] = c + \phi_1 E[R_{t-1}] + ... + \phi_p E[R_{t-p}] + E[A_t] = c + \phi_1 \mu + ... + \phi_p \mu \tag{5.14}$$

$$\Rightarrow E[R_t] = \mu = \frac{c}{\sum_{j=1}^{p} \phi_j}$$

By definition we see that AR models are written in their invertible process. Box and Jenkins (1976) showed that every autoregressive model can be written as a $MA(\infty)$ process, since $\sum_{j=1}^{p} |\phi_j| < \infty$. Moreover, to be stationary the roots of $\phi_p(B) = 0$ must lie outside of the unit circle.

As we did with MA processes we check the characteristics of $AR(1)$ process with $c = 0$ to have a better understanding of the autoregressive models. We start calculating the autocovariance function:

$$\gamma_k = Cov(R_t, R_{t-k}) = E[R_t R_{t-k}] = E[\phi_1 R_{t-1} R_{t-k}] + E[A_t R_{t-k}]$$

$$\Rightarrow \gamma_k = \phi_1 \gamma_{k-1}, \quad k \geq 1$$

and the autocorrelation function becomes

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1 \frac{\gamma_{k-1}}{\gamma_0} = \phi_1 \rho_{k-1} = \phi_1^k, \quad k \geq 1$$

Hence, when $|\phi_1| < 1$ the process is stationary and the absolute value of ACF decays exponentially.

From Equations (5.9)-(5.11) we calculate the PACF:

$$\rho_{kk}^p = \begin{cases} \rho_1 = \phi_1, & k = 1 \\[2mm] \dfrac{\begin{vmatrix} 1 & \phi_1 \\ \phi_1 & \phi_1^2 \end{vmatrix}}{\begin{vmatrix} 1 & \phi_1 \\ \phi_1 & 1 \end{vmatrix}} = 0, & k = 2 \\[4mm] 0, & k > 3 \end{cases}$$

Hence, the PACF of the AR(1) shows a non-zero spike at lag 1, and then cuts off.

In general, we have that an AR(p) process shows infinitely many non zero spikes in the ACF. However, the PACF is vanished after lag p. This property is useful for identifying an AR model and the corresponding order p.

5.1.3 Autoregressive moving average models (ARMA)

Autoregressive moving average models or ARMA(p,q) are defined as a combination of autoregressive models and moving average models introduced in Sections 5.1.2 and 5.1.1 respectively. Then, using ARMA models the time series process R_t are explained as a linear combination of the previous time points $R_{t-1}, ..., R_{t-p}$ and the previous errors $A_{t-1}, ..., A_{t-q}$. We define an **autoregressive moving average model of orders p and q, ARMA(p,q)**:

$$R_t = c + \phi_1 R_{t-1} + ... + \phi_p R_{t-p} + A_t - \theta_1 A_{t-1} - ... - \theta_q A_{t-q} \tag{5.15}$$

The model defined in Equation (5.15) can be rewritten using the backshift operator as follow:

$$\phi_p(B) R_t = c + \theta_q(B) A_t$$

being $\phi_p(B) = (1 - \phi_1 B - ... - \phi_p B^p)$ and $\theta_q(B) = (1 - \theta_1 B - ... - \theta_q B^q)$. For the process to be invertible, we require that the roots of $\theta_q(B) = 0$ lie outside the unit circle. To be stationary, we require that the roots of $\phi_p(B) = 0$ lie outside the unit circle. Also, we assume that $\theta_q(B) = 0$ and $\phi_p(B) = 0$ share no common roots. Using the same methodology than in Equation (5.14) and using $E[A_{t-i}] = 0 \ \forall i \in \{0, 1, ..., q\}$ we obtain the mean of the process:

$$E[R_t] = \mu = \frac{c}{\sum_{j=1}^{p} \phi_j}$$

Wei (2005) proved the main characteristics of the basic ARMA(1,1) model with $\mu = 0$. Then variance and ACF are given by:

$$Var(R_t) = \gamma_0 = \frac{(1 + \theta_1^2 - 2\phi 1 \theta_1)}{(1 - phi_1^2)} \sigma_A^2$$

$$\rho_k = \begin{cases} \frac{(\phi_1 - \theta_1)(1 - \phi_1 \theta_1)}{1 + \theta_1^2 - 2\phi_1 \theta_1}, & k = 1 \\ \phi_1 \rho_{k-1}, & k \geq 2 \end{cases}$$

Note that the autocorrelation function of ARMA(1,1) model combines characteristics of AR(1) and MA(1) models, following the same pattern as the autocorrelation function of an AR(1). Turning to the PACF, one can show that the PACF of an ARMA(1,1) model does not cut off at any finite lag either. It behaves very much like PACF of an MA(1) model.

This characteristics can be extended to the general ARMA(p,q) model. In Wei (2005) it is shown that ACF and PACF of ARMA(p,q) models have infinitely many non zero spikes. To properly select the order of an ARMA model, it is recommended to start with low orders and analyse the ACF and PACF of the residuals, a_t, and increment the orders until obtain a white noise for the residuals is plausible.

ARIMA or autoregressive integrated moving average models are an extension of ARMA models. These models were defined in Box and Jenkins (1976) and characterised by not being stationary but being prone to becoming stationary with the transformation $(1 - B)R_t$. This correspond to have one root of $\phi_p(B) = 0$ in the unit circle. Then, the ARIMA(p,1,q) is given by

15

$$\phi_p(B)(1-B)R_t = c + \theta_q(B)A_t \tag{5.16}$$

Identification of these models is also done with ACF and PACF, which have infinitely many non zero spikes, then equivalent method to ARMA should be done.

5.1.4 ARMA models estimation

Once we have chosen the best model and order for fitting the available data $\boldsymbol{r} = (r_1, ..., r_T)$ we calculate the corresponding parameter estimation. The objective is to obtain the best values of $\{\widehat{c}, \widehat{\phi}_1, ..., \widehat{\phi}_p, \widehat{\theta}_1, ..., \widehat{\theta}_q, \widehat{\sigma}_A\}$ for a general ARMA(p,q) model. Also, a distribution for the residuals time series, $\{a_t, t \in T\}$, should be chosen. The most common are normal or t-Student distributions. In this case, normal distribution is considered during the section.

Following Wei (2005) three different methods can be used for parameter estimation. The **method of moments** use the sample moments to obtain the Yule-Walker equations, this method is useful for AR models. Also, **conditional and unconditional maximum likelihood methods** are introduced. The difference between them is the assumption of known initial values $\boldsymbol{a_*} = (a_{1-q}, ..., a_{-1}, a_0)$ and $\widehat{\boldsymbol{r}}_* = (\widehat{r}_{1-p}, ..., \widehat{r}_{-1}, \widehat{r}_0)$ which could also be obtained with a backward model.

Then, we show the conditional maximum likelihood method on a general ARMA(p,q) model with normal residuals. Thus, the joint probability density of $\mathbf{a} = (a_1, ..., a_T)$ is given by

$$P(\boldsymbol{a}|\phi, \mu, \theta, \sigma_A^2) = (2\pi\sigma_A^2)^{-T/2} exp\left[\frac{-1}{2\sigma_A^2}\sum_{i=1}^{T} a_i^2\right]$$

Defining $\bar{r}_t = r_t - \mu$ and using Equation (5.15) we obtain:

$$a_t = \theta 1 a_{t-1} + ... + \theta q a_{t-q} + \bar{r}_t - \phi_1\bar{r}_{t-1} - ... - \phi_p\bar{r}_{t-p}$$

Thus, using the already introduced the known initial values, $\boldsymbol{a_*}$ and $\widehat{\boldsymbol{r}}_*$, join with the observed values, $\boldsymbol{r}$; the conditional log-likelihood function is

$$ln(L_*(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \sigma_A^2)) = -\frac{T}{2}ln(2\pi\sigma_A^2) - \frac{S_*(\phi, \mu, \theta)}{2\sigma_A^2}$$

where

$$S_*(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}) = \sum_{i=1}^{T} a_i^2(\phi, \mu, \theta|\widehat{\boldsymbol{r}}_*, \boldsymbol{r}, \boldsymbol{a_*})$$

is the conditional sum of squares function. The values $\widehat{\boldsymbol{\phi}}$, $\widehat{\mu}$ and $\widehat{\boldsymbol{\theta}}$ which maximise $ln(L_*(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \sigma_A^2))$ also minimise the conditional sum of squares function, $S_*(\boldsymbol{\phi}, \mu, \boldsymbol{\theta})$.

Once that $\widehat{\boldsymbol{\phi}}$, $\widehat{\mu}$ and $\widehat{\boldsymbol{\theta}}$ have been calculated we estimate the variance of the residuals, $\widehat{\sigma}_A^2$

$$\widehat{\sigma}_A^2 = \frac{S_*(\widehat{\boldsymbol{\phi}}, \widehat{\mu}, \widehat{\boldsymbol{\theta}})}{degree\ of\ freedom}$$

where the degree of freedom is the number of terms used in $S_*(\widehat{\boldsymbol{\phi}}, \widehat{\mu}, \widehat{\boldsymbol{\theta}})$ minus the number of parameters estimated. In our case the degree of freedom is $T - (p + q + 1)$. After estimating the parameters we are ready to forecast from the estimated model.

16

5.1.5 ARMA models forecasting

When we talk about forecasting we pretend to predict the values of the time series using the info available until that moment. Then for forecasting an ARMA(p,q) model we use the observed data, $r_1, ..., r_T$; and the corresponding estimated parameters, $\widehat{c}, \widehat{\phi}_1 ... , \widehat{\phi}_p, \widehat{\theta}_1, ..., \widehat{\theta}_q$. With this information we calculate the 1-step ahead forecast, $\widehat{\mu}_{t+1}$:

$$\widehat{\mu}_{T+1} = \widehat{c} + \sum_{i=1}^{p} \widehat{\phi}_i r_{T+1-i} - \sum_{j=1}^{q} \widehat{\theta}_j (r_{T+1-i} - \widehat{\mu}_{T+1-i}) \tag{5.17}$$

where we have estimated the errors in MA part with the corresponding forecast $\widehat{a}_t = r_t - \widehat{\mu}_t$. Considering also k-steps ahead, we forecast:

$$\widehat{\mu}_{T+k} = \widehat{c} + \sum_{i=1}^{p} \widehat{\phi}_i \widehat{\mu}_{T+k-i} - \sum_{j=1}^{q} \widehat{\theta}_j \widehat{a}_{T+k-i}$$

being $\widehat{\mu}_{T+k-i} = r_{T+k-i}$ if $k - i \geq 0$ and $\widehat{a}_{T+k-i} = 0$ if $k - i > 0$. Since we have stationary models, the long term forecasts tend to the estimated mean of the model, $\widehat{\mu}$.

5.2 GARCH models

In this section we describe the generalised autoregressive conditional heteroscedasticity models. These models were developed to model the volatility of time series, which make them very useful for financial time series. The basic model was introduced by Engle (1982), to model a white noise process with conditional variance effects. Considering a time series of returns, $\{r_t, t \in T\}$, after fitting the corresponding ARMA model we obtain the residual time series, $\{a_t, t \in T\}$, such that from Equation (5.15) $a_t = r_t - c - \phi_1 r_{t-1} - ... - \phi_p r_{t-p} + \theta_1 a_{t-1} + ... + \theta_q a_{t-q} = r_t - \mu_t$. Then, as shown in Section 5.1 $\{a_t, t \in T\}$ is a white noise, moreover they usually show volatility clustering and autocorrelation on the squared returns, introduced in Section 3.

To include all these characteristics Engle (1982) modelled a white noise process $\{A_t, t \in T\}$ as follow:

$$A_t = \sigma_t \epsilon_t \tag{5.18}$$

being $\{\sigma_t, t \in T\}$ a stationary stochastic process; and $\{\epsilon_t, t \in T\}$ a Gaussian standardised independent white noise process, such that $E(\epsilon_t) = 0$ and $Var(\epsilon_t) = 1 \ \forall t$. Moreover, σ_t and ϵ_t are also independent. Since A_t are white noise, from Definition 5.3 we have $E(A_t) = 0$ and $Var(A_t) = \sigma_A^2$.

Given $\mathcal{F}_t$, the $\sigma - algebra$ $\mathcal{F}_t = \sigma(A_i | i < t)$, the main characteristics of (5.18) are:

$$E(A_t) = E(\sigma_t)E(\epsilon_t) = 0 \tag{5.19}$$

$$E(A_t | \mathcal{F}_{-1}) = E(\sigma_t | \mathcal{F}_t)E(\epsilon_t) = 0 \tag{5.20}$$

$$Var(A_t) = E(A_t^2) - E(A_t)^2 = E(\epsilon_t^2 \sigma_t^2) = E(\epsilon_t^2)E(\sigma_t^2) = E(\sigma_t^2) = \sigma_A^2 \tag{5.21}$$

$$Var(A_t | \mathcal{F}_{-1}) = E(\epsilon_t^2)E(\sigma_t^2 | \mathcal{F}_{-1}) = \sigma_t^2 \tag{5.22}$$

where $\{\sigma_t^2, t \in T\}$ represents the conditional variance or volatility of the time series. Being σ_t and ϵ_t independent and ϵ_t independent process, we calculate the autocovariance of the process $\{A_t^2, t \in T\}$:

$$E(A_t A_{t-k}) = E(\sigma_t \epsilon_t \sigma_{t-k} \epsilon_{t-k}) = E(\epsilon_t \epsilon_{t-k}) E(\sigma_t \sigma_{t-k}) = 0 \qquad (5.23)$$

These characteristics show the model proposed by Engle (1982) meets the characteristics introduced in Section 3. Then we continue modelling the conditional variance using ARCH and GARCH models in Sections 5.2.1 and 5.2.2 respectively.

5.2.1 Autoregressive conditional heteroscedasticity models (ARCH)

In this section the define the autoregressive conditional heteroscedasticity models, which were introduced in Engle (1982). In these models the volatility, σ_t^2, is explained as a linear regression of the squared white noise, A_t^2.

Then, given a process of white noise, $\{A_t, t \in T\}$ we write an **autoregressive conditional heteroscedasticity model of order r, ARCH(r)** as follow:

$$A_t = \sigma_t \epsilon_t$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 A_{t-1}^2 + ... + \alpha_r A_{t-r}^2 \qquad (5.24)$$

with $\alpha_0 > 0$ and $\alpha_1, ..., \alpha_r$ non negative parameters to ensure that the volatility is positive. Since an $\{A_t, t \in T\}$ ARCH(r) process is stationary, $E(A_t^2) = E(A_{t-k}^2) = \sigma_A^2 \; \forall k$, using Equation (5.22) and $Var(A_t|\mathcal{F}_{t-1}) = E(A_t^2|\mathcal{F}_{t-1}) - E(A_t|\mathcal{F}_{t-1})^2 = E(A_t^2|\mathcal{F}_{t-1}) = \sigma_t^2$ from Equation (5.20); then the marginal variance of the process is given by:

$$\sigma_A^2 = Var(A_t) = E(A_t^2) = E(E(A_t^2|\mathcal{F}_{t-1})) = \alpha_0 + \sum_{i=1}^{r} \alpha_i E(A_{t-i}^2) = \alpha_0 + \sum_{i=1}^{r} \alpha_i \sigma_A^2$$

$$\Rightarrow \sigma_A^2 = \frac{\alpha_0}{1 - \alpha_1 - ... - \alpha_r}$$

having $\sum_{i=1}^{r} \alpha_i < 1$ for being consistent.

5.2.2 Generalised autoregressive conditional heteroscedasticity models (GARCH)

Bollerslev (1986) extended ARCH models into generalised autoregressive conditional heteroscedasticity models. In this case the volatility, σ_t^2 is given as a linear regression over the previous volatilities and squared white noise, A_t^2. This extension is equivalent of the ARMA models on AR models described in Section 5.1.

Considering $\{A_t, t \in T\}$ be a white noise process, then we write a **generalised autoregressive conditional heteroscedasticity model of orders r and s, GARCH(r,s)**:

$$A_t = \sigma_t \epsilon_t$$
$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{r} \alpha_i A_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2 \qquad (5.25)$$

with $\alpha_0 > 0$, $\alpha_i \geq 0$ for $1 \geq i \geq r$, and $\beta_j \geq 0$ for $1 \geq j \geq s$. Tsay (2010) showed that the marginal variance of a GARCH(r,s) process is:

$$Var(A_t^2) = \sigma_A^2 = \frac{\alpha_0}{1 - \sum_{i=1}^{max\{r,s\}}(\alpha_i + \beta_i)}$$

18

being $\sum_{i=1}^{max\{r,s\}}(\alpha_i + \beta_i) < 1$.

Another important characteristic of GARCH processes is that can be rewritten as a ARMA process in terms of A_t^2, given a GARCH(r,s) process, then we rewrite:

$$A_t^2 = \alpha_0 + \sum_{i=1}^{max\{r,s\}}(\alpha_i + \beta_i)A_{t-i}^2 + v_t - \sum_{j=1}^{s}\beta_j v_{t-j}$$

for $v_t = A_t^2 - \sigma_t^2$ white noise.

This characteristic is important for the proper model and orders identification. Bollerslev (1986) showed that by using the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the squared white noises, A_t^2, the correct order is selected using the same method that for ARMA models developed by Box and Jenkins (1976) and described in Section 5.1. Then, we select an ARCH(r) model when we observe infinity non zero spikes in the ACF of A_t^2, but only the first r spikes of the PACF are different from zero. In case of the GARCH(r,s) models both ACF and PACF show infinity non null spikes, being recommended to start with small orders and increase until obtain a residual white noise.

Equivalent to ARIMA models, Bollerslev (1986) also considered the option of having a unitary root on the GARCH model. This leads to **integrated generalised autoregressive conditional heterocedasticity model or IGARCH**. The basic model is given by IGARCH(1,1), which is characterised by $\alpha_1 + \beta_1 = 1$. Then the IGARCH(1,1) is defined as follow

$$A_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + (1 - \beta_1)A_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{5.26}$$

IGARCH models are the non stationary version of GARCH models, but are also prone to become stationary.

5.2.3 Estimation and forecasting

As we have shown GARCH models can be rewritten as ARMA process, which is useful for parameter estimation. Then, considering the observed squared residuals of the ARMA process $\hat{a}_t^2 = (r_t - \hat{\mu}_t)^2$ the parameters are estimated with the maximum likelihood method, using the ARMA estimator $\hat{\mu}_t$ given in Equation (5.17).

For forecasting, we also obtain equivalent results to ARMA models. Using the estimated parameters, $\hat{\alpha}_0, ..., \hat{\alpha}_r, \hat{\beta}_1, ..., \hat{\beta}_s$, the one step ahead volatility, $\hat{\sigma}_{T+1}^2$ of a GARCH(r,s) process is obtained

$$\hat{\sigma}_{T+1}^2 = \hat{\alpha}_0 + \sum_{i=1}^{r}\hat{\alpha}_i \hat{a}_{T+1-i}^2 + \sum_{j=1}^{s}\hat{\beta}_j \hat{\sigma}_{T+1-j}^2$$

5.3 Diagnostic checking

In this section we use different tools for checking the chosen model and estimated parameters. For this we analyse the standardised residuals, ϵ_t. As we called in Section 5.2 we have $\hat{a}_t = r_t - \hat{\mu}_t$, where $\hat{\mu}_t$ represents the forecast ARMA(p,q) model at time t given in Equation (5.17); join with Equation (5.18) we obtain the estimated standardised residuals

$$\widehat{\epsilon}_t = \frac{\widehat{a}_t}{\widehat{\sigma}_t} = \frac{r_t - \widehat{\mu}_t}{\widehat{\sigma}_t} \tag{5.27}$$

To have a correct model assessment the standardised residuals should be white noise, with null autocorrelations, and fitting the used distribution density. Also, the autocorrelations effects on the squared of the standardised residuals ϵ_t^2 should disappear.

To check the distribution choice Q-Q plots are an useful tool. A Q-Q plot is a scatter plot of the empirical sample quantiles versus the theoretical ones. A good fit would be obtained when the points obtained follow the 45 degree line $x = y$. We continue checking that the standardised residuals are white noise; null mean and unitary variance is easily assessed plotting the obtain time series $\{\widehat{\epsilon}_t, t \in T\}$, which should fluctuate around zero in a band with width $[-3, 3]$.

Finally, autocorrelations of $\widehat{\epsilon}_t$ and $\widehat{\epsilon}_t^2$ can be assessed using the sample autocorrelation function (ACF) from Definition 5.4 or the Ljung-Box test.

In case of using the sample ACF is expected to be no significantly different from zero, the values are obtained

$$\widehat{\rho}_k = \frac{\widehat{\gamma}_k}{\widehat{\gamma}_0} = \frac{\sum_{i=1}^{T-k}(\widehat{\epsilon}_i - \bar{\epsilon})(\widehat{\epsilon}_{i+k} - \bar{\epsilon})}{\sum_{i=1}^{T}(\widehat{\epsilon}_i - \bar{\epsilon})^2}$$

where $\bar{\epsilon} = \frac{1}{T}\sum_{i=1}^{T}\widehat{\epsilon}_i$. Bartlett (1946) proved that the 95% confidence interval for the autocorrelation of a Gaussian white noise is:

$$CI_{0.05}(\rho_k) = \left(-\frac{1.96}{\sqrt{N}}, \frac{1.96}{\sqrt{N}}\right) \quad \forall k$$

being N the sample size.

To check for zero autocorrelations the Ljung-Box test is a useful tool. The test compares under the null hypothesis that the first K autocorrelations are zero, $H_0 : \rho_1 = ... = \rho_K = 0$ versus $H_1 : not H_0$. The $Ljung - Box - Pierce\, Q - statistic$ is given by

$$Q(K) = T(T+2)\sum_{k=1}^{K} \frac{\widehat{\rho}_k}{T-k}$$

We have under H_0 that $Q(K) \sim \chi_K^2$. Thus, the null hypothesis, H_0, is rejected when $Q(K)$ exceed the $(1 - \lambda) - quantile$ of the χ_K^2.

In this section we have modelled and forecast stationary univariate time series using ARMA-GARCH models. In the next section we introduce vine copulas models to measure and model multivariate dependence between univariate time series, very useful to forecast the log returns of a portfolio composed of several financial time series.

6 Vine copula

In this section we introduce vine copulas. These models are used to model the dependence and tail asymmetries of multivariate variables. The main idea is to construct multivariate copulas using bivariate copulas. For this section we mainly follow Czado (2019). In Section 6.1 we define the bivariate copula with their characteristics. We continue generalising to higher dimensions with R-, D- and C-vine copulas in Section 6.2. Finally, in Sections 6.3 and 6.4 we introduce the model selection, parameter estimation and simulation for vine copula models. We begin defining a copula.

Definition 6.1. *A **d-dimensional copula, C**, is a multivariate distribution function on the d-dimensional hypercube $[0,1]^d$ with uniformly distributed marginals. Also, the corresponding **copula density, c**, for an absolutely continuous copula can be obtained by partial differentiation, i.e., $c(u_1, ...u_d) := \frac{\partial^d}{\partial u_1 ... \partial u_d} C(u_1, ..., u_d)$ for all $\boldsymbol{u}$ in $[0,1]^d$.*

The basic idea of copulas is to transform the univariate variables to the uniform scale using their corresponding marginal distribution function for characterising the dependence. This allows to model multivariate variables with different marginal distributions. This idea was introduced in Sklar (1959) with the following theorem.

Theorem 6.1. *Let $\mathbf{X}$ be a d-dimensional random vector with joint distribution function F and marginal distribution functions $F_i, i = 1, ..., d$ then the joint distribution function can be expressed as*

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d)). \tag{6.28}$$

with associated density or probability mass function

$$f(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d))f_1(x_1)...f_d(x_d). \tag{6.29}$$

for some d-dimensional copula C with copula density c. For absolutely continuous distributions, the copula C is unique.

The inverse also holds, the copula corresponding to a multivariate distribution function F with marginal distribution functions $F_i, i = 1, ...d$ can be expressed as

$$C(u_1, ..., u_d) = F(F_1^{-1}(u_1), ..., F_d^{-1}(u_d)). \tag{6.30}$$

and its copula density or probability mass function is determined by

$$c(u_1, ..., u_d) = \frac{f(F_1^{-1}(u_1), ..., F_d^{-1}(u_d))}{f_1(F_1^{-1}(u_1))...f_d(F_d^{-1}(u_d))}. \tag{6.31}$$

Proof. Proofs can be found in Nelsen (2006). □

6.1 Bivariate copula

The use of bivariate copulas is important in the later extension in vine copulas for higher dimensional variables, since the vine copula construction is based in a copula construction. Thus, in this section we introduce some definitions, characteristics and the most used bivariate copula families. Considering a bivariate variable $\mathbf{X} = (X_1, X_2)$ and using Theorem 6.1 we obtain the **conditional density** and **distribution functions** as follow

$$f_{1|2}(x_1|x_2) = c_{12}(F_1(x_1), F_2(x_2))f_2(x_2).$$

$$F_{1|2}(x_1|x_2) = \frac{\partial}{\partial u_2}C_{12}(F_1(x_1), u_2)|_{u_2=F_2(x_2)} = \frac{\partial}{\partial F_2(x_2)}C_{12}(F_1(x_1), F_2(x_2)).$$

We can also obtain the conditional distribution function of a bivariate copula distribution. These are called **h-functions** and are given by

$$h_{1|2}(u_1|u_2) := \frac{\partial}{\partial u_2}C_{12}(u_1, u_2). \tag{6.32}$$

$$h_{2|1}(u_2|u_1) := \frac{\partial}{\partial u_1}C_{12}(u_1, u_2). \tag{6.33}$$

Moreover, we introduce the **tail dependence coefficients** which measure the extreme dependence values. For a given a bivariate variable with copula C we define the **upper tail dependence coefficient**

$$\lambda^{upper} = \lim_{t \to 1^-} \mathbb{P}(X_2 > F_2^{-1}(t)|X_1 > F_1^{-1}(t)) = \lim_{t \to 1^-} \frac{1 - 2t + C(t,t)}{1 - t}.$$

Similarly, we define the **lower tail dependence coefficient**

$$\lambda^{lower} = \lim_{t \to 0^+} \mathbb{P}(X_2 \leq F_2^{-1}(t)|X_1 \leq F_1^{-1}(t)) = \lim_{t \to 0^+} \frac{C(t,t)}{t}.$$

We continue introducing the main bivariate copula families. Thus, the simplest bivariate copula is the **independence copula**, which has no parameters and is given by

$$C(u_1, u_2) = u_1 u_2.$$

From elliptical distributions we define elliptical copulas using the Sklar's theorem. The main representatives are the Gaussian and Student's copulas. The **bivariate Gaussian copula** using a bivariate normal distribution with zero mean vector, unit variances, Φ, and correlation ρ is

$$C(u_1, u_2; \rho) = \Phi_2(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \rho).$$

where Φ_2 is the bivariate normal distribution. In case of the **bivariate Student's copula** we have

$$C(u_1, u_2; R, \nu) = T_{R,\nu}(T_\nu^{-1}(u_1), T_\nu^{-1}(u_2)).$$

where $T_{R,\nu}$ is the bivariate standard Student's t distribution with scale parameter matrix $R \in [-1, 1]^{2x2}$ and $\nu > 0$ degree of freedom. T_ν^{-1} denotes the inverse of the univariate standard Student's t distribution with ν degree of freedom.

The set of **bivariate Archimedian copulas** defined in Czado (2019) represent a family of parametric copulas which use the same structure

$$C(u_1, u_2) = \varphi^{[-1]}(\varphi(u_1) + \varphi(u_2)).$$

with φ a continuous, strictly monotone decreasing and convex function with $\varphi(1) = 0$, and

$$\varphi^{[-1]}(t) := \begin{cases} \varphi^{-1}(t) & , 0 \leq t \leq \varphi(0) \\ 0 & , \varphi(0) \leq t \leq \infty \end{cases}$$

The most important bivariate Archimedian copulas with a single parameter are resume in Table 6.1.

Family	Copula	Parameter	Tail dependence
Clayton	$(u_1^{-\delta} + u_2^{-\delta} - 1)^{-\frac{1}{\delta}}$	$0 < \delta < \infty$	Lower
Gumbel	$exp[-\{(-ln\ u_1)^\delta + (-ln\ u_2)^\delta\}^{\frac{1}{\delta}}]$	$\delta \geq 1$	Upper
Frank	$-\frac{1}{\delta}ln\left(\frac{1}{1-e^{-\delta}}[(1-e^{-\delta}) - (1-e^{-\delta u_1})(1-e^{-\delta u_2})]\right)$	$\delta \in [-\infty, \infty]\backslash\{0\}$	None
Joe	$1 - \left((1-u_1)^\delta + (1-u_2)^\delta - (1-u_1)^\delta(1-u_2)^\delta\right)^{\frac{1}{\delta}}$	$\delta \geq 1$	Upper

Table 6.1: Bivariate Archimedian copulas with single parameter

Also, some rotations of the copula are usually considered to accommodate larger ranges of dependence

- 90º: $c_{90}(u_1, u_2) := c(1 - u_2, u_1)$

- 180º: $c_{180}(u_1, u_2) := c(1 - u_1, 1 - u_2)$

- 270º: $c_{270}(u_1, u_2) := c(u_2, 1 - u_1)$

6.2 R-, D- and C-vine copula

In this section we want to generalise the bivariate copula theory to higher dimensions. To obtain the joint distribution a pair copula decomposition is used. In case of the three dimensions, Czado (2019) proved using Theorem 6.1 that the density function is

$$
\begin{aligned}
f(x_1, x_2, x_3) =& c_{13;2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2); x_2) \\
& c_{23}(F_2(x_2), F_3(x_3))\ c_{12}(F_1(x_1), F_2(x_2)) \\
& f_3(x_3)f_2(x_2)f_1(x_1)
\end{aligned}
$$

where $c_{13;2}(\cdot, \cdot; x_2)$ is the bivariate copula associated to (X_1, X_3) given $X_2 = x_2$. The corresponding three dimensional copula is given by

$$
\begin{aligned}
c(u_1, u_2, u_3) =& c_{13;2}(C_{1|2}(u_1|u_2), C_{3|2}(u_3|u_2); x_2) \\
& c_{23}(u_2, u_3)c_{12}(u_1, u_2)
\end{aligned}
$$

Generalising to d random variables $(X_1, ..., X_d)$. Being $D = \{1, ..., d\}$ excluding i and j, the copula associated with the bivariate conditional distribution (X_i, X_j) is denoted $C_{ij;D}(\cdot, \cdot; \boldsymbol{x}_D)$ for $\boldsymbol{X}_D = \boldsymbol{x}_D$. Also, we introduce the abbreviation $c_{i,j;D} := c_{i,j;D}(F_{i|D}(x_i|\boldsymbol{x}_D), F_{j|D}(x_j|\boldsymbol{x}_D); \boldsymbol{x}_D)$ for $i < j$.

We continue introducing two common structures used in vine copulas. We define a **drawable vine (D-vine) density** if the joint density $f_{1,...,d}$ can be constructed as

$$
f_{1,...,d}(x_1, ..., x_d) = \left[\prod_{j=1}^{d-1}\prod_{i=1}^{d-j} c_{i,(i+j);(i+1),...,(i+j-1)}\right]\left[\prod_{k=1}^{d} f_k(x_k)\right]
$$

Also, we define a **canonical vine (C-vine) density** when the joint density can be decomposed

23

$$f_{1,...,d}(x_1,...,x_d) = \left[\prod_{j=1}^{d-1}\prod_{i=1}^{d-j} c_{j,(i+j);1,...,(j-1)}\right]\left[\prod_{k=1}^{d} f_k(x_k)\right]$$

In general a vine density can be specified by a set of linked graphs. We define a graph as a pair $G = (N, E)$, such that $E \subseteq \{\{x, y\} : x, y \in N\}$ are called edges of the graphs, and N are the nodes. The number of neighbours of a node $v \in N$ is the *degree of* v, denoted by $d(v)$. Given a graph is called **path** if we have $P = \{N, E\}$ with nodes $N = \{v_0, v_1, ..., v_k\}$ and edges $E = \{\{v_0, v_1\}, \{v_1, v_2\}, ..., \{v_{k-1}, v_k\}\}$. A **cycle** is a path with $v_0 = v_k$. Then, we say that $T = (N, E)$ is a tree if any two nodes of T are connected by a unique path in T, which means that every node has at least one connection to some node and there is not any cycle.

From this concepts we define a **regular vine (R-vine) tree sequence** as a set of trees $\mathcal{V} = (T_1, ..., T_{d-1})$ on d elements if:

- Each tree $T_j = (N_j, E_j)$ is connected, i.e. for all nodes $a, b \in T_j, j = 1, ..., d-1$, there exist a path $n_1, ..., n_k \subset N_j$ with $a = n_1, b = n_k$.

- T_1 is a tree with node set $N_1 = \{1, ..., d\}$ and edge set E_j.

- For $j \geq 2$, T_j is a tree with node set $N_j = E_{j-1}$ and edge set E_j.

- Proximity condition. For $j = 2, ..., d-1$ and $\{a, b\} \in E_j$ it must hold that $|a \cap b| = 1$.
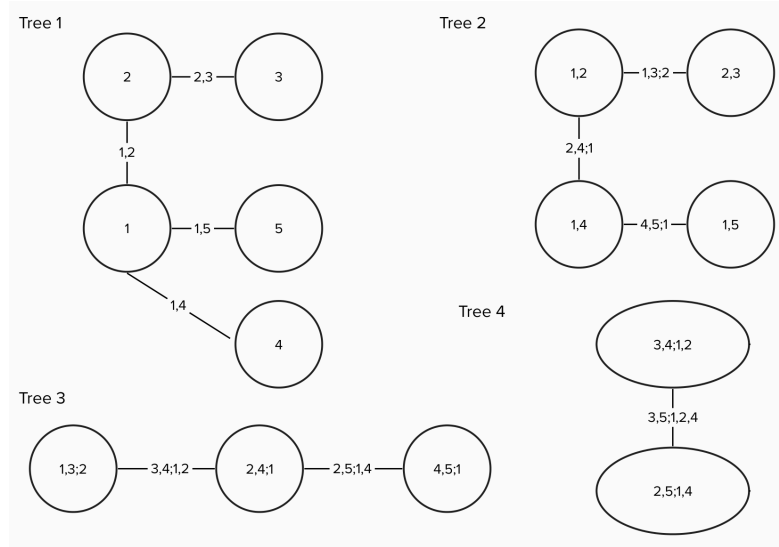
Figure 6.2: Example of R-vine tree sequence with five elements.

We introduce some notation related to the vine tree structure. For any edge $e \in E_i$ we define the **complete union A_e** as the set

$$A_e := \{j \in N_1 | \exists e_1 \in E_1, ..., e_{i-1} \in E_{i-1} \text{ such that } j \in e_1 \in ... \in e_{i-1} \in e\}$$

The **conditioning set D_e** of an edge $e = \{a, b\}$ is $D_e := A_a \bigcap A_b$. Then the edge is abbreviated to $e = (e_a, e_b; D_e)$, with $e_a := A_a \backslash D_e$ and $e_b := A_b \backslash D_e$.

24

Also, D-vine and C-vine can be characterised in terms of their tree sequence. Thus, given a tree sequence $\mathcal{V} = (T_1, ..., T_{d-1})$ is called

- **D-vine tree sequence** if for each node $n \in N_i$ we have $|\{e \in E_i | n \in e\}| \leq 2$.

- **C-vine tree sequence** if for each tree T_i there is one node $n \in N_i$ such that $|\{e \in E_i | n \in e\}| = d - i$. Such a node is called the root node of tree T_i.
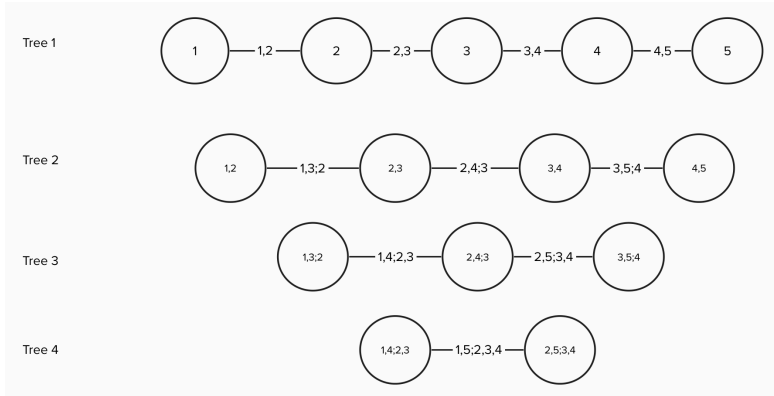
Figure 6.3: Example of D-vine tree sequence with five elements.
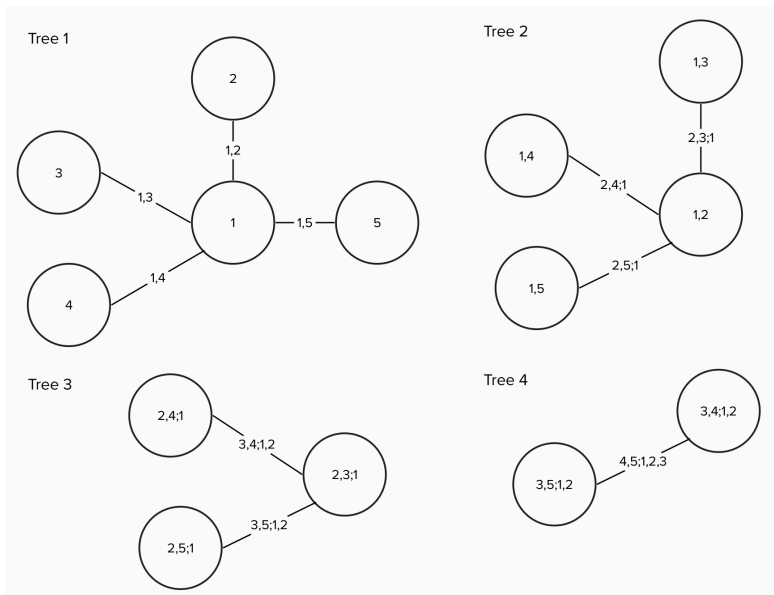
Figure 6.4: Example of C-vine tree sequence with five elements.

In general, considering F the joint distribution for the d dimensional random vector $\boldsymbol{X} = (X_1, ..., X_d)$ has a **regular vine distribution**, if we can specify a triplet $(\mathcal{F}, \mathcal{V}, \mathcal{B})$ such that:

1. Marginal distributions: $\mathcal{F} = (F_1, ..., F_d)$ is a vector of continuous invertible marginal distribution functions.

2. Regular vine tree sequence: $\mathcal{V}$ is an R-vine tree sequence on d elements.

3. Bivariate copulas: The set $\mathcal{B} = \{C_e | e \in E_i; i = 1, .., d-1\}$, where C_e is a symmetric bivariate copula with density.

After defining the different classes of vine copulas and how are they constructed we continue with the vine copula tree selection and parameter estimation from data.

6.3   Selection and parameter estimation of vine copula

In this section we introduce a process to select the best vine model. We focus in the general case of a regular vine. The selection is done using the Dißmann algorithm based on the top-down selection, introduced in Dißmann et al. (2013).

The method is based in maximising the pair weights for each tree, starting with T_1. Then for each tree the algorithm is based in different steps.

1. First we calculate the weight $W_{i,j}$ for all possible index pairs $\{i,j\}, 1 \leq i < j \leq n$. The first tree is arbitrary but the following trees should also fulfil the proximity condition.

2. We select the maximum spanning tree and estimate the corresponding parameters for the selected copula, C_{a_e,b_e}.

3. Generate pseudo observations $C_{a_e|b_e}(u_{k,b_e}, \widehat{\boldsymbol{\theta}}_{a_e,b_e})$ and $C_{b_e|a_e}(u_{k,a_e}, \widehat{\boldsymbol{\theta}}_{a_e,b_e})$ for $k = 1, ...n$.

4. Finally, we use the generated pseudo observations for determining the following tree edges weight $w_{a_e,b_e;D_e}$.

To estimate the parameters $\boldsymbol{\theta}_e$ for edge $e = (a_e, b_e; D_e)$ in tree T_i is based on the pseudo-observations. In particular, $\boldsymbol{\theta}_e$ is estimated by maximising

$$\prod_{k=1}^{n} c_{a_e,b_e;D_e}\big(u_{k,a_e|D_e,\widehat{\boldsymbol{\theta}}(T_{1,...,i-1})}; \theta_e\big) \tag{6.34}$$

To compare the possible models information criteria are used. Akaike (1973) and Schwarz (1978) defined the Akaike (AIC) and Bayesian (BIC) information criteria, respectively. This well known model information criteria use the log-likelihood, penalising the number of parameter used. Then, considering a R-vine $(\mathcal{V}, \mathcal{B})$ with parameter vector $\boldsymbol{\theta}$ and sample $\boldsymbol{u}$ of size n the information criteria are given by

$$AIC_{RV} = -2 \sum_{k=1}^{n} ln(l_k(\widehat{\boldsymbol{\theta}}; \boldsymbol{u}_k)) + 2K$$

$$BIC_{RV} = -2 \sum_{k=1}^{n} ln(l_k(\widehat{\boldsymbol{\theta}}; \boldsymbol{u}_k)) + ln(n)K$$

where K is the number of model parameters and $l_k(\widehat{\boldsymbol{\theta}}; \boldsymbol{u}_k)$ is the corresponding likelihood of the R-vine

$$l_k(\boldsymbol{\theta}; \boldsymbol{u}_k) := \prod_{i=1}^{d-1} \prod_{e \in E_i} c_{a_e,b_e;D_e}(C_{a_e|D_e}(u_{k,a_e}|\boldsymbol{u}_{k,D_e}), C_{b_e|D_e}(u_{k,b_e}|\boldsymbol{u}_{k,D_e}))$$

Apart from the above information criteria, Vuong (1989) proposed the Vuong test for comparing different models. These tests compares the likelihood of the two competing statistical models. Then, the null hypothesis is $H_0 : E_0[ln f_1(\boldsymbol{X}|\boldsymbol{\theta}_1)] = E_0[ln f_2(\boldsymbol{X}|\boldsymbol{\theta}_2)]$ versus $H_1 : not\ H_0$.

Given i.i.d. observations $\boldsymbol{x}_k$, $k = 1, ..., n$ from the true density h_0 with $\widehat{\boldsymbol{\theta}}_j$, $j = 1, 2$ parameter estimates resulting from fitting the two models. Then we define the log-likelihood ratio contribution

$$m_k(\boldsymbol{x}_k) := ln \left[\frac{f_1(\boldsymbol{x_k}|\widehat{\boldsymbol{\theta}}_1)}{f_2(\boldsymbol{x_k}|\widehat{\boldsymbol{\theta}}_2)} \right] \quad k = 1, ..., n$$

and the likelihood ratio statistic is

$$LR_n(\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2)(\boldsymbol{x}_1, ..., \boldsymbol{x}_n) := \sum_{k=1}^{n} m_k(\boldsymbol{x}_k)$$

Then under the null hypothesis it holds that

$$v(\boldsymbol{X}_1, ..., \boldsymbol{X}_n) := \frac{LR_n(\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2)(\boldsymbol{X}_1, ..., \boldsymbol{X}_n)}{\sqrt{n\widehat{\omega}(\boldsymbol{X}_1, ..., \boldsymbol{X}_n)}} \rightarrow N(0, 1)$$

being $\widehat{\omega}(\boldsymbol{X}_1, ..., \boldsymbol{X}_n)$ the empirical variance of the likelihood ratio statistics $LR_n(\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2)(\boldsymbol{X}_1, ..., \boldsymbol{X}_n)$.

We reject H_0 if and only if $|v(\boldsymbol{x}_1, ..., \boldsymbol{x}_n)| > \Phi^{-1}(1 - \frac{\alpha}{2})$. Then, we select model 1 if $v(\boldsymbol{x}_1, ..., \boldsymbol{x}_n) > \Phi^{-1}(1 - \frac{\alpha}{2})$ and select model 2 if $v(\boldsymbol{x}_1, ..., \boldsymbol{x}_n) < -\Phi^{-1}(1 - \frac{\alpha}{2})$.

6.4 Simulating from vine copula

Here we give the general procedure to simulate from a general R-vine copula. In case of being a C- or D-vine the approach is similar, but can be found in Czado (2019). The idea is to recursively simulate using the univariate conditional distributions derived from the joint R-vine copula. Then, a sample $u_1, ..., u_d$ is obtained as follow

$$
\begin{aligned}
&\textbf{First} : \text{Sample i.i.d. } w_j \sim U[0, 1], \ j = 1, ..., d \\
&\textbf{Then} : u_1 := w_1 \\
&\qquad u_2 := C_{2|1}^{-1}(w_2|u_1) \\
&\qquad \vdots \\
&\qquad u_d := C_{d|d-1,...,1}^{-1}(w_d|u_{d-1}, ..., u_1)
\end{aligned}
\tag{6.35}
$$

The univariate conditional copula distribution are generalised from Equations (6.32) and (6.33) to higher orders

$$C_{e_a|e_b;D_e}(w_1|w_2; \theta_{e_a,e_b;D_e}) := \frac{\partial}{\partial w_2} C_{e_a,e_b;D_e}(w_1, w_2; \theta_{e_a,e_b;D_e}) \tag{6.36}$$

$$C_{e_b|e_a;D_e}(w_2|w_1; \theta_{e_a,e_b;D_e}) := \frac{\partial}{\partial w_1} C_{e_a,e_b;D_e}(w_1, w_2; \theta_{e_a,e_b;D_e}) \tag{6.37}$$

depending also on the estimated pair copula parameters.

7 Stationary vine copula

In this section we introduce the class of stationary vine copula models, which model the cross-sectional and temporal dependence together using vine copula based approach. Most of this section follows Nagler et al. (2022). We begin in Section 7.1 with the translation invariance, needed for S-vines. We continue defining the stationary vine copulas in Section 7.2. In Section 7.3 is given the Markov property and Section 7.4 is the parameter estimation and simulation for stationary vine models.

In contrast to the previous ARMA-GARCH and vine copula models the stationary vine copula includes the time dependence in vine copula. Since copulas work on the u-scale we first need to transform the original time series process, $\{R_{t,i}, t \in T\}$ $\forall i \in \{1, ..., d\}$. To transform the series we assume strict stationarity as defined in Definition 5.1. Thus, given the marginal distributions $F_1, ..., F_d$ we obtain the u-scale time series process $\{U_{t,i}, t \in T\}$ $\forall i \in \{1, ...d\}$ such that $U_{t,i} = F_i(R_{t,i})$. In this case we have two sub-indices in each random variable $U_{t,i}$, the first refers to the time index and i to the variable. From now, in this section we use the notation (t, i) referred to the random variable $U_{t,i}$. Therefore we consider a vine structure on $T \times d$ dimensions.

The idea to model the multivariate time series with stationary vine copulas is to use a vine for capturing the cross-sectional dependence of $\boldsymbol{U}_t \in \mathbb{R}^d$ for all time points t. The cross-sectional structure are linked by one edge connecting a vertex from the structure at t with a vertex of the structure at $t + 1$. Since we are assuming stationary time series, it is reasonable to assume that the cross-sectional structure and the linked edge are time invariant. There are three different models depending the cross sectional structures and linked edges:

- D-vine for multivariate stationary vines introduced in Smith (2015). The cross-sectional structure is a D-vine and two cross-sectional structures are connected with border vertices generating a big D-vine structure through cross-sectional and time dependence structures.

- M-vine introduced in Beare and Seo (2015). In this case the cross-sectional structure is also a D-vine. However, they are connected at one vertex that lies on the same border of the D-vine.

- COPAR introduced in Brechmann and Czado (2015). The cross-sectional structure is a C-vine and the first trees of two C-vines at time points t and $t + 1$ are connected at the root node.
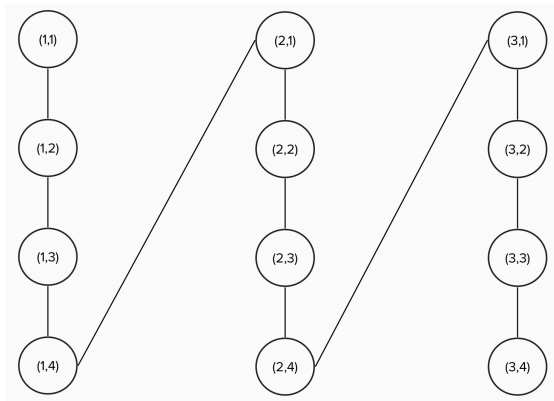
Figure 7.5: Example of the first tree of a four dimensional D-vine for multivariate stationary vines on three time points.
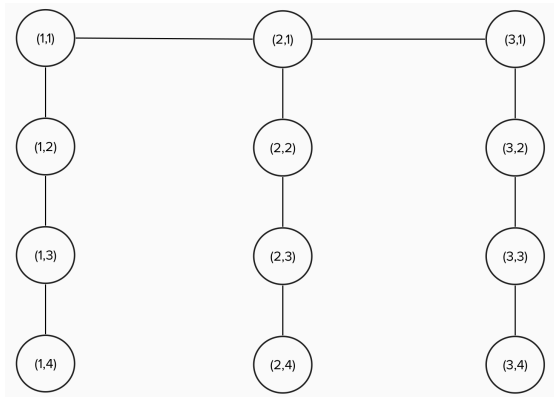
Figure 7.6: Example of the first tree of a four dimensional M-vine on three time points.
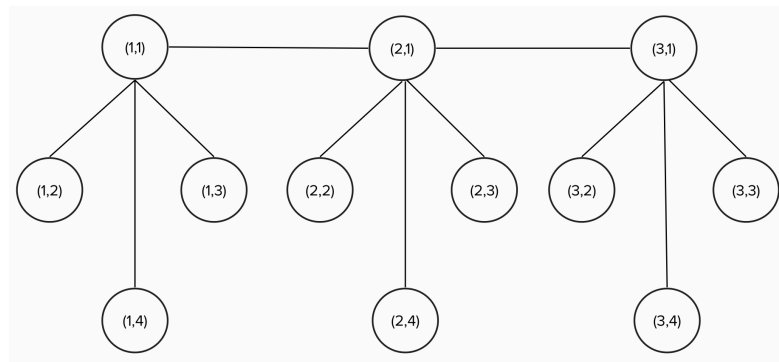
Figure 7.7: Example of the first tree of a four dimensional COPAR on three time points.

7.1 Translation invariance

We have assumed strict stationarity on the original univariate time series $\{R_{t,i}, t \in T\}$ $\forall i \in \{1, ..., d\}$. However, this concept should be extended to the joint distribution. Then, we say that

the time series $\boldsymbol{U}_1, ..., \boldsymbol{U}_T \in \mathbb{R}^d$ is strictly stationary if and only if $\boldsymbol{U}_{t_1}, ..., \boldsymbol{U}_{t_m}$ and $\boldsymbol{U}_{t_{1+\tau}}, ..., \boldsymbol{U}_{t_{m+\tau}}$ have the same joint distribution for all $1 \leq t_1 < t_2 < ... < t_m \leq T$, $1 \leq \tau \leq T - max_{j=1}^m t_j$, and $1 \leq m \leq T$. This condition is included in the vine copula model with the translation invariance concept.

Definition 7.1. *Given a vine copula $(\mathcal{F}, \mathcal{V}, \mathcal{B})$ on the set of nodes $V_1 = \{1, ...T\} \times \{1, ..., d\}$ is* **translation invariant** *if $c_{a_e, b_e; D_e} = c_{a_{e'}, b_{e'}; D_{e'}}$ holds for all edges e, $e' \in \cup_{k=1}^{Td-1} E_k$ for which there exist $\tau \in \mathbb{Z}$ such that*

$$a_e = a_{e'} + (\tau, 0), \quad b_e = b_{e'} + (\tau, 0), \quad D_e = D_{e'} + (\tau, 0) \tag{7.38}$$

with $D_e = \{v + (\tau, 0) : v \in D_{e'}\}$.

By definition translation invariance is a necessary condition for stationarity. However, in general it is not a sufficient condition for stationarity. Hence, the useful vine structures are those that translation invariance is a sufficient condition. Figure 7.8 shows the effect of the translation invariance. In the figure, the labels with same colour in the same tree represent the same pair copula.
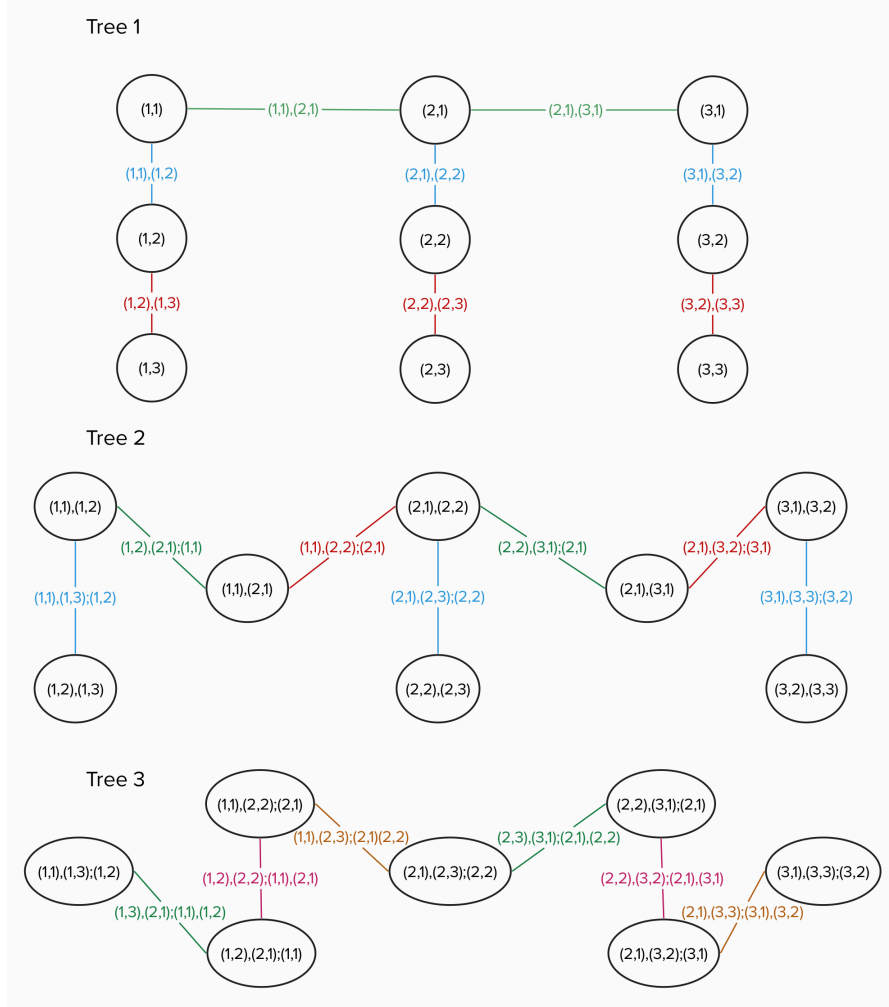
Figure 7.8: Example of the first three trees of a three dimensional M-vine on three time points.

7.2 Stationary vine

Before defining a stationary vine we need to introduce a couple of concepts.

Definition 7.2. *Consider a vine sequence* $\mathcal{V} = (V_k, E_k)_{k=1}^{Td-1}$ *on* $\{1, ...T\} \times \{1, ..., d\}$ *and* $V_1' = \{t, ...t + m\} \times \{1, ..., d\}$ *for some* t, m *with* $1 \leq t \leq T$, $0 \leq m \leq T - t$. *Given* $\binom{V_k'}{2}$ *the set of all nodes of the tree* k. *Thus, we define* $E_k' = E_k \cap \binom{V_k'}{2}$ *and* $V_{k+1}' = E_k'$ *for all* $k \geq 1$. *Then the sequence of graphs* $\mathcal{V}_{t,t+m} = (V_k', E_k')_{k=1}^{(m+1)d-1}$ *is called* ***restriction of*** $\mathcal{V}$ *on the points* $t, ..., t + m$.

A vine restriction refers to delete all edges and vertices where the time indices in the range $[t, t + m]$ does not appear. A restriction of vine, $\mathcal{V}$, is not always a vine. For example consider the vine $(1, i) - (3, i) - (2, i)$; the restriction on $\{1, 2\} \times \{i\}$ is given by the two single nodes $(1, i)$ and $(2, i)$. This is not a vine for being disconnected.
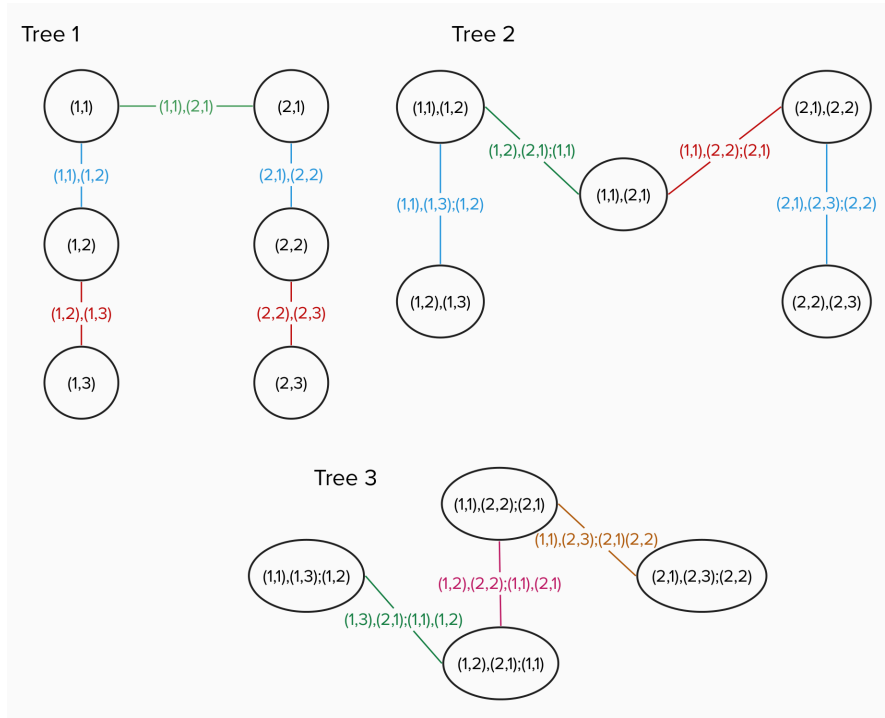
Figure 7.9: First three trees of the M-vine given in Figure 7.8 restricted to $\{1, 2\} \times \{1, 2, 3\}$.

The first tree of the restriction of a vine $\mathcal{V}$ is composed by all vertices in V_1' and the corresponding edges. In the remaining trees, edges and vertices are affected by the deletion in the previous trees. Figure 7.9 shows an example of restriction vine plot in Figure 7.8.

Definition 7.3. *Let $m \geq 0$, and $\mathcal{V} = (V_k, E_k)_{k=1}^{(m+1)d-1}$ be a vine on $\{t, ..., t+m\} \times \{1, ..., d\}$ and $\mathcal{V}' = (V_k', E_k')_{k=1}^{(m+1)d-1}$ be a vine on $\{s, ..., s+m\} \times \{1, ..., d\}$. We say that $\boldsymbol{\mathcal{V}}$ **is a translation of $\boldsymbol{\mathcal{V}'}$** (denoted by $\mathcal{V} \sim \mathcal{V}'$) if for all $k = 1, ..., d-1$ and edges $e \in E_k$, there is an edge $e' \in E_k'$ such that $e = e' + (t-s, 0)$.*

The translation of a vine $\mathcal{V}$ is obtained by shifting all the vertices and edges in time. An example is shown in Figure 7.10, which represents a translation from $\{1, 2\} \times \{1, 2, 3\}$ to $\{5, 6\} \times \{1, 2, 3\}$.

Theorem 7.1. *Let $\mathcal{V}$ be a vine on the set $V_1 = \{1, ..., T\} \times \{1, ..., d\}$. Then the following statements are equivalent:*

1. *The vine copula model $(\mathcal{F}, \mathcal{V}, \mathcal{B})$ is stationary for all translation invariant choices of $\mathcal{B}$.*

2. *There are vines $\mathcal{V}^{(m)}, m = 0, ...T-1$, defined on $\{1, ..., m+1\} \times \{1, ..., d\}$, such that for all $m = 0, ..., T-1$, $1 \leq t \leq T-m$,*

$$\mathcal{V}_{t, t+m} \sim \mathcal{V}^{(m)}$$

Definition 7.4. *A vine $\mathcal{V}$ on the set $V_1 = \{1, ..., T\} \times \{1, ..., d\}$ is **stationary** if it satisfies condition 2 of Theorem 7.1.*
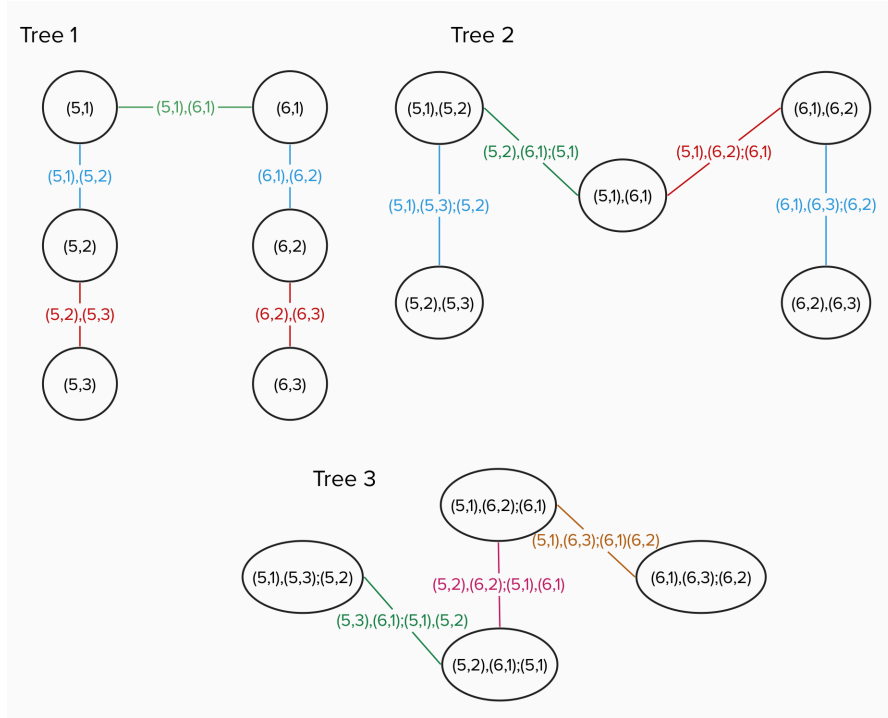
Figure 7.10: Vine translation of M-vine Figure 7.9 from $\{1, 2\} \times \{1, 2, 3\}$ to $\{5, 6\} \times \{1, 2, 3\}$.

The S-vines are characterised by $\mathcal{V}^{(0)}$ which contains the cross-sectional structure of $\mathcal{V}$. $\mathcal{V}^{(1)}$ correspond to two nested $\mathcal{V}^{(0)}$ structures. In general, $\mathcal{V}^{(m)}$ is the structure of $\mathcal{V}^{(m-1)}$ with a new cross-sectional structure and the corresponding edge between them. Nagler et al. (2022) show that M-vines and D-vines for multivariate stationary vine copulas are stationary, but CO-PAR model is not. However, S-vines do not require the cross-sectional structure to be D-vines, moreover we have some degree of freedom in how we connect variables across time points.

7.3 Markovian models

Stationarity is very useful and necessary to estimate the model since it reduces the model complexity. While a general vine copula model on $\boldsymbol{U}_1, ..., \boldsymbol{U}_T \in [0, 1]^d$ requires $Td(Td - 1)/2 = O(T^2d^2)$ pair copulas; the number of pair copulas is reduced to $(T - 1)d^2 + d(d - 1)/2 = O(Td^2)$ when they are translation invariance. Usually time series are also constrained with the Markov property.

Definition 7.5. *A time series $\boldsymbol{U}_1, ..., \boldsymbol{U}_T \in [0, 1]^d$ is called a **Markov process of order p** if, for all $\boldsymbol{u} \in \mathbb{R}^d$, it fulfils*

$$\mathbb{P}(\boldsymbol{U}_t \leq \boldsymbol{u} | \boldsymbol{U}_{t-1}, ..., \boldsymbol{U}_1) = \mathbb{P}(\boldsymbol{U}_t \leq \boldsymbol{u} | \boldsymbol{U}_{t-1}, ..., \boldsymbol{U}_{t-p})$$

A vine copula $(\mathcal{F}, \mathcal{V}, \mathcal{B})$ on a stationary vine $\mathcal{V}$ is Markov of order p if and only if $c_e \equiv 1$ for all $e \notin \mathcal{V}_{t,t+p}$, $t = 1, ..., T - p$.

Then, the number of pair copula needed for a stationary vine copula model with Markov order p is reduced to $pd^2 + d(d - 1)/2 = O(pd^2)$. In the example shown in Figure 7.8 with $T = 3$

and $d = 3$ the general vine model without constrains needs $\frac{3*3(3*3-1)}{2} = 36$ pair copulas. In case of being translation invariant the number of copulas correspond with the sum of number of different colours in each tree, in this case $(3-1)3^2 + \frac{3(3-1)}{2} = 21$ pair copulas in total. Finally, if we consider the vine copula to be stationary with Markov order 1, the number of pair copula is the sum of the different colours of vine represented in Figure 7.9, in total $1 * 3^2 + \frac{3(3-1)}{2} = 12$ pair copulas.

7.4 Parameter estimation and model simulation

The method used for the model selection and parameter estimation is equivalent to the one defined in Section 6.3 having to account for the stationarity of the vine copulas and the Markov order. In this case the parameters $\boldsymbol{\theta}_{[e]}$ are also obtained by maximising the log-likelihood function, which is similar to Equation (6.34); being $[e]$ the class of edges with the same copula. In this case the maximum likelihood depends on the pair copulas and parameters estimated in the previous trees, so their values are obtained recursively.

Model selection is done with an algorithm similar to Dißmann et al. (2013) described in Section 6.3. Here, we first focus on cross-sectional structure selecting the strongest dependencies as the algorithm describes using the copula weights. Then, the following trees related to cross-sectional structure are selected using the proximity condition. Also, M- and D- vines structures can be imposed to simplify the structure selection. Finally, the Markov order is selecting comparing different models with the criteria information selection, AIC and BIC described in Section 6.3, which are computed from the maximal pair copula log-likelihood.

Finally, for simulating from a stationary vine the conditional copula distribution and its inverse are used. The method is equivalent to the one described in Section 6.4. Given a independent identically distributed uniform variables the copula sample is obtained recursively with the univariate conditional distributions from Equations (6.36) and (6.37).

8 Fleishman's transformation

Throughout this section we discuss the current method used by BVK for estimating the portfolio risk measures. The method was introduce in Vale and Maurelli (1983), which samples portfolio values based on the first four moments of each time series as described in Fleishman (1978).

The method described in this section uses the mean, variance-covariance matrix, skewness and kurtosis for generate multivariate non-normal samples. The general idea of the method is to obtain multivariate non-normal samples from multivariate standard normal samples. Non-normal samples with the desired correlation matrix are generated using the Fleishman's transformation using normal variables with a specific R_x correlation matrix.

The univariate Fleishman's transformation generates a univariate non-normal random variable, Y, as a linear combination of the first three powers of a standard normal random variable X. Consider Y to be standardised non-normal variable, i.e. $E(Y) = 0$ and $Var(Y) = 1$. The transformation from a standard normal variable X is given by:

$$Y = a + bX + cX^2 + dX^3 \tag{8.39}$$

where a, b, c and d are chosen to provide Y with $E(Y) = 0$, $Var(Y) = 1$, $Skew(Y) = \gamma_1$ and $Kurt(Y) = \gamma_2$. To determine the constants, Fleishman expanded Equation (8.39) to express the first four moments of the non-normal variable Y in terms of the first fourteen moments of X. Since X is standard normally distributed, the first fourteen moments are known constants. Fleishman was able to represent the solution to the constants of Equation (8.39) as a system of nonlinear equations. For a standardised distribution for Y, the constants a, b, c and d are found by solving:

$$b^2 + 6bd + 2c^2 + 15d^2 - 1 = 0 \tag{8.40}$$

$$2c(b^2 + 24bd + 105d^2 + 2) - \gamma_1 = 0 \tag{8.41}$$

$$24[bd + c^2(1 + b^2 + 28bd) + d^2(12 + 48bd + 141c^2 + 225d^2)] - \gamma_2 = 0 \tag{8.42}$$

$$a = -c \tag{8.43}$$

with γ_1 and γ_2 being the desired skew and kurtosis respectively of Y.

For the generalisation for the multivariate case, we first calculate the univariate coefficients a_i, b_i, c_i and d_i, $\forall i \in \{1, \ldots, d\}$ where d is the number of variables. We use $X_1, \ldots, X_d$ independent standard normal random variables in the Fleishman univariate method.

The idea is to generate multivariate normal variables which transformed individually with Equation (8.39) the variables Y are obtained with the desired correlations. To obtain the desired final correlations we calculate an intermediate correlation matrix for the standard normal variables X.

The intermediate correlation matrix for the associated normal random variable is calculated pairwise. For each pair of variables, consider Y_1 and Y_2, we have standard normal X_1, X_2 associated such that

$$Y_i = \mathbf{w_i'P_i} \quad i = \{1, 2\} \tag{8.44}$$

where $\mathbf{w_i'} = [a_i, b_i, c_i, d_i]$ and $\mathbf{P_i'} = [1, X_i, X_i^2, X_i^3]$.

Letting r_{Y_1, Y_2} be the desired correlation between two non-normal variables, Y_1 and Y_2 associated with the normal variables X_1 and X_2, and noting that, since the variables are standardised, the correlation between Y_1 and Y_2 is their expected cross product:

$$r_{Y_1, Y_2} = E[Y_1 Y_2] = E[\mathbf{w_1'} \mathbf{P_1} \mathbf{P_2'} \mathbf{w_2}] = \mathbf{w_1'} \mathbf{R}_{1,2} \mathbf{w_2} \tag{8.45}$$

where $\mathbf{R}_{1,2}$ is the expected matrix product of $\mathbf{P_1}$ and $\mathbf{P_2'}$.

$$\mathbf{R}_{1,2} = E(\mathbf{P_1} \mathbf{P_2'}) = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & \rho_{X_1, X_2} & 0 & 3\rho_{X_1, X_2} \\ 1 & 0 & 2\rho_{X_1, X_2}^2 + 1 & 0 \\ 0 & 3\rho_{X_1, X_2} & 0 & 6\rho_{X_1, X_2}^3 + 9\rho_{X_1, X_2} \end{bmatrix} \tag{8.46}$$

Collecting terms, and using the relationship between a and c given in Equation (8.43), a third-degree polynomial in ρ_{X_1, X_2}, the correlation between the normal variables X_1 and X_2, results:

$$r_{Y_1, Y_2} = \mathbf{w_1'} \mathbf{R}_{1,2} \mathbf{w_2} = \rho_{X_1, X_2}(b_1 b_2 + 3 b_1 d_2 + 9 d_1 d_2) + \rho_{X_1, X_2}^2 (2 c_1 c_2) + \rho_{X_1, X_2}^3 (6 d_1 d_2) \tag{8.47}$$

Solving this polynomial for ρ_{X_1, X_2} provides the pairwise correlation between the two variables X_1 and X_2 required to provide the desired post-transformation correlation r_{Y_1, Y_2}. These correlations can then be assembled into a intermediate correlation matrix. Then, for the desired correlation matrix R_Y of $\mathbf{Y} = (Y_1, ..., Y_d)^t$ with elements r_{Y_1, Y_2} determine ρ_{X_1, X_2} as in Equation (8.47) and the corresponding $\mathbf{w_i} = (a_i, b_i, c_i, d_i)^t$ and $\mathbf{w_j} = (a_j, b_j, c_j, d_j)$.

In a second step a matrix decomposition procedure is used to generate multivariate normal variables $\mathbf{Z}$ with zero mean vector, unit variances and the desire correlation matrix of dimension d, R, from multivariate uncorrelated population $\mathbf{X}$ with zero mean vector and variances one:

$$\mathbf{Z} = F \mathbf{X} \tag{8.48}$$

being F the principal components of R, given by $F = VD$; V being the matrix formed by the eigenvectors of R associated to the eigenvalues, λ. D is the diagonal matrix with the singular values of R, $\sqrt{\lambda}$, in the diagonal. Then, $\mathbf{Z} \sim N_d(0, R)$.

Finally, the non-normal standardised dependent random vector $\tilde{\mathbf{Y}} = (Y_1, ..., Y_d)^t$ are determined individually using Equation (8.39) with the correlated normal random vector $\tilde{\mathbf{Z}} = (Z_1, ..., Z_d)$ as follow:

$$\hat{Y}_i = a_i + b_i Z_i + c_i Z_i^2 + d_i Z_i^3 \quad i = 1, ..., d \tag{8.49}$$

By construction, R is not guaranteed to be a positive definite matrix, since it is obtained pairwise.

9 Backtesting

To check and compare the estimated risk measures with the different methods, we define in this section some measures and backtests. The measures are estimations of the level used for calculating the VaR and ES. Using them also in a rolling window setup allows to analyse in which circumstances the risk measures are better forecasted for the three approaches. Backtesting the results give us also some feedback on the forecast and models with different hypothesis tests. In Section 9.1 we define the VaR backtests, while Section 9.2 are the ES backtests. The backtests have been implemented in R using the packages **GAS** Catania et al. (2022) and **esback** Bayer and Dimitriadis (2020).

We begin estimating the empirical VaR level, $\widehat{\alpha}$, $VaR_\alpha(R_t)$ over a time window. It can be calculated for the whole forecast period or uses a rolling window setup. For this, we first define a binary random variable L_t at time t as follows

$$L_t^\alpha = \begin{cases} 1 & VaR_\alpha(R_t) < -R_t \\ 0 & VaR_\alpha(R_t) \geq -R_t \end{cases} \tag{9.50}$$

From Definition 4.6 and continuity of $\{R_t, t \in T\}$ we have

$$E(L_t^\alpha) = \mathbb{P}(VaR_\alpha(R_t) < -R_t) = \mathbb{P}(inf\{m \in \mathbb{R} : \mathbb{P}(m < -R_t) \leq \alpha\} < -R_t) = \alpha$$

Then, L_t^α is estimated by

$$l_t^\alpha = \begin{cases} 1 & \widehat{VaR}_\alpha(r_t) < -r_t \\ 0 & \widehat{VaR}_\alpha(r_t) \geq -r_t \end{cases} \tag{9.51}$$

and the estimated alpha level, $\widehat{\alpha}$, is

$$\widehat{\alpha} := \frac{1}{T} \sum_{t=1}^{T} l_t^\alpha. \tag{9.52}$$

By definition L_t is a Bernoulli variable for each t with success probability α. Assuming that they are independent we construct a confidence interval for α at λ-level as follow

$$CI_\lambda(\alpha) = \widehat{\alpha} \pm z_{1-\lambda/2}\sqrt{\frac{\widehat{\alpha}(1-\widehat{\alpha})}{n}} \tag{9.53}$$

where $z_{1-\lambda/2}$ is the standard normal $(1 - \lambda/2) - quantile$.

McNeil and Frey (2000) introduce a similar measure for the expected shortfall. From Definition 4.7 the expected shortfall is the expectation of the returns that exceed the VaR. Then, we define a measure β_α of the expected shortfall

$$\beta_\alpha := E[r_t - \widehat{ES}_\alpha(r_t)|r_t > \widehat{VaR}_\alpha(r_t)] \quad \forall t \in \{1, ..., T\} \tag{9.54}$$

Since Equation (9.54) only consider the returns that exceed the value at risk at level α, the expected shortfall is well estimated when the expected value of $r_t - \widehat{ES}_\alpha(r_t)$ is zero. From Definition 4.7 we have that the expected value measure the expected lost in case of exceeding the VaR.

In this case a rolling window and confidence interval is not calculated because β depends on the VaR, so the length of the rolling window would not be the same through the time period.

Apart of the measures which helps to have an idea of the estimated measures, we describe some backtests for checking the forecasted risk measures. In general, we obtain a well fitted model when the null hypothesis $H_0 :$ "The risk measures estimation procedure is correct" is not rejected with a certain confidence level.

9.1 VaR Backtesting

In this section we follow Christoffersen (1998), who introduce two backtests for the value at risk. These test are the unconditional and conditional likelihood ratio tests. The conditional likelihood ratio tests the VaR including a test for the independence of the risk measures at each time point t. The likelihood ratio tests use Equation (9.51) for comparing with the α-VaR level.

Definition 9.1. *Given the process $\{L_t^\alpha, t \in T\}$ for the log returns R_t and estimated risk measure $\widehat{VaR}_\alpha(R_t)$ being independent. Then the **likelihood ratio test of unconditional coverage, LR_{uc}**, tests the hypothesis:*

$$H_0 : \quad E[L_t^\alpha] = \alpha \tag{9.55}$$

$$H_1 : \quad E[L_t^\alpha] \neq \alpha \tag{9.56}$$

The likelihood obtained under the null hypothesis is given by

$$L_0(\alpha; l_1^\alpha, ..., l_T^\alpha) = (1-\alpha)^{\sum_{t=1}^T (1-l_t^\alpha)} \alpha^{\sum_{t=1}^T l_t^\alpha}$$

and under the alternative

$$L_1(\pi; l_1^\alpha, ..., l_T^\alpha) = (1-\pi)^{\sum_{t=1}^T (1-l_t^\alpha)} \pi^{\sum_{t=1}^T l_t^\alpha}$$

The likelihood ratio test is

$$LR_{uc}^\alpha = -2log \left(\frac{L_0(\alpha; l_1^\alpha, ..., l_T^\alpha)}{L_1(\widehat{\pi}; l_1^\alpha, ..., l_T^\alpha)} \right) \sim \chi^2(1) \quad under \ H_0$$

with the maximum likelihood estimator $\widehat{\pi}$ given in Equation (9.52) being $\frac{1}{T}\sum_{t=1}^T l_t^\alpha$. Then, we reject the null hypothesis at level γ if $LR_{uc}^\alpha > \chi_\gamma^2(1)$.

As we have previously indicated this test assume independence of the variables l_t^α. However, this can be included in the test introducing an independent test. Here the time dependence is modelled by a stationary first order Markov Chain. Christoffersen (1998) defined the likelihood ratio test of independence as follow.

Definition 9.2. *The **likelihood ratio independence test, LR_{ind}**, compares the following hypothesis*

$$H_0^{ind} : \quad \mathbb{P}(L_t^\alpha = 0 | L_{t-1}^\alpha = 1) = \mathbb{P}(L_t^\alpha = 1 | L_{t-1}^\alpha = 1) \quad \forall t$$

$$H_1^{ind} : \quad \mathbb{P}(L_t^\alpha = 0 | L_{t-1}^\alpha = 1) \neq \mathbb{P}(L_t^\alpha = 1 | L_{t-1}^\alpha = 1) \quad \forall t$$

Since L_t^α is binary we obtain a binary first-order Markov chain. Assumed to be stationary with transition probability matrix

$$\Pi_1 = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix}$$

where $\pi_{ij} = \mathbb{P}(l_t^\alpha = j | l_{t-1}^\alpha = i)$. The likelihood function under the null hypothesis is

$$L_1(\Pi_1; l_1^\alpha, ..., l_T^\alpha) = (1 - \pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1 - \pi_{11})^{n_{10}} \pi_{11}^{n_{11}}$$

with n_{ij} being the number of observed transition from state i to state j. Then, the maximum likelihood estimator for Π_1, $\widehat{\Pi}_1$, is

$$\widehat{\Pi}_1 = \begin{bmatrix} \frac{n_{00}}{n_{00}+n_{01}} & \frac{n_{01}}{n_{00}+n_{01}} \\ \frac{n_{10}}{n_{10}+n_{11}} & \frac{n_{11}}{n_{10}+n_{11}} \end{bmatrix}$$

Under the null hypothesis the transition matrix is

$$\Pi_0 = \begin{bmatrix} 1 - \pi_0 & \pi_0 \\ 1 - \pi_0 & \pi_0 \end{bmatrix}$$

with $\pi_0 = \mathbb{P}(L_t^\alpha = 0 | l_{t-1}^\alpha = 1) = \mathbb{P}(L_t^\alpha = 1 | l_{t-1}^\alpha = 1)$. The likelihood under the null hypothesis is

$$L_0(\Pi_0; l_1^\alpha, ..., l_T^\alpha) = (1 - \pi_0)^{(n_{00}+n_{10})} \pi_0^{(n_{01}+n_{11})}$$

and the ML estimate $\widehat{\pi_0} = (n_{01} + n_{11})/(n_{00} + n_{10} + n_{01} + n_{11})$ given $\widehat{\Pi}_0 = \begin{bmatrix} 1 - \widehat{\pi}_0 & \widehat{\pi}_0 \\ 1 - \widehat{\pi}_0 & \widehat{\pi}_0 \end{bmatrix}$. Finally, the statistic for the test is given by

$$LR_{ind}^\alpha = -2log \left(\frac{L_0(\widehat{\Pi}_0; l_1^\alpha, ..., l_T^\alpha)}{L_1(\widehat{\Pi}_1; l_1^\alpha, ..., l_T^\alpha)} \right) \sim \chi^2(1) \quad under \ H_0^{ind}$$

We reject the null hypotheses at level γ if $LR_{ind}^\alpha > \chi_\gamma^2(1)$, being $\chi_\gamma^2(1)$ the γ quantile of $\chi^2(1)$.

Further, both test can be combined to obtain the conditional coverage test. The **likelihood ratio with conditional coverage, LR_{cc}** tests the same hypothesis given in Equations (9.55) and (9.56) join with the independence of l_t^α. The statistic for the test is

$$LR_{cc}^\alpha = LR_{uc}^\alpha + LR_{ind}^\alpha \sim \chi^2(2) \quad \text{under } H_0 \text{ and } H_0^{ind}$$

for the process $\{l_t^\alpha, t \in T\}$ being α the level of the value at risk. We reject the null hypothesis at level γ if $LR_{cc}^\alpha > \chi_\gamma^2(2)$, where $\chi_\gamma^2(2)$ is the γ quantile of $\chi^2(2)$.

9.2   ES Backtesting

In case of the expected shortfall we define two backtests, the exceedance residual test and the conditional calibrated test. From Definition 4.7 we have that expected shortfall depends on value at risk.

The **exceedance residual test, ER**, was developed in McNeil and Frey (2000). They introduce the measure given in Equation (9.54) which is expected to have zero mean. Then they proposed the following hypothesis test

$$H_0 : \quad E[r_t - \widehat{ES}_\alpha(r_t) | r_t > \widehat{VaR}_\alpha(r_t)] = 0$$

$$H_1 : \quad E[r_t - \widehat{ES}_\alpha(r_t) | r_t > \widehat{VaR}_\alpha(r_t)] \neq 0$$

To test the hypothesis a bootstrap test from Efron and Tibshirani (1993) is utilised.

The **conditional calibration test, CC**, was defined in Nolde and Ziegel (2017). This test compares value at risk and expected shortfall at the same time. Considering the vector of risk measures (VaR_α, ES_α) at level α, they consider the identification function

$$\boldsymbol{V}((VaR_\alpha, ES_\alpha), x) = \begin{pmatrix} \alpha - \mathbb{1}\{x > VaR_\alpha\} \\ VaR_\alpha - ES_\alpha - \frac{1}{\alpha}\mathbb{1}\{x > VaR_\alpha\}(VaR_\alpha - x) \end{pmatrix}$$

Then, given a series of risk measures estimation, $(\widehat{VaR}_\alpha(X_t), \widehat{ES}_\alpha(X_t))$ on $\{X_t, t \in T\}$ is said calibrated for $(VaR_\alpha(X_t), ES_\alpha(X_t))$ on average if

$$E[V((\widehat{VaR}_\alpha(X_t), \widehat{ES}_\alpha(X_t)), X_t)] = \boldsymbol{0} \quad \forall t \in T$$

Also, it is said to be conditional calibrated if

$$E[V((\widehat{VaR}_\alpha(X_t), \widehat{ES}_\alpha(X_t)), X_t)|\mathcal{F}_{t-1}] = \boldsymbol{0}$$

being $\mathcal{F}_{t-1}$ the σ-algebra containing all the information until time $t-1$.

Since conditional calibration is stronger notion than average calibration, a traditional backtest is developed with null hypothesis

$$H_0: \quad \text{The sequence of predictions } (\widehat{VaR}_\alpha(X_t), \widehat{ES}_\alpha(X_t)) \text{ is}$$
$$\text{conditional calibrated for } (VaR_\alpha(X_t), ES_\alpha(X_t))$$

Nolde and Ziegel (2017) showed that $E[\boldsymbol{V}((\widehat{VaR}_\alpha(X_t), \widehat{ES}_\alpha(X_t)), X_t)|\mathcal{F}_{t-1}] = \boldsymbol{0}$ is equivalent to $E[\boldsymbol{h'_t}V((\widehat{VaR}_\alpha(X_t), \widehat{ES}_\alpha(X_t)), X_t)] = \boldsymbol{0}$ for all $\mathcal{F}_{t-1}$-measurable $\mathbb{R}^2$-valued functions $\boldsymbol{h_t}$. The simple conditional calibrated test is obtained with $\boldsymbol{h_t} = (1, 1)$, which give us the Wald-type test statistic

$$T_1 = T \left(\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{h_t}\boldsymbol{V}((\widehat{VaR}_\alpha(X_t), \widehat{ES}_\alpha(X_t)), X_t) \right) \widehat{\Omega}_T^{-1} \left(\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{h_t}\boldsymbol{V}((\widehat{VaR}_\alpha(X_t), \widehat{ES}_\alpha(X_t)), X_t) \right)$$

where

$$\widehat{\Omega}_T^{-1} = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{h_t}\boldsymbol{V}((\widehat{VaR}_\alpha(X_t), \widehat{ES}_\alpha(X_t)), X_t))^2$$

They showed that under the null hypothesis $T_1 \xrightarrow{d} \chi^2(1)$ as $T \to \infty$.

10   Methods description

In this section we put together the theory from Sections 5, 6, 7 and 8 to describe the different method used for estimating the risk measures. The three methods follow the same structure, forecasting or sampling the portfolio value for calculating the VaR and ES using Equations (4.6) and (4.7). We also define the method used for forecasting in a rolling window.

ARMA-GARCH with copula vine method is described in Figure 10.11. This method models the serial and the cross sectional dependence separately. The serial dependence, which refers to the dependence created in each time series, is model first using the ARMA model for the mean of each portfolio component and GARCH models for each volatility. Once we have removed the serial dependence, we estimate a vine copula model on the u-scale residuals of the GARCH model for the cross sectional dependence, which refers to the dependence between the assets that compose the portfolio. Once we have select the best model and estimate the corresponding parameters we are able to sample from the copula model and transform to the return scale with ARMA-GARCH models. Once, we have a sample of size N for each portfolio asset we obtain the portfolio value using the weights in Equation (3.4).

Stationary vine (S-vine) copula method from Figure 10.12 models serial and cross sectional dependence together. For this, stationary vine copula models are used. To include the serial dependence the translation invariance defined in Section 7.1 is essential. Also, with the Markov order selected, we reduce the model calculations. Then, since copula models work in the u-scale, we should first select and apply marginal distribution for each portfolio component. Once we have transform the data from return to u-scale we model with stationary vine copula and estimate the parameters. Next, we sample from it an transform back to the return scale with the estimated marginal models. Finally, as we do with ARMA-GARCH and copula method we obtain the portfolio values to estimate the risk measures.

Fleishman's transformation method is showed in Figure 10.13. This method is the one used nowadays by BVK for estimating the risk measures. In this case, we do not model the dependence between the time series. The method uses the first four moments of the components of the portfolio for sampling the portfolio values. The base of this method is to calculate a sample of size N for each asset of the portfolio using Equation (8.39) transformation, from normal variables with the corresponding correlation. With this method we are able to easily sample the individual component of the portfolio with the desire mean, variance, skewness and kurtosis and also the corresponding correlation matrix. Once we have those values we use the weights to calculate the portfolio values and estimate value at risk and expected shortfall.

To forecast from the methods described in Figures 10.11, 10.12 and 10.13 we use a rolling window structure illustrated in Figure 10.14. The method is based on the division of the time period T in several windows of size d. Each window uses the same model, but the parameters are recalculated every h observations corresponding to the refit size. Since in each window the forecast size is also h, then we obtain a forecast for each risk measure for each point in time. For being consistent we need $\frac{T-d}{h}$ to be an integer.

$$\begin{pmatrix} r_1^1 & \cdots & r_T^1 \\ \vdots & \ddots & \vdots \\ r_1^d & \cdots & r_T^d \end{pmatrix}$$

Fit ARMA model
$a_t^i \;=\; r_t^i - \widehat{\mu}_t^i$

$$\begin{pmatrix} a_1^1 & \cdots & a_T^1 \\ \vdots & \ddots & \vdots \\ a_1^d & \cdots & a_T^d \end{pmatrix}$$

Move to u-scale
$F_i(\epsilon_t^i) \;=\; u_t^i$

$$\begin{pmatrix} \epsilon_1^1 & \cdots & \epsilon_T^1 \\ \vdots & \ddots & \vdots \\ \epsilon_1^d & \cdots & \epsilon_T^d \end{pmatrix}$$

Fit GARCH model
$\epsilon_t^i \;=\; \dfrac{a_t^i}{\widehat{\sigma}_t^i}$

$$\begin{pmatrix} u_1^1 & \cdots & u_T^1 \\ \vdots & \ddots & \vdots \\ u_1^d & \cdots & u_T^d \end{pmatrix}$$

Fit R-vine copula $C_{1,\dots,d}$
and simulate N samples

$$\begin{pmatrix} \widehat{u}_1^1 & \cdots & \widehat{u}_N^1 \\ \vdots & \ddots & \vdots \\ \widehat{u}_1^d & \cdots & \widehat{u}_N^d \end{pmatrix}$$

Obtain portfolio values
$\widehat{p}_{T+1}^n = \sum_{j=1}^d w_j \widehat{r}_n^j$

$$\begin{pmatrix} \widehat{r}_1^1 & \cdots & \widehat{r}_N^1 \\ \vdots & \ddots & \vdots \\ \widehat{r}_1^d & \cdots & \widehat{r}_N^d \end{pmatrix}$$

Transform to
return scale
$\widehat{r}_n^i = F_i^{-1}(\widehat{u}_n^i)\widehat{\sigma}_{T+1}^i + \widehat{\mu}_{T+1}^i$

$$\begin{pmatrix} \widehat{p}_{T+1}^1 \\ \vdots \\ \widehat{p}_{T+1}^N \end{pmatrix}$$

Calculate risk measures

$$\begin{pmatrix} \widehat{VaR}_\alpha(p_{T+1}) \\ \widehat{ES}_\alpha(p_{T+1}) \end{pmatrix}$$

Figure 10.11: Risk measures estimation one step ahead using ARMA-GARCH models with R-vine copula method for a portfolio with d assets.

$$\begin{pmatrix} r_1^1 & \cdots & r_T^1 \\ \vdots & \ddots & \vdots \\ r_1^d & \cdots & r_T^d \end{pmatrix}$$

Move to uscale
$$u_t^i = F_i(r_t^i)$$

$$\begin{pmatrix} u_1^1 & \cdots & u_T^1 \\ \vdots & \ddots & \vdots \\ u_1^d & \cdots & u_T^d \end{pmatrix}$$

Simulate N samples
from S-vine copula
$\widehat{C}_{\{T-p,\ldots,T\}\times\{1,\ldots,d\}}$.
$Length(\boldsymbol{u}_n^i) = p$

Fit S-vine copula
with Markov order p
$C_{\{T-p,\ldots,T\}\times\{1,\ldots,d\}}$

$$\begin{pmatrix} \widehat{\boldsymbol{u}}_1^1 & \cdots & \widehat{\boldsymbol{u}}_N^1 \\ \vdots & \ddots & \vdots \\ \widehat{\boldsymbol{u}}_1^d & \cdots & \widehat{\boldsymbol{u}}_N^d \end{pmatrix}$$

Transform sample
to return scale
$\widehat{r}_n^{i,j} = F_i^{-1}(\widehat{u}_n^{i,j})$
with $j \in \{1,\ldots,p\}$
and $i \in \{1,\ldots,d\}$

$$\begin{pmatrix} \widehat{\boldsymbol{r}}_1^1 & \cdots & \widehat{\boldsymbol{r}}_N^1 \\ \vdots & \ddots & \vdots \\ \widehat{\boldsymbol{r}}_1^d & \cdots & \widehat{\boldsymbol{r}}_N^d \end{pmatrix}$$

Calculate risk measures

$$\begin{pmatrix} \widehat{p}_{T+1}^1 & \cdots & \widehat{p}_{T+p}^1 \\ \vdots & \ddots & \vdots \\ \widehat{p}_{T+1}^N & \cdots & \widehat{p}_{T+p}^N \end{pmatrix}$$

Calculate portfolio values
$\widehat{p}_{T+j}^n = \sum_{i=1}^d w_i \widehat{r}_n^{i,j}$

$$\begin{pmatrix} \widehat{VaR}_\alpha(\widehat{p}_{T+1}) & \cdots & \widehat{VaR}_\alpha(\widehat{p}_{T+p}) \\ \widehat{ES}_\alpha(\widehat{p}_{T+1}) & \cdots & \widehat{ES}_\alpha(\widehat{p}_{T+p}) \end{pmatrix}$$

Figure 10.12: Risk measures estimation p steps ahead using stationary copula models with Markov order p for a portfolio with d assets.

$$\begin{pmatrix} r_1^1 & \cdots & r_T^1 \\ \vdots & \ddots & \vdots \\ r_1^d & \cdots & r_T^d \end{pmatrix}$$

Standardise the log returns $y_t^i = \frac{r_t^i - \mu_i}{\sigma_i}$

$$\begin{pmatrix} y_1^1 & \cdots & y_T^1 \\ \vdots & \ddots & \vdots \\ y_1^d & \cdots & y_T^d \end{pmatrix}$$

Sample Nxd i.i.d standard normal variables

Estimate $\widehat{a}_i, \widehat{b}_i, \widehat{c}_i, \widehat{d}_i$ $i = 1, ..., d$ and the inter-correlations matrix R

$$\begin{pmatrix} x_1^1 & \cdots & x_N^1 \\ \vdots & \ddots & \vdots \\ x_1^d & \cdots & x_N^d \end{pmatrix}$$

Obtain $N_d(0, R)$ sample with R correlation matrix. $\boldsymbol{z_j} = F\boldsymbol{x_j}$ $j = 1, ..., N$ $\boldsymbol{z_j}, \boldsymbol{x_j} \in \mathbb{R}^d$. F principal components of R.

$$\begin{pmatrix} z_1^1 & \cdots & z_N^1 \\ \vdots & \ddots & \vdots \\ z_1^d & \cdots & z_N^d \end{pmatrix}$$

Calculate portfolio values $\widehat{p}_{T+1}^n = \sum_{i=1}^d w_i \widehat{r}_n^i$

$$\begin{pmatrix} \widehat{r}_1^1 & \cdots & \widehat{r}_N^1 \\ \vdots & \ddots & \vdots \\ \widehat{r}_1^d & \cdots & \widehat{r}_N^d \end{pmatrix}$$

Transform to return scale $\widehat{y}_j^i = \widehat{a}_i + \widehat{b}_i(z_j^i) + \widehat{c}_i(z_j^i)^2 + \widehat{d}_i(z_j^i)^3$ and $\widehat{r}_j^i = \widehat{y}_j^i \sigma_i + \mu_i$

$$\begin{pmatrix} \widehat{p}_{T+1}^1 \\ \vdots \\ \widehat{p}_{T+1}^N \end{pmatrix}$$

Calculate risk measures

$$\begin{pmatrix} \widehat{VaR}_\alpha(p_{T+1}) \\ \widehat{ES}_\alpha(p_{T+1}) \end{pmatrix}$$

46

Figure 10.13: Risk measures estimation one step ahead using Fleishman's transformation method for a portfolio with d assets.
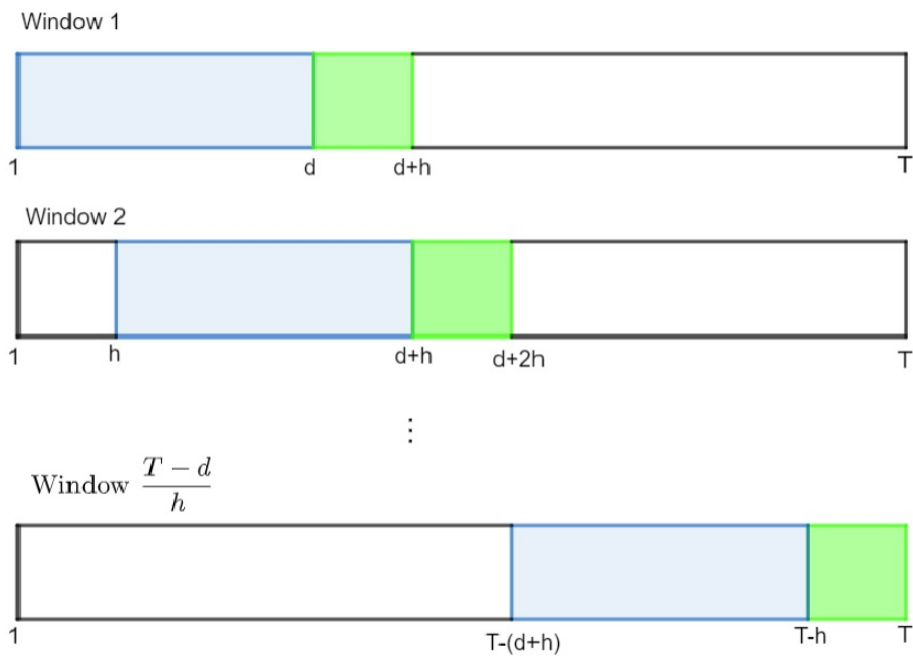
Figure 10.14: Example risk measures estimation structure with a rolling window of size d. The forecasting and parameter recalculation is done every h time points.

11 Data description

In the BVK database we analyse 15 indices used as benchmarks for different strategies in the company. The time series considered run from 20/06/2005 until 11/11/2022 in daily basis, 4340 observation for each time series. The indices can be classified depending on the type of assets that they are benchmarking. Table 11.2 shows the classification by department with the corresponding abbreviation:

	Department	Index Name	Abbreviation
1	Hedge Fund	HFRX ED: Merger Arbitrage	HF Merger
2	Hedge Fund	SG CTA Index	HF CTA
3	Real Estate	MSCI US REIT Index	REITS MSCI
4	Real Estate	REITS Europe	REITS Eur
5	Private Equity	MSCI World Infrastructure	Infrastructure
6	Private Equity	Russell 2000	Russell 2000
7	Private Equity	Russell 3000	Russell 3000
8	Equity	MSCI Europe large cap	MSCI EUR LC
9	Equity	MSCI Emerging Markets	MSCI EM
10	Equity	MSCI US Small Cap Growth	MSCI US SC
11	Equity	MSCI EUR Small Cap Net Return	MSCI EUR SC
12	Equity	MSCI ACWI	MSCI ACWI
13	Fixed Income	Global Aggregate Corporate Total Return	FI Global Agg
14	Fixed Income	Global Capital Securities Total Return	FI Global Cap
15	Fixed Income	ICE BofA BB-B Rated Developed Markets	FI ICE

Table 11.2: Department, name and abbreviation of the indices used in the case study.

The daily closing price has been downloaded from Bloomberg (2023), to assess the different risk measures, we consider an equivalent BVK portfolio, an equal weighted portfolio and a portfolio based on the market capitalisation. The market capitalisation (Market cap) proxies the size of each index. With respect to the currency, we have downloaded each index in its original currency, to not have a currency bias.

Since we have indices from different regions, the calendar day also differ. To have a common calendar, when there is an empty value in some index, the day is deleted for all the time series.

In Table 11.3 we have a general description of the different indices, including the currency or market capitalisation.

For a better understanding of each index, in the following section we summarise the characteristics of each index.

	Index	Currency	Market Capitalisation (Millions)	Number Constituents
1	HF Merger	USD	2000	Na
2	HF CTA	USD	10000	20
3	REITS MSCI	USD	984849.54	131
4	REITS Eur	EUR	111404	46
5	Infrastructure	USD	3357282.11	125
6	Russell 2000	USD	2806774.78	1954
7	Russell 3000	USD	40787893.59	2963
8	MSCI EUR LC	EUR	8753553.38	200
9	MSCI EM	USD	17316534.49	1384
10	MSCI US SC	USD	4729267.87	1950
11	MSCI EUR SC	EUR	1591259.16	1023
12	MSCI ACWI	USD	71570677.53	2891
13	FI Global Agg	EUR	10123263.92	15231
14	FI Global Cap	EUR	672419.84	886
15	FI ICE	USD	1549652.87	Na

Table 11.3: Currency, market capitalisation and number of constituents at 15/11/2022.

11.1   Indices description

In this section we describe the indices briefly, including the geographical and sector distribution. The constituents are not publicly available for some indices.

HFRX ED: Merger Arbitrage

Merger Arbitrage strategies employ an investment process focused on opportunities in equity and equity related instruments of companies which are currently engaged in a corporate transaction. Merger Arbitrage involves primarily announced transactions, typically with limited or no exposure to situations in which no formal announcement is expected to occur. Opportunities are frequently presented in cross border, collared and international transactions which incorporate multiple geographic regulatory institutions, with typically involve minimal exposure to corporate credits.

SG CTA Index

The SG CTA Index calculates the net daily rate of return for a group of 20 Commodity Trading Advisors (CTAs) selected from the largest managers open to new investment. The SG CTA Index is equal-weighted and reconstituted annually with re-balancing on January 1st.

MSCI US REIT Index

The MSCI US REIT Index is comprised of equity Real Estate Investment Trusts (REITs). The Index is based on the MSCI USA Investable Market Index (IMI) which captures the large, mid and small cap segments of the USA market. With 129 constituents, it represents about 99% of the US REIT universe and securities are classified under the Equity REITs Industry. The MSCI US REIT is a free float-adjusted market capitalisation weighted index. This means that the market capitalisation is calculated using the free float shares of the market.

REITS Europe

The constituents of the Euronext IEIF Europe Index are a selection of the most representative Real Estate Investment Trusts (REITs) in Europe, chosen for their market capitalisation and liquidity. The Euronext IEIF Europe Index selects companies whose market capitalisation is more than 0.4% of the Universe (Property companies listed on European regulated markets that have opted for a tax-transparency regime), with a minimum free float of 20% and with a minimum daily average turnover representing 0.2%.
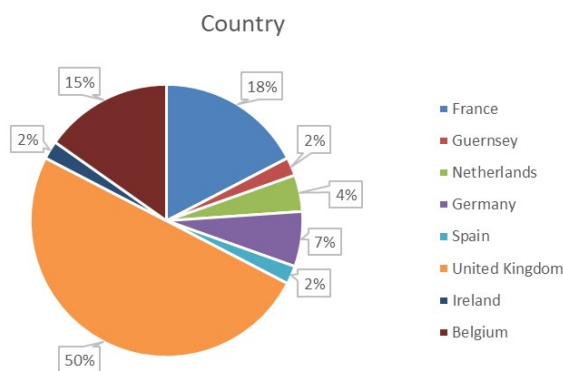
Figure 11.15: Geographic distribution of REIT Eur.

Apart from United Kingdom with 50% of the Index, the REIT Europe is distributed in the main European countries. Respect to the sector, all constituents belongs to Real Estate.

MSCI World Infrastructure

The MSCI Infra Index is a free float -adjusted market cap weighted index. Its members are infrastructure owners and operators who tend to demonstrate highly inelastic demand patterns, stable, predictable returns and inflation-linked pricing power. Includes companies in the telecom, utilities, energy, transportation and social infrastructure sector.
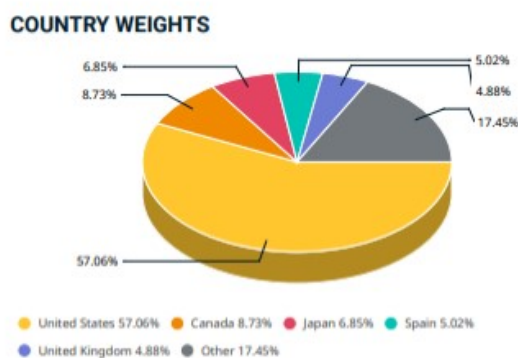
Figure 11.16: Geographic distribution of Infrastructure.

Russell 3000

The Russell 3000 is composed of 3000 large U.S. companies, as determined by market capitalisation. This portfolio of securities represents approximately 98% of the investable U.S. equity market. The Russell 3000 Index is constructed to provide a comprehensive, unbiased and stable barometer of the broad market and is completely reconstituted annually to ensure new and growing equities are included.
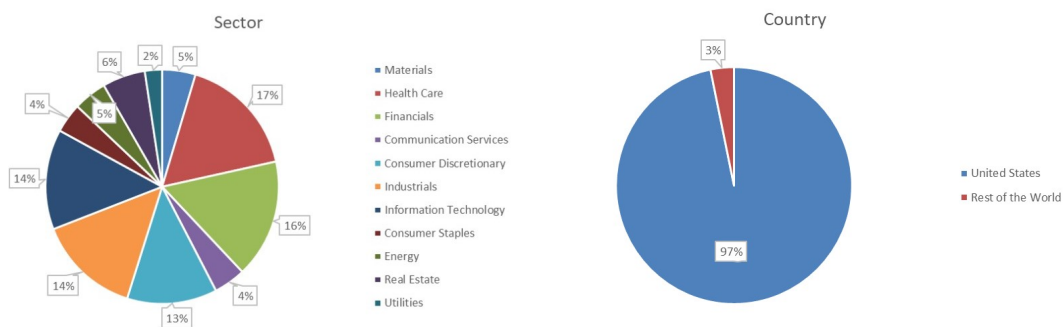
(a) Sector distribution of Russell 3000.

(b) Geographic distribution of Russell 3000.

Figure 11.17: Sector and geographic distribution of Russell 3000.

Figure 11.17a shows the main U.S. equity market is widely distributed across different sectors. The main ones are Health Care, Financial and Industrial.

Russell 2000

The Russell 2000 Index is comprised of the smallest 2000 companies in the Russell 3000 Index based on a combination of their market cap and current index membership, representing approximately 8% of the Russell 3000 total market capitalisation. The Russell 2000 is constructed to provide a comprehensive and unbiased small-cap barometer and is completely reconstituted annually.
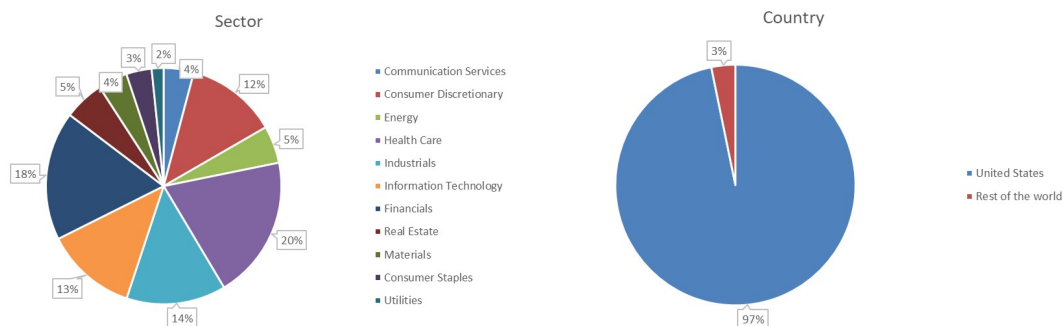
(a) Sector distribution of Russell 2000.

(b) Geographic distribution of Russell 2000.

Figure 11.18: Sector and geographic distribution of Russell 2000.

We get similar results to the Russell 3000 with large distribution between the sectors. The biggest sectors are again Health Care, Financial and Industrial.

MSCI Europe Large Cap

The MSCI Europe Large Cap Index captures large cap representation across 15 Developed Markets countries in Europe. With 199 constituents, the index covers approximately 70% of the free float-adjusted market capitalisation across European Developed Markets equity universe.

(a) Sector distribution of MSCI Eur LC. (b) Geographic distribution of MSCI Eur LC.
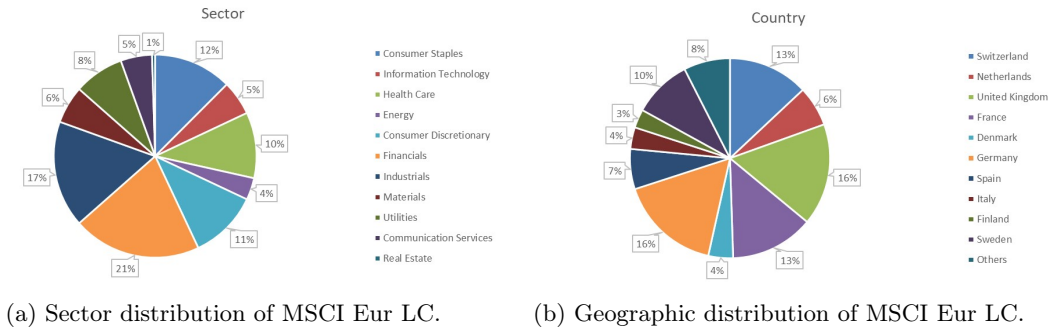
Figure 11.19: Sector and geographic distribution of MSCI Eur LC.

We find a wide variety on sector and countries. The most important sectors are Financial, Industrial and Consumer Staples. Respect to the countries, the biggest are United Kingdom, Germany and Switzerland.

MSCI Emerging Markets

The MSCI EM (Emerging Markets) Index is a free-float weighted equity index that captures large and mid cap representation across 24 Emerging Markets countries. With 1377 constituents, the index covers approximately 85% of the free float-adjusted market capitalisation in each country.

(a) Sector distribution of MSCI EM. (b) Geographic distribution of MSCI EM.
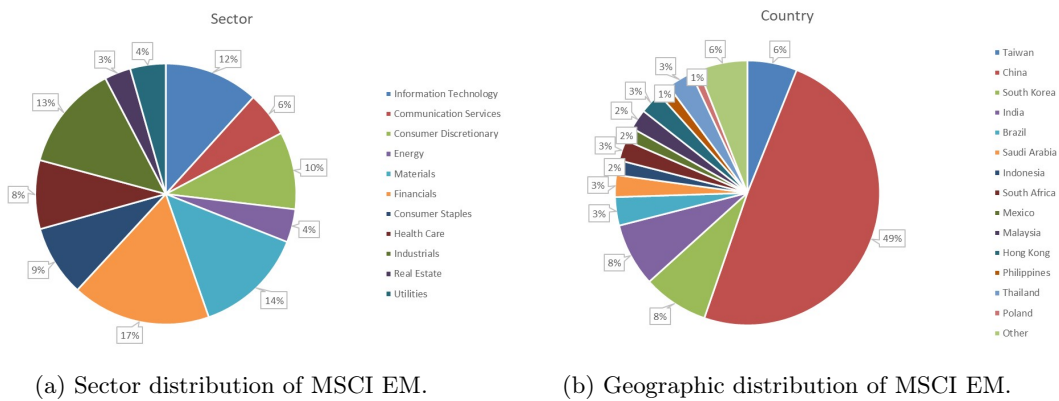
Figure 11.20: Sector and geographic distribution of MSCI EM.

In this case we should remark the size of China, with almost 50% of the Index. In terms

of the sector, is more distributed, the most important are Financial, Materials and Information Technology.

MSCI US Small Cap Growth

The MSCI USA Small Cap Index is designed to measure the performance of the small cap segments of the US market. With 1938 constituents, the index covers approximately 14% of the free float-adjusted market capitalisation in the US.

(a) Sector distribution of MSCI US SC.

(b) Geographic distribution of MSCI US SC.
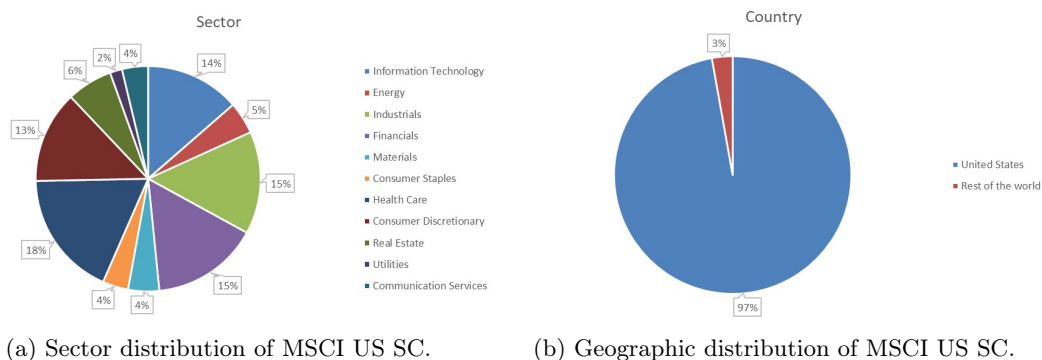
Figure 11.21: Sector and geographic distribution of MSCI US SC.

The main sectors are Health Care, Industrial and Financial.

MSCI EUR Small Cap Net Return

The MSCI Europe Small Cap is a free float-adjusted market capitalisation weighted index that is designed to measure the equity market performance of the small cap size segment in the MSCI Europe Investable Market Index (IMI). With 991 constituents, the index covers approximately 14% of the free float-adjusted market capitalisation in the European equity universe.

(a) Sector distribution of MSCI Eur SC.

(b) Geographic distribution of MSCI Eur SC.
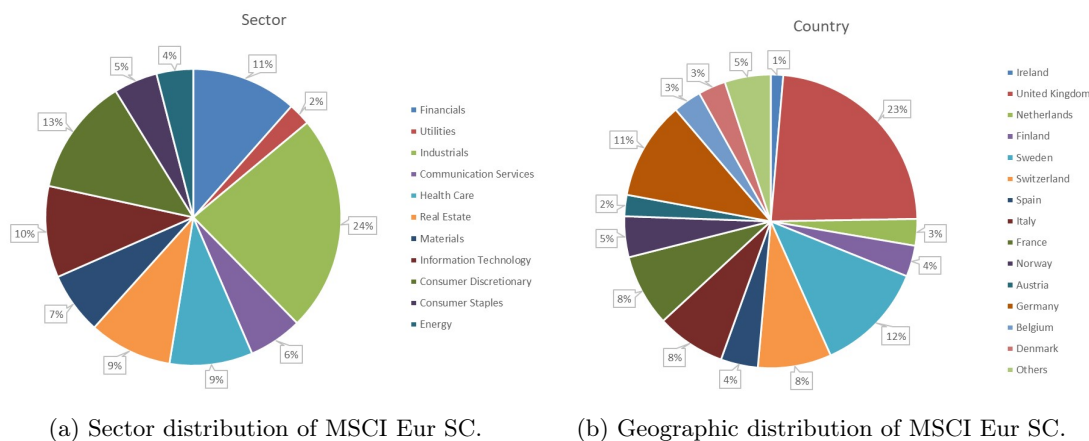
Figure 11.22: Sector and geographic distribution of MSCI Eur SC.

The most important small cap sectors in Europe are Industrial, Consumer Discretionary and Financial; similar to the US small cap market. By country United Kingdom stands out with 28%.

MSCI ACWI

The MSCI ACWI Index is designed to represent performance of the full opportunity set of large and mid cap stocks across 23 developed and 24 emerging markets. It covers more than 2933 constituents across 11 sectors and approximately 85% of the free float-adjusted market capitalisation in each market. The index is built taking into account variations reflecting conditions across regions, market cap sizes, sectors, style segments and combinations.

(a) Sector distribution of MSCI ACWI.   (b) Geographic distribution of MSCI ACWI.
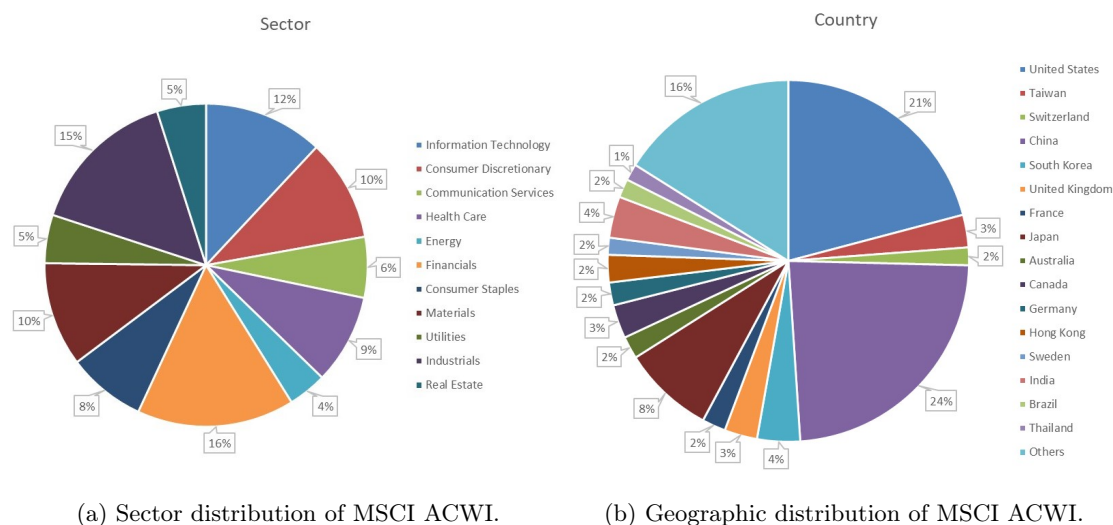
Figure 11.23: Sector and geographic distribution of MSCI ACWI.

We must highlight two main countries which are China and United Estates, followed by Japan. Sectors are more spread out, the main ones being Financial and Information Technology.

Global Aggregate Corporate Total Return

The Bloomberg Global Aggregate-Corporate Index is a flagship measure of global investment grade, fixed-rate corporate debt. This multi-currency benchmark includes bonds from developed and emerging markets issuers within the industrial, utility and financial sectors.

Global Capital Securities Total Return

The investment objective of the fund is to maximise total returns over the long term. The fund invests primarily in global investment grade bonds. The fund also invests in high yield bonds and other fixed income securities including mortgage backed securities and asset backed securities.

ICE BofA BB-B Rated Developed Markets

The index includes original issue zero coupon bonds, Eurodollar bonds, 144 securities and pay-in-kind securities with BB-B rated and have a developed markets country of risk.Developed

markets is defined as an FX-G10 member, a Western European nation, or a territory of the US. The FX-G10 includes all Euro members, the US, Japan, the UK, Canada, Australia, New Zealand, Switzerland, Norway and Sweden. Index constituents are capitalisation-weighted, based on their current amount outstanding.

11.2 Data analysis

In this section we explore the data available. We want to examine the joint data behaviour. For this, we start plotting the prices and log returns of the time series and portfolios considered.
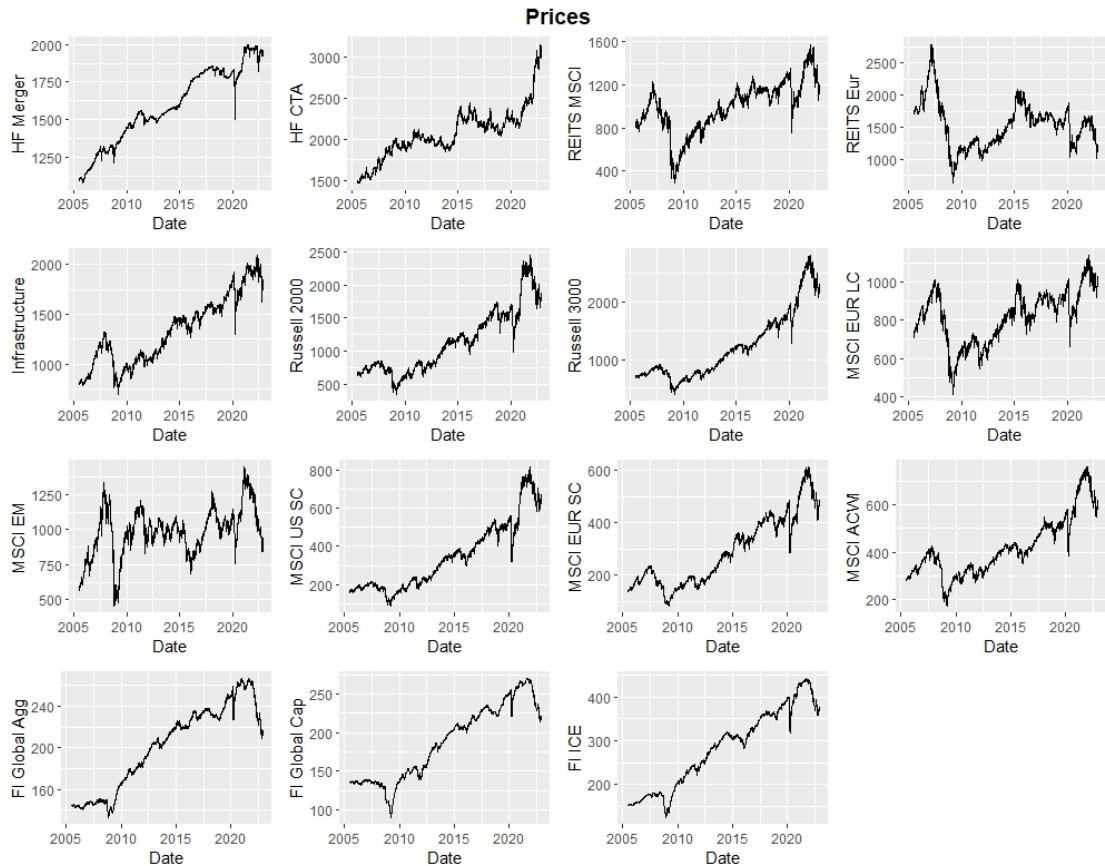
Figure 11.24: Prices from the 15 indices of Table 11.2.

Figure 11.24 shows most of the time series have a similar evolution, with a stable growth and some important falls. However, there are some indices with different performance, the most remarkable are the REITS Eur and MSCI EM (Emerging Markets).

In general, we observe in Figure 11.25 some periods with higher volatility around 2008 due to the financial crisis and 2020 due to the COVID-19. Also, conditional heteroscedasticity (GARCH) effects are observed in the graphs. While the mean of the time series is around zero, the volatility is grouped by periods of high and low values. This is a typical characteristic of financial time series. For modelling these effects we can use ARMA-GARCH models or include in a stationary
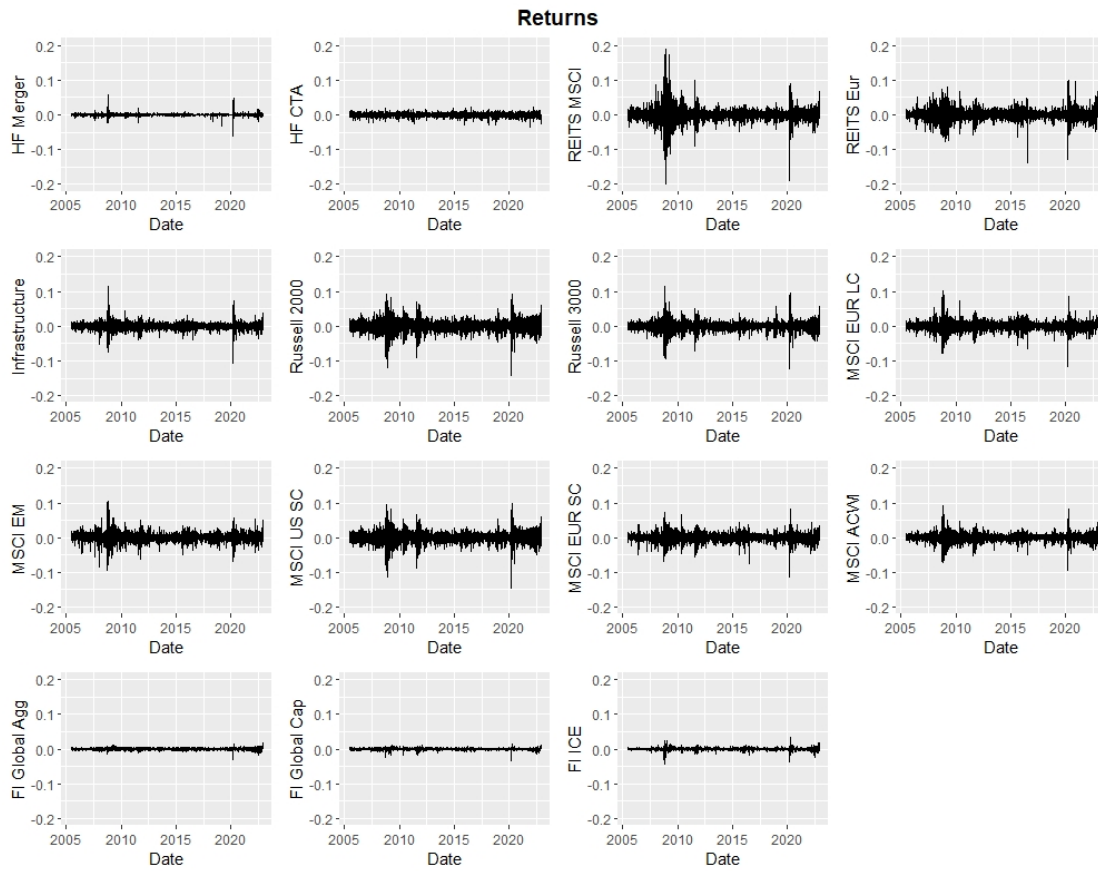
Figure 11.25: Returns from the 15 indices of Table 11.2.

copula model. In the following section we try to capture those effects with different models for a better VaR and ES forecast.

In the next sections we apply marginal and copula models, stationary vine copula models and the Fleishman approximation to three portfolios composed of the indices in Table 11.2 to sample the portfolio returns and forecast the VaR and ES. The portfolios are an equal weighted portfolio (EW), a portfolio based on the BVK currently strategy (BVK) and market capitalisation portfolio (MC).

When applying the Fleishman approximation in a rolling window to the three portfolios we obtain some window where the method does not converge. Then, the study is split in two cases with three portfolios each. In Section 12 we apply marginal and copula models and stationary vine copula models to the complete portfolios, including a comparison of the results obtained. Afterwards in Section 13 we sample the reduced portfolios using the three methods.

12 Complete portfolio

As we mention in the previous section, in this section we select the best model for marginal and copula models and stationary vine copula models; using them in a rolling window setup for forecasting the risk measures of the portfolios considered in Table 12.4.

	Index	BVK Weight	Market Cap Weight	Equal Weight
1	HF Merger	0.023964676	1.216790e-05	0.0666666
2	HF CTA	0.044415709	6.083952e-05	0.0666666
3	REITS MSCI	0.159135120	5.991778e-03	0.0666666
4	REITS Eur	0.159135120	6.777766e-04	0.0666666
5	Infrastructure	0.053570032	2.042554e-02	0.0666666
6	Russell 2000	0.054959743	1.707628e-02	0.0666666
7	Russell 3000	0.054959743	2.481516e-01	0.0666666
8	MSCI EUR LC	0.056061704	5.325620e-02	0.0666666
9	MSCI EM	0.045708941	1.053530e-01	0.0666666
10	MSCI US SC	0.029519269	2.877264e-02	0.0666666
11	MSCI EUR SC	0.029519269	9.681145e-03	0.0666666
12	MSCI ACWI	0.049542248	4.354326e-01	0.0666666
13	FI Global Agg	0.057876476	6.158946e-02	0.0666666
14	FI Global Cap	0.047896057	4.090970e-03	0.0666666
15	FI ICE	0.133735893	9.428014e-03	0.0666666

Table 12.4: Currency, market capitalisation, number of constituents, and weights in equal weighted, BVK and market capitalisation portfolios of the indices used in the complete study.

The portfolios prices and returns from Table 12.4 are shown in Figure 12.26.
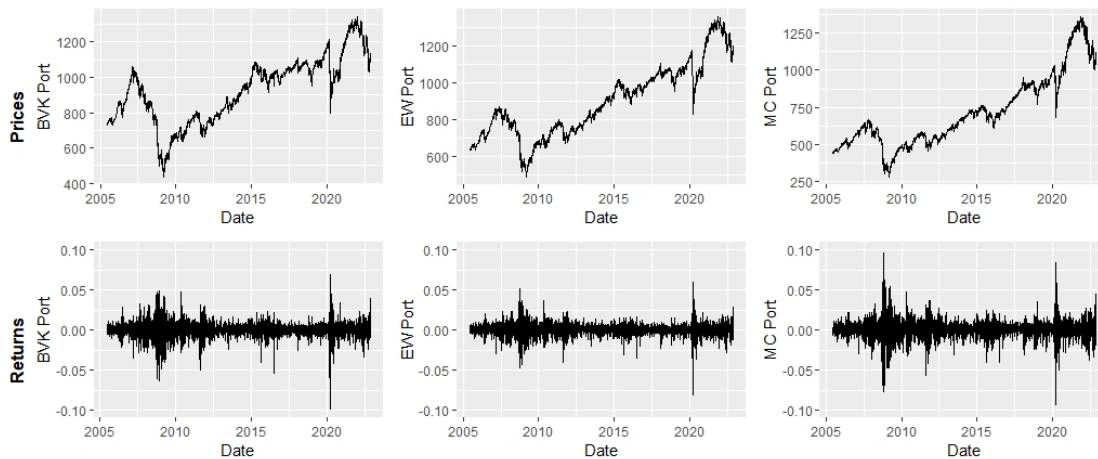
Figure 12.26: Prices and returns of BVK, equal weighted and market capitalisation portfolios described in Table 12.4.

Figure 12.26 shows the prices behave similar to most of the indices prices shown in Figure

11.24. Also, the GARCH effects are inherited from the indices GARCH effects that compose the portfolio.

In Sections 12.1 and 12.2 we develop the ARMA-GARCH with vine copula models and stationary vine copula models respectively for the portfolios in Table 12.4. Finally, in Section 12.3 we compare the results obtain between both methods.

12.1 Marginal and copula models

In this section we fit ARMA-GARCH and copula models to the data in Table 12.4 and then to be used for sampling portfolio returns and forecasting the risk measures. In this approach we should first apply univariate ARMA-GARCH models to each time series for removing serial dependence in Section 12.1.1 and then, use a vine copula model for assessing the cross sectional dependence in Section 12.1.2. Before implementing the method we determine the ARMA-GARCH order for each index. Then, we choose the best vine copula model and forecast the different risk measures on the portfolio level.

12.1.1 Marginal model analysis

In this section we select the models that better fit the univariate time series data. Then, we should select appropriate ARMA and GARCH orders and the innovation distribution which better model the data adequately.

In finance, the returns series are usually modelled with ARMA(1,1)-GARCH(1,1). Also, Skewed-Student innovations are often considered, since it is heavy tailed and non symmetric. Then, we check if this model fit the univariate time series adequately.

Therefore, we check if the ARMA(1,1)-GARCH(1,1) model with Skewed-Student innovations fit the returns data using the QQ-plots, defined in Section 5.3, for the residual distribution in Figure 12.27 and the Ljung-Box test, described in Section 5.3, for the ARMA-GARCH order in Figure 12.29.

In the Figure 12.27 we observe the tails of HF Merger, HF CTA, REITS Eur, Infrastructure, Russell 2000, Russell 3000, MSCI Eur LC, MSCI US SC, MSCI Eur SC, MSCI ACWI, FI Global Agg, FI Global Cap and FI ICE are not correctly fitted. To select the best distribution we compare the QQ-plots of the standardised residuals for different innovation distributions. The distribution compared are the Normal Inverse Gaussian distribution (nig), Generalised Hyperbolic (ghyp), Johnson's SU distribution (jsu) and Skewed-Student (sstd). The t-Student (std) and skew-normal distribution (snorm) have been also considered but do not fit well, which is to be expected with Skewed student being more generic.

We take REITS Eur as an example of the univariate time series. The others univariate time series QQ-plots are in Appendix A. We select the innovation distribution which better fits the diagonal line in the QQ-plot which is ghyp innovation distribution. In Table 12.5 we present the chosen innovation distributions after this exercise.

We continue the analysis with the ARMA and GARCH orders for the univariate time series. We first check the most common model ARMA(1,1)-GARCH(1,1) applying the Ljung-Box test to the standardised residuals of each index.

The grey colours in Figure 12.29 represent a $p-value < 0.05$. This means there are significant autocorrelation effects of the corresponding lag on the standardised residuals.
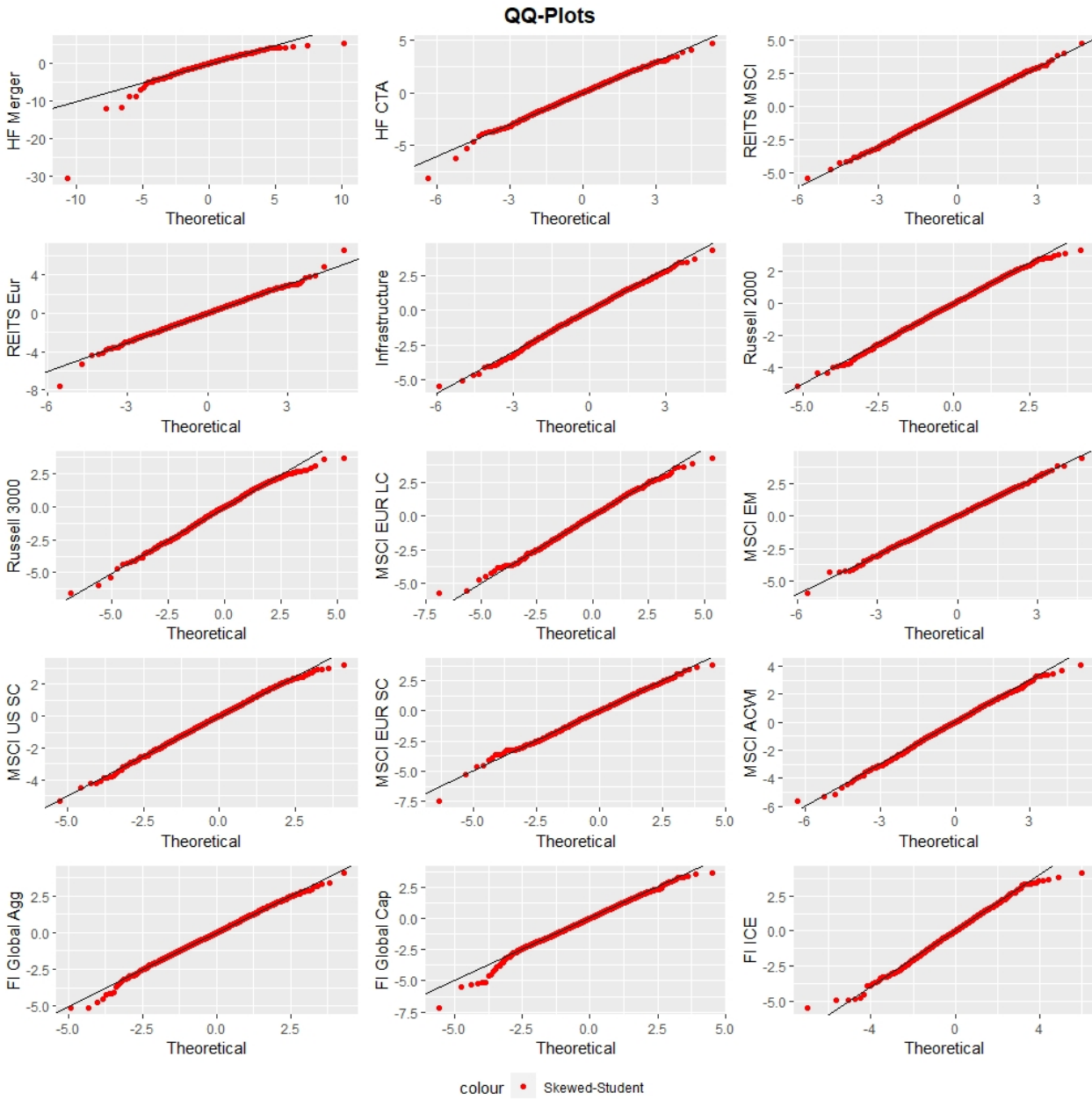
Figure 12.27: QQ-plots of the standardised residuals after fitting an ARMA(1,1)-GARCH(1,1) model with Skewed-Student innovations to each of the 15 time series.
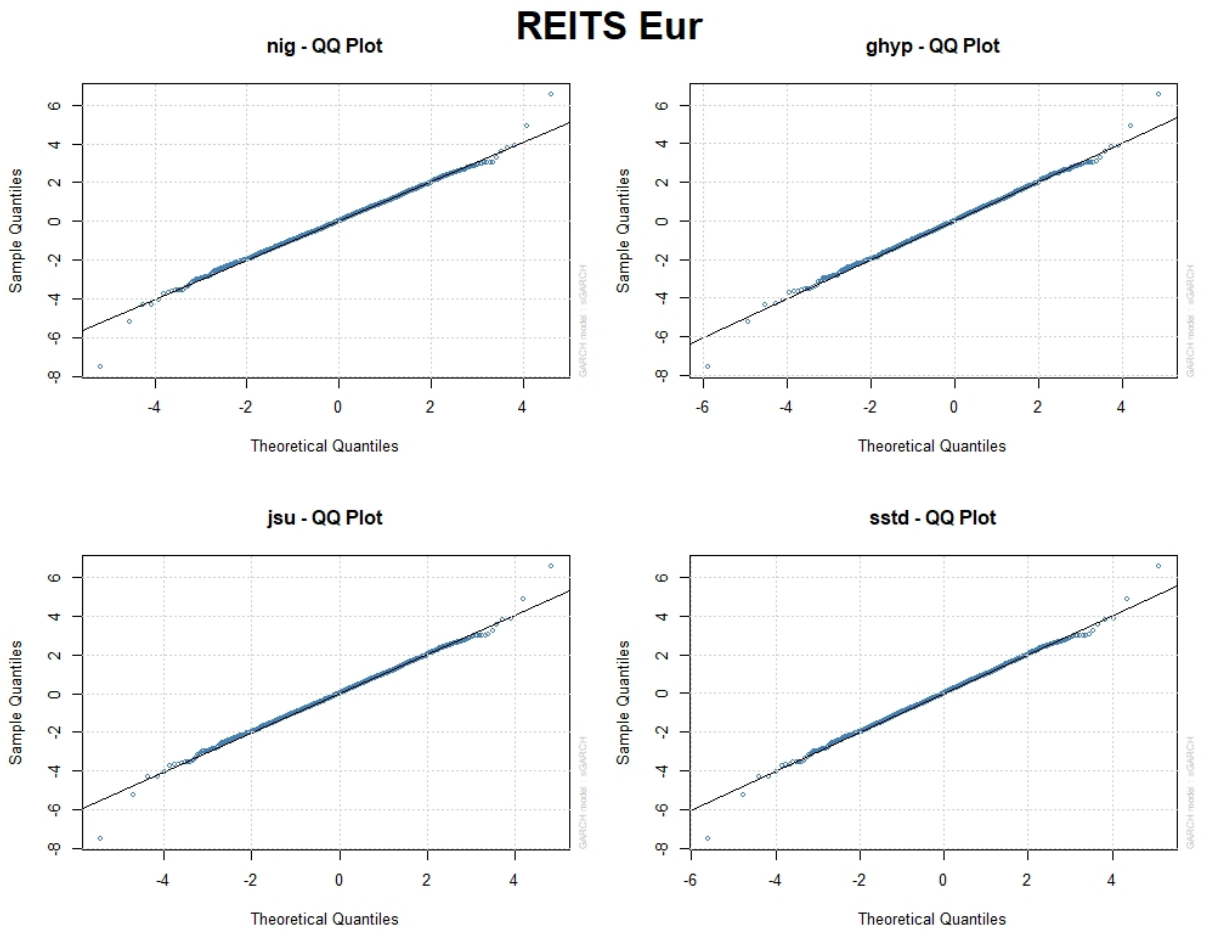
Figure 12.28: REITS Eur standardised residuals QQ-plots after fitting an ARMA(1,1)-GARCH(1,1) with different innovation distributions.

	Index	Innovation distribution
1	HF Merger	nig
2	HF CTA	ghyp
3	REITS MSCI	sstd
4	REITS Eur	ghyp
5	Infrastructure	sstd
6	Russell 2000	ghyp
7	Russell 3000	ghyp
8	MSCI EUR LC	nig
9	MSCI EM	sstd
10	MSCI US SC	ghyp
11	MSCI EUR SC	sstd
12	MSCI ACWI	nig
13	FI Global Agg	ghyp
14	FI Global Cap	nig
15	FI ICE	nig

Table 12.5: Chosen innovation distributions for each index in a ARMA(1,1)-GARCH(1,1) setup.
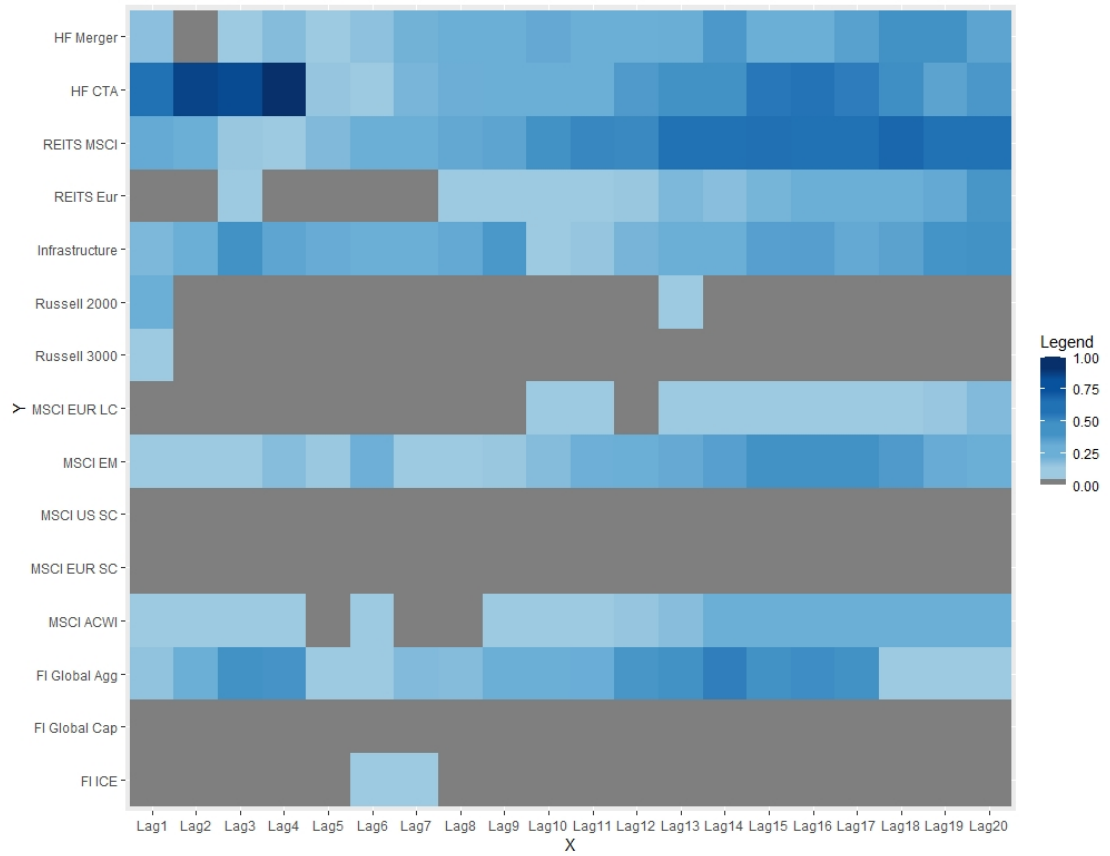
Figure 12.29: P-values of the Ljung-Box test statistics for different lags after fitting an ARMA(1,1)-GARCH(1,1) model for each index with innovation distributions specified in Table 12.5.

We have some index time series with non zero autocorrelation effects on the standardised residuals. Therefore, we should correct the orders selected. We now review the ARMA orders of the following time series: REITS Eur, Russell 2000, Russell 3000, MSCI Eur LC, MSCI US SC, MSCI Eur SC, FI Global Cap and FI ICE.

We base the choice of the order on the autocorrelation function (ACF), the partial autocorrelation function (PACF) and the Bayesian Information Criterion (BIC) after fitting different ARMA(p,q) models without GARCH effects using the innovation distributions of Table 12.5.

REITS Eur

AR	MA	Mean	BIC
0	1	0	-5.83458820431736
1	0	0	-5.83454315349623
2	1	0	-5.83286768261158
2	0	0	-5.83272834426793
0	2	0	-5.8326787857292
1	1	0	-5.83267148578969
0	1	1	-5.83266222315687
1	0	1	-5.83261716432284
3	2	0	-5.83165079440322
2	2	0	-5.83156108864491
3	1	0	-5.83131799550836
1	3	0	-5.83116221573652
3	0	0	-5.83111420557001
0	0	1	-5.83107712611561
0	3	0	-5.83100733871227
2	1	1	-5.83094218934115
2	0	1	-5.83080244753165
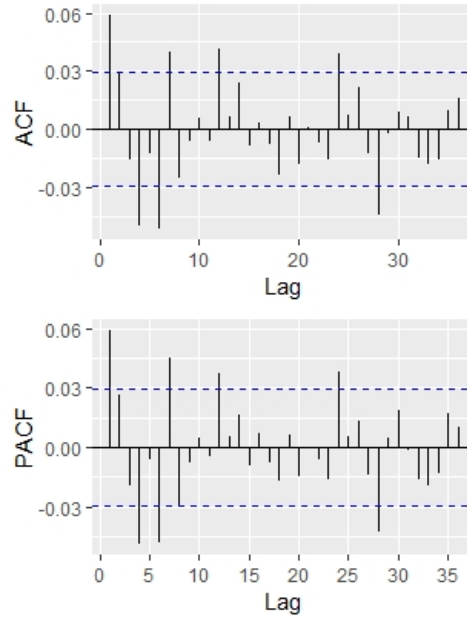
Figure 12.30: Performance of different ARMA orders for REITS Eur's returns with Generalised Hyperbolic innovations and ACF and PACF of REITS Eur's returns.

Similar as for the innovation distribution choice, we focus our ARMA and GARCH order selection study on REITS Eur. We then use the same order selection arguments for the remaining indices. The corresponding plots considered for the selection of the ARMA order are in Appendix B.

In case of the REITS Eur, from Figure 12.30 we select the model with lower BIC. The ACF and PACF mainly have the first lag being significant. Then the ARMA(0,1) model corrects this autocorrelation effect.

The selected ARMA orders for all indices are given in Table 12.6.

We continue choosing the best GARCH orders. For this we utilise theoretical properties of GARCH models. The GARCH order study is done on the squared residuals after fitting the ARMA model of Table 12.6. From Section 5.2.2 we know GARCH models can be rewritten as a ARMA model for the squared residuals A_t^2. Then, can be proved that these can be rewritten for:

	Index	Innovation distribution	Model
1	HF Merger	nig	ARMA(1,1)
2	HF CTA	ghyp	ARMA(1,1)
3	REITS MSCI	sstd	ARMA(1,1)
4	REITS Eur	ghyp	ARMA(0,1)
5	Infrastructure	sstd	ARMA(1,1)
6	Russell 2000	ghyp	ARMA(0,1)
7	Russell 3000	ghyp	ARMA(1,0)
8	MSCI EUR LC	nig	ARMA(0,1)
9	MSCI EM	sstd	ARMA(1,1)
10	MSCI US SC	ghyp	ARMA(0,1)
11	MSCI EUR SC	sstd	ARMA(0,1)
12	MSCI ACWI	nig	ARMA(1,1)
13	FI Global Agg	ghyp	ARMA(1,1)
14	FI Global Cap	nig	ARMA(2,1)
15	FI ICE	nig	ARMA(1,1)

Table 12.6: Selected ARMA orders and innovation distributions for each index of Table 11.2.

- $GARCH(p,q)$ as an $ARMA(r,q)$ model using $r = max(p,q)$.

- $IGARCH(1,1)$ as an $ARIMA(0,1,1)$.

- $IGARCH(p,q)$ as an $ARMA(r,q)$ model using $r = max(p,q)$, $q > 0$ and $p + q > 1$.

Therefore, we base the GARCH order selection on the ACF and PACF of the squared residuals, the BIC of different ARIMA(p,i,q) orders applied on the squared residuals after applying ARMA models in Table 12.6, and the Ljung-Box test.

REITS Eur

AR	MA	Mean	ARFIMA	BIC
0	0	1	0	-15.617455596703
0	0	1	1	-15.616751505363
1	0	1	1	-15.616193018540
2	0	1	0	-15.616124734864
0	1	1	1	-15.616067958450
1	1	1	0	-15.615769560271
0	2	1	1	-15.614394274111
2	1	1	0	-15.614083741014
0	1	1	0	-15.614074178135
1	1	1	1	-15.613475626648
2	1	1	1	-15.612323423320
1	2	1	0	-15.612207166078
2	2	1	0	-15.612193625019
2	0	1	1	-15.610759649065
2	2	1	1	-15.610418361050
1	2	1	1	-15.609719945452
1	0	1	0	-15.555898820515

Figure 12.31: Performance of different ARMA orders for REITS Eur's squared residuals after fitting ARMA(0,1) with Generalised Hyperbolic innovations. ACF and PACF of REITS Eur's squared residuals after fitting ARMA(0,1) with Generalised Hyperbolic innovations.

67

From Figure 12.31 we select model GARCH(1,1) for REITS Eur since we have many significant lags in the ACF and PACF. We could consider the ARIMA(0,1,1), corresponding to IGARCH(1,1), or ARIMA(1,0,1) corresponding to GARCH(1,1). Since both of them have similar BIC as we see in Figure 12.31, we finally choose GARCH(1,1) based on the Ljung-Box test.

Appendix C shows the corresponding graphs used for the selection of the GARCH order of the remaining univariate time series. It should be remarked that in general, we observe the models with lowest BIC not include GARCH effects (ARIMA(0,1,0)). However, we have seen in Figure 11.25 the existence of those effects in all univariate time series.

Finally, Table 12.7 summarise the chosen marginal time series models for the indices of Table 11.2.

	Index	Innovation distribution	Model
1	HF Merger	nig	ARMA(1,1)-GARCH(1,1)
2	HF CTA	ghyp	ARMA(1,1)-GARCH(1,1)
3	REITS MSCI	sstd	ARMA(1,1)-GARCH(1,1)
4	REITS Eur	ghyp	ARMA(0,1)-GARCH(1,1)
5	Infrastructure	sstd	ARMA(1,1)-GARCH(1,1)
6	Russell 2000	ghyp	ARMA(0,1)-GARCH(1,1)
7	Russell 3000	ghyp	ARMA(1,0)-GARCH(1,1)
8	MSCI EUR LC	nig	ARMA(0,1)-GARCH(1,1)
9	MSCI EM	sstd	ARMA(1,1)-GARCH(1,1)
10	MSCI US SC	ghyp	ARMA(0,1)-GARCH(1,1)
11	MSCI EUR SC	sstd	ARMA(0,1)-GARCH(1,1)
12	MSCI ACWI	nig	ARMA(1,1)-GARCH(1,1)
13	FI Global Agg	ghyp	ARMA(1,1)-GARCH(1,1)
14	FI Global Cap	nig	ARMA(2,1)-GARCH(1,2)
15	FI ICE	nig	ARMA(1,1)-GARCH(2,1)

Table 12.7: Selected ARMA and GARCH orders and innovation distributions for each index.

We plot again in Figure 12.32 the p-values corresponding to Ljung-Box test after applying the models specified in Table 12.7.

We still observe some significant autocorrelation dependence. The time series with higher autocorrelation are MSCI Eur SC, FI Global Cap and FI ICE. Since the Ljung-Box test compares in the null hypotheses having the first i lags independent, then some null hypotheses are rejected because of the effect of previous lags. We can check it with the AFC.

Figure 12.33 shows most of the lags are not significant. However, the effect of the first significant lags is carried over the following Ljung-Box test. Therefore, we maintain the ARMA-GARCH orders selected in Table 12.7.

In Figure 12.34 we show that the standardised residuals behave like white noise. Despite of having GARCH effects removed, we still observe some outlier. The largest one is in HF Merger. The outliers observed are usually caused due to some important price movement during a low volatility period. In case of the HF Merger the outlier occurs on the 28/02/2019 when the price drops from 1842,57 to 1780,54 ($-3,37\%$), during a period of very low volatility.

Index	AR(1)	AR(2)	MA(1)	GARCH(1)	GARCH(2)	ARCH(1)	ARCH(2)
HF Merger	-0.121		0.079	0.181		0.818	
HF CTA	-0.092		0.174	0.060		0.886	
REITS MSCI	0.782		-0.816	0.099		0.898	
REITS Eur			0.048	0.109		0.885	
Infrastructure	-0.234		0.305	0.096		0.889	
Russell 2000			-0.061	0.089		0.903	
Russell 3000	-0.067			0.128		0.869	
MSCI EUR LC			-0.034	0.126		0.864	
MSCI EM	0.1005		0.069	0.089		0.898	
MSCI US SC			-0.045	0.097		0.896	
MSCI EUR SC			0.058	0.143		0.849	
MSCI ACWI	-0.073		0.198	0.115		0.880	
FI Global Agg	0.986		-0.972	0.057		0.937	
FI Global Cap	1.061	-0.101	-0.899	0.113		0.540	0.331
FI ICE	0.6116		-0.341	0.130	0.005	0.862	

Table 12.8: Estimated parameters for ARMA-GARCH models selected in Table 12.7 for each index.

Once we have selected the best marginal models we continue in Section 12.1.2 with choosing the best vine structure.
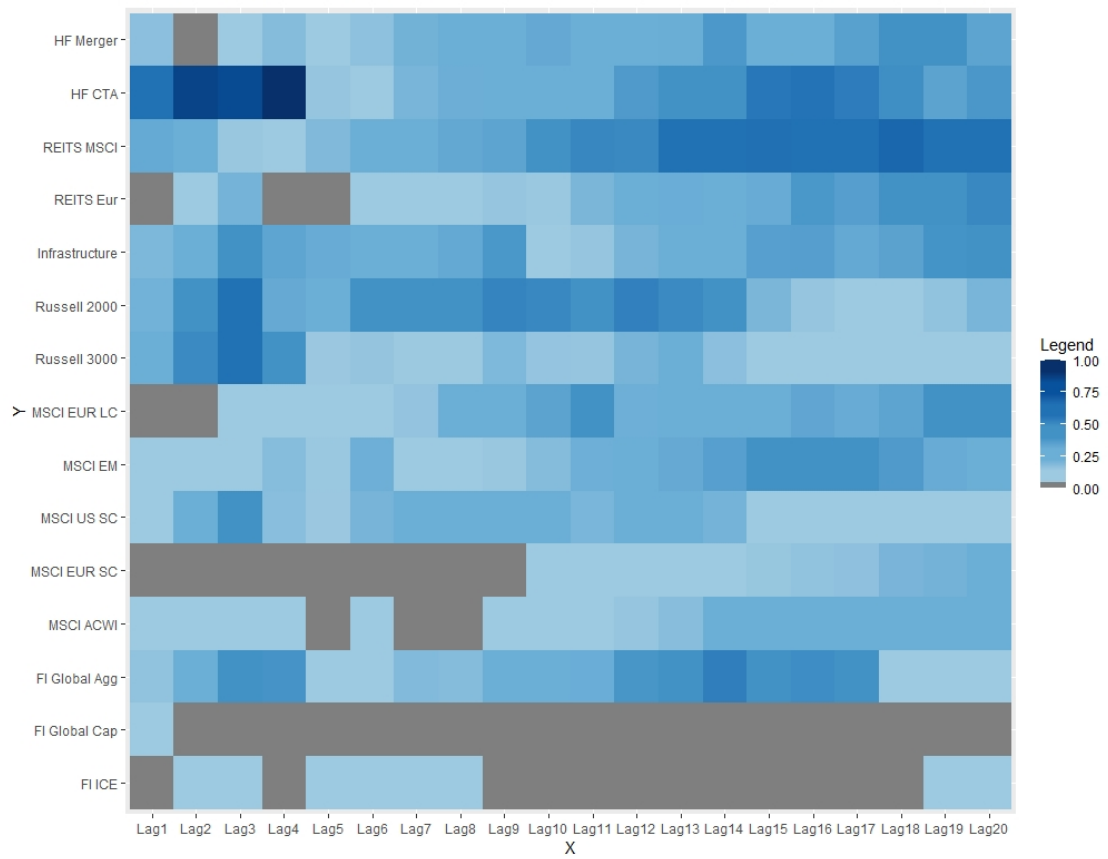
Figure 12.32: P-values of Ljung-Box test statistics for different lags on standardised residuals after fitting the models specified on Table 12.7 for each index of Table 11.2.
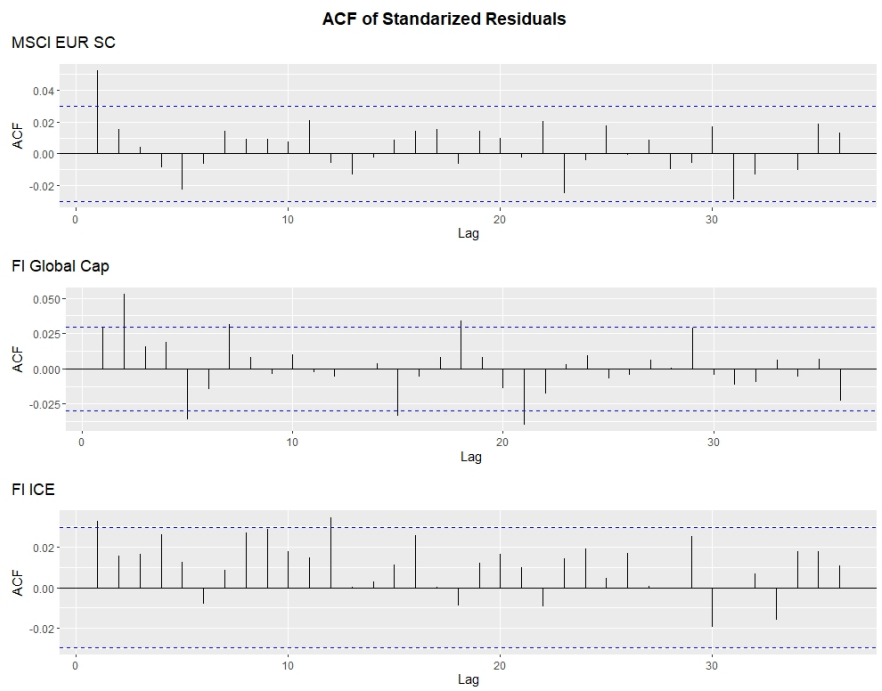
Figure 12.33: ACF of standardised residuals after fitting the models of Table 12.7 corresponding to MSCI Eur SC, FI Global Cap, FI ICE.

71
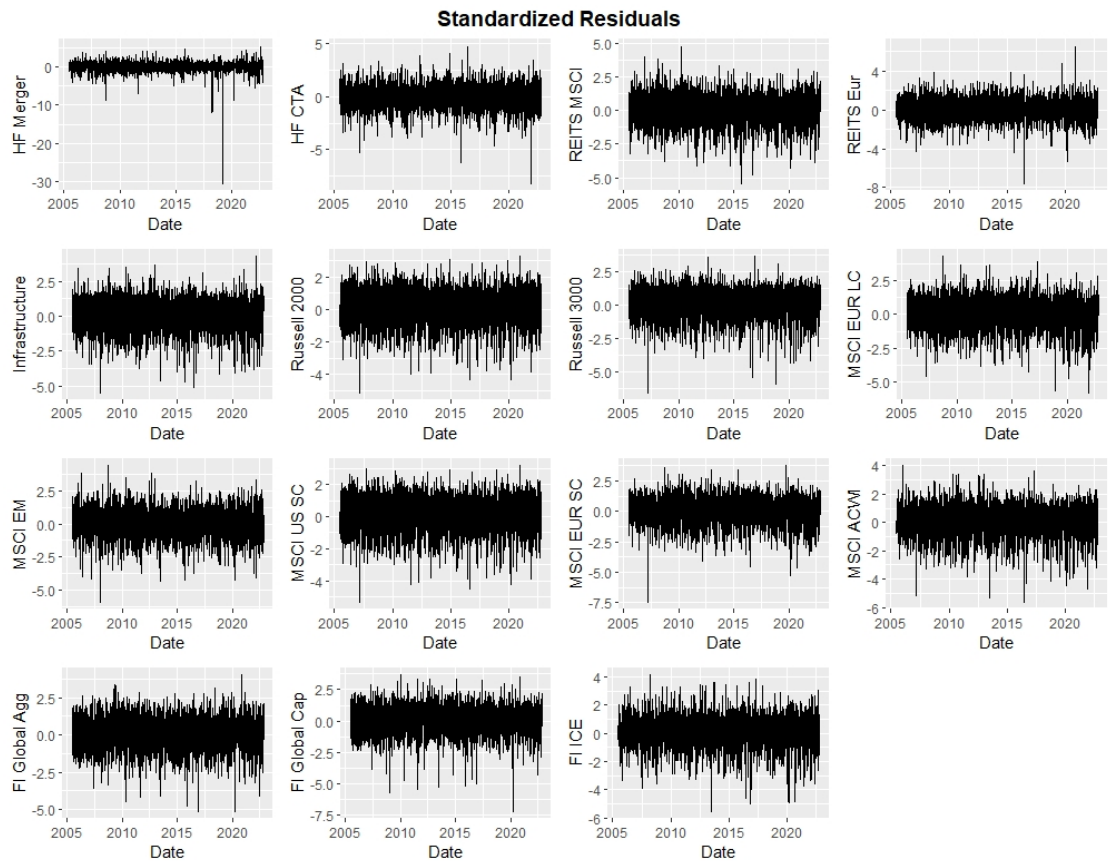
Figure 12.34: Standardised residual time series after fitting the univariate ARMA-GARCH models specified in Table 12.7.

12.1.2 Vine copula based analysis

Once that we have remove the serial dependence we can analyse the cross sectional dependence with the copula analysis. We start transforming the data to u-scale.
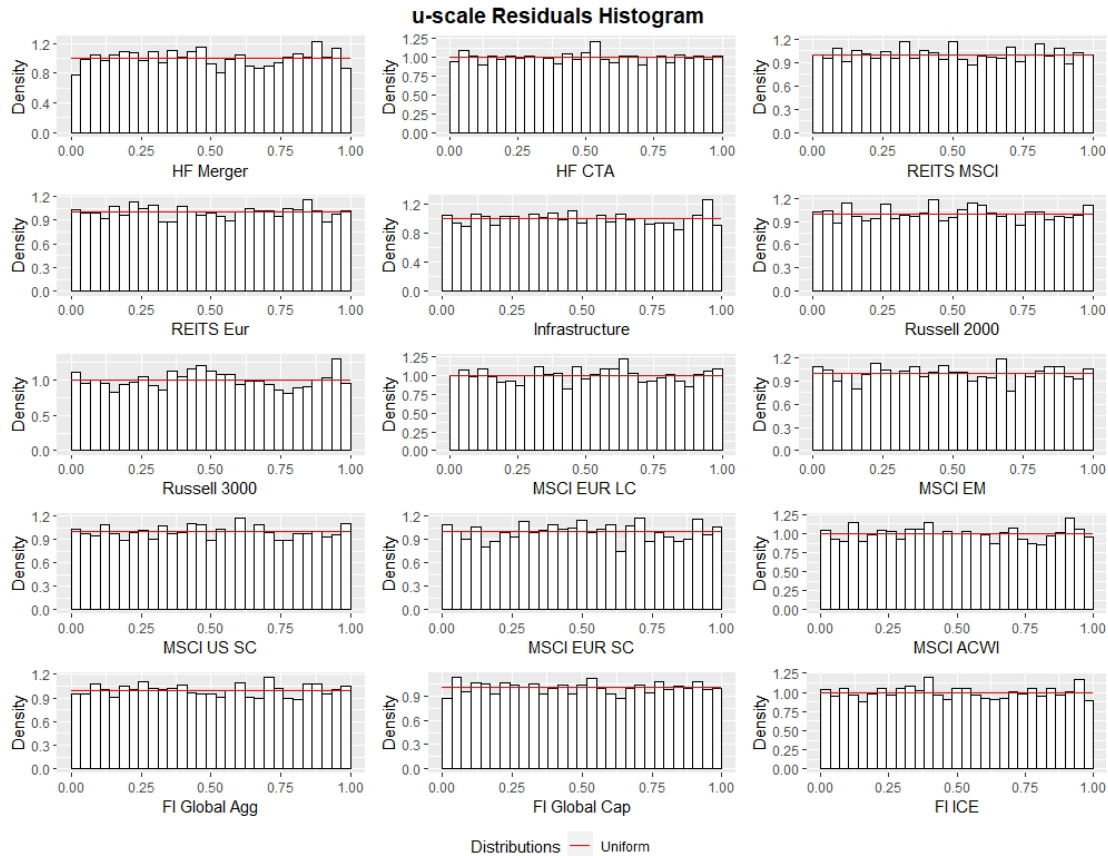
Figure 12.35: Marginal histograms of the pseudo copula data of the standardised residuals after fitting the models specified in Table 12.7 to the returns of the indices in Table 11.2.

Once that we have transformed the data to the u-scale we can analyse the cross sectional dependence. First, we check the dependence between the variables, by showing the marginally normalised contours for each pair of variables in the lower triangular panel of Figure 12.36, in the diagonal we have the histograms of the residuals on the u-scale, and finally, in the upper triangular panels we give the pairwise scatter plots of the pseudo copula data with the corresponding empirical Kendall's τ.

Figure 12.36 shows different type of dependence between the variables. Some of them show a high dependence, like Russell 2000 and MSCI US SC, or Russell 3000 and MSCI US SC. Most of the contours look symmetrical, could be modelled by Gaussian copula (MSCI EM and Infrastructure); or Student's t copula, if they show symmetric tail dependence and diamond shape, for example, MSCI ACWI and Russell 3000.

In financial series usually left tails are bigger than right ones, also we should expect a higher

lower tail dependence, because people react more to crisis and bearish markets than to bullish markets. In our case, most of the pairs have not lower tail dependence, this is cause that we are considering large indices based on many assets. Therefore, they are more difficult to be influenced by the markets. In addition, some indices consist on totally different types of assets, so we could expect them to be not highly dependent. This shows that lower tail dependence structures are also plausible for this data set.

Before analysing the copula model, we select the best vine copula structure. In the following table we compare the AIC, BIC and log likelihood for the R-vine, C-vine and D-vine structures:

Type	AIC	BIC	LogLik	Number of parameters
R-vine	-77169.70	-76175.46	38740.85	156
C-vine	-76618.58	-75592.47	38470.29	161
D-vine	-76713.76	-75662.16	38521.88	165

Table 12.9: AIC, BIC, LogLik and number of parameters of R-, C- and D-vine copulas estimated on transformed standardised residuals after fitting models of Table 12.7.

Table 12.9 shows that the best fitting vine structure is the R-vine based on AIC, BIC or log-likelihood of the three models fitted. Anyway, we apply the Vuong test to check that R-vine is significantly better than C- and D-vine structures. Since different number of parameters have been estimated in the three vine copula structures, we apply the Voung test with Akaike and Schwarz corrections. Results are in Table 12.10.

Test	Akaike Statistic	Akaike p-value	Schwarz Statistic	Schwarz p-value
R-vine vs C-vine	7.044	1.87e-12	7.46	8.94e-14
R-vine vs D-vine	6.299	3.00e-10	7.14	9.53e-13
C-vine vs D-vine	-1.194	0.232	-0.93	0.354

Table 12.10: P-value and statistics of Vuong test with Akaike and Schwarz corrections for comparing R-, C- and D-vine copulas fitted in Table 12.9.

Table 12.10 confirms R-vine as the best vine copula structure. The p-values obtained are lower than 0.05 when comparing R-vine against C- or D-vine.

We now investigate the fitted R-vine used in Table 12.9 in more detail.

The first R-vine tree shows that MSCI ACWI and MSCI US SC are the most important indices, by showing high correlations with several indices. Also, we remark that, apart from Fixed Income and Equity, the indices related to the same department are not close in R-vine tree, as we might have expected.

From Table 12.11 we obtain that most of the fitted copula distributions are symmetric with both tail dependence different from zero, the case of Student's t and BB1 copula. Also, we remark that there are more lower tail dependence estimated than upper tail dependence.

Now we estimate the value at risk and expected shortfall in a 2 years rolling window setup for the portfolios described in Table 12.4 in Section 12.1.3.

	Family	Usage	Tail dependence
1	Gaussian	7	None
2	t-Student	44	Both
3	Clayton (rot 0)	4	Lower
4	Joe (rot 0)	0	Upper
5	Frank	8	None
6	Gumbel (rot 0)	1	Upper
7	BB1	13	Both
8	BB7	1	Both
9	BB8	4	None
10	Independent	10	None
11	Clayton (rot 180)	2	Upper
12	Joe (rot 180)	2	Lower
13	Gumbel (rot 180)	4	Lower

Table 12.11: Usage and type of tail dependence of the different family distribution implemented in the R-vine copula fitted in Table 12.9.
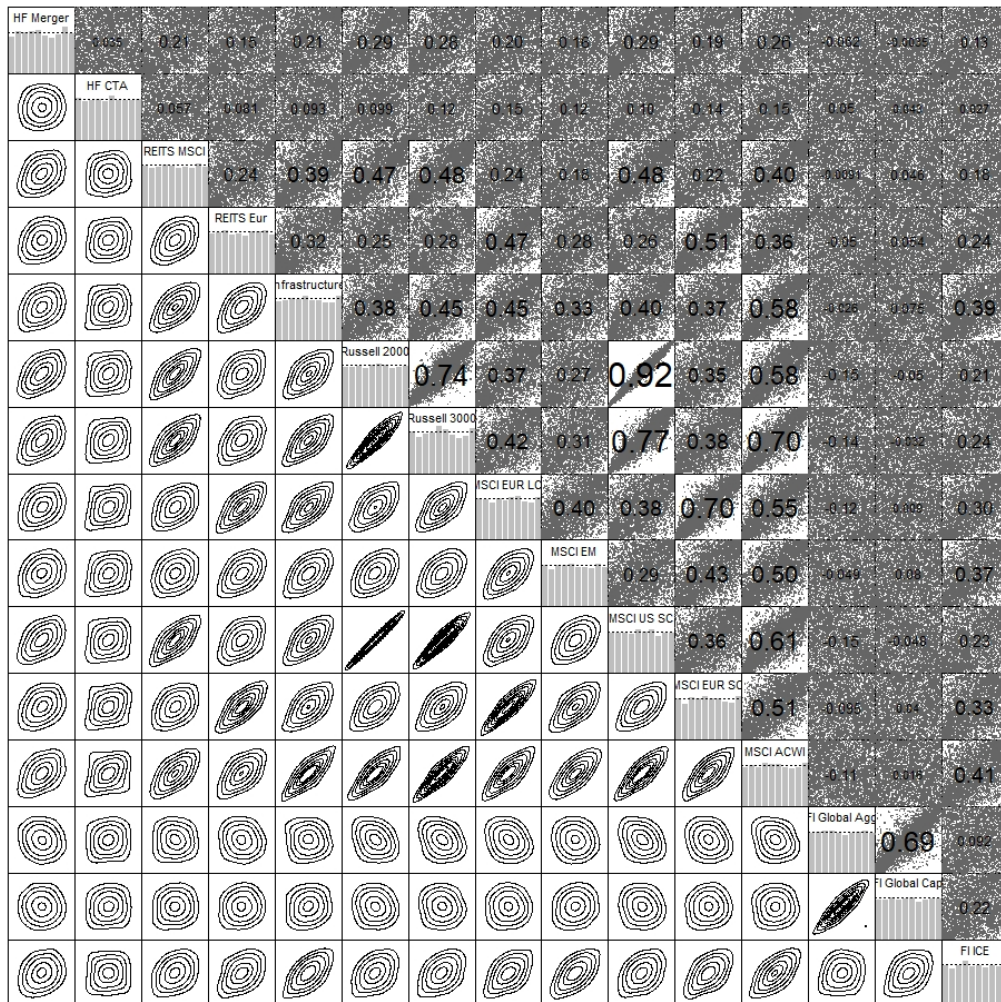
Figure 12.36: Marginally normalised contours for each pair of variables in the lower triangular. Histograms of the standardised residuals after fitting models given in Table 12.7 on the u-scale in the diagonal. Pairwise scatter plots of the pseudo copula data with the corresponding empirical Kendall's τ in the upper triangular.
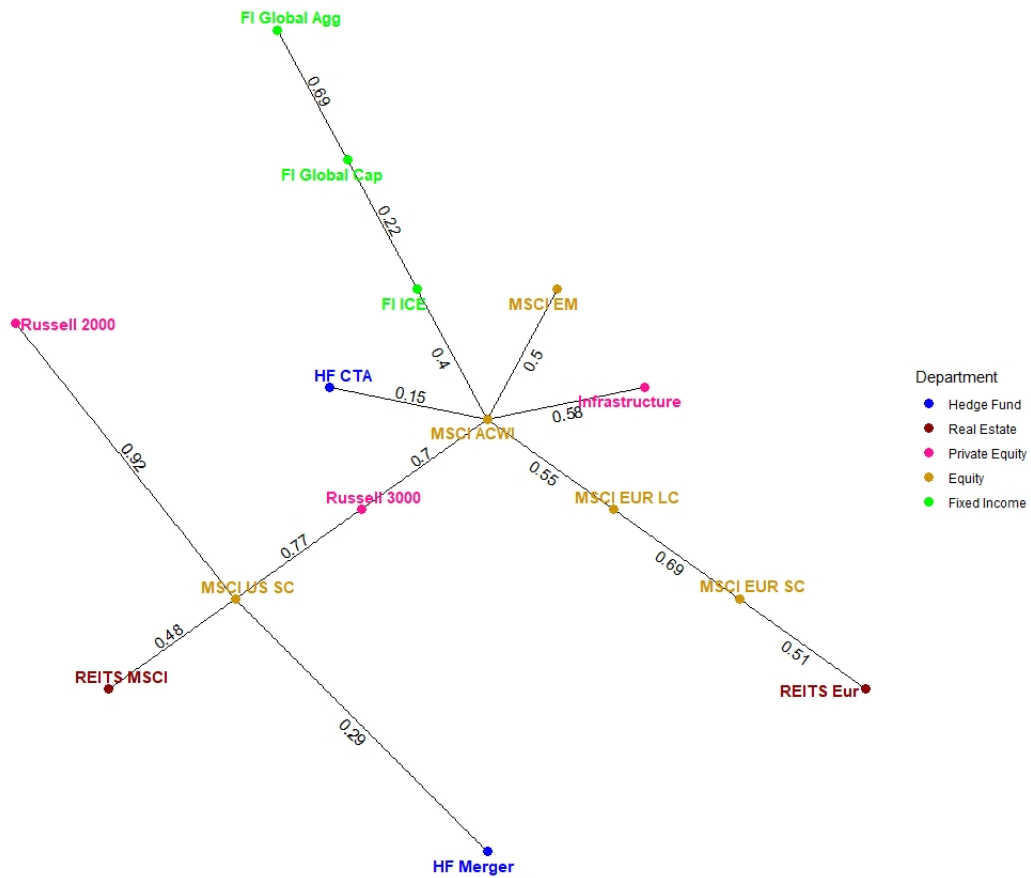
Figure 12.37: First R-vine tree of the R-vine copula fitted in Table 12.9.

12.1.3 Risk measures forecast

In this section we finally forecast the different risk measures. For the final implementation of the model we use all the model characteristic analysed in the previous study in Sections 12.1.1-12.1.2. Then, we sample portfolios values from estimated variables of ARMA-GARCH marginal models displayed in Table 12.7 and the fitted R-vine modelling the dependence for each window. For the estimation we use a 2 year rolling window (500 observations) for the marginal models and the copula model, obtaining around 15 years estimation. We refit the model parameters every 2 months (40 observations), for marginal and vine copula models. Also, we use a daily portfolio sample size of 5000 for the quantile VaR and ES estimations.

Figure 12.38: Estimated value at risk (VaR) at levels 0.05 and 0.01 of equal weighted, BVK and market capitalisation portfolios in Table 12.4 using the vine copula model specified in Section 6 together with ARMA-GARCH margins specified in Table 12.7.

The risk measures estimated for the three portfolios show good results. In Figures 12.38 and 12.39 we observe how VaR and ES move similar to the observed returns time series. Also, the estimated confidence interval in Figure 12.40 adjust the volatility to the observed one.
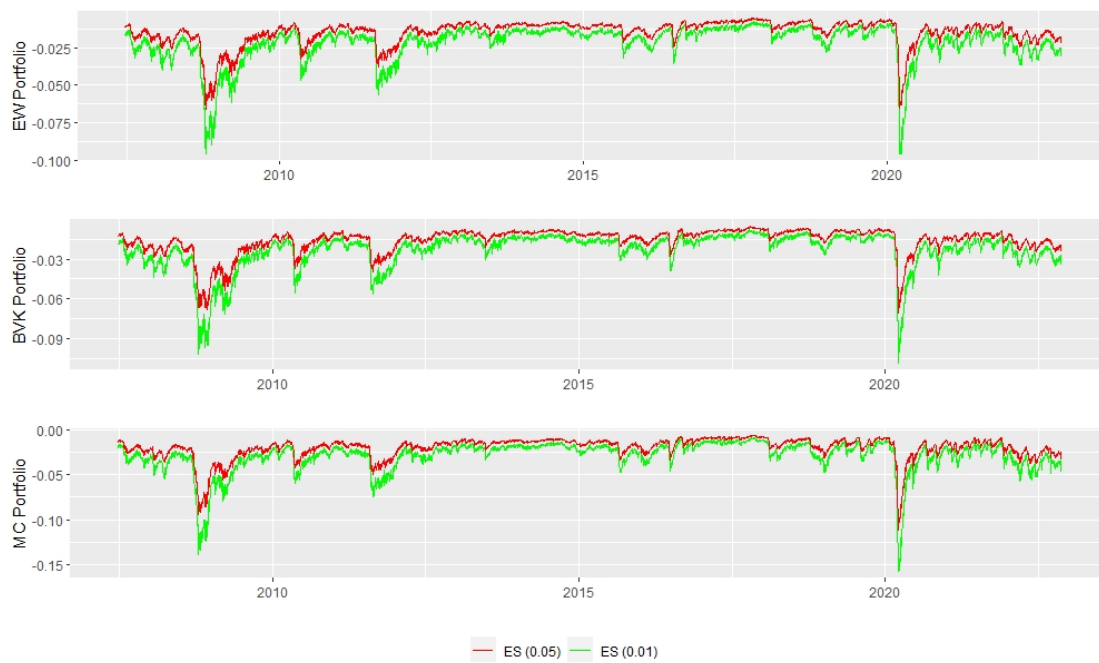
Figure 12.39: Estimated expected shortfalls (ES) at levels 0.05 and 0.01 of equal weighted, BVK and market capitalisation portfolios in Table 12.4 using the vine copula model specified in Section 6 together with ARMA-GARCH margins specified in Table 12.7.
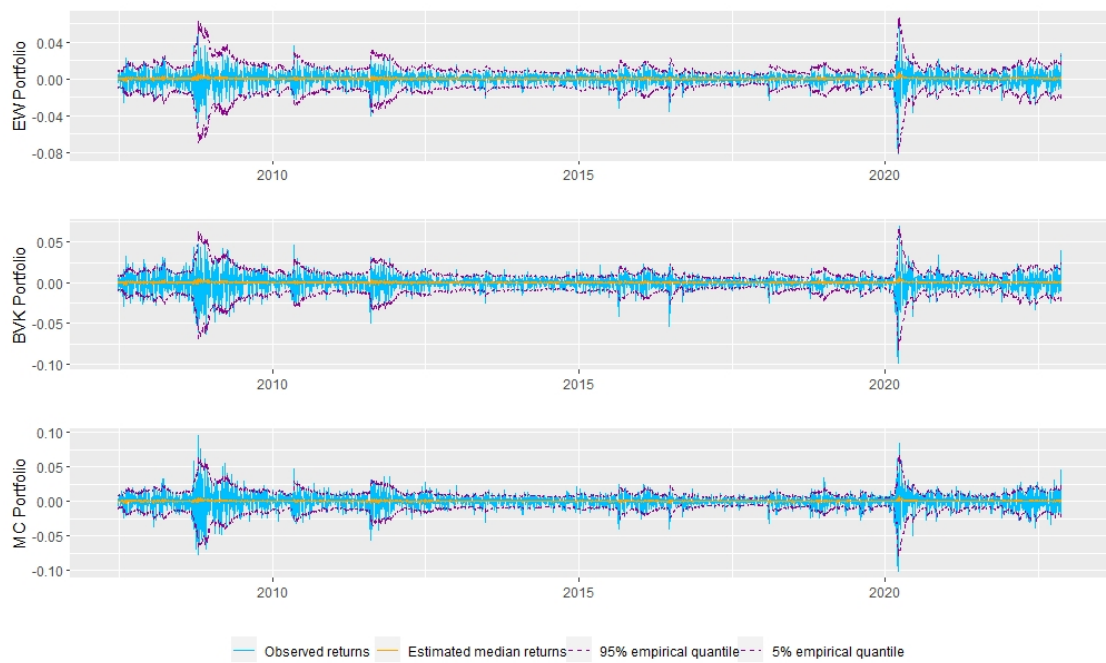
Figure 12.40: Observed, estimated and 90% confidence level returns of equal weighted, BVK and market capitalisation portfolio in Table 12.4 using the vine copula model specified in Section 6 together with ARMA-GARCH margins specified in Table 12.7.

12.2 Stationary vine copula models

With the stationary vine copula we model the serial and cross sectional dependence at the same time. In this section we apply the stationary vine copulas to capture the whole dependence structure. The model is applied for the three original portfolios BVK port, EW port and MC port given in Table 12.4.

First of all, we transform the original returns data to the u-scale for each time series. To select the best distribution for each return time series we compare Skewed-Student, Generalised Hyperbolic, Normal Inverse Gaussian and Johnson´s SU with QQ-plots for each time series.
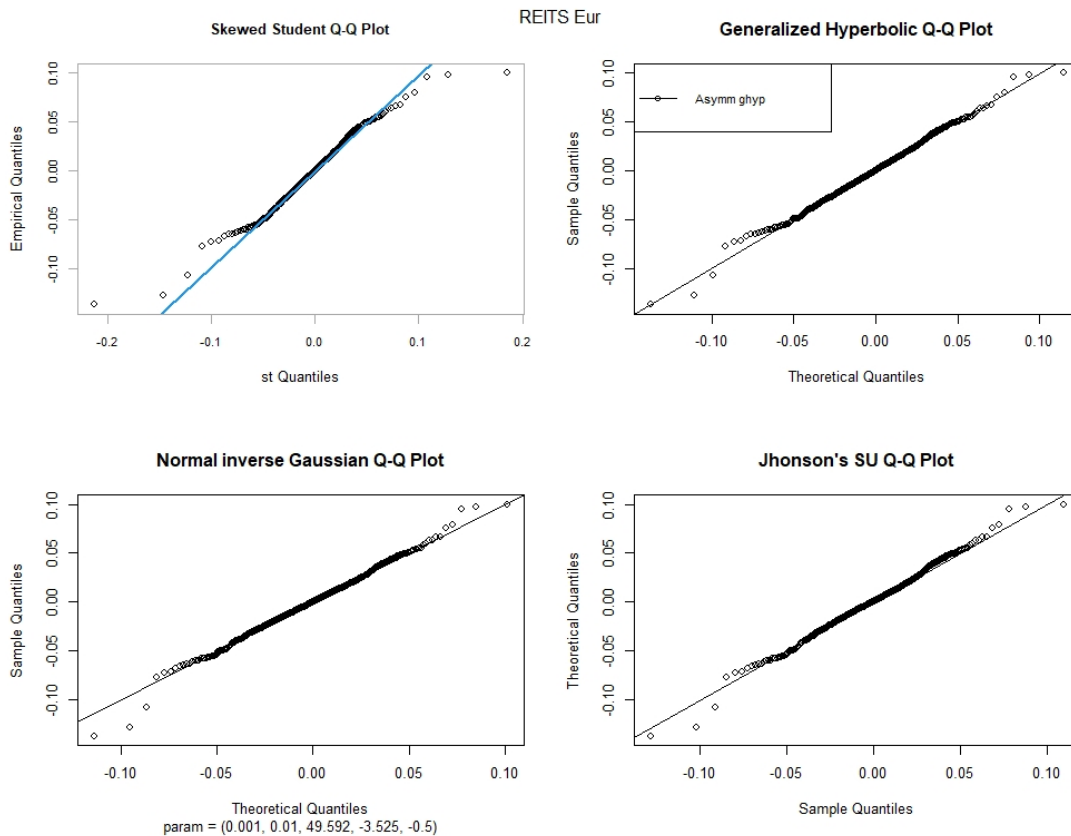
Figure 12.41: REITS Eur returns QQ-plots with sstd, ghyp, nig and jsu distributions.

Figure 12.41 shows the QQ-plots for REITS Eur. Generalised Hyperbolic and Johnson´s SU show similar results, being the best choices, we select the Generalised Hyperbolic in this case. For the remaining time series, QQ-plots are shown in Appendix D.

The chosen returns distribution for each time series are summarised in Table 12.12.

We check the distribution selection with the u-scale histograms shown in Figure 12.42.

Once we have the data in u-scale we select the best cross sectional structure of the vine copula. Thus we distinguish:

	Index	Return distribution
1	HF Merger	ghyp
2	HF CTA	sstd
3	REITS MSCI	nig
4	REITS Eur	ghyp
5	Infrastructure	jsu
6	Russell 2000	ghyp
7	Russell 3000	ghyp
8	MSCI EUR LC	ghyp
9	MSCI EM	nig
10	MSCI US SC	ghyp
11	MSCI EUR SC	ghyp
12	MSCI ACWI	jsu
13	FI Global Agg	ghyp
14	FI Global Cap	ghyp
15	FI ICE	ghyp

Table 12.12: Chosen return distributions for each index.

– $S - vine$. Arbitrary R-vine structure is allowed for the cross sectional model.

– $D - vine$. The cross sectional structure is a D-vine. Each cross sectional structure is joined to the following one by the end of the D-vine. Finally we obtain a long D-vine over the time as shown in Figure 7.5.

– $M - vine$. The cross sectional structure is a D-vine too. In this case, D-vine are joined together by the initial point of the D-vines. We don´t obtain a D-vine structure over all. Check Figure 7.6.

Table 12.13 give the AIC, BIC and log-likelihood of each model considered.

Type	AIC	BIC	LogLik	Number of parameters
S-vine	-94438.68	-91435.76	47690.34	472
D-vine	-93658.36	-90508.78	47323.18	504
M-vine	-93879.68	-90781.12	47425.84	479

Table 12.13: AIC, BIC and LogLik of S-, D- and M-vine copulas estimated on transformed data with return distributions of Table 12.12.

From Table 12.13 we obtain the best model structure is the S-vine for the three measures considered. We continue the analysis showing the characteristics of the S-vine copula estimated in Table 12.13.

Figure 12.43 shows the dependence structure at two contiguous points in time. The first time point is represented with blue edges, while the second time moment is marked with red labels. The same R-vine copula structure is model for each cross sectional time point. Similar to Figure 12.37 we have fixed income indices together; while hedge funds, real estate, private equity and
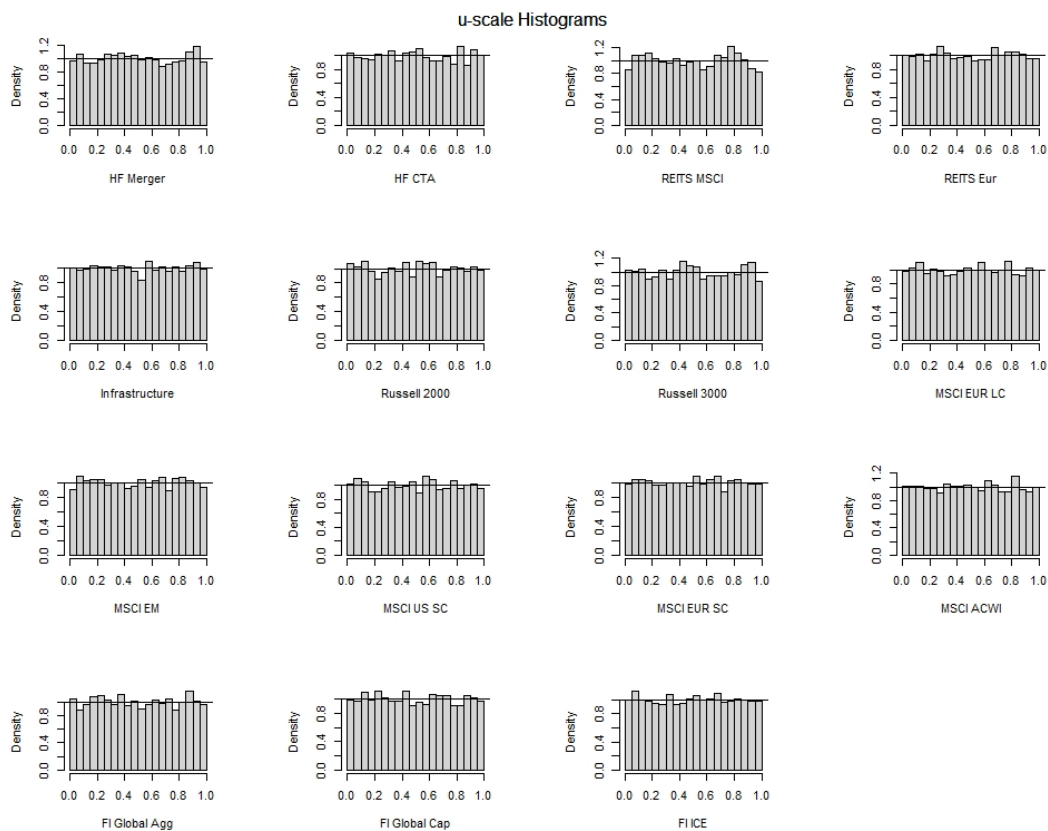
Figure 12.42: Histograms of the u-scale returns after transforming with return distributions of Table 12.12.

equity are not together in the S-vine tree, similar to the marginal and copula method show in Figure 12.37. Also we should remark the serial dependence is model with a bivariate copula between Russell 3000 and MSCI EM represented with a green edge between both nodes.

Table 12.14 shows the most used family is the t-Student which is symmetric, having lower and upper tail dependence. Similar number of lower tail dependence and upper tail dependence are model.
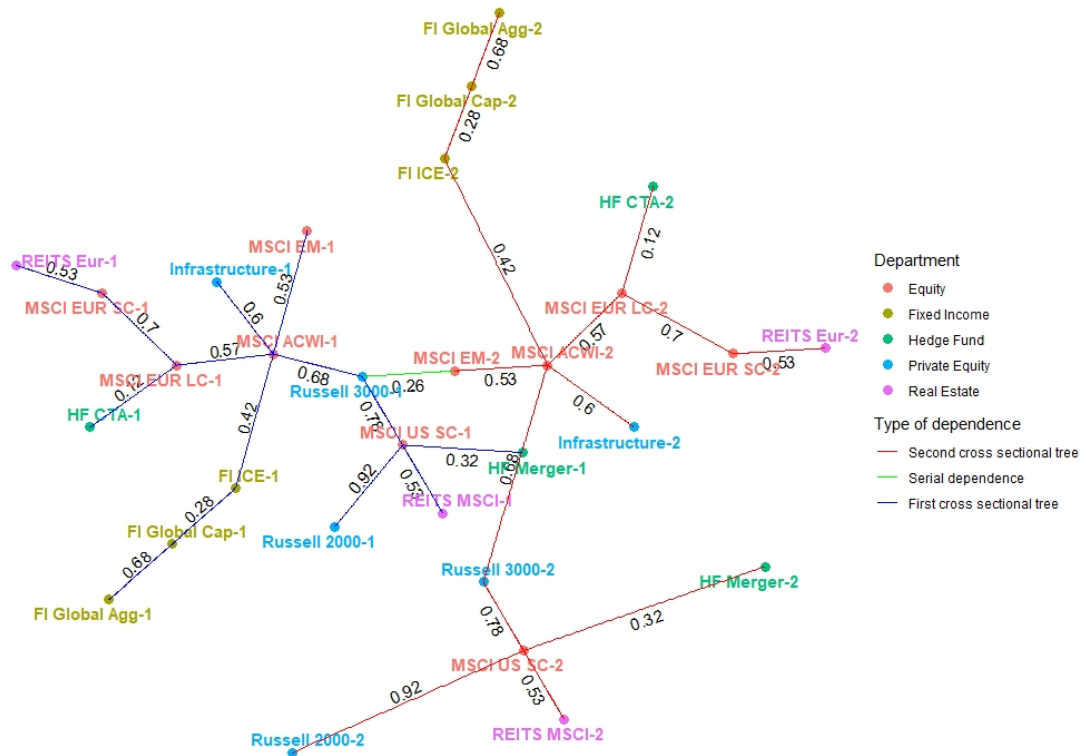
Figure 12.43: First S-vine tree of the S-vine copula estimated in Table 12.13.

	Family	Usage	Tail dependence
1	Gaussian	9	None
2	t-Student	160	Both
3	Clayton (rot 0)	9	Lower
4	Joe (rot 0)	11	Upper
5	Frank	19	None
6	Gumbel (rot 0)	5	Upper
7	BB1	4	Both
8	BB7	10	Both
9	BB8	15	None
10	Independent	48	None
11	Clayton (rot 180)	7	Upper
12	Joe (rot 180)	8	Lower
13	Gumbel (rot 180)	5	Lower

Table 12.14: Usage and type of tail dependence of the different family distribution implemented in the S-vine copula fitted in Table 12.13.

12.2.1 Risk measures forecast

In this section we present the results obtained when applying the stationary vine copula models to the three portfolios in Table 12.4. For portfolio sampling and risk measures forecast we use the same specifications as in Section 12.1.3. We apply a 2 year rolling window (500 observation) for the S-vine copula model estimation, refitting the model parameters of the S-vine every 40 observations (2 months). For each date we sample 5000 values of each of the three portfolios.

Figure 12.44: Estimated value at risk (VaR) at 0.05 and 0.01 levels of equal weighted, BVK and market capitalisation portfolios in Table 12.4 using S-vine copula models.

The results estimated with the stationary vine copula method show the importance of the window size selected. In Figure 12.44 we observe the estimated VaR take coherent values but does not follow the observed returns like ARMA-GARCH vine copula approach shown in Figure 12.38. The window size affects the estimation of the marginal parameter distribution, which are used to move from the estimated u-scale copula to the return scale. With respect to the expected shortfall, Figure 12.45 shows big outliers at the same time point for the three portfolios. This extreme values are observed during a period of low volatility, but the effect of COVID-19 in 2020 with high volatility is still reflected in the estimations of the marginal parameters. Finally, Figure 12.46 show similar results as the VaR in Figure 12.44, plotting smoother estimations than ARMA-GARCH vine copula approach, because of the effect of marginal distributions.

Figure 12.45: Estimated expected shortfall (ES) at 0.05 and 0.01 levels of equal weighted, BVK and market capitalisation portfolios in Table 12.4 using S-vine copula models.
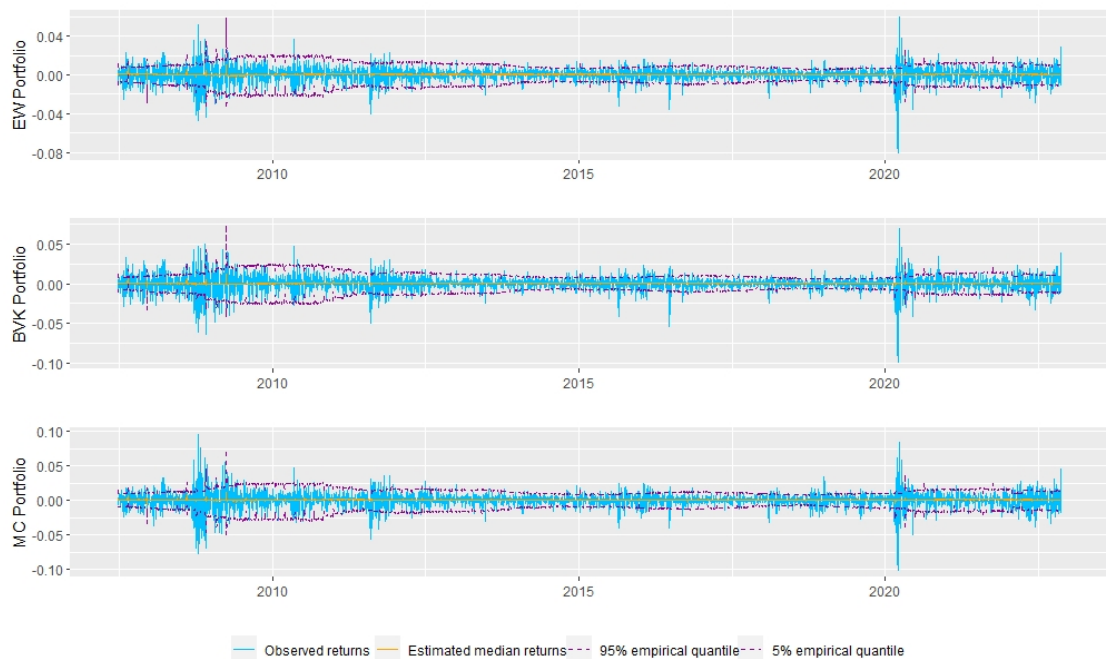
Figure 12.46: Observed, estimated and 90% confidence level returns of equal weighted, BVK and market capitalisation portfolios in Table 12.4 using S-vine copula models.

12.3 Models comparison

In this section we compare the result obtained for the complete portfolio using the ARMA-GARCH model join with the vine copula and the stationary vine copula approach. We first compare the estimated VaR and ES in Figures 12.47 and 12.48. The results obtained for the VaR show similar results, where the stationary vine is smoother than ARMA-GARCH vine copula. In case of the ES the effect of the extreme values estimated with stationary vine copula makes difficult to compare. Apart of this outliers, both method show values near to zero.

Figure 12.47: Forecasted VaR at 0.05 and 0.01 levels of equal weighted, BVK and market capitalisation portfolios described in Table 12.4 with ARMA-GARCH vine copula approach and S-vine copula approach.

Figures 12.49 and 12.50 show the estimated VaR level, $\widehat{\alpha}$, in a 100 rolling window with the corresponding confidence interval, using (9.53). Figures 12.49 and 12.50 show equivalent results for the three portfolios. The ARMA-GARCH vine copula shows smoother results around the α-level. In general, both methods underestimate the VaR during high volatility periods and overestimate it in low volatility moments. This effect is easily observable in Figure 12.50 during the 2008 and COVID-19 crisis where the $\widehat{\alpha}$ values are much higher than the theoretical one, 0.01. Also, since stationary copula vines are smoother we expect them to be more underestimated than ARMA-GARCH with R-vine copula during high volatility periods, because the proportion of returns exceeding the estimated VaR is higher.

Figure 12.48: Forecasted ES at 0.05 and 0.01 levels of equal weighted, BVK and market capital-isation portfolio described in Table 12.4 with ARMA-GARCH vine copula approach and S-vine copula approach.

In Table 12.15 we resume the statistical measures introduce in Section 9 during the whole period, from 06/2007 to 11/2022. The table shows both methods underestimate the VaR level, which is risky during high volatility market periods. However, despite of underestimating, ARMA-GARCH vine copula approach estimate the VaR better, with a higher percentage of $\widehat{\alpha}$ in the corresponding confidence interval. With respect to the expected shortfall, we obtain similar results for both methods, having values of β near to zero for all the portfolios. Since expected shortfall depends on the value at risk is more difficult to obtain good conclusions. Since bold values represent best value for each measure, we obtain better results for ARMA-GARCH vine copula approaches.

Figure 12.49: In the graphs above we show the estimated $\widehat{\alpha}$ at 0.05 VaR level using ARMA-GARCH vine copula and S-vine copula in a 100 observation rolling window of the equal weighted, BVK and market capitalisation portfolios described in Table 12.4. The rows show the equal weighted, BVK and market capitalisation portfolio returns specified in Table 12.4.

We also assess the used models with the backtests described in Section 9.1 and 9.2. Table 12.16 presents the p-values for the corresponding hypothesis tests. Bold values correspond to accepted H_0 null hypothesis. As we observe more S-vine test accept the hypothesis null of well risk measure assessment, specially for expected shortfall. This could be because the ES depends on VaR.
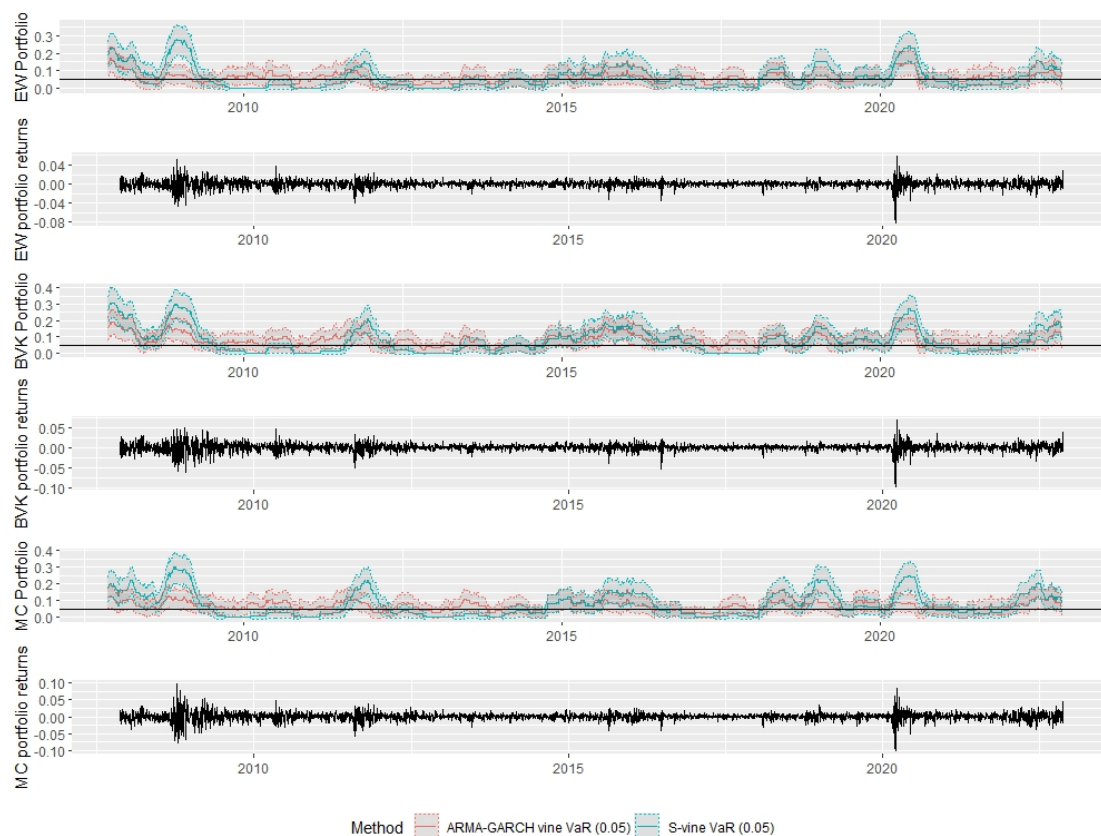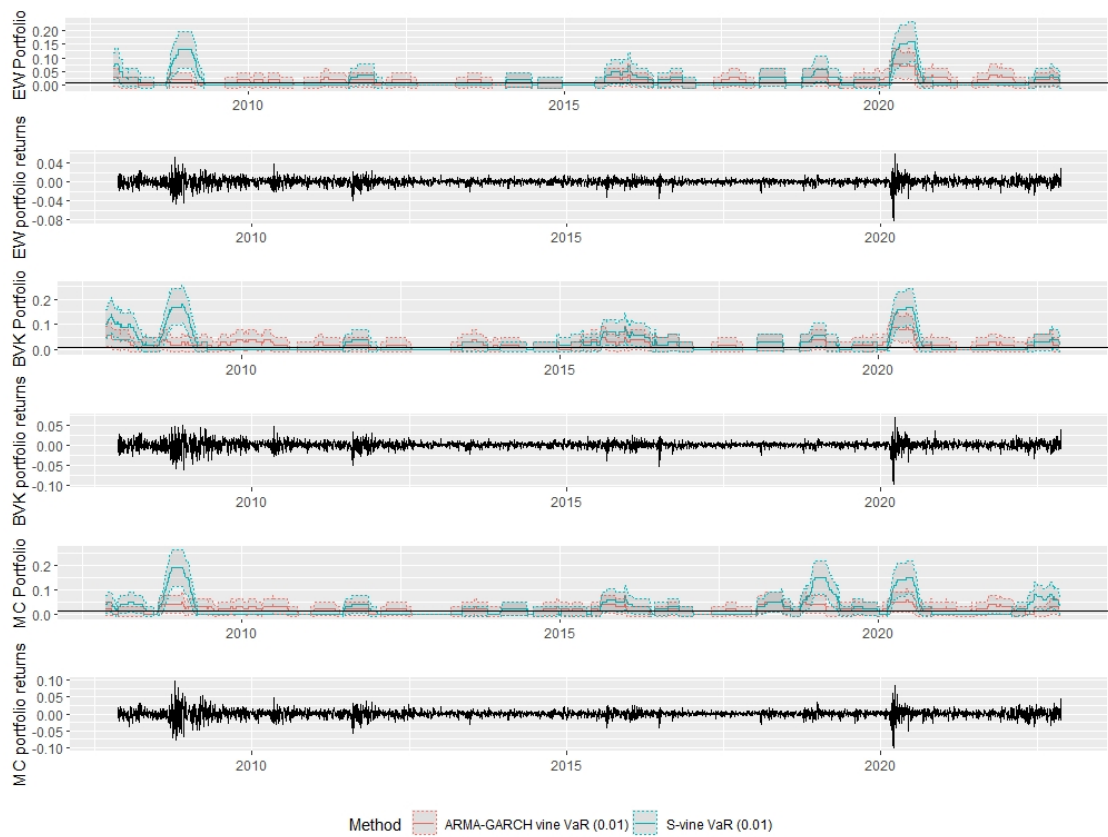
Figure 12.50: In the graphs above we show the estimated $\widehat{\alpha}$ at 0.01 VaR level using ARMA-GARCH vine copula and S-vine copula in a 100 observation rolling window of the equal weighted, BVK and market capitalisation portfolios described in Table 12.4. The rows show the equal weighted, BVK and market capitalisation portfolio returns.

Portfolio	Level α	Measure	ARMA-GARCH/ vine copula	S-vine
Equal weighted	0.05	$\hat{\alpha}$	**0.06145**	0.06588
		β	-0.000769	**0.000309**
		% of $\hat{\alpha}$ in $CI_{0.05}(\alpha)$	**90.05**	42.40
	0.01	$\hat{\alpha}$	**0.01562**	0.018229
		β	**-0.00166**	-0.00226
		% of $\hat{\alpha}$ in $CI_{0.05}(\alpha)$	**72.19**	37.35
BVK	0.05	$\hat{\alpha}$	0.08177	**0.08020**
		β	**-0.00112**	-0.00116
		% of $\hat{\alpha}$ in $CI_{0.05}(\alpha)$	**83.47**	44.35
	0.01	$\hat{\alpha}$	**0.021093**	0.02526
		β	**-0.0024**	-0.00297
		% of $\hat{\alpha}$ in $CI_{0.05}(\alpha)$	**82.16**	38.77
Market capitalisation	0.05	$\hat{\alpha}$	**0.06953**	0.07786
		β	-0.00116	**-0.000826**
		% of $\hat{\alpha}$ in $CI_{0.05}(\alpha)$	**95.42**	40.45
	0.01	$\hat{\alpha}$	**0.019010**	0.02578
		β	**-0.00112**	-0.00169
		% of $\hat{\alpha}$ in $CI_{0.05}(\alpha)$	**86.52**	42.86

Table 12.15: Estimated $\hat{\alpha}$ and β estimations at 0.05 and 0.01 levels between 13/11/2007 and 11/11/2022 using ARMA-GARCH vine copula, and S-vine copula on equal weighted, BVK and market capitalisation portfolios described in Table 13.18. Also, the percentage of $\hat{\alpha} \in CI_{0.05}(\alpha)$ estimated in a 100 rolling window in Figures 12.49 and 12.50 for $\alpha = 0.05$ and $\alpha = 0.01$ respectively. Bold numbers represent the best value of each row comparing ARMA-GARCH vine copulas and S-vines.

Since the estimated period is very long we divide it in 7 different periods of equal length, around 2 years and 2 months. The results are shown in Figures 12.51 and 12.52 for 0.05 and 0.01 α-values for the risk measures calculations. As we define in Section 9 the backtests LR_{uc} and LR_{cc} correspond to VaR and ER and CC to ES or both. In Figures 12.51 and 12.52 the red circles correspond to the portfolios obtained with ARMA-GARCH join with the vine copula method, while the green ones are the S-vine portfolio. Since the circles are the p-values obtained in the backtest, considering a α 0.05 test level for all hypothesis test we accept the null hypothesis of good risk measures assessment for those which are above the 0.05 line. From the figures we remark that the models estimated in the period 06/2018-09/2020 are not good, this period is characterised because of the start of the COVID-19, when we pass from low volatility period to a high volatility one. In general, most of the models used estimate well the risk measures, there is not a big different between the ARMA-GARCH vine copula and S-vine models. The 7 period analysis also shows that the performance of the backtests depends on the time windows selected.

Portfolio	Level α	Backtest	ARMA-GARCH/ vine copula	S-vine
Equal weighted	0.05	LR_{uc}	0.001	4.7e-04
		LR_{cc}	4.0e-06	9.2e-09
		ER	0.031	**0.879**
		CC	0.011	**0.321**
	0.01	LR_{uc}	0.036	0.025
		LR_{cc}	0.012	2.7e-04
		ER	0.008	**0.840**
		CC	**0.053**	**0.624**
BVK	0.05	LR_{uc}	7.7e-16	3.6e-15
		LR_{cc}	0.00	0.003
		ER	0.02	**0.424**
		CC	2.6e-07	**0.445**
	0.01	LR_{uc}	8.1e-10	3.8e-09
		LR_{cc}	6.3e-09	5.6e-12
		ER	0.002	**0.832**
		CC	0.001	**0.790**
Market capitalisation	0.05	LR_{uc}	1.4e-09	1.2e-11
		LR_{cc}	6.5e-12	0.000
		ER	0.003	**0.820**
		CC	4.3e-06	**0.663**
	0.01	LR_{uc}	1.5e-07	1.1e-06
		LR_{cc}	2.1e-07	2.5e-10
		ER	0.033	**0.97**
		CC	0.0042	**0.751**

Table 12.16: P-values for LR unconditional and conditional test, exceedance residual test and conditional calibration test for VaR and ES at 0.05 and 0.01 levels estimated using ARMA-GARCH vine copula and stationary vine copula for equal weighted, BVK and market capitalisation portfolios showed in Table 12.4.

Figure 12.51: P-values corresponding to the backtests developed in Sections 9.1 and 9.2 for VaR and ES at 0.05 level estimated using ARMA-GARCH vine copula and S-vine copula for portfolios described in Table 12.4 during the periods described in the graph.

Figure 12.52: P-values corresponding to the backtests developed in Sections 9.1 and 9.2 for VaR and ES at 0.01 level estimated using ARMA-GARCH vine copula and S-vine copula for portfolios described in Table 12.4 during the periods described in the graph.

13 Reduced portfolio

In this section we perform an equivalent study than the Section 12 but now including the Fleishman's transformation method. Since Fleishman's transformation method does not converge in some rolling window when applying it to the whole portfolio, we consider a reduced portfolio. Then, we follow a similar study, beginning with marginal and copula method in Section 13.1, we continue with the study and forecast of stationary vine copula method in Section 13.2, in Section 13.3 we introduce and model the Fleishman's transformation and, finally, we compare the obtained results in Section 13.4.

The reduced portfolio used throughout this section have been selected from indices in Table 11.2. We try to maintain at least one representative of each sector in Table 11.2. As in the Section 12, we build an equal weighted portfolio, other based on BVK strategy and a market capitalisation portfolio. The indices selected for the reduced portfolio join with the corresponding weights are shown in Tables 13.17 and 13.18.

	Department	Index	Currency	Market Capitalisation (Millions)
1	Hedge Fund	HF CTA	USD	10000
2	Real Estate	REITS MSCI	USD	984849.54
3	Real Estate	REITS Eur	EUR	111404.00
4	Private Equity	Russell 2000	USD	2806774.78
5	Private Equity	Russell 3000	USD	40787893.59
6	Equity	MSCI US SC	USD	4729267.87
7	Fixed Income	FI ICE	USD	1549652.87

Table 13.17: Department, indices, currency and market capitalisation of the indices used in the reduced portfolio study.

	Index	BVK Weight	Market Cap Weight	Equal Weight
1	HF CTA	0.06838039	0.000196156	0.1428571
2	REITS MSCI	0.15913512	0.019318411	0.1428571
3	REITS Eur	0.15913512	0.002185256	0.1428571
4	Russell 2000	0.08174476	0.055056560	0.1428571
5	Russell 3000	0.08174476	0.800078844	0.1428571
6	MSCI US SC	0.21035143	0.092767408	0.1428571
7	FI ICE	0.23950843	0.030397365	0.1428571

Table 13.18: BVK, Market cap and Equal weighted portfolios considered in the study.

Figure 13.53 shows similar behave of the three portfolios, including GARCH effects and similar periods of high volatility around 2008 and 2020.

13.1 Marginal and copula models

In this section we select the best ARMA-GARCH and vine copula model for fitting the data corresponding to Table 13.18. We begin with the marginal model selection in Section 13.1.1,
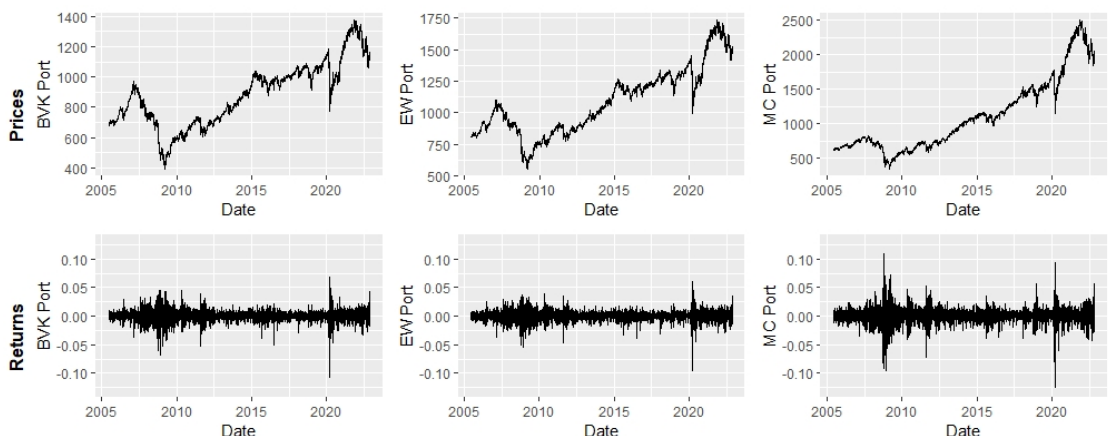
Figure 13.53: Prices and returns of BVK, equal weighted and market cap described in Table 13.18.

then we choose the best copula model in Section 13.1.2 and we finally forecast the VaR and ES in Section 13.1.3.

13.1.1 Marginal model analysis

In this section we fit the best ARMA-GARCH order to each index in Table 13.18. Since the analysis is individual for each index we obtain the same results as in Section 12.1.1. Then, taken the chosen model from Table 12.7 we fit the models of Table 13.19.

	Index	Innovation distribution	Model
1	HF CTA	ghyp	ARMA(1,1)-GARCH(1,1)
2	REITS MSCI	sstd	ARMA(1,1)-GARCH(1,1)
3	REITS Eur	ghyp	ARMA(0,1)-GARCH(1,1)
4	Russell 2000	ghyp	ARMA(0,1)-GARCH(1,1)
5	Russell 3000	ghyp	ARMA(1,0)-GARCH(1,1)
6	MSCI US SC	ghyp	ARMA(0,1)-GARCH(1,1)
7	FI ICE	nig	ARMA(1,1)-GARCH(2,1)

Table 13.19: Selected ARMA and GARCH orders and innovation distributions for each index in Table 13.18.

13.1.2 Vine copula based analysis

In this section we select the best vine copula model for the standardised residuals after fitting model describe in Table 13.19.

In Figure 13.54 we have different kind of dependence between the variables. Some of them are really high correlated, like Russell 2000, Russell 3000 and MSCI US. However, we also observe that the FI ICE is almost uncorrelated with the rest of the indices.

96

To select the best vine structure we compare the different criteria selection in Table 13.20. The R-vine fitted model has the best statistics values. However, we perform the Vuong test from Section 6.3 to check it.

Type	AIC	BIC	LogLik	Number of parameters
R-vine	-33210.51	-32968.33	16643.26	38
C-vine	-33192.56	-32937.62	16636.28	40
D-vine	-33178.15	-32929.59	16628.07	39

Table 13.20: AIC, BIC, LogLik and number of parameters of Regular, C- and D-vine copulas estimated on transformed standardised residuals after fitting models of Table 13.19.

Test	Akaike Statistic	Akaike p-value	Schwarz Statistic	Schwarz p-value
R-vine vs C-vine	1.33	0.18	2.27	0.02
R-vine vs D-vine	1.20	0.23	1.43	0.15
C-vine vs D-vine	0.54	0.59	0.30	0.76

Table 13.21: P-value and statistics of Vuong test with Akaike and Schwarz corrections for comparing R-, C- and D-vine copulas fitted in Table 13.20.

Table 13.21 shows that using Akaike criteria, for a $p - value = 0.05$ the three models are equivalent, since we have $Akaike\, p - values > 0.05$ for the three tests. Therefore, the null hypotheses is accepted. Considering the Schwarz criteria, we have similar results as Akaike criteria, the only difference is comparing R- and C-vine, where the $Schwarz\, p - value = 0.023 < 0.05$, indicating that R-vine model is better than C-vine model. Then in this situation we could choose between fitting an R- or D-vine, we select R-vine because is more general and shows better statistics in Table 13.20.

From the tree shown in Figure 13.55 we deduce that Russell 3000 and MSCI US SC are the most important indices, being directly model with three and four indices. Also, private equity and real estate indices are not directly modelled.

	Family	Usage	Tail dependence
1	t-Student	13	Both
2	BB1	4	Both
3	Joe (rot 180)	1	Lower

Table 13.22: Usage and tail dependence of the different family distribution implemented in the R-vine copula estimated on u-scale transformed data from the standardised residuals after applying models in Table 13.19.

In this case the number of fitted copulas are mostly from t-Student family. In the following section we show the predicted risk measures for the three portfolios.
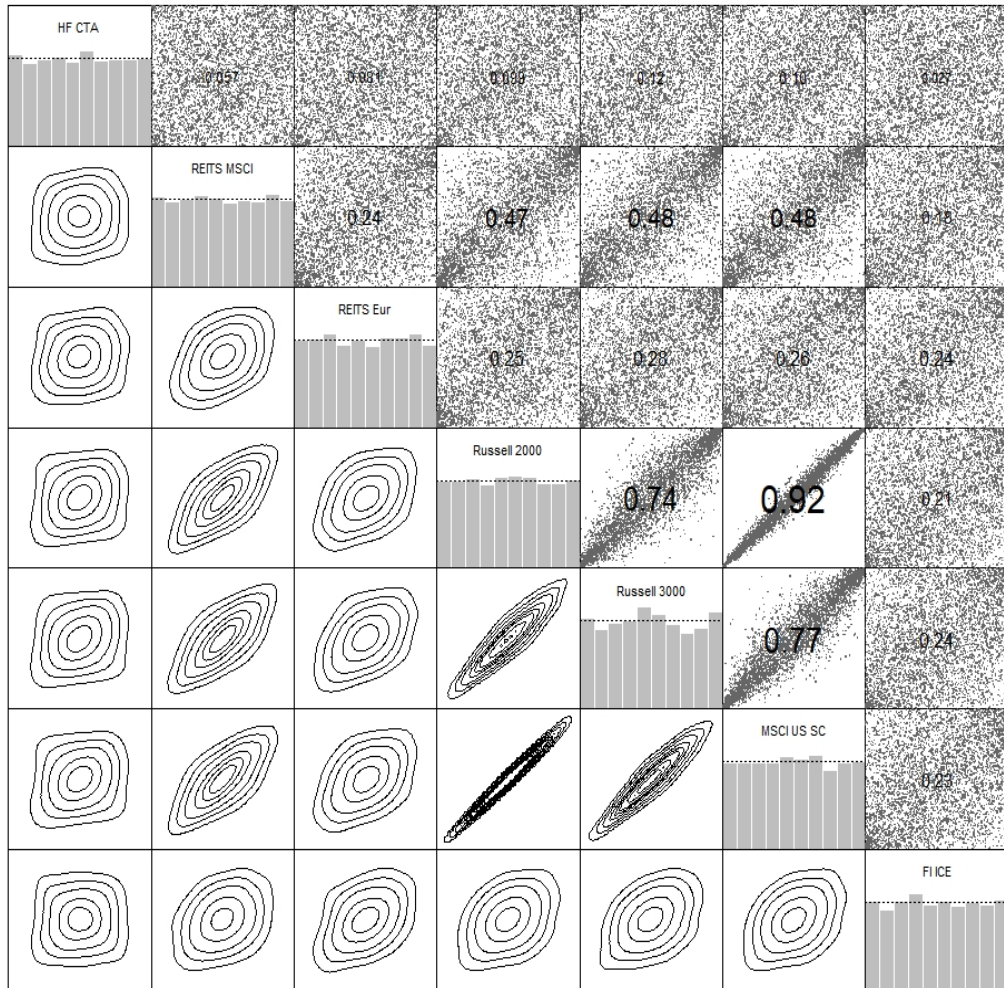
Figure 13.54: Marginally normalised contours for each pair of variables in the lower triangular. Histograms of the standardised residuals after fitting models given in Table 13.19 on the u-scale in the diagonal. Pairwise scatter plots of the pseudo copula data with the corresponding empirical Kendall's τ in the upper triangular.
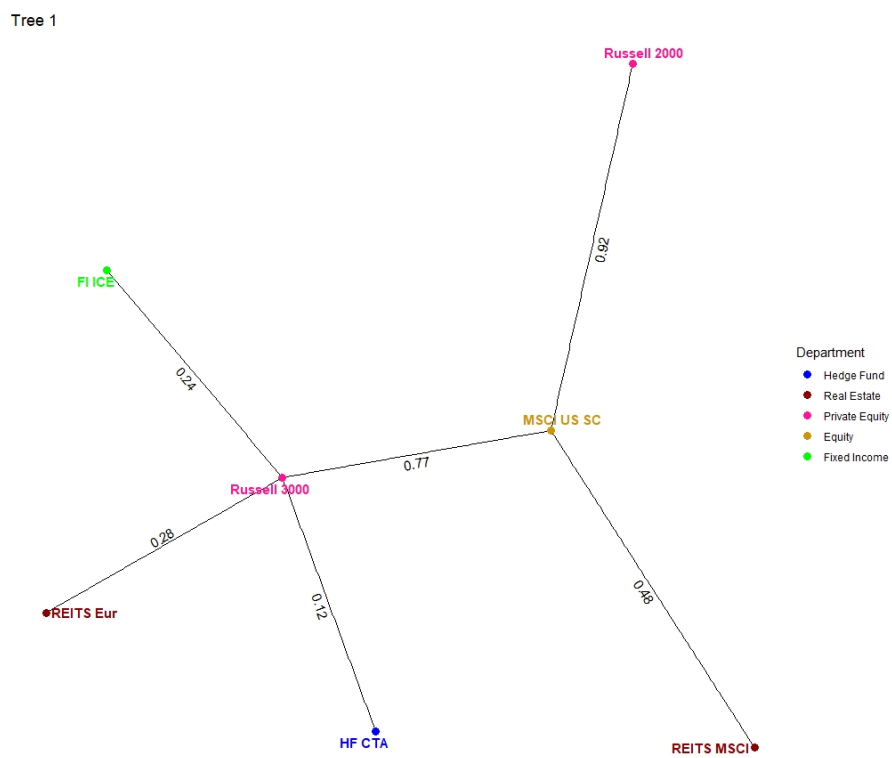
Figure 13.55: First R-vine tree of the R-vine copula estimated in Table 13.20.

13.1.3   Risk measures forecast

In this section we forecast the VaR and ES of the portfolios described in Table 13.18. We also consider the same rolling window characteristics as in Section 12. Then, we use a 500 observation rolling window, refitting the model parameters every 40 observations. Sampling 5000 portfolio values for each time moment, which are used for forecasting the risk measures.

Figure 13.56: Estimated value at risk at 0.05 and 0.01 levels of equal weighted, BVK and market capitalisation portfolios in Table 13.18 using the R-vine copula model specified in Section 6 together with ARMA-GARCH margins specified in Table 13.19.

The figures in this section show equivalent results to the ones obtained for the complete portfolio in Section 12.1.3. Figures 13.56 and 13.57 show the VaR and ES forecast which describe the movement of the observed returns for the three portfolios. Also, Figure 13.58 show a good estimation of the confidence interval of the returns, showing a well estimation of the evolution of the time series.

Figure 13.57: Estimated expected shortfall at 0.05 and 0.01 levels of equal weighted, BVK and market capitalisation portfolios in Table 13.18 using the R-vine copula model specified in Section 6 together with ARMA-GARCH margins specified in Table 13.19.
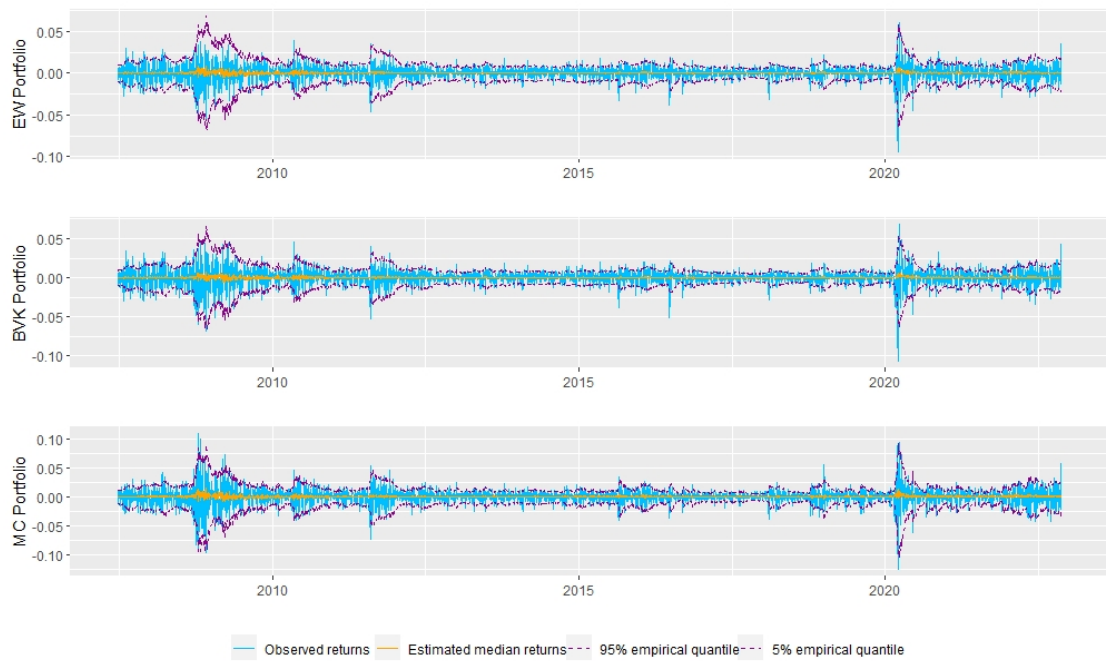
Figure 13.58: Observed, estimated and 90% confidence level returns of equal weighted, BVK and market capitalisation portfolios in Table 13.18 using the R-vine copula model specified in Section 6 together with ARMA-GARCH margins specified in Table 13.19.

13.2 Stationary vine copula models

We continue the study of the reduced portfolio introducing the stationary vine copula models. In this section we work in a equivalent way to Section 12.2. Then, we start transforming the original returns time series to the u-scale with the return distributions of Table 12.12.
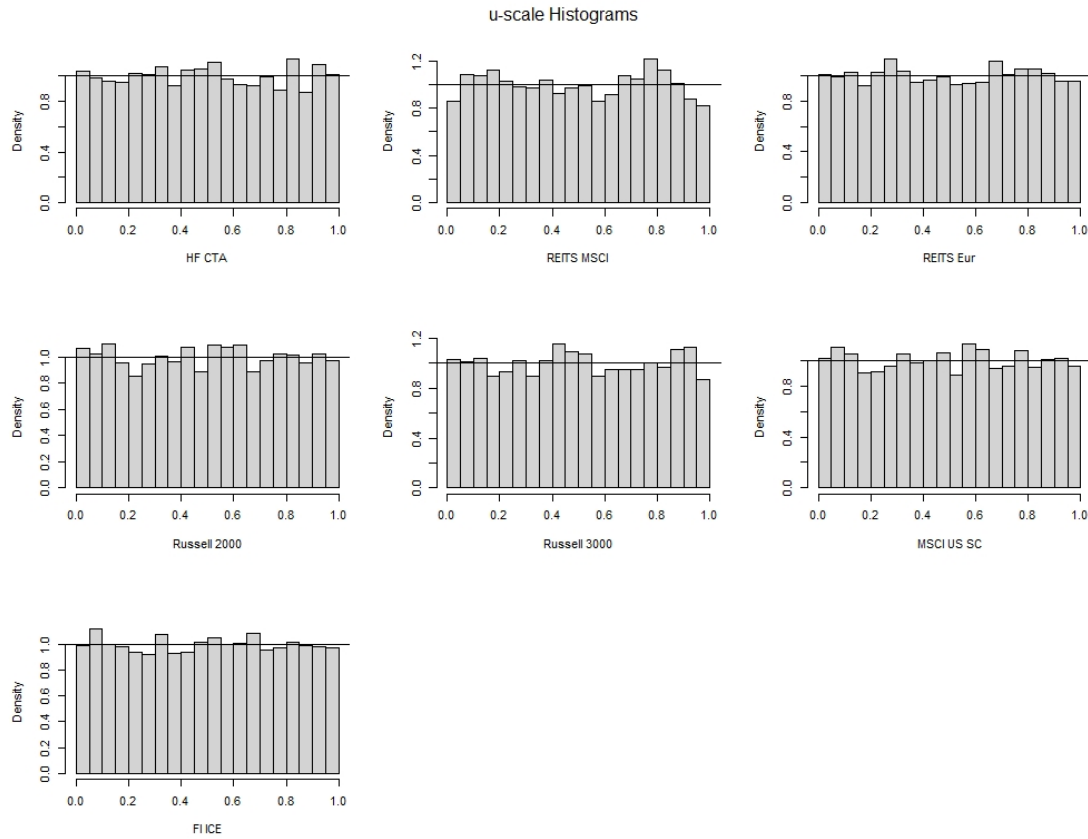

Figure 13.59: Histograms of the u-scale transformed data with innovations distributions of Table 12.12.

Once we have the u-scale data, we compare the three possible vine structure models, S-, D-, and M-vine. Table 13.23 shows the best fitted model is S-vine for the Akaike information criteria, Bayesian information criteria and log-likelihood.

Figure 13.60 plots the first vine tree corresponding to the S-vine copula fitted. The cross sectional R-vine structure is similar to the structure in the R-vine copula obtained in Section 13.1.2, being Russell 3000 and MSCI US SC the main indices. In this case, the serial dependence is modelled with a bivariate copula between MSCI US SC and FI ICE represented with the green label in the graph.

The S-vine copula fitted shows in Table 13.24 similar results as the previous vine copula fitted, being the t-Student the most used distribution.

We continue in the next Section forecasting the risk measures using a S-vine rolling window

Type	AIC	BIC	LogLik
S-Vine	-40403.54	-39714.97	20309.77
D-Vine	-40254.33	-39508.38	20244.16
M-Vine	-40238.56	-39486.23	20237.28

Table 13.23: AIC, BIC and log-likelihood of S-, D- and M-vine copulas estimated on the u-scale transformed data with return distributions of Table 12.12.

	Family	Usage	Tail dependence
1	Gaussian	1	None
2	t-Student	40	Both
4	Joe (rot 0)	1	Upper
5	Frank	3	None
6	Gumbel (rot 0)	1	Upper
8	BB7	4	Both
9	BB8	3	None
10	Independent	9	None
12	Joe (rot 180)	1	Lower

Table 13.24: Usage and type of tail dependence of the different family distribution implemented in the S-vine copula fitted in Table 13.23.

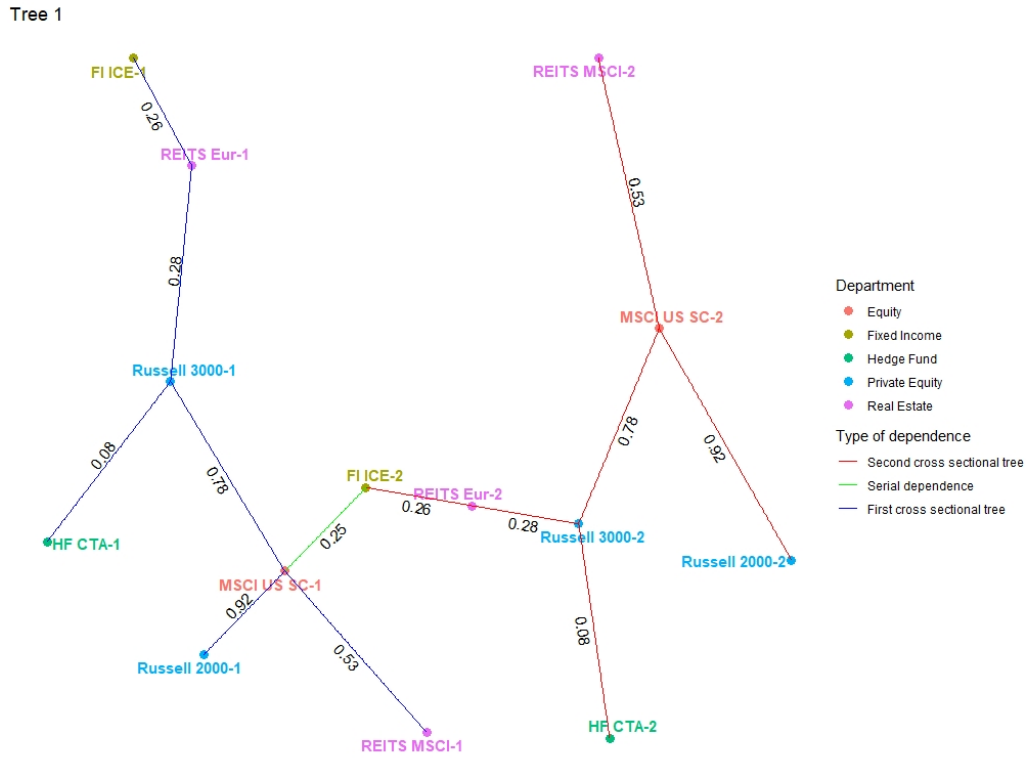on the reduced portfolios described in Table 13.18.

Figure 13.60: First vine tree of the S-vine copula fitted in Table 13.23.

13.2.1 Risk measure forecast

In this section we plot different graphs to show the forecasted risk measures. The previous analysis includes forecast has been made on a 500 rolling window, refitting the model parameters every 40 observations and sampling 5000 portfolio values at each time point.

The results plotted in Figures 13.61 and 13.62 show that VaR values are conditioned by the length on the window considered, affecting to the marginal estimations. In case of the ES, Figure 13.62 we do not observe as extreme values as in Section 12.2.1 but the moments with higher losses do not correspond to the moments with higher volatility. Also, we observe in these figures the effect of the window size in the three portfolios.
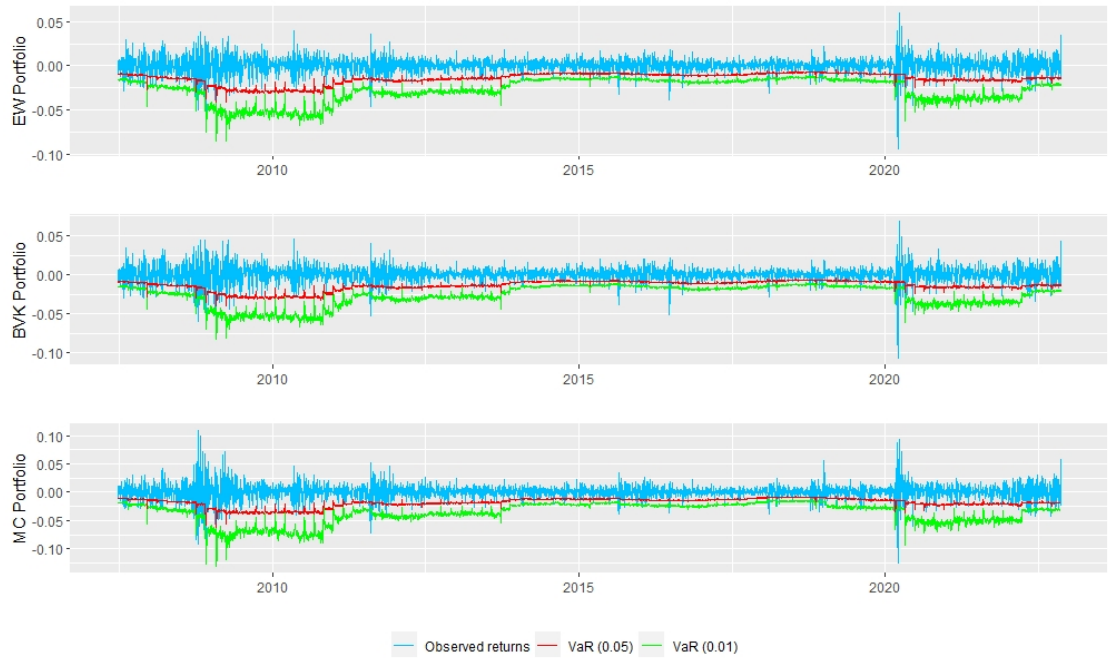
Figure 13.61: Estimated value at risk at 0.05 and 0.01 levels of equal weighted, BVK and market capitalisation portfolio in Table 13.18 using S-vine copula models.

Figure 13.62: Estimated expected shortfall at 0.05 and 0.01 levels of equal weighted, BVK and market capitalisation portfolio in Table 13.18 using S-vine copula models.
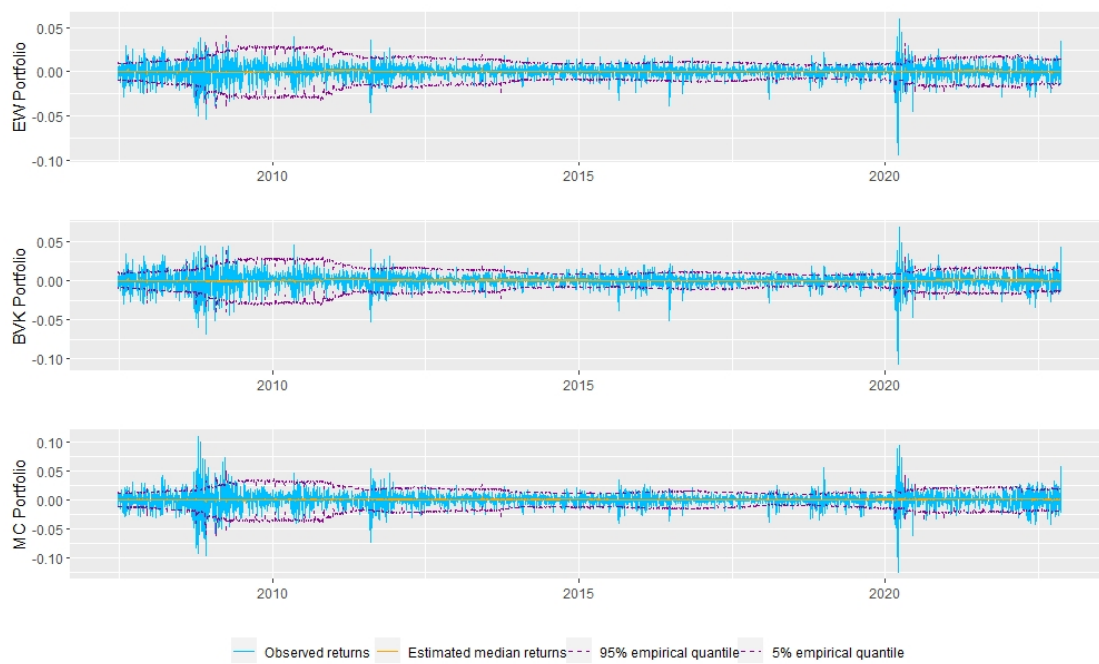
106

Figure 13.63: Observed, estimated and 90% confidence level returns of equal weighted, BVK and market capitalisation portfolios in Table 13.18 using S-vine copula models.

13.3 Fleishman's transformation method

Throughout this section we apply the Fleishman's transformation method introduced in Section 8 to the reduced portfolios from Table 13.17 to sampling the portfolio returns values and estimating the risk measures. In this case we use similar conditions as for the previous methods. We estimate the risk measures in a two years rolling window (500 observations). However, since we can not forecast for not being time dependent method, we always refit the data in the training rolling window. As we do with the previous methods we use a sample size of 5000 for the risk measure estimation.
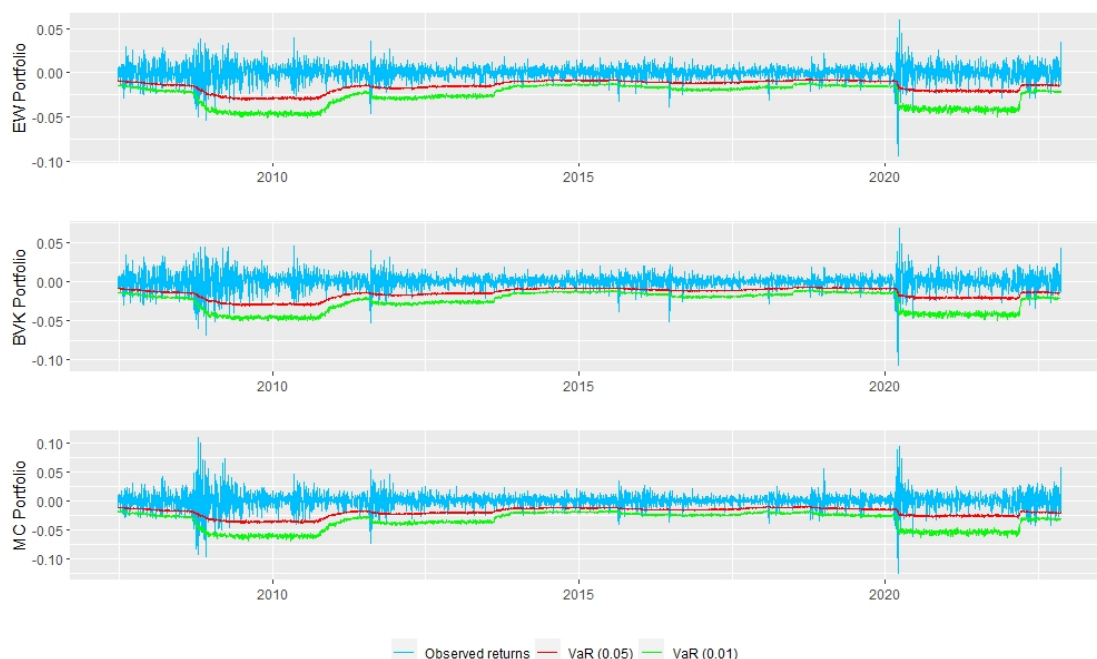
13.3.1 Risk measure forecast

Figure 13.64: Estimated value at risk at 0.05 and 0.01 levels of equal weighted, BVK and market capitalisation portfolios in Table 13.18 using Fleishman's transformation specified in Section 8.

The risk measures estimated with Fleishman's transformation are also restricted by the chosen window size. Figure 13.64 shows smooth VaR estimations. The window size affects on the results overestimating the risk measure during low volatility periods. Expected shortfall in Figure 13.65 show also smooth results, with some periods of higher estimated lost because of the 2008 or COVID-19 crisis.

Figure 13.66 shows the confidence interval estimated for the returns observed for each portfolio. We observe some periods of lower and higher volatility but they do not correspond exactly with the original returns time series.

In the next section we make a better comparison of all the methods for all reduced portfolios considered.

Figure 13.65: Estimated expected shortfall at 0.05 and 0.01 levels of equal weighted, BVK and market capitalisation portfolios in Table 13.18 using Fleishman's transformation specified in Section 8.
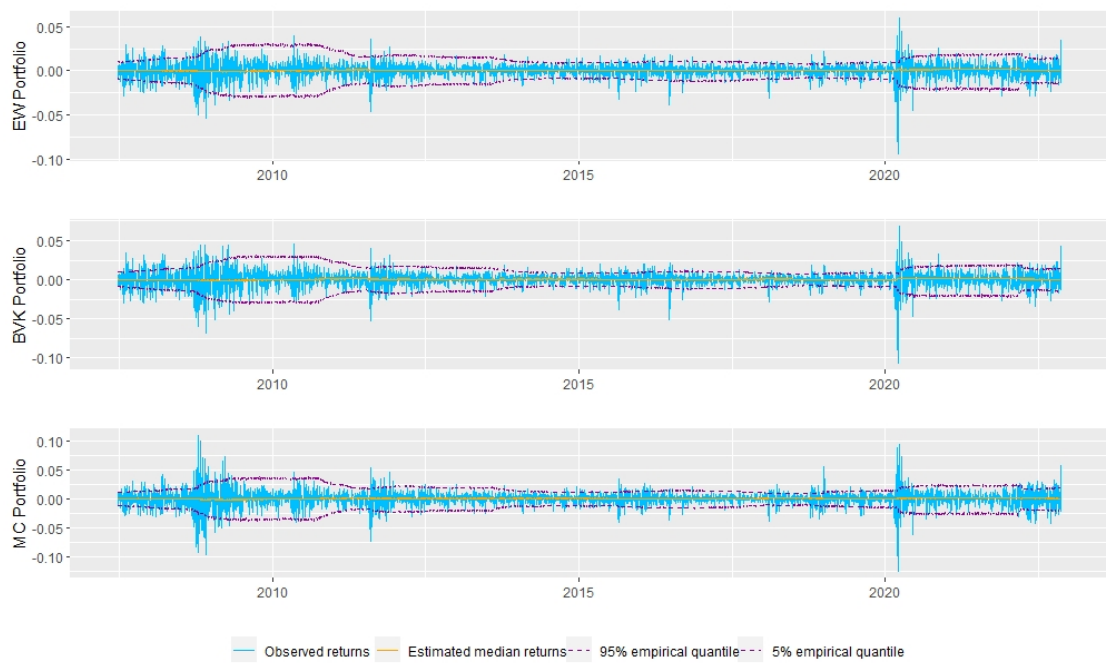
Figure 13.66: Observed, estimated and 90% confidence level returns of equal weighted, BVK and market capitalisation portfolios in Table 13.18 using Fleishman's transformation specified in Section 8.

13.4 Models comparison

In this section use the measures and backtests introduced in Section 9 to compare the ARMA-GARCH with vine copula, stationary copula and the Fleishman's transformation methods.

We begin comparing the VaR forecasted in Sections 13.1.3, 13.2.1 and 13.3.1 for the three methods, which is plotted in Figure 13.67. First, we remark that the results obtained from Fleishman's method and stationary copulas are very similar for both α-levels considered, being smoother than for the ARMA-GARCH vine copula approach. In case of the expected shortfall, Figure 13.68 shows some concrete extreme values estimated by stationary vine copulas during periods where the other two methods estimate small losses. Despite of showing similar results, stationary vine copulas model forecasts more extreme loses than Fleishman's transformation approach.
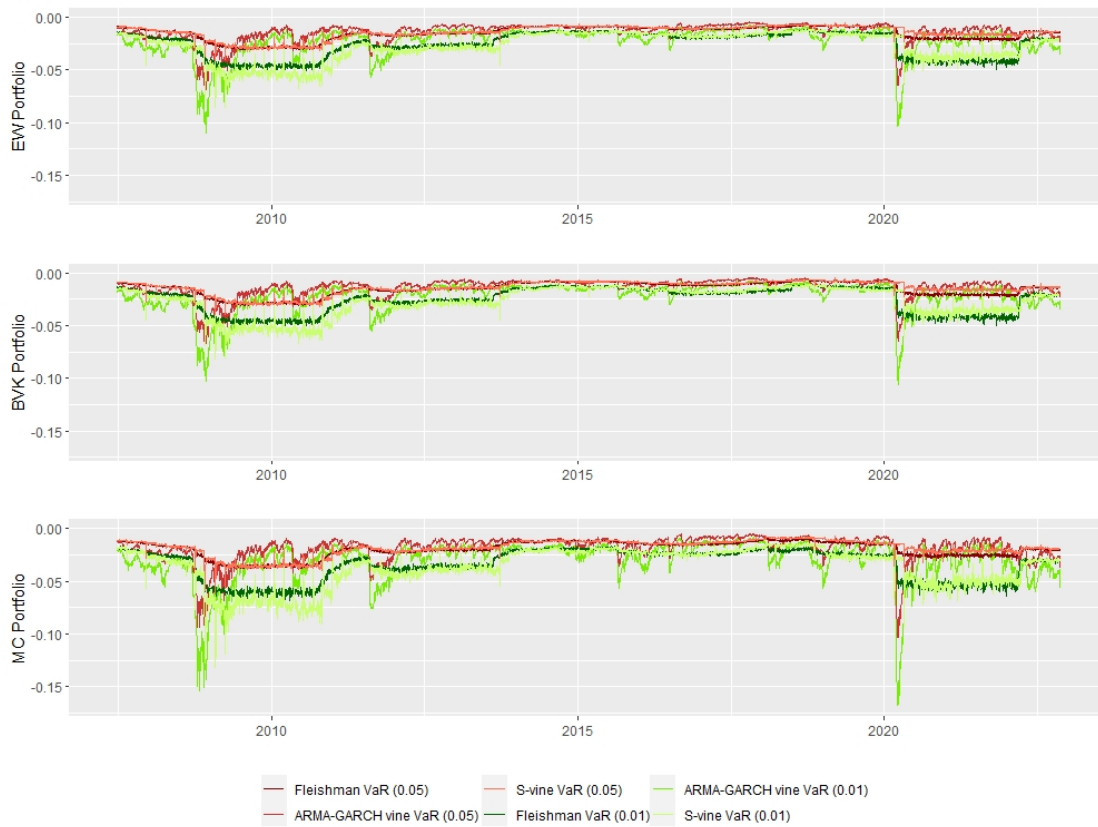

Figure 13.67: Forecasted VaR at 0.05 and 0.01 levels of equal weighted, BVK and market capitalisation portfolios described in Table 13.18 with ARMA-GARCH vine copula, Fleishman's transformation, and S-vine copulas.

We continue evaluating the performance of the methods estimating the VaR level, $\widehat{\alpha}$, using Equation (9.52) in a 100 length rolling window. Figure 13.69 shows the α rolling window at level 0.05 and Figure 13.70 for $\alpha = 0.01$, both of them in a 95% confidence interval. Similarly to the complete portfolio, the VaR is underestimated during high volatility periods. In general we also observe a similar evolution of $\widehat{\alpha}$ for Fleishman and stationary vine copulas, with similar values
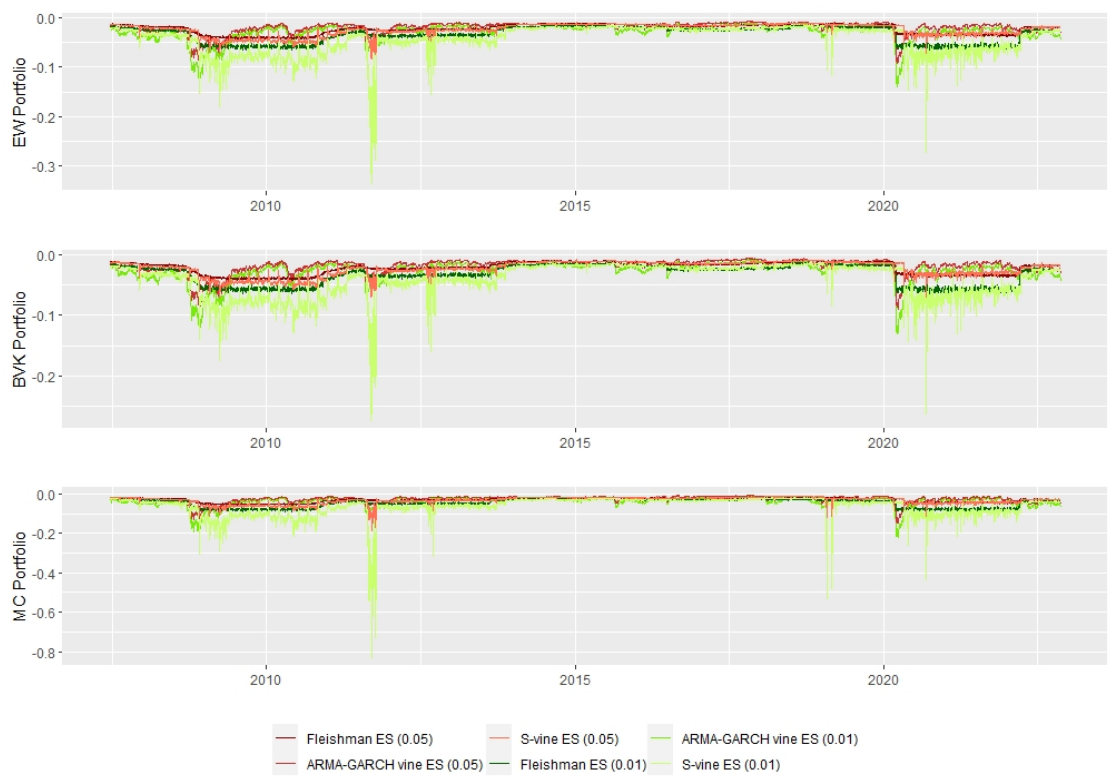
111

Figure 13.68: Forecasted ES at 0.05 and 0.01 levels of equal weighted, BVK and market capitalisation portfolio described in Table 13.18 with ARMA-GARCH vine copula, Fleishman's transformation, and S-vine copulas.

during the whole period. Here ARMA-GARCH with vine copula is more stable, moving around the theoretical values.

Table 13.25 resume the measures from Section 9 calculated for the three portfolios. Bold numbers represent the best value of the row. Clearly the model which better estimates the rolling window is the ARMA-GARCH vine copula, having the highest percentage of theoretical α-levels in the 95%confidence interval. Stationary vine copulas show better results than Fleishman, but not very far, both methods tends to under and overestimate during high and low volatility periods. With respect to the estimated $\widehat{\alpha}$ in the whole period we have some differences between the portfolios; for equal weighted portfolio all methods estimate $\widehat{\alpha}$ better than for BVK and market capitalisation portfolios, anyway the VaR is underestimated for the three portfolios at $\alpha = 0.05$ and $\alpha = 0.01$ levels. In case of the β given in Equation (9.54) all the values are very near to zero, showing the ARMA-GARCH models with vine copulas the model with lower $|\beta|$ values.

Portfolio	Level α	Measure	ARMA-GARCH/ vine copula	S-vine	Fleishman
Equal weighted	0.05	$\hat{\alpha}$	0.053645	0.05625	**0.05**
		β	-0.000647	**-0.000157**	-0.002138
		% of $\hat{\alpha}$ in $CI_{0.05}(\alpha)$	**85,61**	50,99	43,53
	0.01	$\hat{\alpha}$	**0.013281**	0.015625	0.016145
		β	**-0.000805**	-0.002403	-0.0040374
		% of $\hat{\alpha}$ in $CI_{0.05}(\alpha)$	**65,64**	40,24	40,29
BVK	0.05	$\hat{\alpha}$	0.06510	0.06796	**0.06458**
		β	**-0.000958**	-0.00151	-0.00294
		% of $\hat{\alpha}$ in $CI_{0.05}(\alpha)$	**89.14**	41.25	40.48
	0.01	$\hat{\alpha}$	**0.01796**	0.02005	0.02213
		β	**-0.00141**	-0.00257	-0.00479
		% of $\hat{\alpha}$ in $CI_{0.05}(\alpha)$	**79.55**	38.88	40.29
Market capitalisation	0.05	$\hat{\alpha}$	**0.063541**	0.07031	0.06510
		β	-0.00043	0.000709	-0.0031080
		% of $\hat{\alpha}$ in $CI_{0.05}(\alpha)$	94.54	38.48	39.49
	0.01	$\hat{\alpha}$	**0.01875**	0.02291	0.0234
		β	0.000984	**-0.000252**	-0.002577
		% of $\hat{\alpha}$ in $CI_{0.05}(\alpha)$	**82.62**	31.10	37.22

Table 13.25: $\hat{\alpha}$ and β estimations at 0.05 and 0.01 levels between 13/11/2007 and 11/11/2022 using ARMA-GARCH vine copula, Fleishman's transformation and S-vine copula on equal weighted, BVK and market capitalisation portfolios described in Table 13.18. Also, the percentage of $\widehat{\alpha} \in CI_{0.05}(\alpha)$ estimated in a rolling window in Figures 13.69 and 13.70 for $\alpha = 0.05$ and $\alpha = 0.01$ respectively.

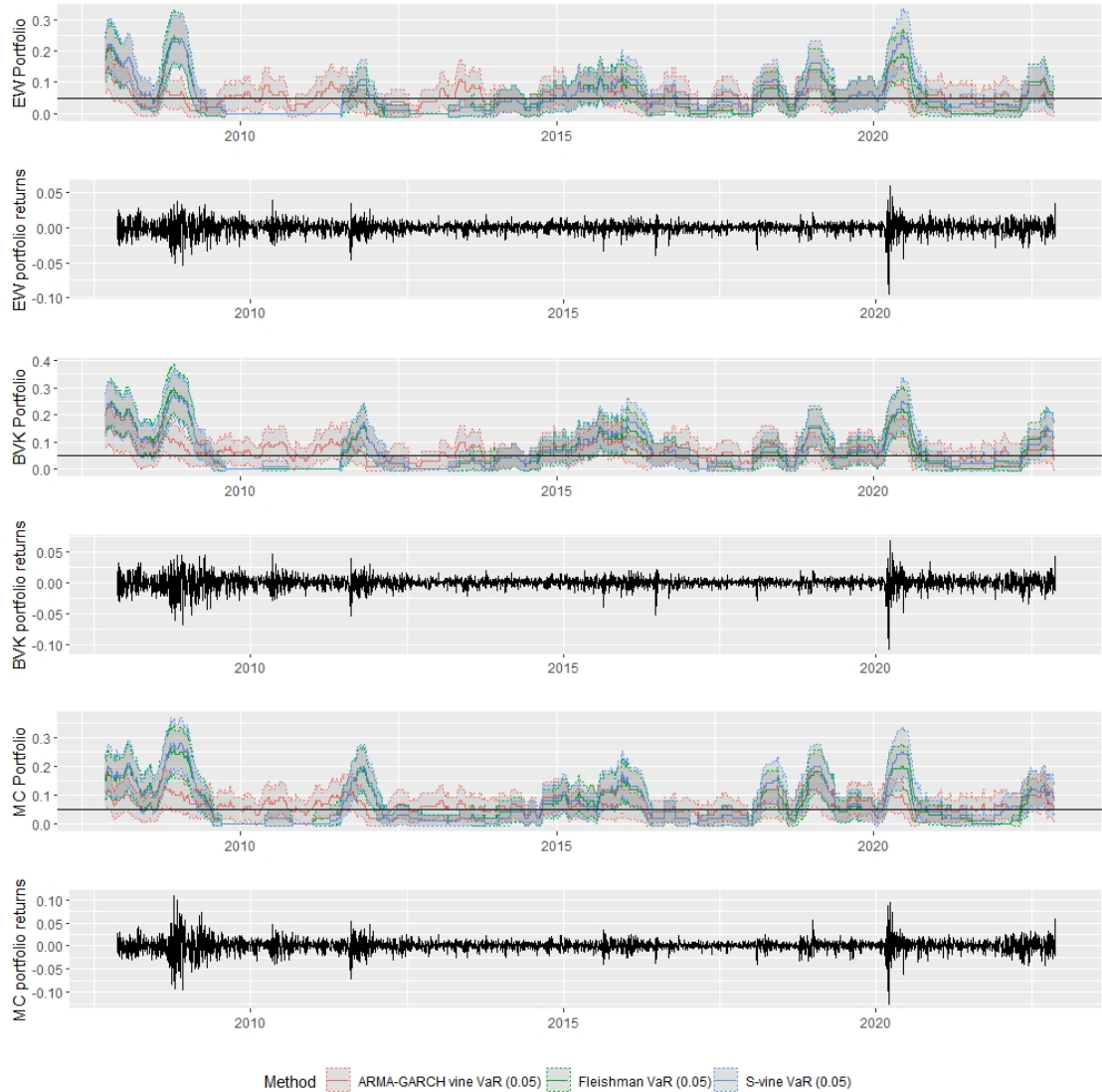Figure 13.69: In the graphs above we show the estimated $\widehat{\alpha}$ in a 0.05 VaR level and the corresponding 95% confidence interval. $\widehat{\alpha}$ is obtained from VaR estimated with ARMA-GARCH vine copula, Fleishman's transformation and S-vine copula in a 100 observation rolling window of the equal weighted, BVK and market capitalisation portfolios described in Table 13.18 using 9.53. The rows show the equal weighted, BVK and market capitalisation portfolio returns specified in Table 13.18.
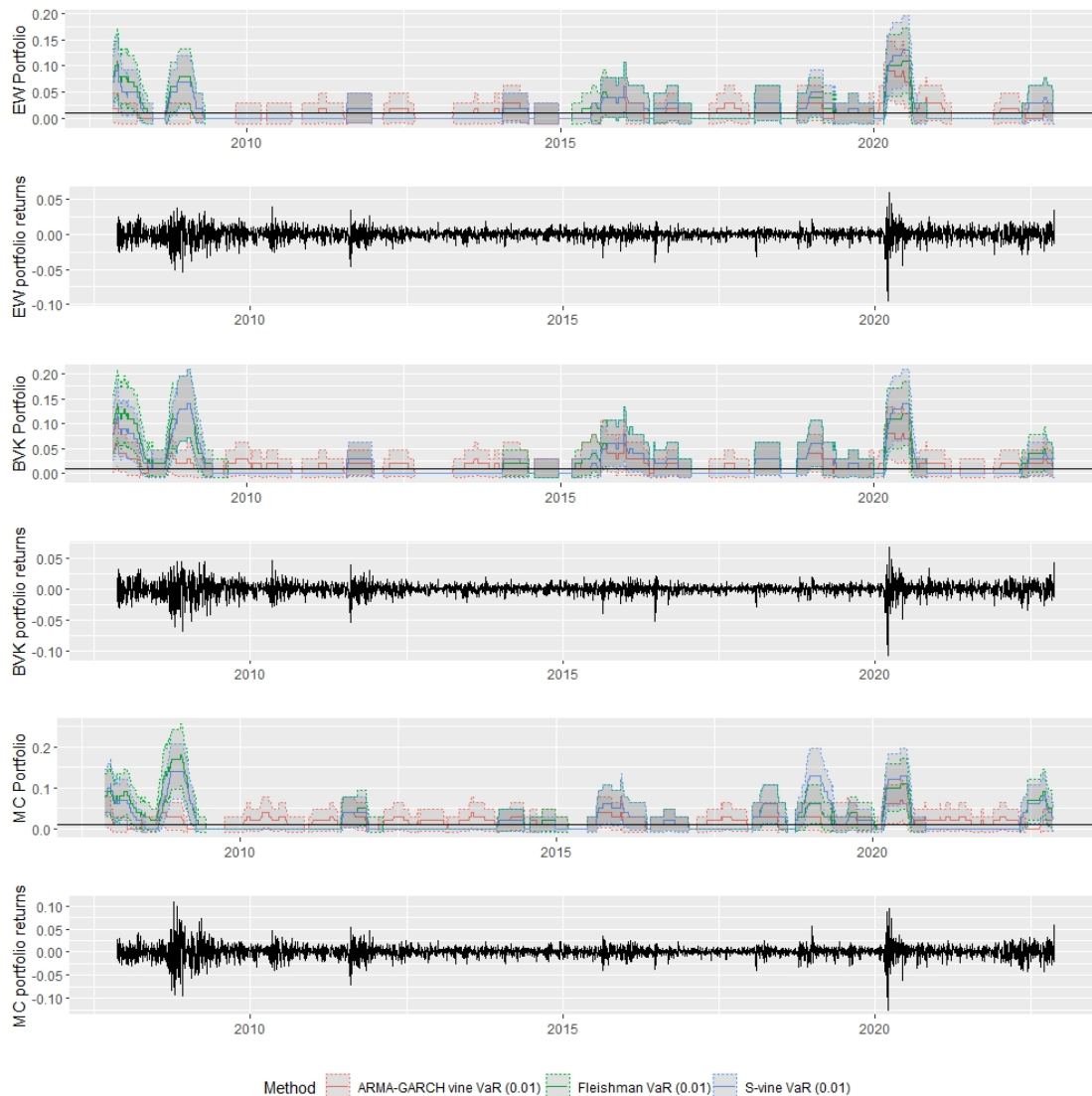
Figure 13.70: In the graphs above we show the estimated $\widehat{\alpha}$ in a 0.01 VaR level and the corresponding 95% confidence interval. $\widehat{\alpha}$ is obtained from VaR estimated with ARMA-GARCH vine copula, Fleishman's transformation and S-vine copula in a 100 observation rolling window of the equal weighted, BVK and market capitalisation portfolios described in Table 13.18 using 9.53. The rows show the equal weighted, BVK and market capitalisation portfolio returns specified in Table 13.18.

After calculating different measures for the estimated VaR and ES, we check the backtest of Sections 9.1 and 9.2. In Table 13.26 are shown the p-values of the performed tests for the three considered methods and portfolios defined in Table 13.18. The p-values are calculated for all the forecast period, from 06/2007 to 11/2022. Considering a 95% test level we accept the null hypothesis for those test with $p - value > 0.05$ representing a good model estimation. In general, we have the risk measures are better estimated when using ARMA-GARCH vine copula approach. However, most of the tests reject the null hypothesis, especially for BVK and market capitalisation portfolios. This could happen because those portfolios are less diversified than the equal weighted portfolio, which could imply more changes of volatility making it more difficult to estimate. This is checked in Table 13.18 were Russell 3000 accumulates more than the 80% of the total portfolio weight. Also, since we are analysing a long time period the wellness of risk measures estimation change over the time. To check it we consider the same backtests for shorter periods.

Portfolio	Level α	Backtest	ARMA-GARCH/ vine copula	S-vine	Fleishman
Equal weighted	0.05	LR_{uc}	**0.941**	**0.081**	**0.766**
		LR_{cc}	0.019	3.9e-11	0.000
		ER	**0.107**	**0.854**	0.002
		CC	**0.934**	0.008	**0.740**
	0.01	LR_{uc}	**0.464**	0.001	4.4e-04
		LR_{cc}	**0.615**	9.1e-08	5.9e-08
		ER	**0.148**	**0.093**	0.008
		CC	**0.606**	0.001	0.010
BVK	0.05	LR_{uc}	9.3e-05	4.3e-07	6.1e-04
		LR_{cc}	2.9e-06	2.8e-15	2.1e-12
		ER	0.012	**0.198**	0.000
		CC	6.6e-04	5.0e-06	0.001
	0.01	LR_{uc}	0.005	1.5e-04	1.6e-10
		LR_{cc}	0.019	2.2e-06	1.5e-15
		ER	0.013	**0.093**	0.000
		CC	0.025	3.05e-04	0.025
Market capitalisation	0.05	LR_{uc}	7.0e-05	1.2e-06	7.0e-05
		LR_{cc}	1.6e-04	1.5e-08	4.2e-08
		ER	0.034	0.001	0.000
		CC	7.8e-04	5.8e-07	4.9e-04
	0.01	LR_{uc}	4.3e-06	4.4e-04	3.2e-11
		LR_{cc}	2.2e-05	4.4e-06	1.6e-13
		ER	**0.435**	0.009	**0.324**
		CC	**0.05**	2.7e-10	1.5e-06

Table 13.26: P-values for LR unconditional and conditional test, exceedance residual test and conditional calibration test for VaR and ES at 0.05 and 0.01 levels estimated using ARMA-GARCH vine copula, stationary vine copula and Fleishman approaches for equal weighted, BVK and market capitalisation portfolios showed in Table 13.18.

Figures 13.71 and 13.72 collect the p-values of the backtests from Sections 9.1 and 9.2 calculated on the estimated risk measures for a period of two years and two months, obtained after dividing the whole estimation period in seven parts. The methods are represented on differ-

ent colours which are expected to be above the 0.05 line, representing the wellness of the risk measures estimations.
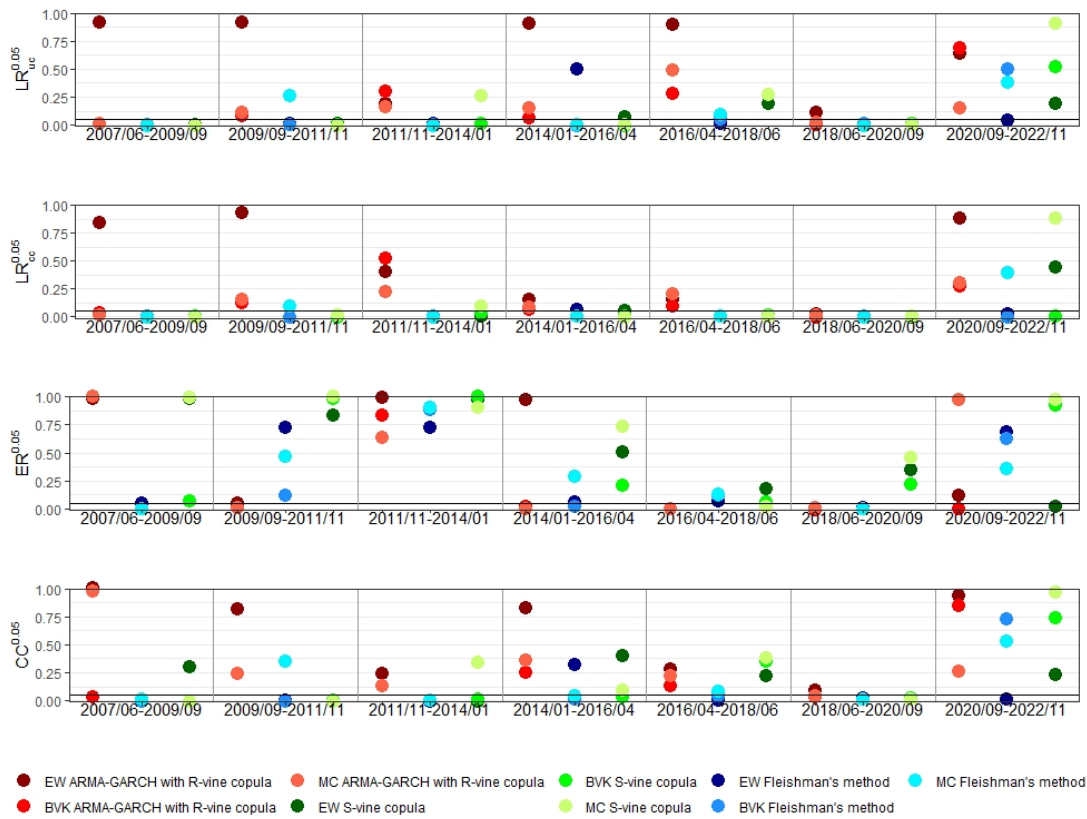

Figure 13.71: P-values corresponding to the backtests developed in Sections 9.1 and 9.2 for VaR and ES at 0.05 level estimated using ARMA-GARCH vine copula, S-vine copula and Fleishman's transformation for portfolios described in Table 13.18 during the periods described in the graph.
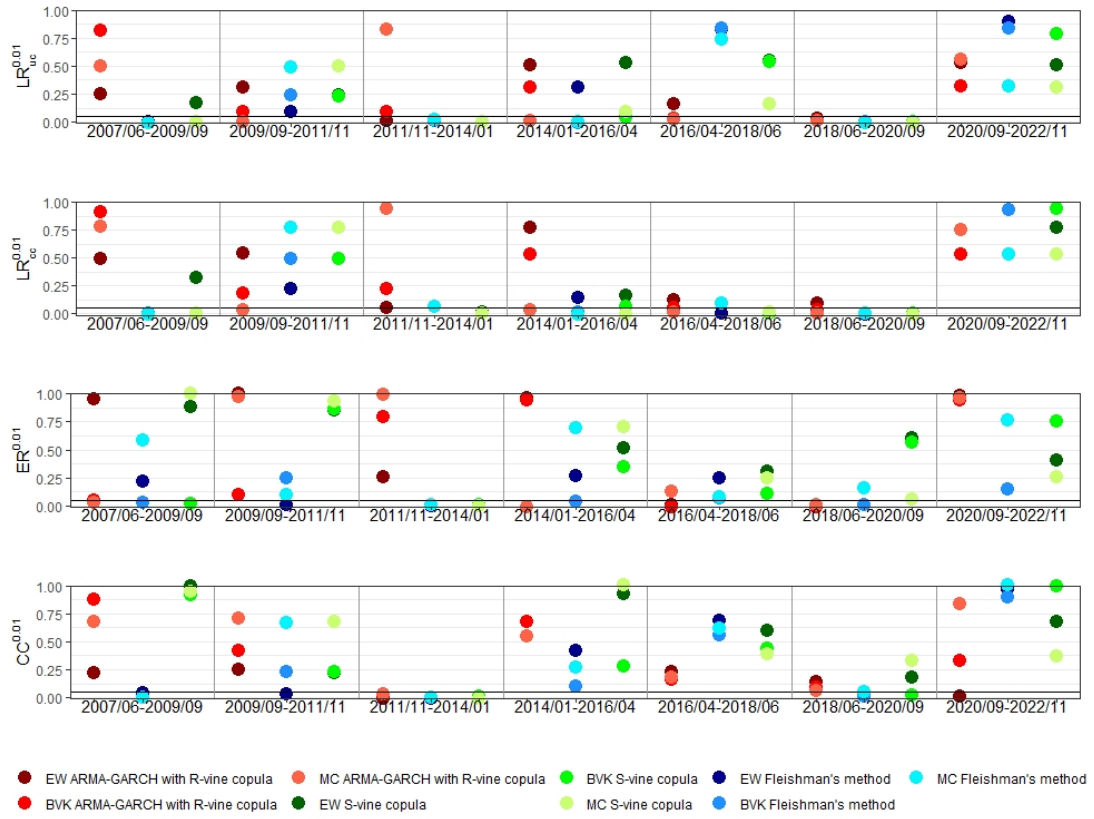
117

Figure 13.72: P-values corresponding to the backtests developed in Sections 9.1 and 9.2 for VaR and ES at 0.01 level estimated using ARMA-GARCH vine copula, S-vine copula and Fleishman's transformation for portfolios described in Table 13.18 during the periods described in the graph.

14 Conclusion

In this master thesis, we have compared three methods for estimating the VaR and the expected shortfall. To do so, we considered two markets composed of 15 and 7 assets respectively, over a period of more than 17 years, from 06/2005 to 11/2022. This period includes times of high volatility, such as the 2008 financial crisis and the COVID-19 pandemic in 2020, as well as periods of lower volatility. Due to this variety, we have a comprehensive study.

For each market, we considered three portfolios: equal weighted, BVK, and market capitalisation. The risk measures were estimated using a rolling window of 500 observations. The parameters were recalculated every 40 observations in the case of ARMA-GARCH with vine copulas and stationary vine. In the case of Fleishman, the parameters were recalculated on a daily basis due to the non-predictive nature of the method.

Although the backtests do not show good results when analysing the entire 15-year period, jointly backtesting the results every two years does demonstrate a good estimation of the risk measures with all three methods. As mentioned before, the period 06/2018-09/2020 shows the worst estimation due to the sudden change in volatility. In these cases, it may be worth considering refitting the complete model instead of just reestimating the parameters.

Furthermore, we have observed that all methods tend to underestimate the risk measures during periods of higher volatility, which should be taken into account since it can lead to great losses.

Regarding the models, we have observed clear differences between them. The model that best estimates the risk measures is the ARMA-GARCH margins combined with vine copulas. The other methods produce smoother results for VaR. Additionally, stationary vine copulas show some outliers during periods of volatility change, which can lead to a poor evaluation of expected shortfall. In general, the ranking for the approaches are ARMA-GARCH margins together with vine copulas, followed by stationary vine and then the Fleishman's approach.

Regarding the different portfolios, we also observe some clear differences. These differences are more significant in the case of the reduced portfolios. We see that the equal-weighted portfolio achieves better results, as it is less volatile than the others. In the case of the market capitalisation portfolio, we show that it is influenced by specific assets, leading to higher volatility and poorer estimation. The BVK portfolio falls somewhere in between. This leads us to believe that greater diversification helps us obtain better estimates of the risk measures. One area where we do not see many differences is in relation to the two markets. Despite one market being composed of eight more assets than the other, the results are similar.

In addition to the results shown by each approach, there are other unobserved variables that must be taken into account when choosing the most suitable method. In ARMA-GARCH marginal models, the choice of marginal distributions and the model order is crucial. In the case of vine copulas, the selection of the best structure is important. For stationary vine copulas, we also need to select the marginal distributions, cross-sectional structure, and the Markov order. In contrast, for the Fleishman method, no prior analysis is required as it is based on the first four moments of observed log returns.

Another difference is the estimation and sampling time, although it has not been measured. ARMA-GARCH with vine copulas is the most time and resource-consuming method to compute, followed by stationary vine copulas and the Fleishman's transformation method. Although the latter needs to be recalculated for each observation. It must be added that the reduced portfolio has been considered due to the limitations of the Fleishman method, which does not always

converge. Therefore, in portfolios with a large number of assets, its implementation might be infeasible.

For future research, the methods can be compared under different scenarios or more specific markets. Additionally, other methods can be introduced in the study. For example, estimating risk measures with stress scenarios or including conditioned vine copulas to account for different market conditions.
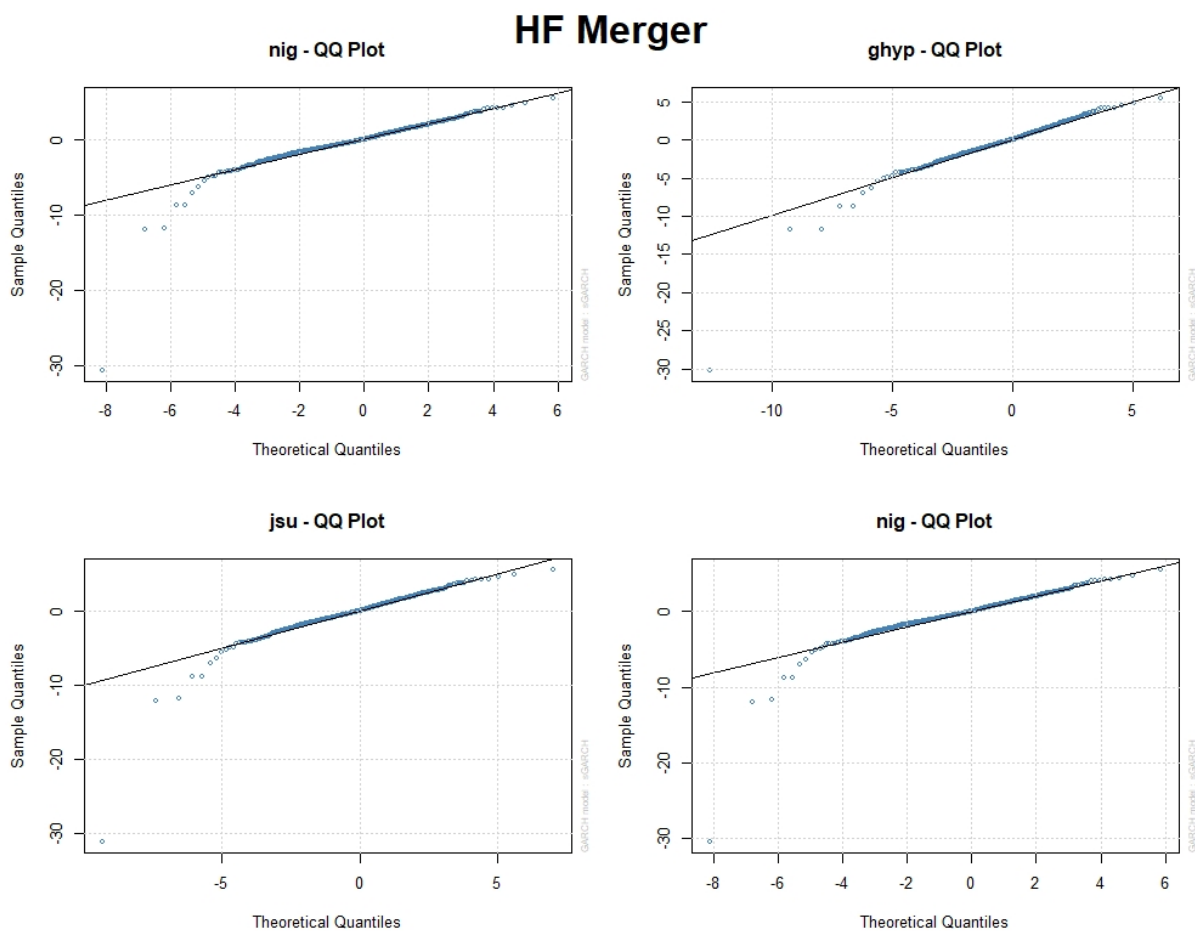
A Innovation distributions QQ-plots

Figure A.73: HF Merger standardized residuals QQ-plots after fitting an ARMA(1,1)-GARCH(1,1) with different innovation distributions
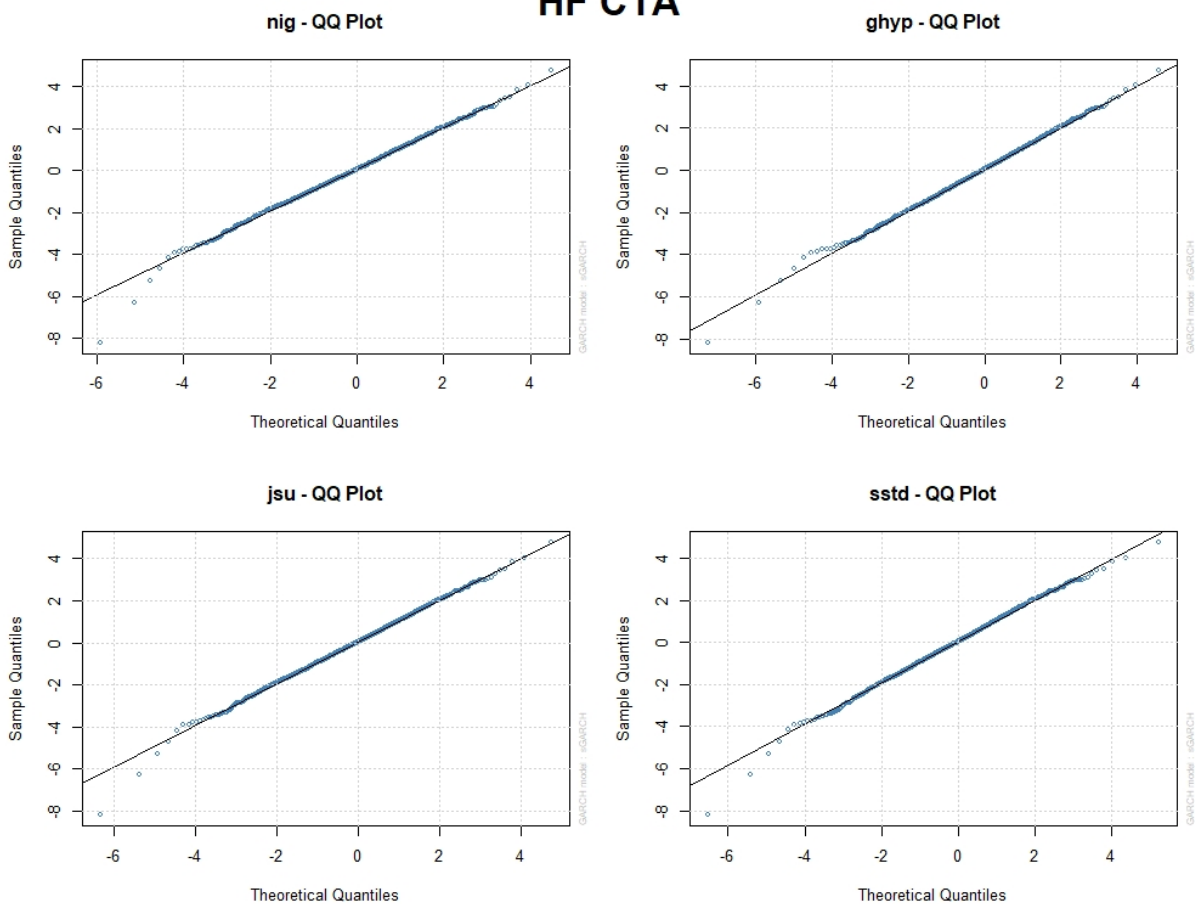
Figure A.74: HF CTA standardized residuals QQ-plots after fitting an ARMA(1,1)-GARCH(1,1) with different innovation distributions
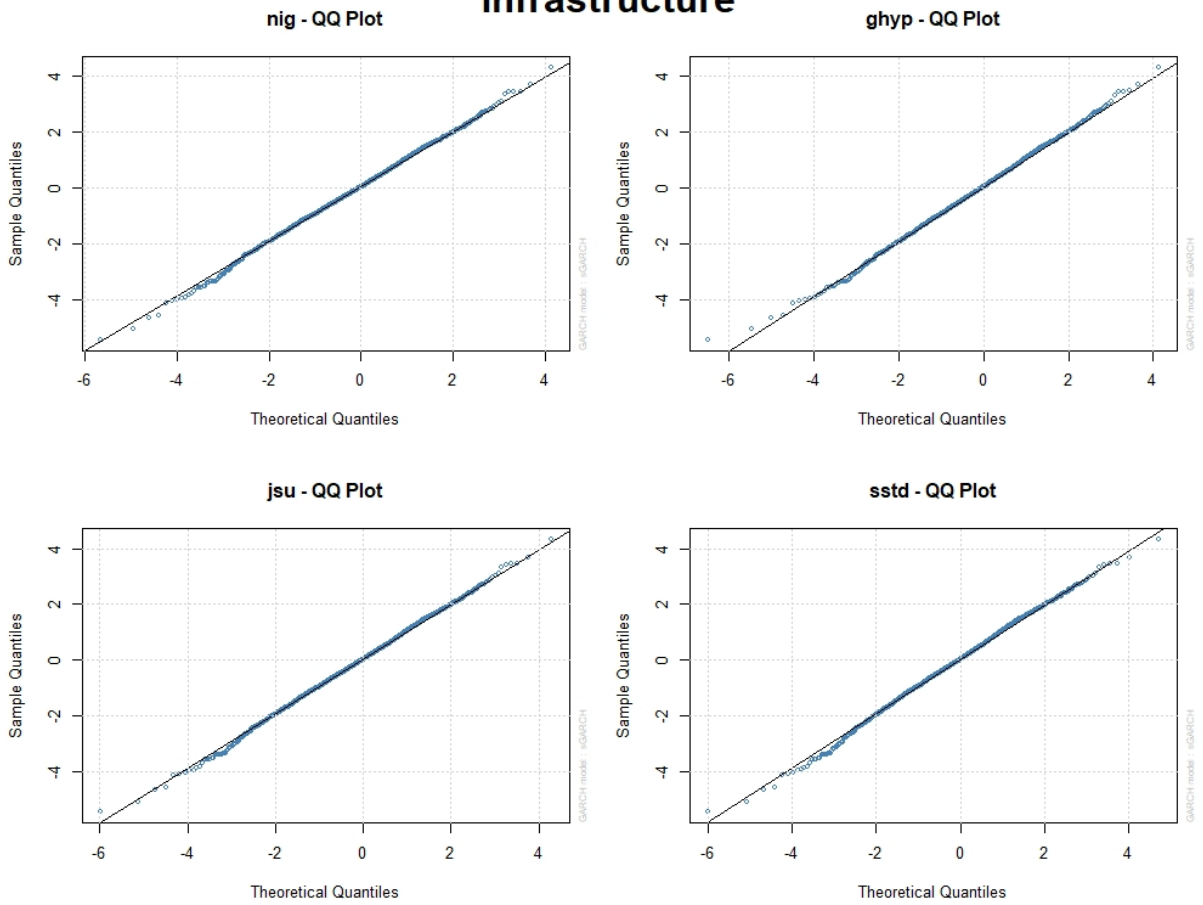
Figure A.75: Infrastructure standardized residuals QQ-plots after fitting an ARMA(1,1)-GARCH(1,1) with different innovation distributions
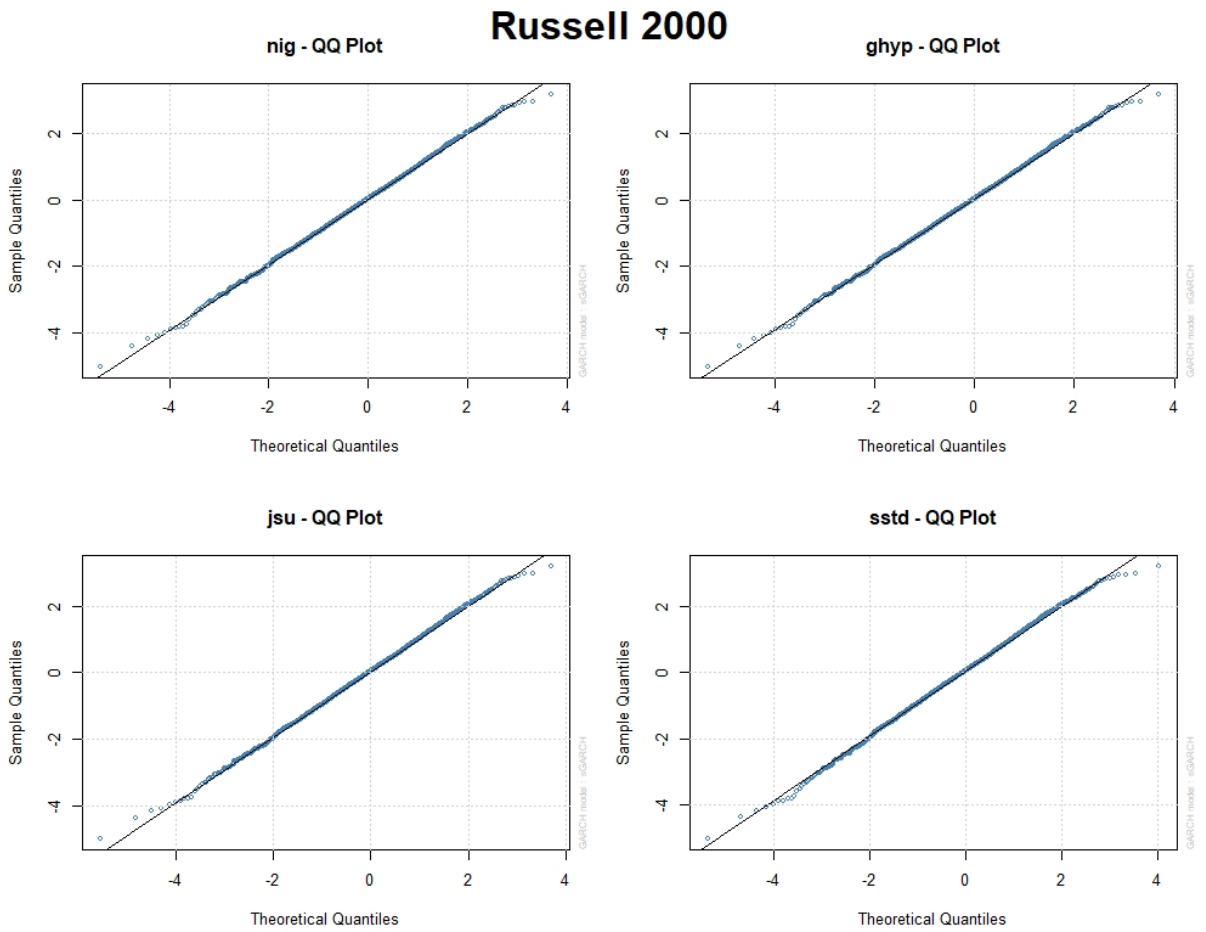
Figure A.76: Russell 2000 standardized residuals QQ-plots after fitting an ARMA(1,1)-GARCH(1,1) with different innovation distributions
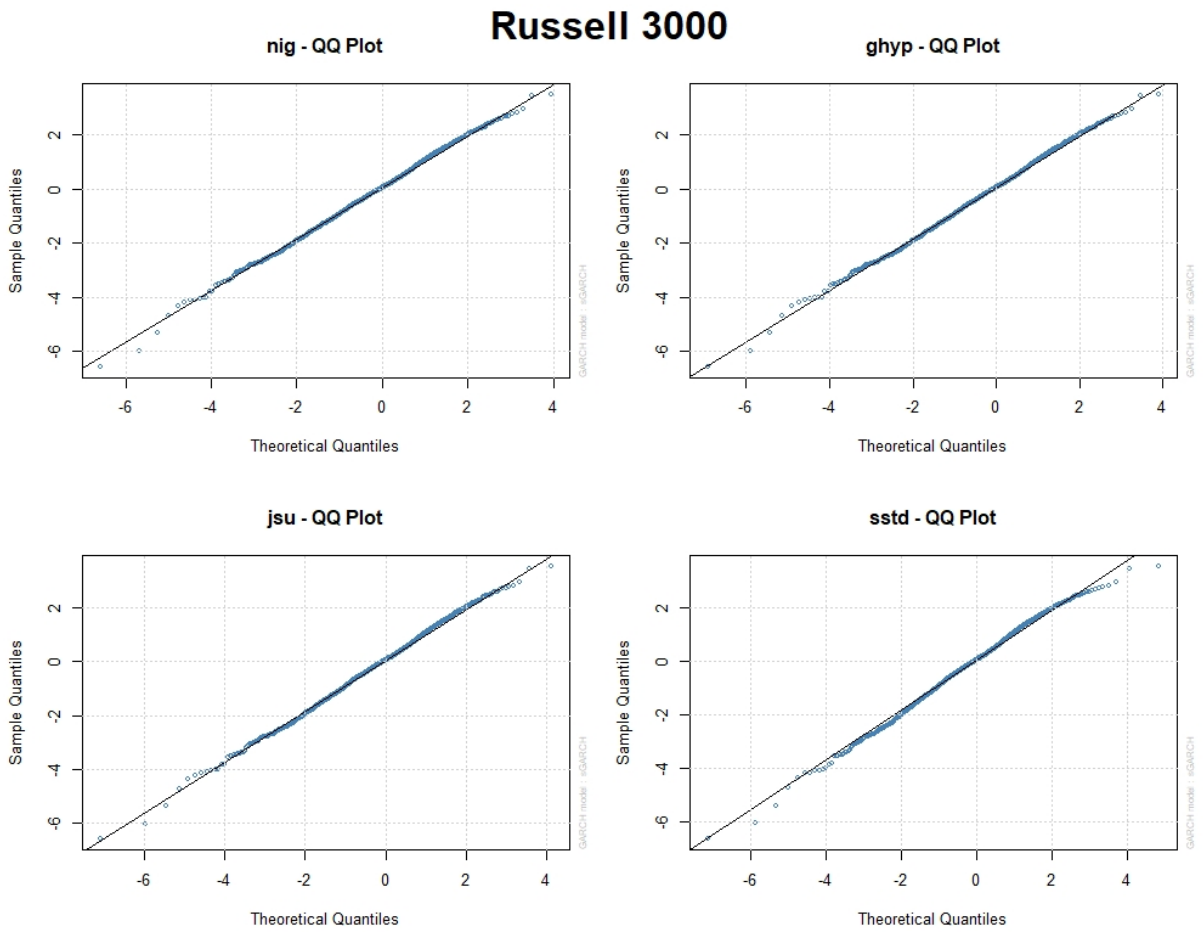
Figure A.77: Russell 3000 standardized residuals QQ-plots after fitting an ARMA(1,1)-GARCH(1,1) with different innovation distributions
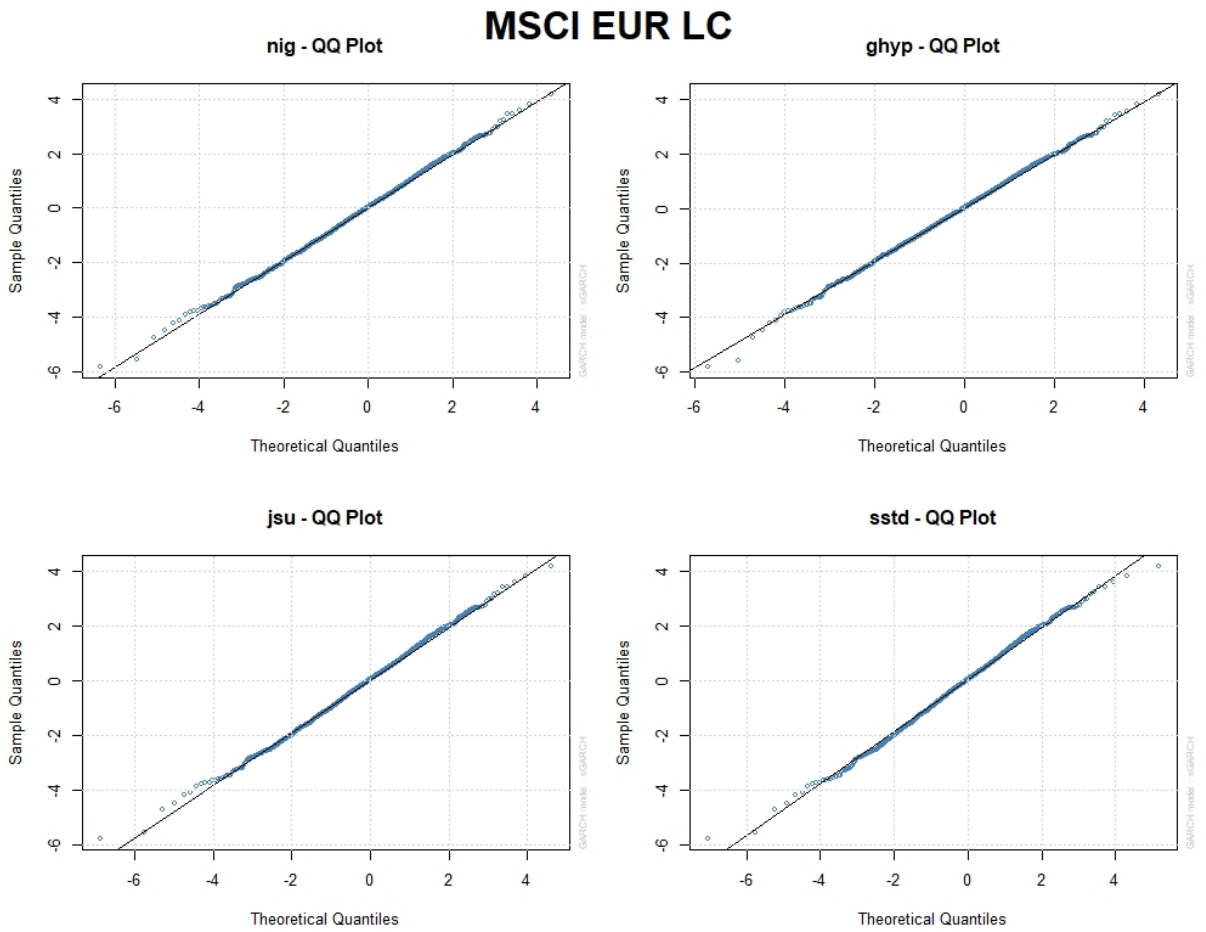
Figure A.78: MSCI Eur LC standardized residuals QQ-plots after fitting an ARMA(1,1)-GARCH(1,1) with different innovation distributions
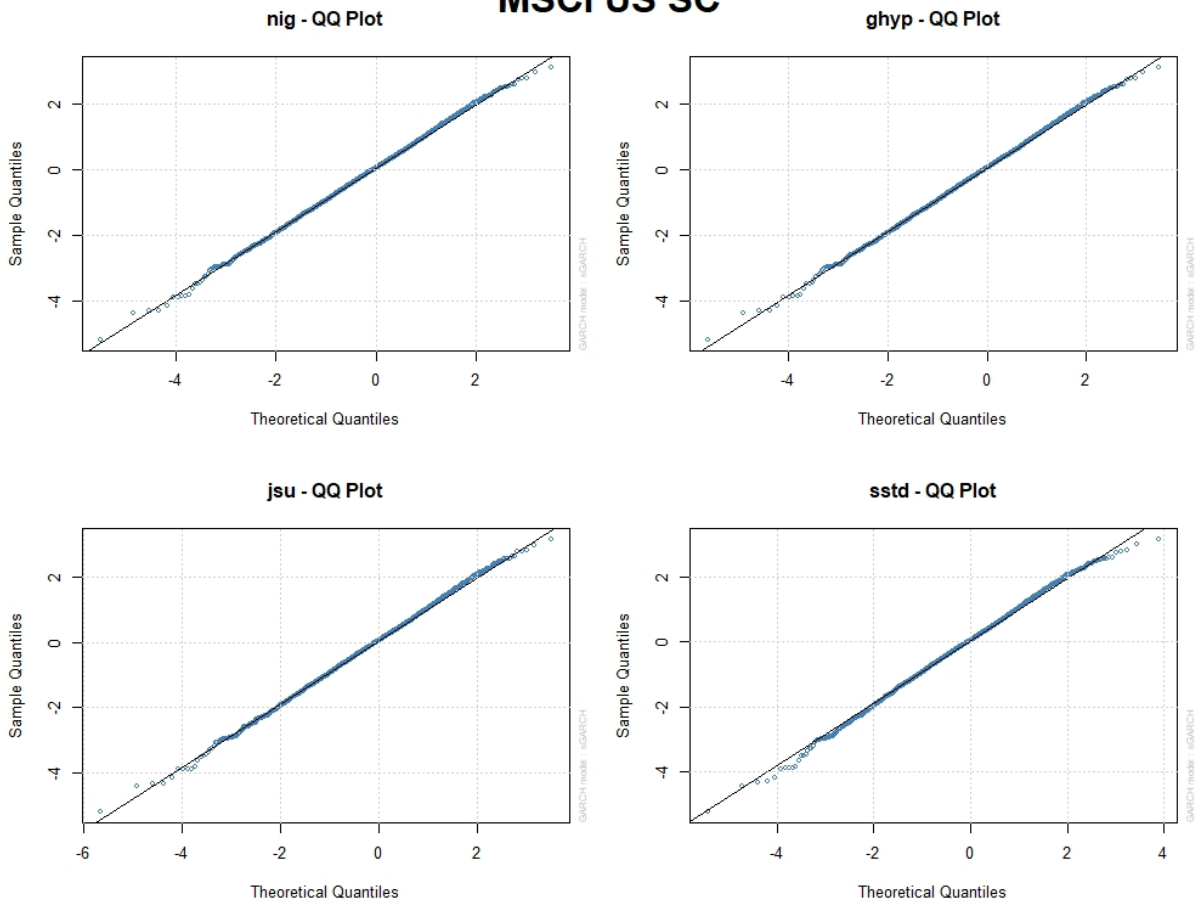
Figure A.79: MSCI US SC standardized residuals QQ-plots after fitting an ARMA(1,1)-GARCH(1,1) with different innovation distributions
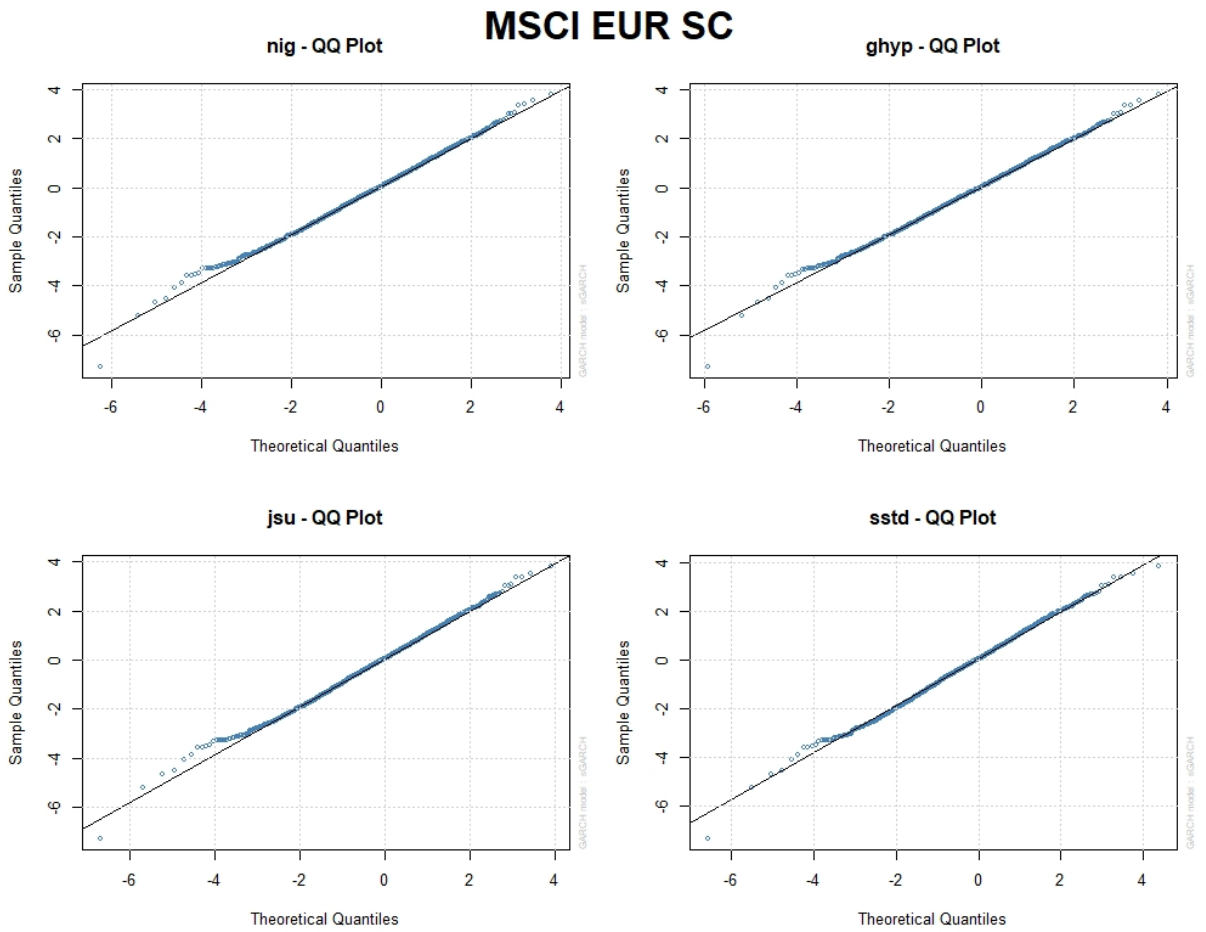
Figure A.80: MSCI Eur SC standardized residuals QQ-plots after fitting an ARMA(1,1)-GARCH(1,1) with different innovation distributions
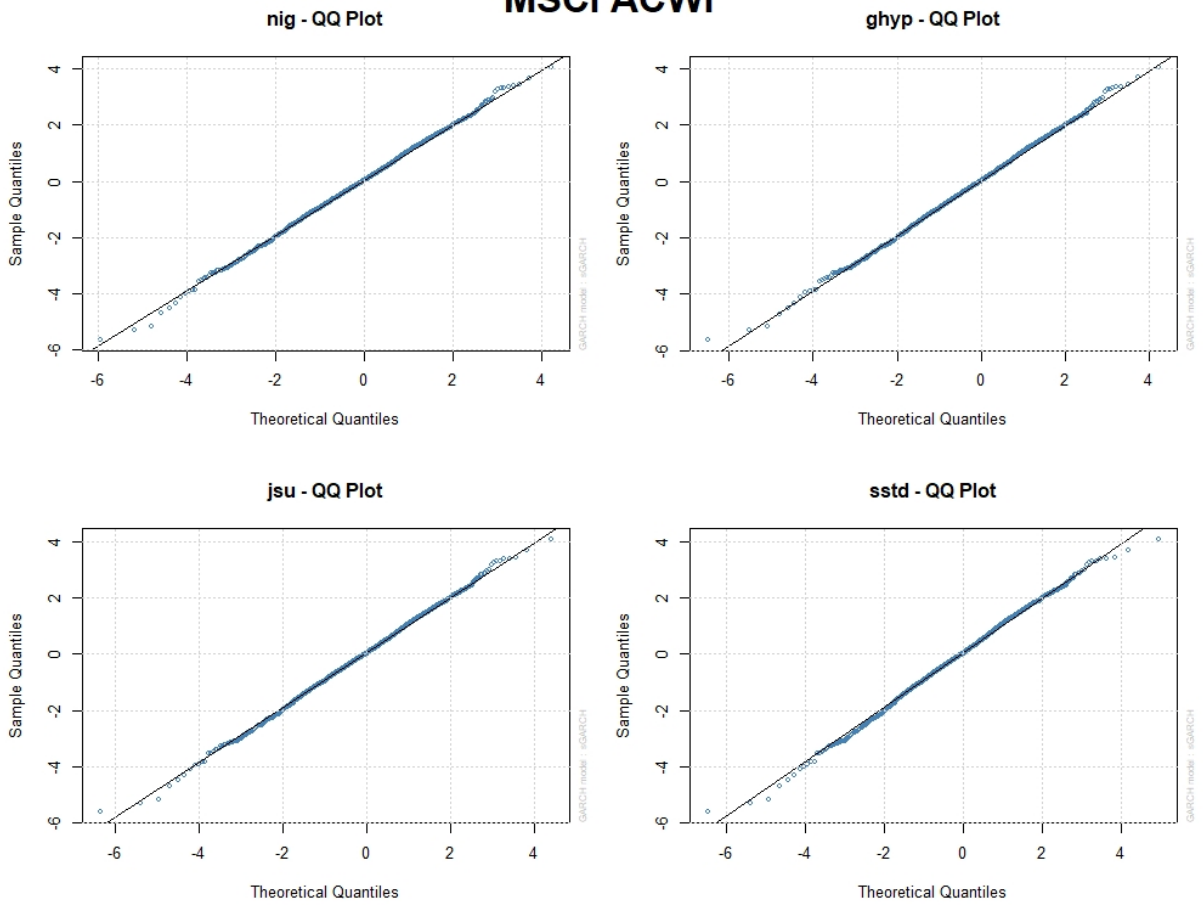
Figure A.81: MSCI ACWI standardized residuals QQ-plots after fitting an ARMA(1,1)-GARCH(1,1) with different innovation distributions

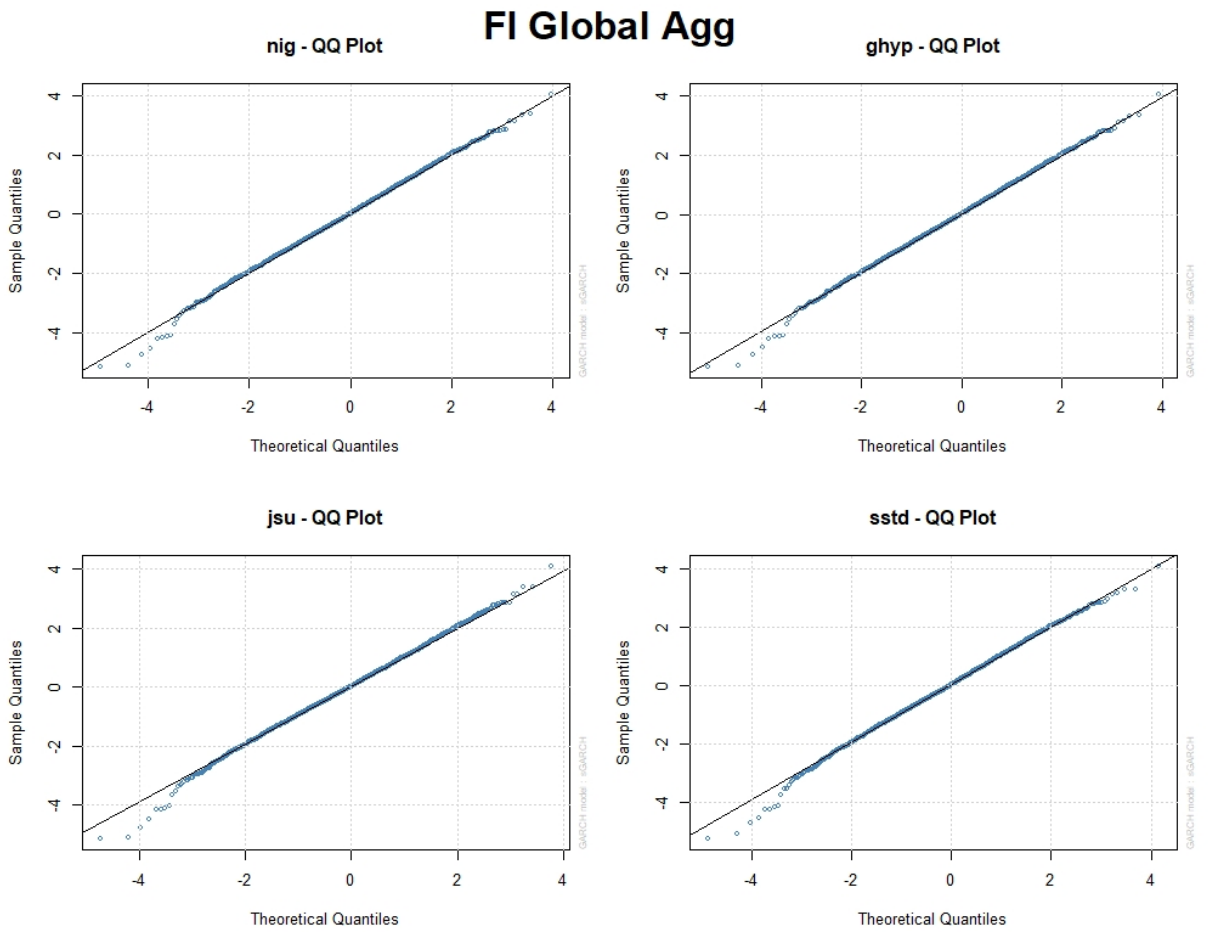Figure A.82: FI Global Agg standardized residuals QQ-plots after fitting an ARMA(1,1)-GARCH(1,1) with different innovation distributions
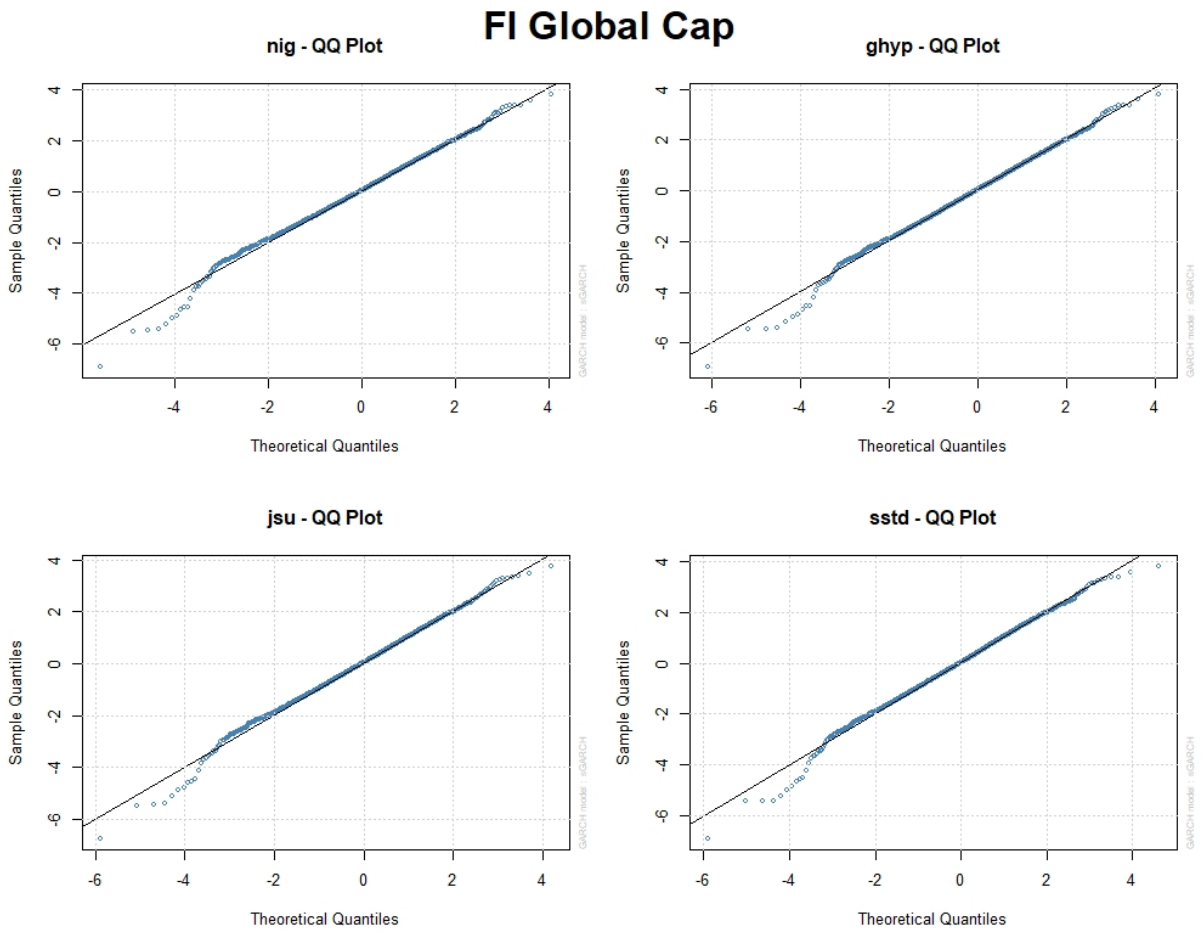
Figure A.83: FI Global Cap standardized residuals QQ-plots after fitting an ARMA(1,1)-GARCH(1,1) with different innovation distributions
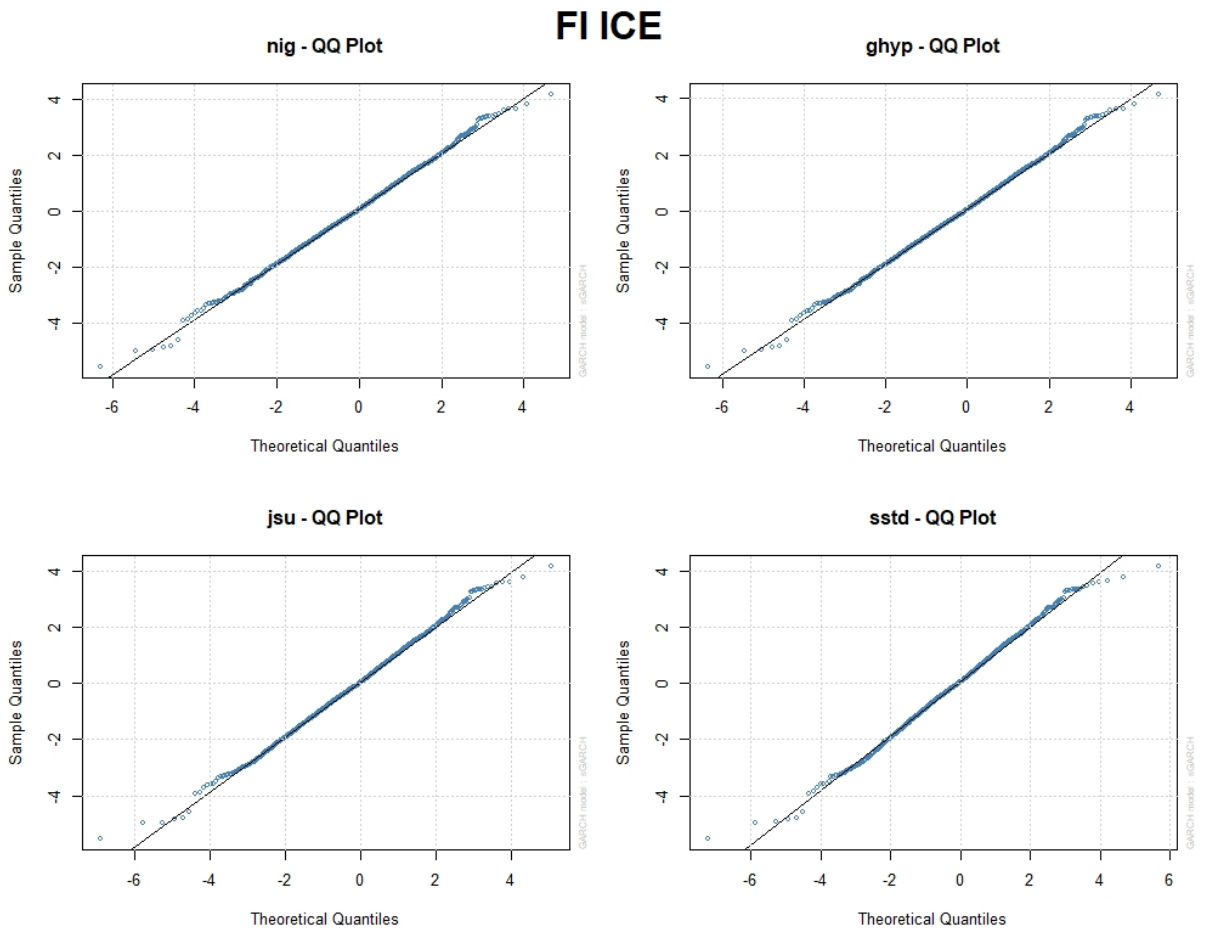
Figure A.84: FI ICE standardized residuals QQ-plots after fitting an ARMA(1,1)-GARCH(1,1) with different innovation distributions

B Performance of different ARMA orders of returns for ARMA selection

Russell 2000

AR	MA	Mean	BIC
1	1	0	-5.63902762407435
0	1	0	-5.63881956327209
1	0	0	-5.63873206101779
1	1	1	-5.63805100141348
0	1	1	-5.63751165362381
1	0	1	-5.63741832169825
0	2	0	-5.63708849901156
2	0	0	-5.63704467270545
3	3	0	-5.63618112172911
0	0	1	-5.63610215516847
0	2	1	-5.63579983526311
2	0	1	-5.63575015723735
1	3	0	-5.6357145154023
3	1	0	-5.63569875912523
2	2	0	-5.63566537223517
2	3	0	-5.63558007405998
3	2	0	-5.63545406129683

Figure B.85: Performance of different ARMA orders for Russell 2000's returns with Generalized Hyperbolic innovations and ACF and PACF of Russell 2000's returns

Russell 3000

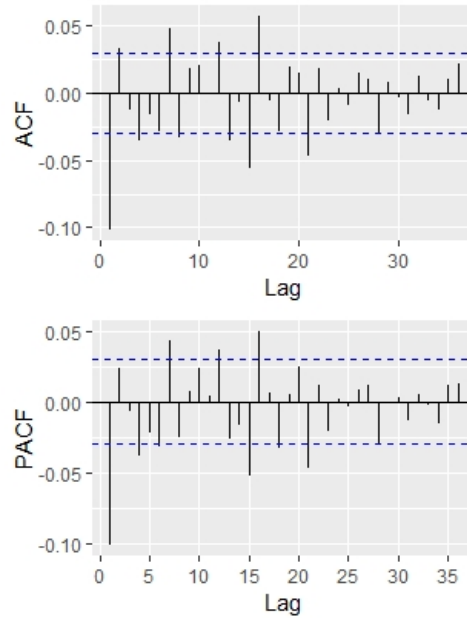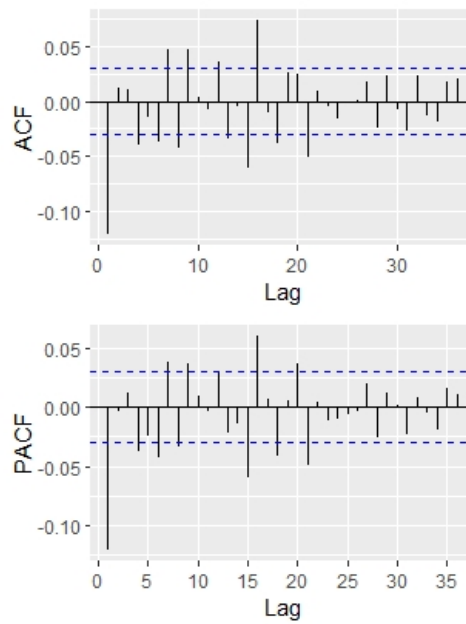| AR	MA	Mean	BIC
1	0	0	-6.23657250543009
0	1	0	-6.23645055651017
1	0	1	-6.23554698621365
0	1	1	-6.23543016270235
0	2	0	-6.23482414936696
2	0	0	-6.23478181471611
1	1	0	-6.23473650786648
2	3	0	-6.23448105185075
3	2	0	-6.23439232564739
0	2	1	-6.23377311074213
2	0	1	-6.23373943423484
1	1	1	-6.23370048156069
2	3	1	-6.23341035396372
3	2	1	-6.23330258250818
3	0	0	-6.23293605359276
0	3	0	-6.23289871301799
1	2	0	-6.23288195067347

Figure B.86: Performance of different ARMA orders for Russell 3000's returns with Generalized Hyperbolic innovations and ACF and PACF of Russell 3000's returns

MSCI EUR LC

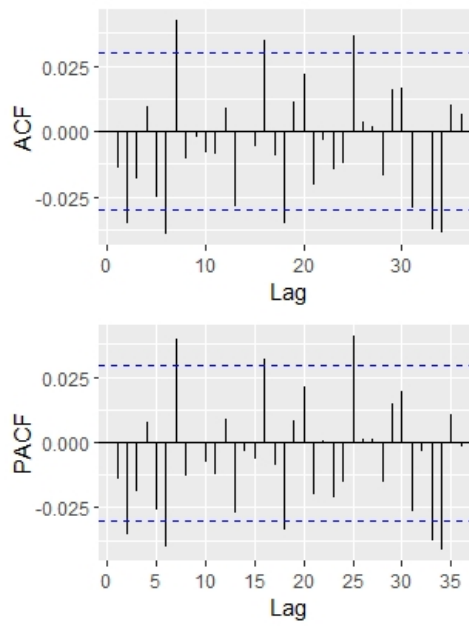| AR	MA	Mean	BIC
1	1	0	-6.25838510282545
1	1	1	-6.25689597702132
1	2	0	-6.25645528164117
2	1	0	-6.25643336705769
2	2	0	-6.25550171602147
3	1	0	-6.25543546325184
1	2	1	-6.25496611627741
2	1	1	-6.25494324955232
2	3	0	-6.25412269856727
3	2	0	-6.254116828426
2	2	1	-6.25400773244689
3	1	1	-6.25395577365244
0	2	0	-6.25369094131556
2	0	0	-6.25353782439107
0	1	0	-6.25341530662967
1	0	0	-6.25339281416837
0	0	1	-6.25332715659373

Figure B.87: Performance of different ARMA orders for MSCI Eur LC's returns with Normal Inverse Gaussian innovations and ACF and PACF of MSCI Eur LC's returns

MSCI US SC

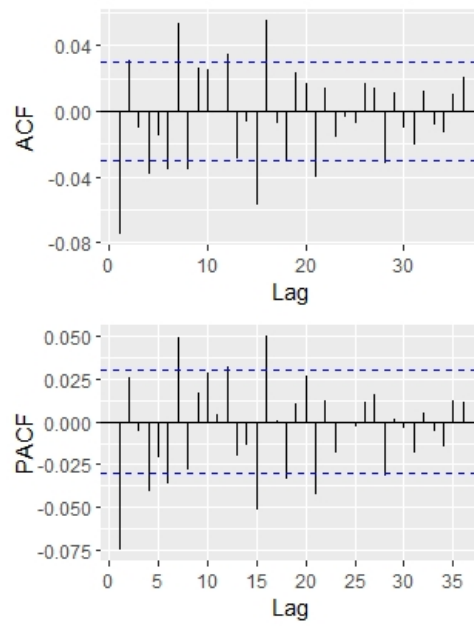| AR	MA	Mean	BIC
1	1	0	-5.77633175411793
0	1	0	-5.77595966549108
1	0	0	-5.77593086399818
1	1	1	-5.77577130613306
0	0	1	-5.77543055261876
0	1	1	-5.77493924011277
1	0	1	-5.77490639552537
0	2	0	-5.77416357632257
2	0	0	-5.77413640305256
0	2	1	-5.77316666274965
2	0	1	-5.77313340468199
3	1	0	-5.77278866452028
1	3	0	-5.77278772373624
3	3	0	-5.77259631527509
2	3	0	-5.77245496030893
3	2	0	-5.77242430377938
0	3	0	-5.7723462635619

Figure B.88: Performance of different ARMA orders for MSCI US SC's returns with Generalized Hyperbolic innovations and ACF and PACF of MSCI US SC's returns

MSCI EUR SC

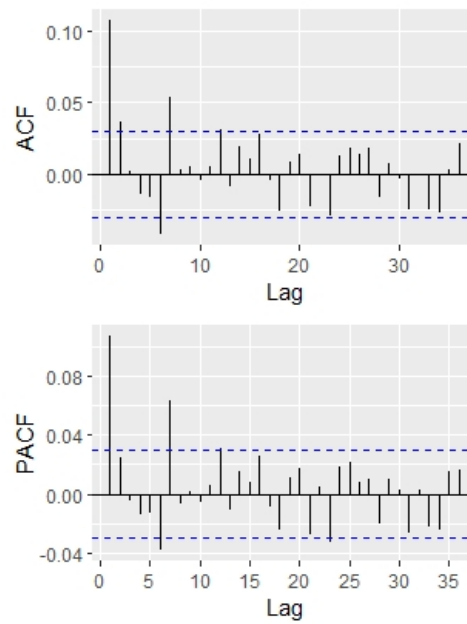| AR	MA	Mean	BIC
0	1	0	-6.27067783571173
1	0	0	-6.27056737918154
0	1	1	-6.26944831587528
1	0	1	-6.26933165130454
2	0	0	-6.26883231287883
0	2	0	-6.26877519399834
1	1	0	-6.26876630117851
2	1	0	-6.26867273086851
2	1	1	-6.26774484628051
2	0	1	-6.26761422885936
0	2	1	-6.26755123145767
1	1	1	-6.26754010014718
2	2	0	-6.26747164583972
3	0	0	-6.26734480762445
0	3	0	-6.26719365594315
3	1	0	-6.26699717303965
1	2	0	-6.26685848668941

Figure B.89: Performance of different ARMA orders for MSCI Eur SC's returns with Skewed-Student innovations and ACF and PACF of MSCI Eur SC's returns

FI Global Cap

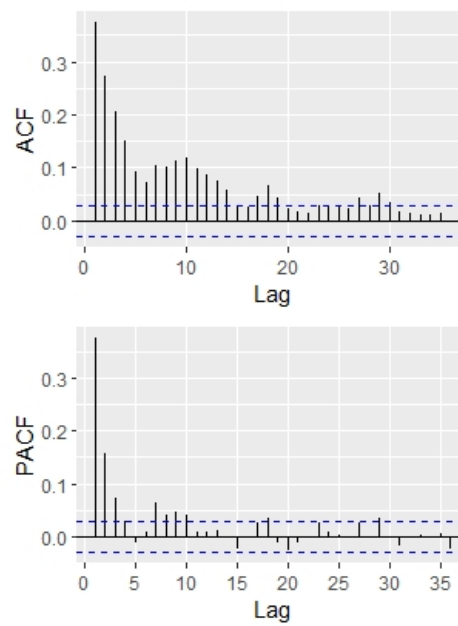| AR	MA	Mean	BIC
2	1	0	-9.40190248217608
2	2	0	-9.40174555455387
1	1	0	-9.40155791134537
3	1	0	-9.4013730275166
1	2	0	-9.40117803146088
1	3	0	-9.40091135726525
2	1	1	-9.40065890482668
1	1	1	-9.40050546875337
2	2	1	-9.40043325770421
3	1	1	-9.40011559307909
1	2	1	-9.39997267685779
2	3	0	-9.39985141895029
3	2	0	-9.39975539186198
1	3	1	-9.39963425775494
2	3	1	-9.39854176873524
3	2	1	-9.3985124743261
3	0	0	-9.39677450594196

Figure B.90: Performance of different ARMA orders for FI Global Cap's returns with Normal Inverse Gaussian innovations and ACF and PACF of FI Global Cap's returns

FI ICE

| AR	MA	Mean	BIC
1	1	1	-9.18111562554716
2	0	1	-9.18097432903011
1	1	0	-9.1809130542954
2	0	0	-9.18068928979622
3	0	1	-9.1793480887335
2	1	1	-9.17932847772071
1	2	1	-9.17918582257626
3	1	0	-9.17911538905657
2	1	0	-9.17911213346945
3	0	0	-9.17911112445735
0	3	1	-9.17910044784807
1	2	0	-9.17898655837192
0	3	0	-9.17873046831562
3	1	1	-9.17861800375553
1	0	1	-9.17826137924918
2	2	1	-9.17777475947028
1	0	0	-9.17771422250522
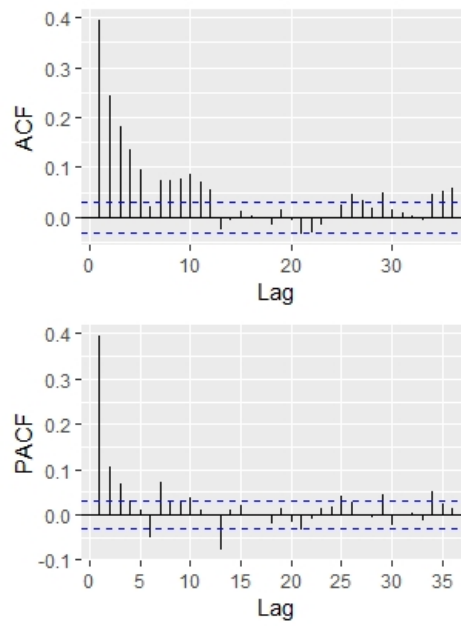
Figure B.91: Performance of different ARMA orders for FI ICE's returns with Normal Inverse Gaussian innovations and ACF and PACF of FI ICE's returns

C Performance of different ARMA orders of ARMA models residuals for GARCH selection

Russell 2000

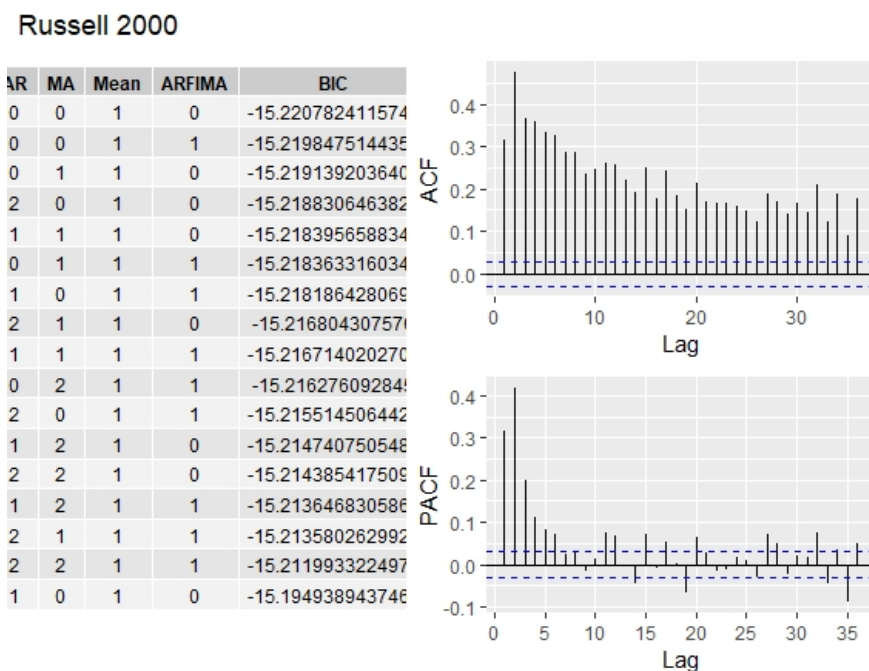| AR	MA	Mean	ARFIMA	BIC
0	0	1	0	-15.220782411574
0	0	1	1	-15.219847514435
0	1	1	0	-15.219139203640
2	0	1	0	-15.218830646382
1	1	1	0	-15.218395658834
0	1	1	1	-15.218363316034
1	0	1	1	-15.218186428069
2	1	1	0	-15.216804307570
1	1	1	1	-15.216714020270
0	2	1	1	-15.216276092849
2	0	1	1	-15.215514506442
1	2	1	0	-15.214740750548
2	2	1	0	-15.214385417509
1	2	1	1	-15.213646830586
2	1	1	1	-15.213580262992
2	2	1	1	-15.211993322497
1	0	1	0	-15.194938943746

Figure C.92: Performance of different ARMA orders for Russell 2000's squared residuals after fitting ARMA(0,1) with Generalized Hyperbolic innovations. ACF and PACF of Russell 2000's squared residuals after fitting ARMA(0,1) with Generalized Hyperbolic innovations

Russell 3000

AR	MA	Mean	ARFIMA	BIC
0	0	1	0	-16.360984789801
0	0	1	1	-16.359659109030
0	1	1	0	-16.358980885271
0	1	1	1	-16.358968215623
1	1	1	0	-16.358171381349
0	2	1	0	-16.357811042925
2	0	1	0	-16.357455993778
1	1	1	1	-16.356195977967
2	1	1	0	-16.356114545937
2	0	1	1	-16.355974269154
2	2	1	0	-16.35454745670
0	2	1	1	-16.354297424165
2	1	1	1	-16.352518735267
1	2	1	1	-16.352485968735
1	2	1	0	-16.351755109329
2	2	1	1	-16.351520200718
1	0	1	1	-16.342773195901
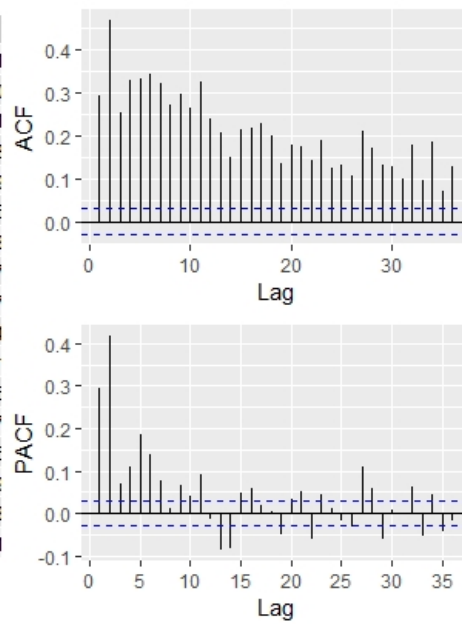
Figure C.93: Performance of different ARMA orders for Russell 3000's squared residuals after fitting ARMA(1,0) with Generalized Hyperbolic innovations. ACF and PACF of Russell 3000's squared residuals after fitting ARMA(1,0) with Generalized Hyperbolic innovations
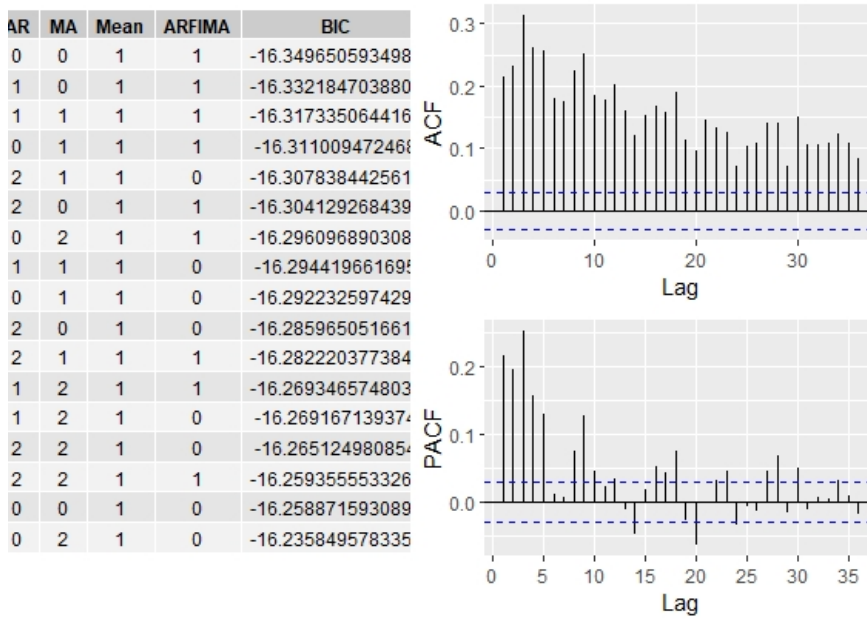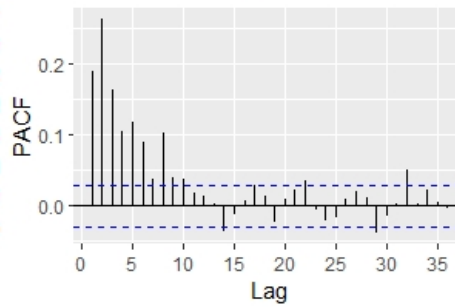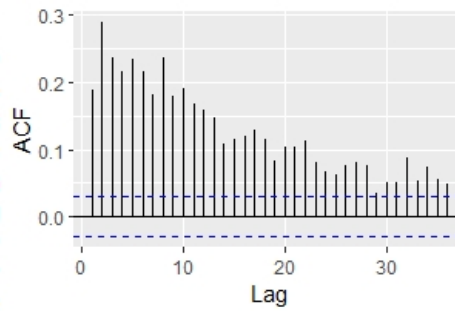
MSCI EUR LC

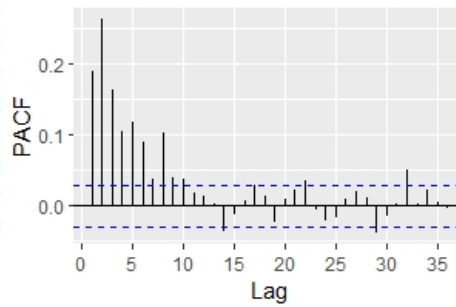| AR	MA	Mean	ARFIMA	BIC
0	0	1	1	-16.349650593498
1	0	1	1	-16.332184703880
1	1	1	1	-16.317335064416
0	1	1	1	-16.31100947246
2	1	1	0	-16.307838442561
2	0	1	1	-16.304129268439
0	2	1	1	-16.296096890308
1	1	1	0	-16.29441966169
0	1	1	0	-16.292232597429
2	0	1	0	-16.285965051661
2	1	1	1	-16.282220377384
1	2	1	1	-16.269346574803
1	2	1	0	-16.26916713937
2	2	1	0	-16.26512498085
2	2	1	1	-16.259355553326
0	0	1	0	-16.258871593089
0	2	1	0	-16.235849578335

Figure C.94: Performance of different ARMA orders for MSCI Eur LC's squared residuals after fitting ARMA(0,1) with Normal Inverse Gaussian innovations. ACF and PACF of MSCI Eur LC's squared residuals after fitting ARMA(0,1) with Normal Inverse Gaussian innovations

MSCI EUR SC

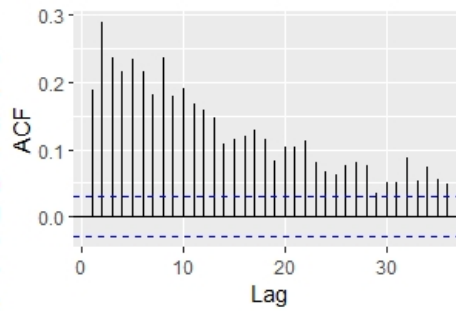| AR	MA	Mean	ARFIMA	BIC
0	1	1	1	-16.291339400677
2	0	1	1	-16.287127342594
1	1	1	1	-16.278802292608
0	2	1	1	-16.272229783724
1	0	1	1	-16.270567437500
1	2	1	1	-16.260072542771
0	0	1	1	-16.259755764407
0	2	1	0	-16.252350164187
1	2	1	0	-16.249017809223
2	2	1	0	-16.242491345052
0	0	1	0	-16.236787980109
2	1	1	0	-16.234259359663
2	1	1	1	-16.229346876914
2	2	1	1	-16.225556379756
0	1	1	0	-16.189879994221
1	1	1	0	-16.141710737269
1	2	0	1	-15.500629836255

Figure C.95: Performance of different ARMA orders for MSCI US SC's squared residuals after fitting ARMA(0,1) with Generalized Hyperbolic innovations. ACF and PACF of MSCI US SC's squared residuals after fitting ARMA(0,1) with Generalized Hyperbolic innovations

xxiv

MSCI EUR SC

AR	MA	Mean	ARFIMA	BIC
0	1	1	1	-16.291339400677
2	0	1	1	-16.287127342594
1	1	1	1	-16.278802292608
0	2	1	1	-16.272229783724
1	0	1	1	-16.270567437500
1	2	1	1	-16.260072542771
0	0	1	1	-16.259755764407
0	2	1	0	-16.252350164187
1	2	1	0	-16.249017809223
2	2	1	0	-16.242491345052
0	0	1	0	-16.236787980109
2	1	1	0	-16.234259359663
2	1	1	1	-16.229346876914
2	2	1	1	-16.225556379756
0	1	1	0	-16.189879994221
1	1	1	0	-16.141710737269
1	2	0	1	-15.500629836255

Figure C.96: Performance of different ARMA orders for MSCI Eur SC's squared residuals after fitting ARMA(0,1) with Skewed-Student innovations. ACF and PACF of MSCI Eur SC's squared residuals after fitting ARMA(0,1) with Skewed-Student innovations

FI Global Cap

| AR	MA	Mean	ARFIMA	BIC
0	0	1	1	-22.52180042560
0	1	1	1	-22.505677489730
2	0	1	1	-22.501763009468
1	2	1	0	-22.49772131136!
2	1	1	1	-22.485763165901
0	1	1	0	-22.482629645618
1	0	1	1	-22.478599296964
1	1	1	0	-22.47406866544
2	1	1	0	-22.46863305719
1	2	1	1	-22.465855636328
2	2	1	0	-22.463531462912
2	0	1	0	-22.458028141342
0	0	1	0	-22.45779147280!
2	2	1	1	-22.455926879817
1	1	1	1	-22.441784373898
1	0	1	0	-22.413471370898
0	0	0	1	-21.827697760236
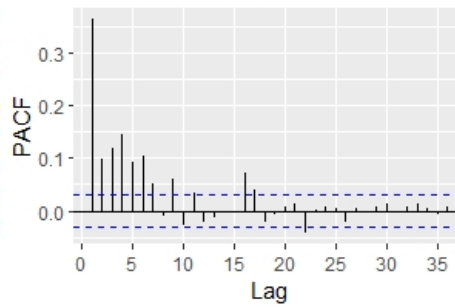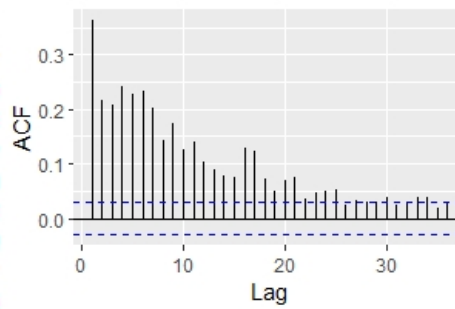
Figure C.97: Performance of different ARMA orders for FI Global Cap's squared residuals after fitting ARMA(2,1) with Normal Inverse Gaussian innovations. ACF and PACF of FI Global Cap's squared residuals after fitting ARMA(2,1) with Normal Inverse Gaussian innovations

FI ICE

AR	MA	Mean	ARFIMA	BIC
0	1	1	1	-22.16879906551
1	2	1	1	-22.152191846303
2	1	1	1	-22.148703783529
0	2	1	1	-22.145612422215
2	2	1	1	-22.143879318336
1	0	1	1	-22.13722993778
0	0	1	1	-22.120099505548
1	1	1	1	-22.117076323720
2	0	1	1	-22.10658795782
2	1	1	0	-22.08027036515
0	2	1	0	-22.072585067251
2	2	1	0	-22.072382237900
1	2	1	0	-22.06756784047
1	1	1	0	-22.045434564531
2	0	1	0	-22.04474813309
0	0	1	0	-22.039968588443
0	1	1	0	-22.019915106912
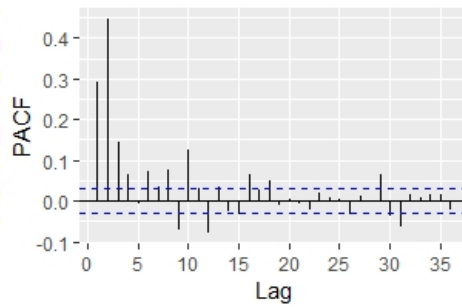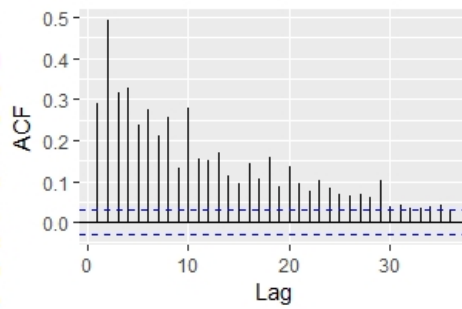
Figure C.98: Performance of different ARMA orders for FI ICE's squared residuals after fitting ARMA(1,1) with Normal Inverse Gaussian innovations. ACF and PACF of FI ICE's squared residuals after fitting ARMA(1,1) with Normal Inverse Gaussian innovations

D   Marginal return distribution selection with QQ-plots
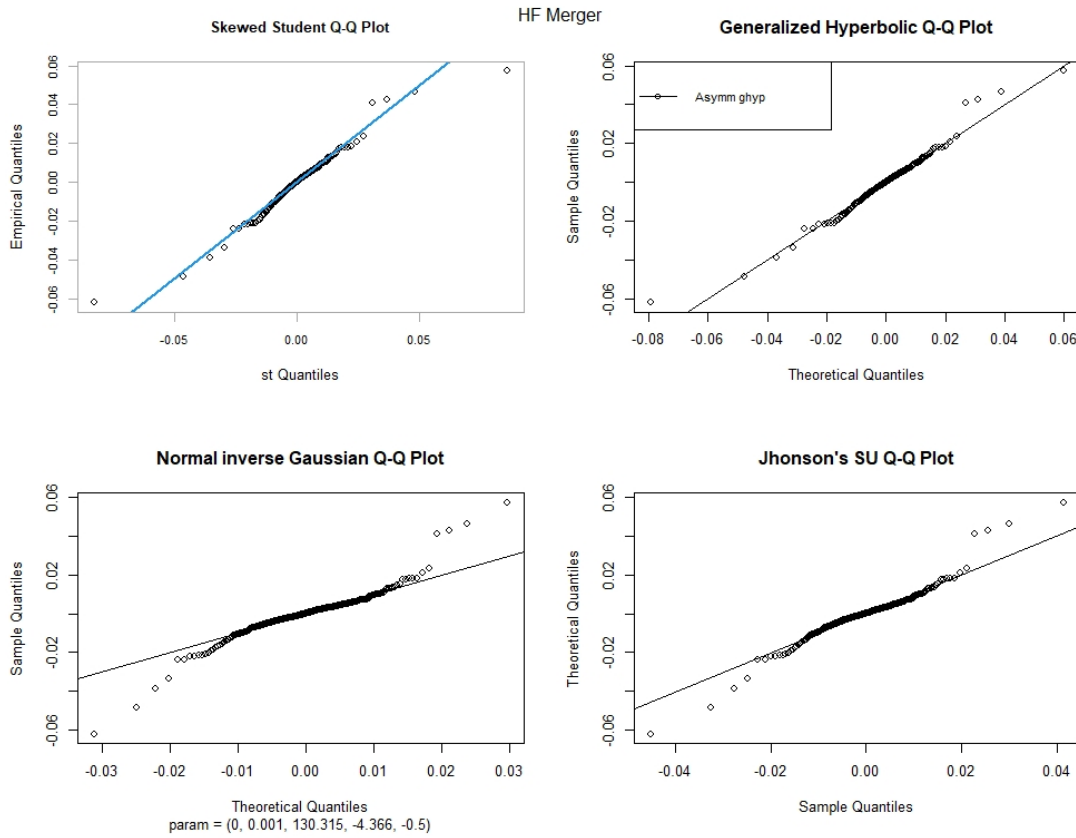
Figure D.99: HF Merger returns QQ-plots with sstd, ghyp, nig and jsu distributions.
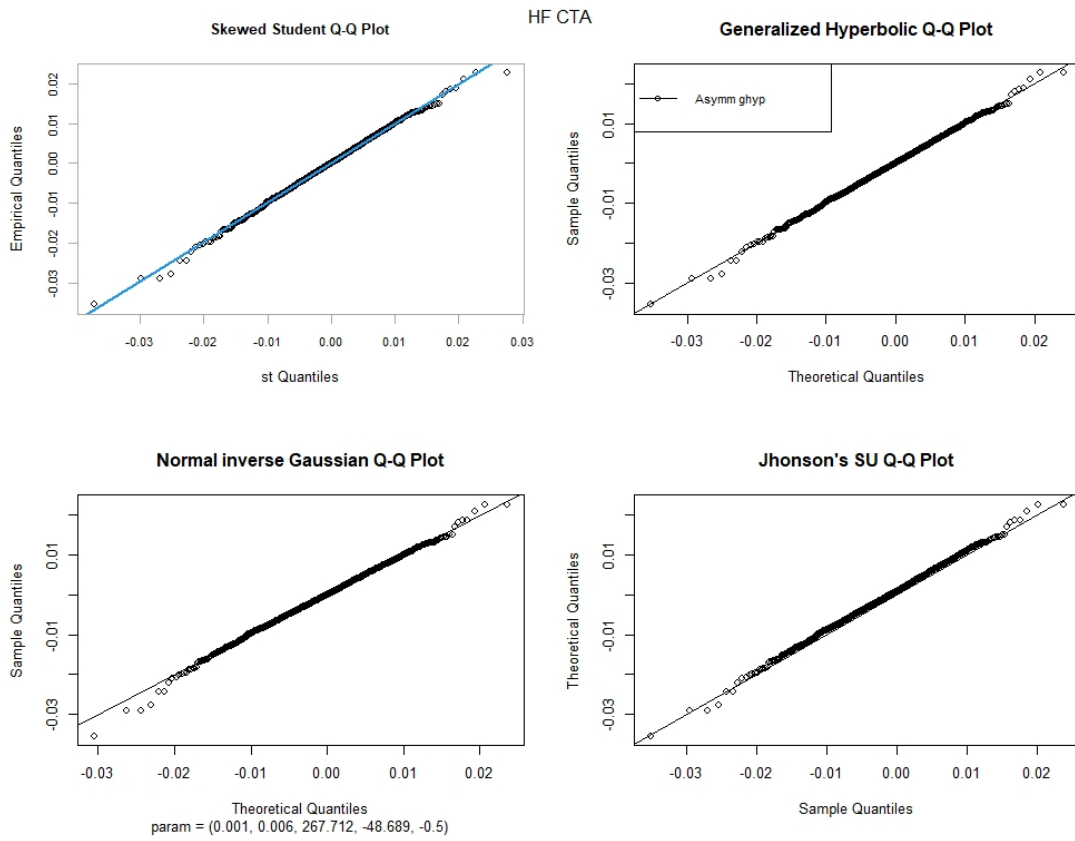
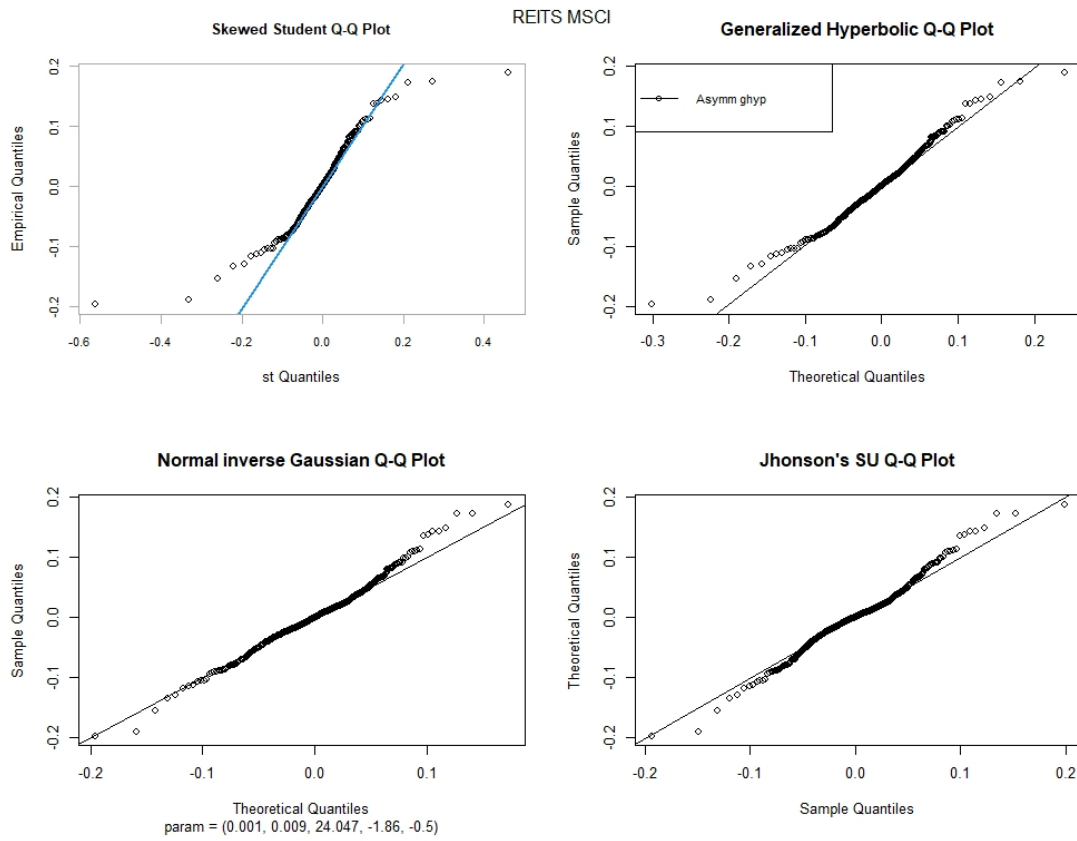Figure D.100: HF CTA returns QQ-plots with sstd, ghyp, nig and jsu distributions.

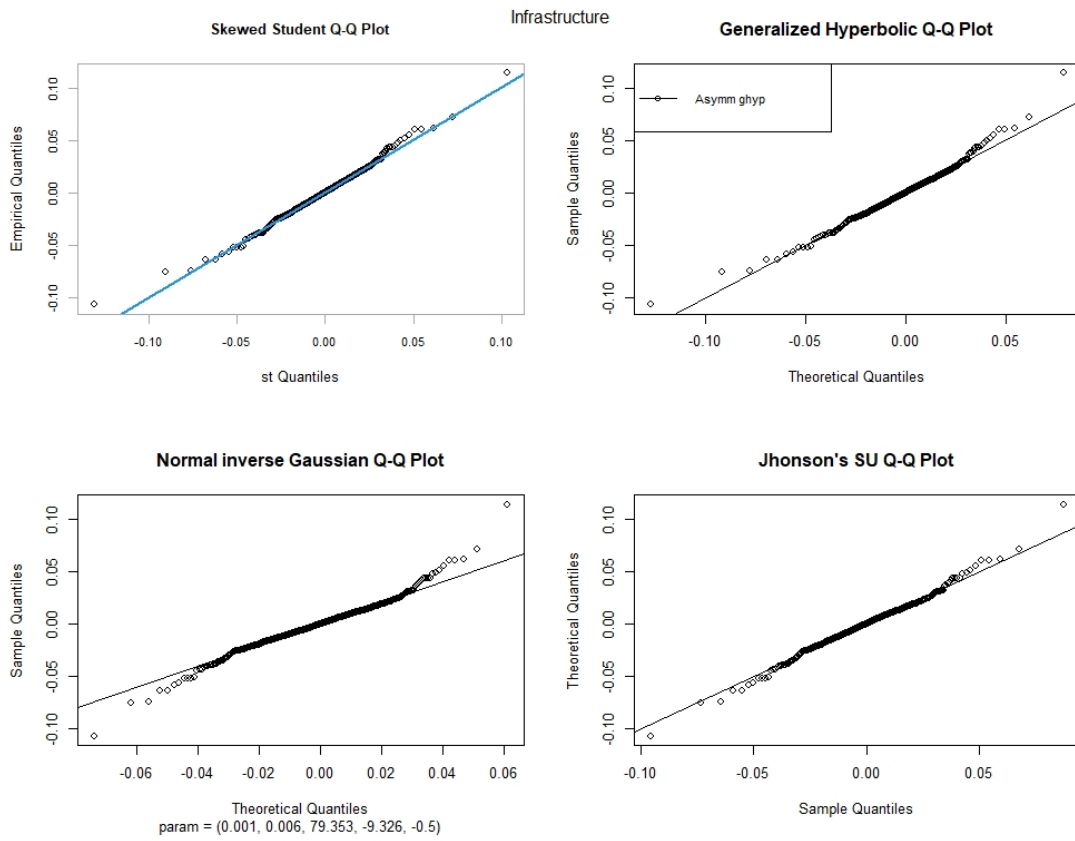Figure D.101: REITS MSCI returns QQ-plots with sstd, ghyp, nig and jsu distributions.

Figure D.102: Infrastructure returns QQ-plots with sstd, ghyp, nig and jsu distributions.
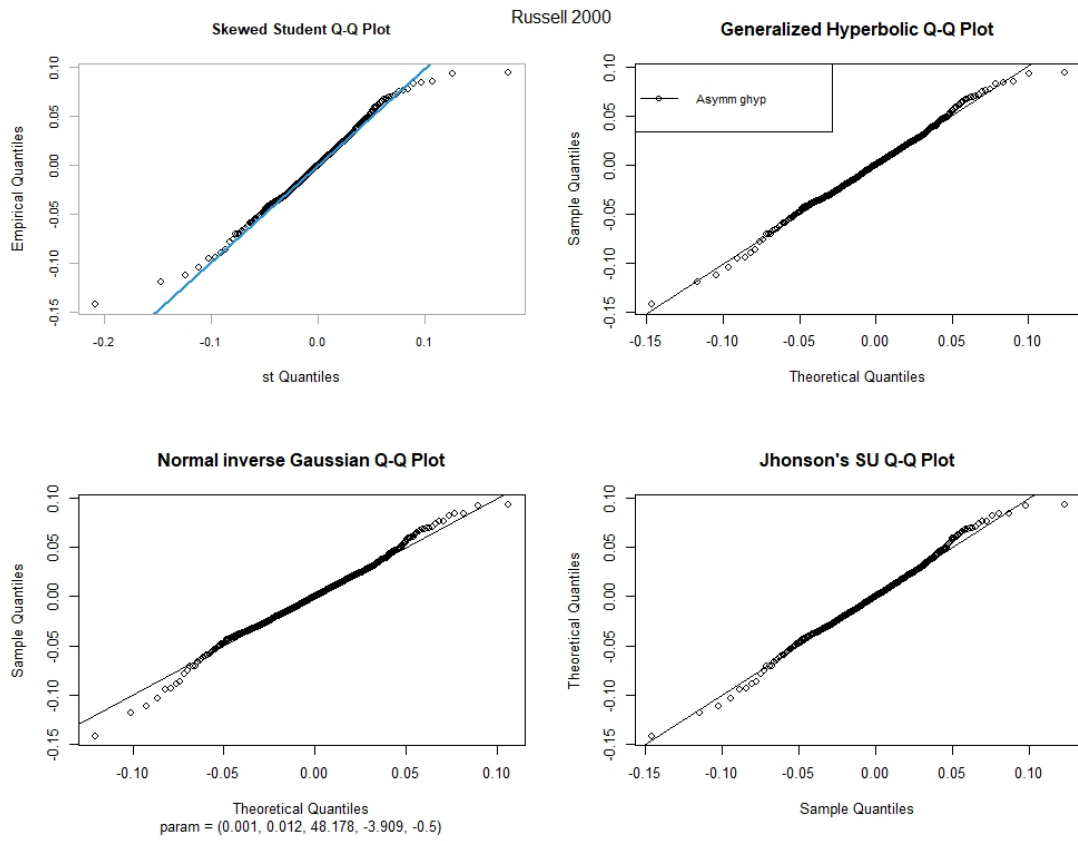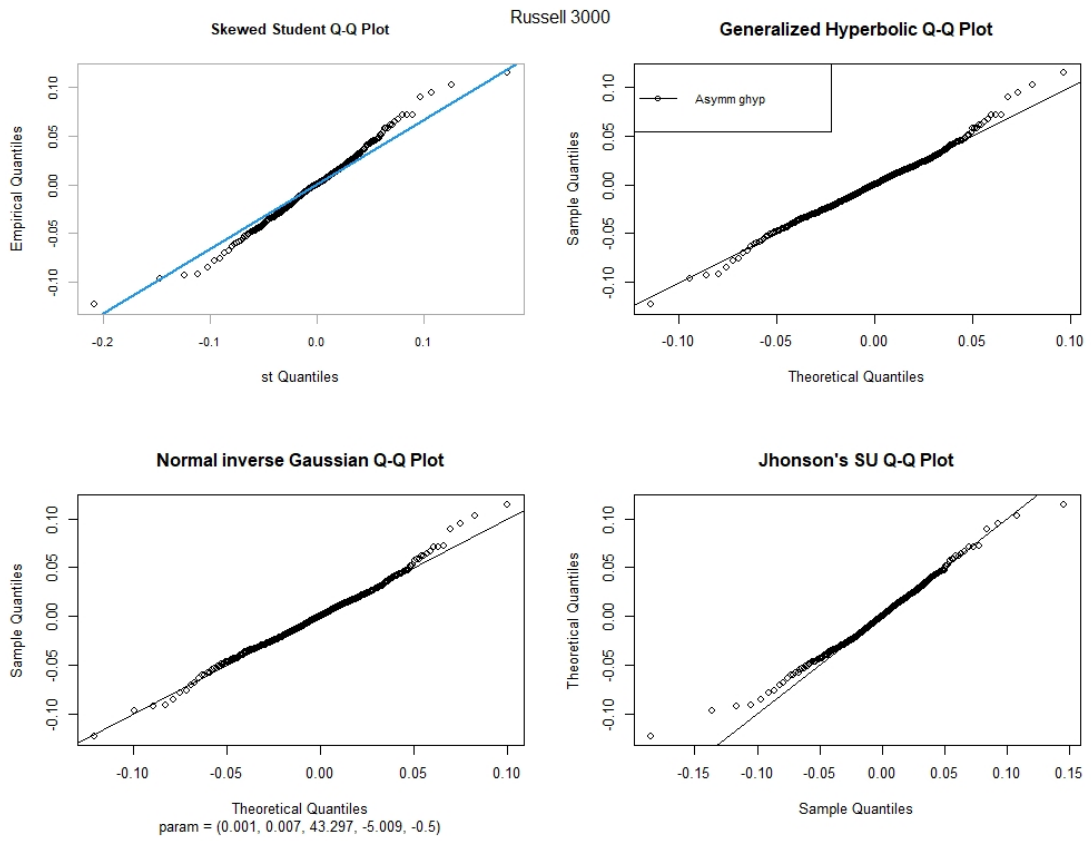
Figure D.103: Russell 2000 returns QQ-plots with sstd, ghyp, nig and jsu distributions.

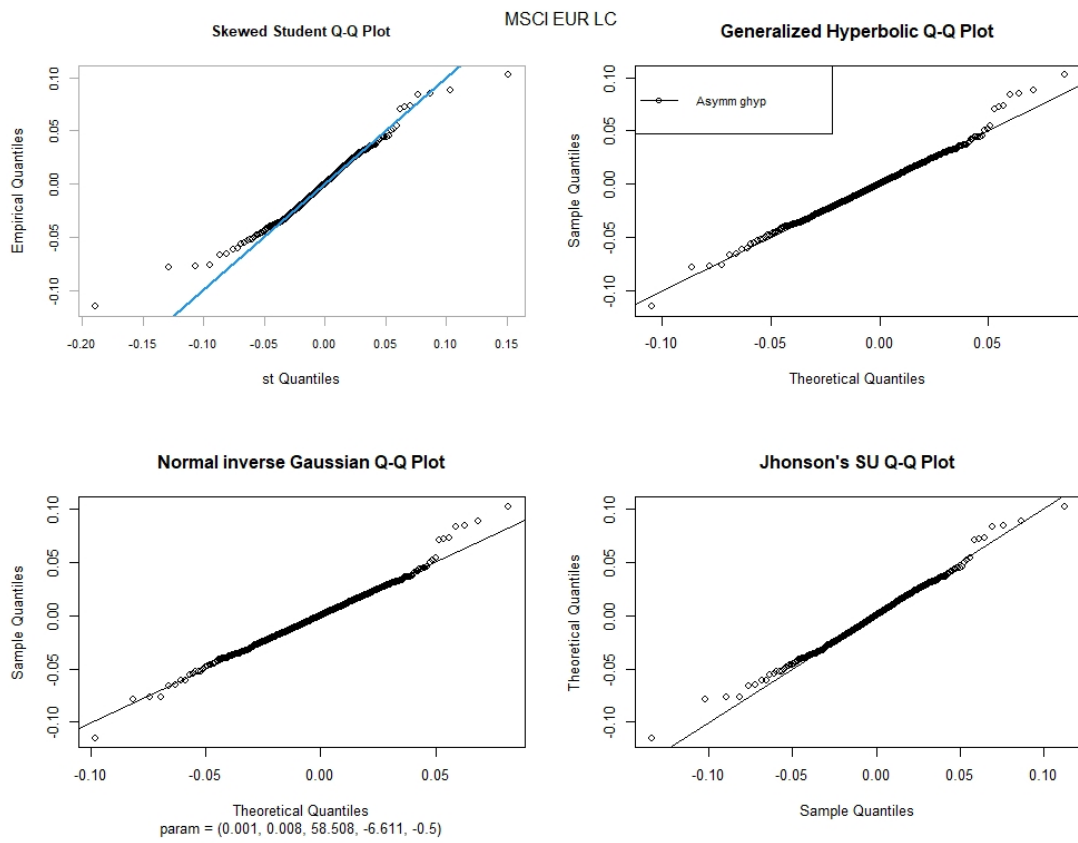Figure D.104: Russell 3000 returns QQ-plots with sstd, ghyp, nig and jsu distributions.

Figure D.105: MSCI Eur LC returns QQ-plots with sstd, ghyp, nig and jsu distributions.
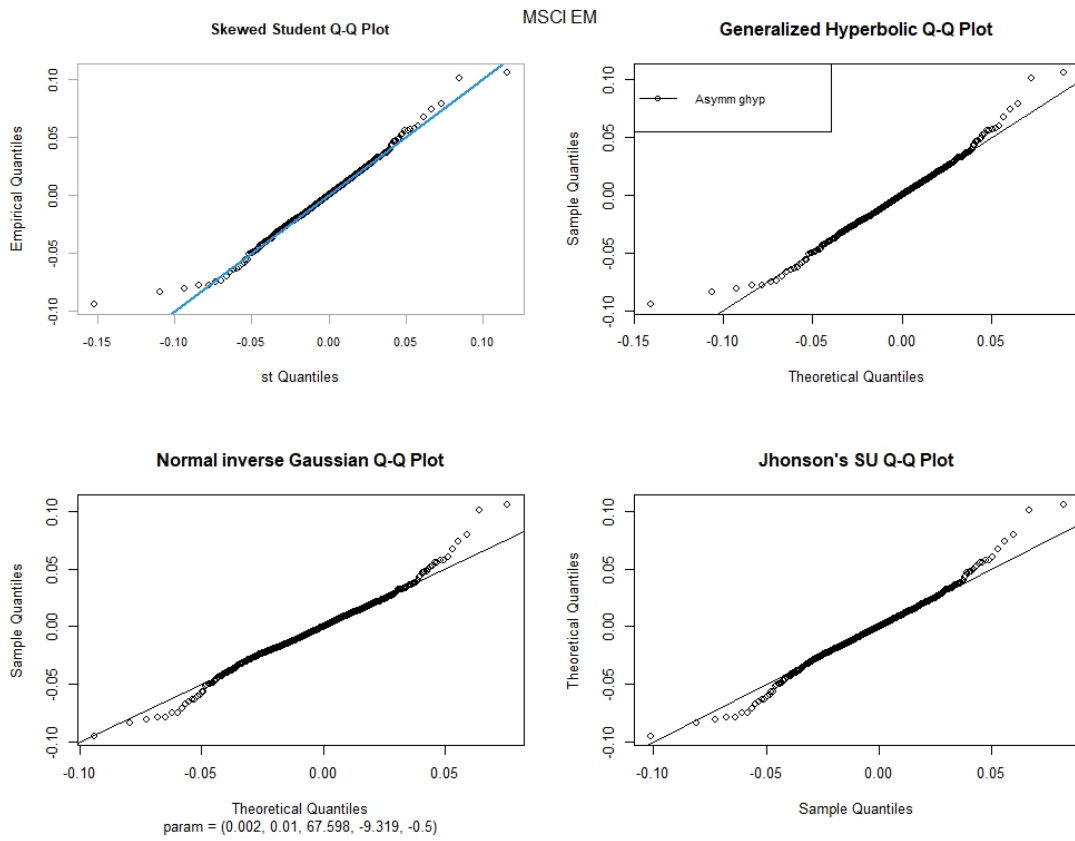
Figure D.106: MSCI EM returns QQ-plots with sstd, ghyp, nig and jsu distributions.
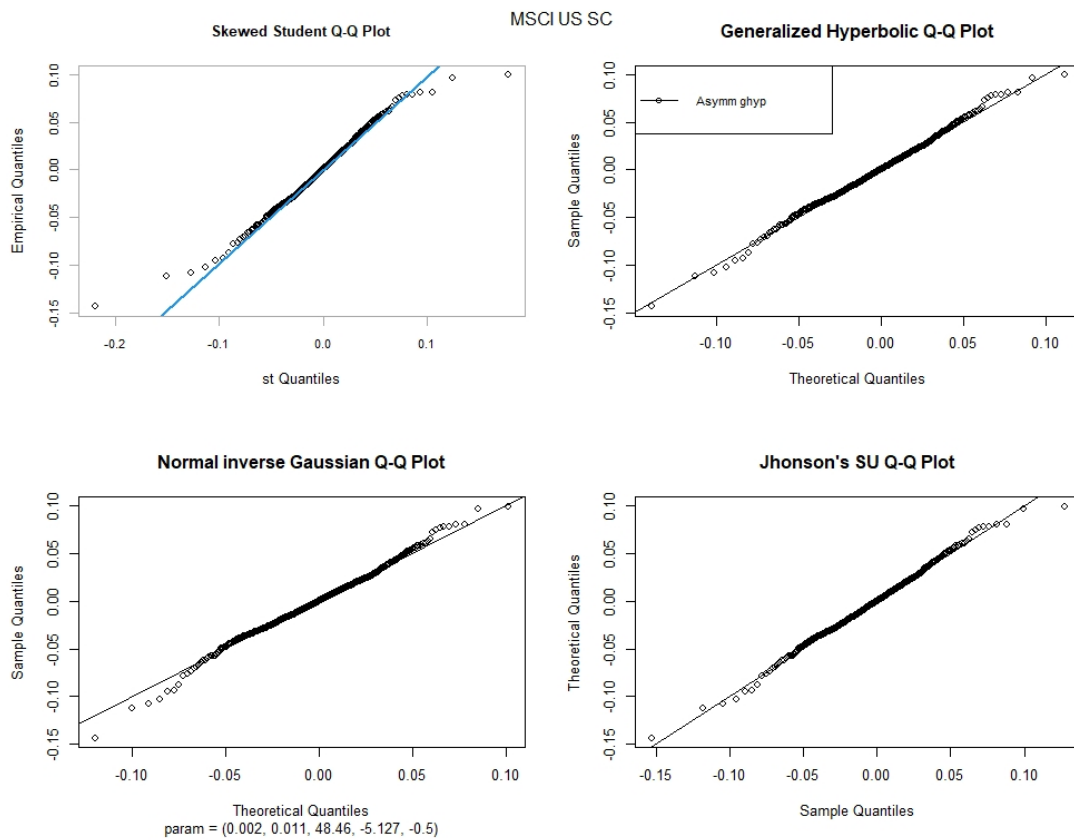
Figure D.107: MSCI US SC returns QQ-plots with sstd, ghyp, nig and jsu distributions.
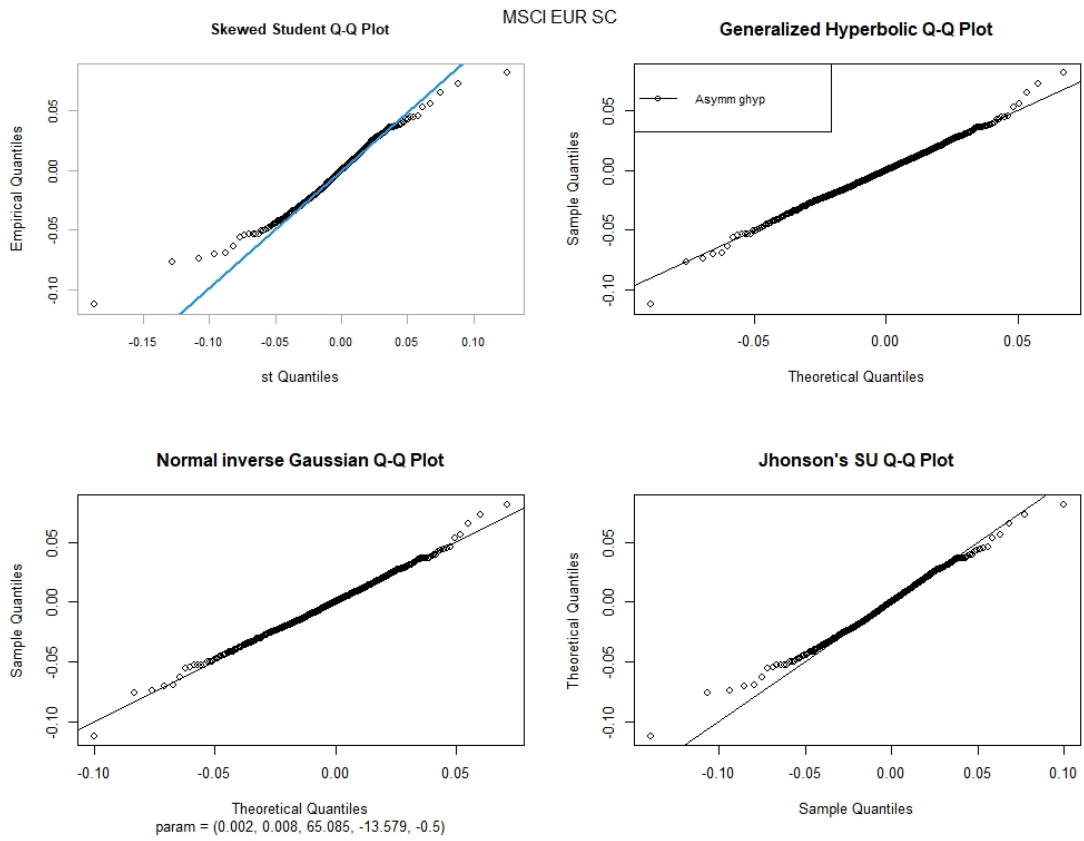
Figure D.108: MSCI Eur SC returns QQ-plots with sstd, ghyp, nig and jsu distributions.
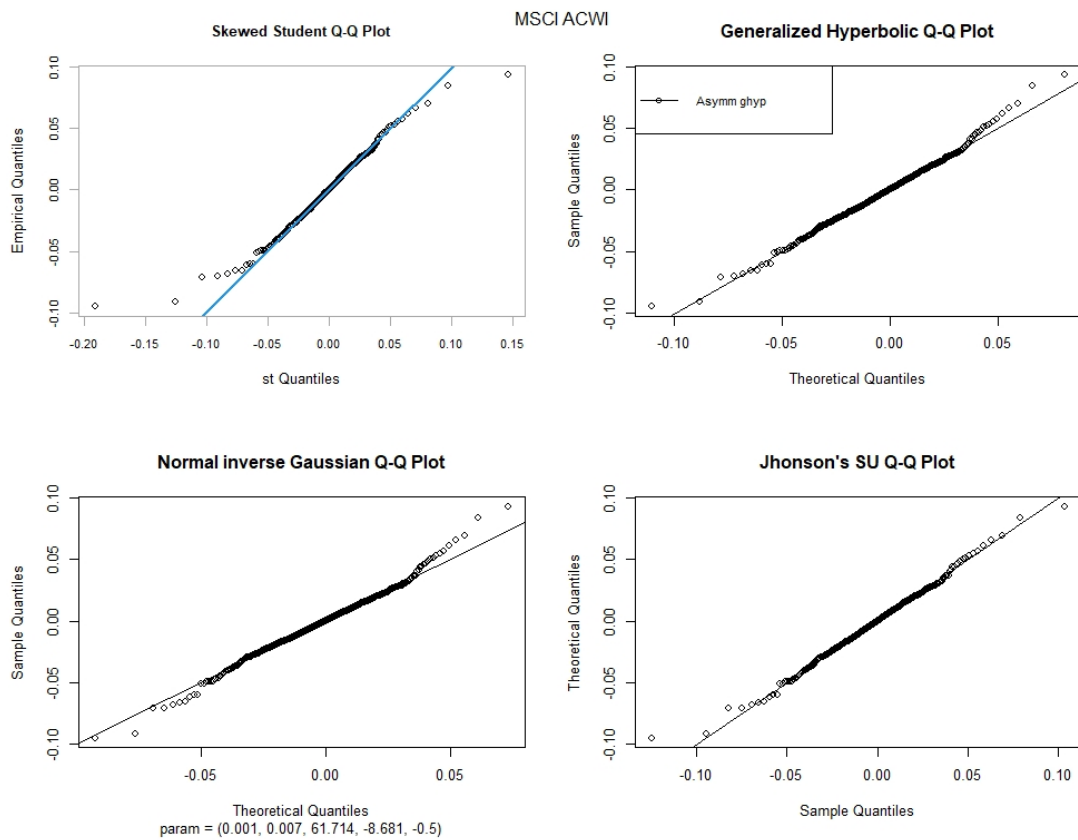
Figure D.109: MSCI ACWI returns QQ-plots with sstd, ghyp, nig and jsu distributions.
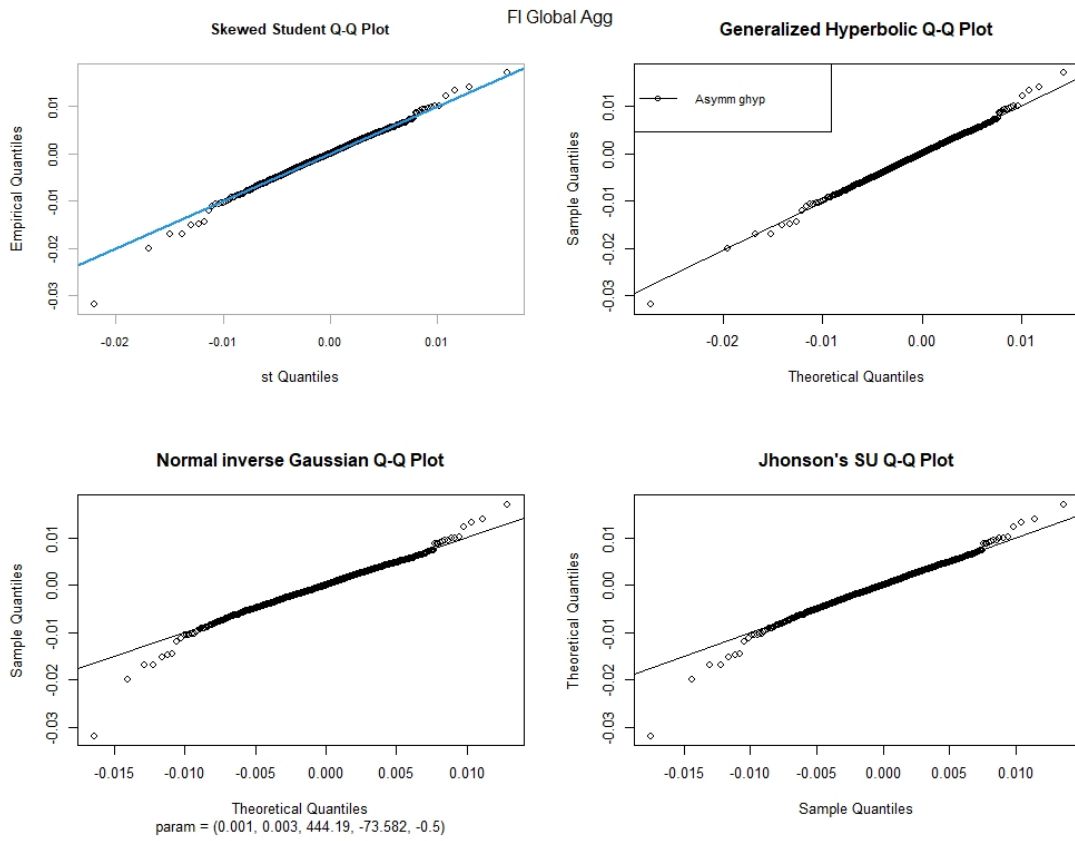
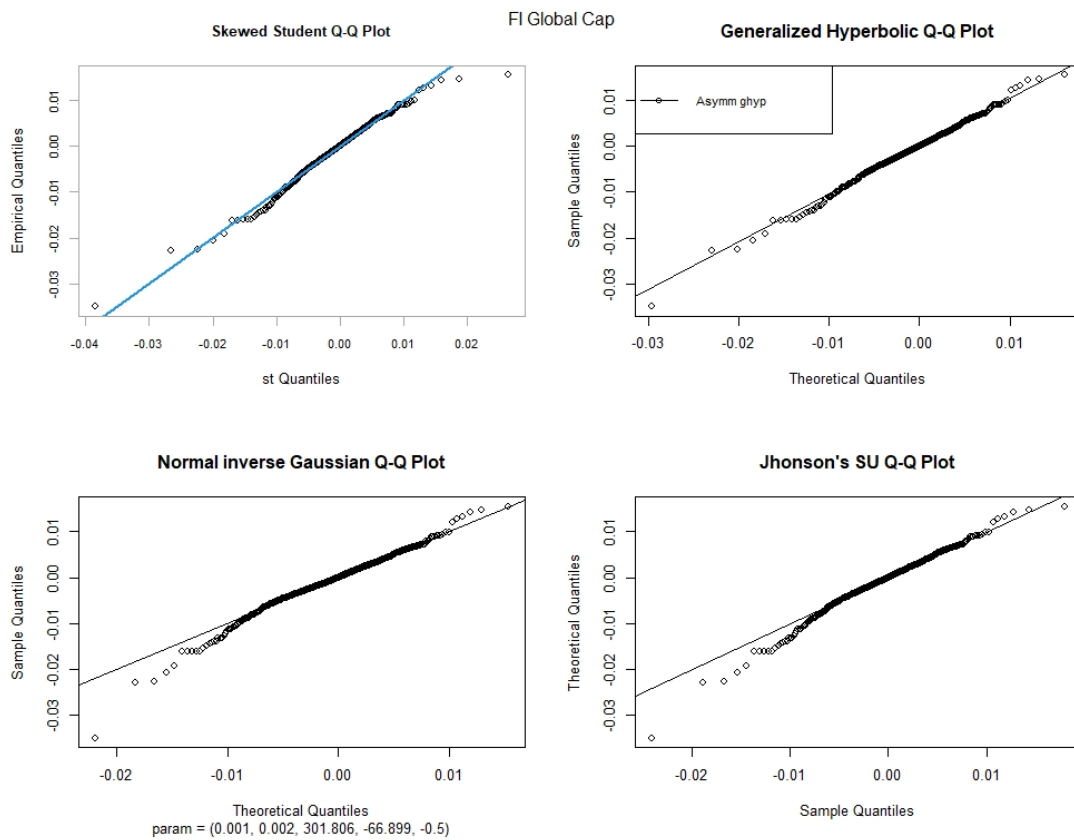Figure D.110: FI Global Agg returns QQ-plots with sstd, ghyp, nig and jsu distributions.

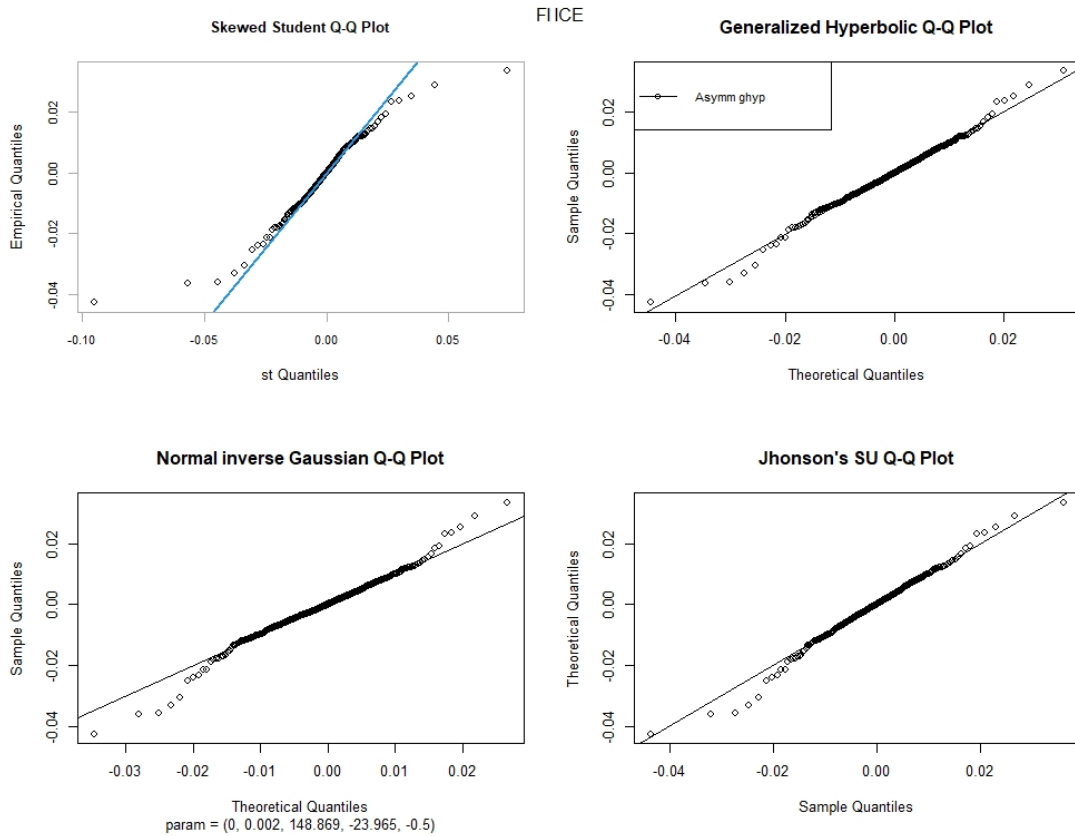Figure D.111: FI Global Cap returns QQ-plots with sstd, ghyp, nig and jsu distributions.

Figure D.112: FI ICE returns QQ-plots with sstd, ghyp, nig and jsu distributions.

References

Akaike, Hirotogu (1973). "Information theory and an extension of the maximum likelihood principle". In: *Proceeding of the Second International Symposium on Information Theory*, pp. 267–281.

Angus, John E. (1994). "The probability integral transform and related results". In: *SIAM Review* 36(4), pp. 652–654.

Artzner, Philippe et al. (1999). "Coherent measures of risk". In: *Mathematical Finance* 9(3), pp. 203–292.

Bartlett, M. S. (1946). "On the theoretical specification and sampling properties of autocorrelated time series". In: *Journal of the Royal Statistical Society* 8(1), pp. 27–41.

Bayer, Sebastian and Timo Dimitriadis (2020). *esback: Expected Shortfall Backtesting*. R package. Version 0.3.0.

Beare, Brendan K. and Juwon Seo (2015). "Vine copula specifications for stationary multivariate Markov Chains". In: *Journal of Time Series Analysis* 36(2), pp. 228–246.

Bloomberg, L.P. (2023). *Bloomberg*. URL: https://www.bloomberg.com/.

Bollerslev, Tim (1986). "Generalized autoregressive conditional heteroskedasticity". In: *Journal of Econometrics* 31, pp. 307–327.

Box and Gwilym M. Jenkins (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day.

Brechmann, Eike Christian and Claudia Czado (2015). "COPAR—multivariate time series modeling using the copula autoregressive model". In: *Applied Stochastic Models in Business and Industry* 31(4), pp. 495–514.

Catania, Leopoldo et al. (2022). *GAS: Generalized Autoregressive Score Models*. R package. Version 0.3.4.

Christoffersen, Peter F. (1998). "Evaluating interval forecasts". In: *International Economic Review* 39(4), pp. 841–862.

Cont, Rama (2001). "Empirical properties of asset returns: stylized facts and statistical issues". In: *Quantitative Finance* 1, pp. 223–236.

Czado, Claudia (2019). *Analyzing dependent data with vine copulas: A practical guide with R*. Springer. ISBN: 9783030137847.

Dißmann, J. et al. (2013). "Selecting and estimating regular vine copulae and application to financial returns". In: *Computational Statistics and Data Analysis* 59, pp. 52–69.

Efron, Bradley and Robert J. Tibshirani (1993). *An introduction to the bootstrap*. Chapman Hall.

Embrechts, Paul and Ruodu Wang (2015). "Seven proofs for the subadditivity of expected Shortfall". In: *DE GRUYTER OPEN* (3), pp. 126–140.

Engle, Robert F. (1982). "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation". In: *Econometrica* 50(4), pp. 987–1008.

Euronext (2023). *Euronext IEIF REIT Europe Index*. URL: https://live.euronext.com/en/product/indices/QS0011070230-XPAR/market-information.

Fleishman, Allen I. (1978). "A method for simulating non-normal distributions". In: *Psychometrika* 43(4), pp. 521–532.

Galanos, Alexios (2022). *rugarch: Univariate GARCH Models*. R package. Version 1.4-8.

Group, Capital (2023). *Capital Group Global Total Return Bond Fund (LUX)*. URL: https://www.capitalgroup.com/intermediaries/gb/en/investments/fund-centre.CGGTRLU.html#?id.

Hofert, Marius et al. (2022). *copula: Multivariate Dependence with Copulas*. R package. Version 1.1-0.

Inc, Hedge Fund Research (2023). *HFRX Index*. URL: https://www.hfr.com/hfrx-index-methodology.

Inc, MSCI (2023a). *MSCI ACWI Index*. URL: https://www.msci.com/documents/10199/a71b65b5-d0ea-4b5c-a709-24b1213bc3c5.

Inc, MSCI (2023b). *MSCI Emerging Market Index*. URL: https://www.msci.com/documents/10199/c0db0a48-01f2-4ba9-ad01-226fd5678111.

Inc, MSCI (2023c). *MSCI Europe Large Cap Index*. URL: https://www.msci.com/documents/10199/949c4734-7267-448b-aa28-a8012cb48b54.

Inc, MSCI (2023d). *MSCI Europe Small Cap Index*. URL: https://www.msci.com/documents/10199/a2bd7d9f-6c01-4056-bbf6-f1d9074366e0.

Inc, MSCI (2023e). *MSCI US REIT Index*. URL: https://www.msci.com/documents/10199/08f87379-0d69-442a-b26d-46f749bb459b.

Inc, MSCI (2023f). *MSCI USA Small Cap Index*. URL: https://www.msci.com/documents/10199/8038650a-0e6f-43d5-bdb0-1f8f3063e565.

Inc, MSCI (2023g). *MSCI World Infrastructure Index*. URL: https://www.msci.com/documents/10199/91cfa8b6-4734-4c68-b14f-a820098d2d58.

Kaiser, Henry F. and Kern Dickman (1962). "Sample and population score matrices and sample correlation matrices from an arbitrary population correlation matrix". In: *Psychometrika* 27(2), pp. 179–182.

Ljung, G. M. and G. E. P. Box (1978). "On a measure of lack of fit in time series models". In: *Biometrika* 65(2), pp. 297–303.

Maarouf (2021). "Backtesting Value-at-Risk of Financial Data Using Vine Copulas". MA thesis. Technical University of Munich.

Mandelbrot, Benoit (1963). "The variation of certain speculative prices". In: *Chicago journals* 36(4), pp. 394–419.

McNeil, Alexander J. and Rüdiger Frey (2000). "Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach". In: *Journal of Empirical Finance* 7, pp. 271–300.

Nagler, Thomas (2022). *svines: Stationary Copula Models*. R package. Version 0.1.4.

Nagler, Thomas and Thibault Vatter (2022). *rvinecopulib: High Performance Algorithms for Vine Copula Modeling*. R package. Version 0.6.2.1.1.

Nagler, Thomas et al. (2022). "Stationary vine copulas models for multivariate time series". In: *Journal of Econometrics* 227, pp. 305–324.

Nelsen, Roger B. (2006). *An introduction to copulas*. Springer. ISBN: 0387286594.

Nolde, Natalia and Johanna F. Ziegel (2017). "Elicitability and backtesting: perspectives for banking regulation". In: *The Annals of Applied Statistics* 11(4), pp. 1833–1874.

Russell, FTSE (2023). *FTSE Russell factsheets*. URL: https://www.ftserussell.com/analytics/factsheets/home/search.

Schwarz, Gideon (1978). "Estimating the dimension of a model". In: *The Annals of Statistics* 6(2), pp. 461–464.

Shumway, Robert H. and David S. Stoffer (2017). *Time Series Analysis and its Applications with R Examples*. Springer.

Sklar, A. (1959). "Fonctions de répartition à n dimensions et leurs marges". In: 8, pp. 229–231.

Smith, Michael Stanley (2015). "Copula modelling of dependence in multivariate time series". In: *International Journal of Forecasting* 31(3), pp. 815–833.

Sommer (2022). "An unconditional and conditional rolling window approach". MA thesis. Technical University of Munich.

Sommer, Emanuel (2022). *portvine:Vine Based (Un)Conditional Portfolio Risk Measure Estimation*. R package. Version 1.0.1.

Sommer, Emanuel et al. (2023). *Vine Copula based portfolio level conditional risk measure forecasting*. URL: https://arxiv.org/abs/2208.09156.

Tsay, Ruey S. (2010). *Analysis of financial time series*. Wiley.

Vale, C. David and Vincent A. Maurelli (1983). "Simulating multivariate nonnormal distributions". In: *Psychometrika* 48(3), pp. 465–471.

Vuong, Quang H. (1989). "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses". In: *Econometrica* 57(2), pp. 307–333.

Wei, William W. S. (2005). *Time series analysis: univarite and multivariate methods*. Pearson Education. ISBN: 0321322169.