

# COPULA APPROACH TO FITTING BIVARIATE TIME SERIES

JUN WANG

Master's thesis  
2023:E56



LUND UNIVERSITY

Faculty of Science  
Centre for Mathematical Sciences  
Mathematical Statistics

# Abstract

We apply the GARCH-copula method to estimate Value at Risk (VaR) for European and Swedish stock indices. First, marginal distributions are estimated by the ARMA-GARCH model with normal, Student's  $t$ , and skewed  $t$  distributions. Then we investigate the tails of innovations of ARMA-GARCH models using the Peaks over thresholds method and find that the distributions of stock returns are asymmetric with heavier left tails than right tails. In order to model the dependence between the time series, we try elliptical copulas (Gaussian, Student- $t$ ) and Archimedean copulas (Gumbel, Frank, and Clayton) to measure the dependence structure between two time series' returns. Parameter estimation is based on the so-called inference for margins, which is a two-stage method. Moreover, we adopt backtesting to test the goodness-of-fit of different copulas using Monte-Carlo simulations.

Our empirical results show that ARMA(1,1)-GARCH(1,1)- $t$  distribution is proven to be the best fits for margins and Student's  $t$  copula gives the highest log-likelihood of the model and best VaR estimation.

**Keywords:** VaR, Copula, ARMA-GARCH, Extreme Value Theory, GPD, Hill estimator

## Acknowledgements

I would like to thank my supervisor, Nader Tajvidi. I could not have undertaken the journey without his support and patience. I appreciate his help in finding an interesting topic and leading me to quantitative finance.

I am also grateful to my parents and friends, who gave me great encouragement during the entire process.

# Contents

<b>1 Introduction</b>	<b>1</b>
1.1 Background	1
1.2 Structure of the Paper	2
<b>2 Theoretical background</b>	<b>3</b>
2.1 Modeling the Margins	3
2.1.1 Ljung-Box Test	4
2.1.2 Goodness-of-fit Test	4
2.2 Generalized Pareto Distribution	4
2.2.1 Peaks Over Threshold (POT)	5
2.2.2 Modeling the left tail with GPD	6
2.2.3 Model Diagnostics	7
2.3 Copulas	8
2.3.1 Definition and properties	8
2.3.2 Elliptical Copula	9
2.3.3 Archimedean Copula	9
2.3.4 Dependence Measures	11
2.4 Parameter Estimation	12
2.4.1 Full Maximum Likelihood (FML)	12
2.4.2 Inference for Margins (IFM)	13
2.4.3 Pseudo Maximum Likelihood (PML)	13
2.5 Value at Risk	14
2.5.1 Introduction	14
2.5.2 Historical Simulation	14
2.5.3 Variance-Covariance Method	14
2.5.4 Monte Carlo Simulation	15
2.5.5 Backtesting	15
<b>3 Empirical Results</b>	<b>17</b>
3.1 Data	17
3.2 Marginal Modeling	19
3.3 Fitting the tails of residuals	24
3.4 Copula Modelling	28
3.5 Backtesting of VaR	30
3.6 VaR prediction	32
<b>4 Summary</b>	<b>34</b>
4.1 Conclusions and further analysis	34
<b>A ACF plots of residuals after marginal modelings</b>	<b>37</b>
<b>B Mean residual life plots and threshold range plots</b>	<b>40</b>

<b>C Diagnostic plots for GPD fitting</b>	45
<b>D Hill Estimator simulation</b>	49

# 1 Introduction

## 1.1 Background

Risk management has always been an essential part of financial industries. For example, banks are required to hold a certain amount of buffer to protect them from potential losses due to credit risk; investors need quantitative tools to measure market risks of portfolios or assets. Value at Risk (VaR) is one of the most commonly used risk measurements by financial analysts. It is defined as the largest expected loss of an asset under certain conditions. A daily VaR of asset portfolios given a 95% confidence level describes that VaR estimate covers all but the highest 5% losses overtime period of a day [1]. VaR estimates for an individual asset are not difficult to calculate since there are plenty of univariate distributions to be fitted to the corresponding time series of an asset. However, VaR estimation of a portfolio consisting of more than two assets requires the joint multivariate distribution of the portfolio. The dependence measure between assets makes it difficult to estimate VaR. Abundant literature focuses on three estimation methods including Historical Simulation, the Variance-Covariance approach, and the Monte Carlo method.

The limitation of the historical simulation method is its high dependence on information provided by historical data. Problems would arise when the number of data is not large enough or historical extreme events are unavailable [2]. Moreover, the assumption of a portfolio following a multivariate normal distribution and of the linear relationship between asset returns proposed by the Variance-Covariance approach is unrealistic for financial time series [2].

Monte-Carlo method implements parametric models to construct multivariate distributions for portfolios. Copula theory allows us to decompose a joint distribution into marginal distributions and a copula function which is used to model the dependence structure among margins. It was first introduced by Sklar [3] and applied to financial risk management by Embrechts et al [4]. Copulas allow us to build and describe complex dependence structures without limitation of normality and linear relationship.

In order to model marginal distributions, two well-known univariate conditional volatility models, auto-regressive conditional heteroskedasticity (ARCH) proposed by Engle [5] and Bollerslev's generalized ARCH (GARCH) model [6], are widely used to capture volatility clustering of financial data, meaning high(low) volatility accompanied with high(low) volatility for marginal modeling. Huang et al. [7], Jondeau and Rockinger [8] implemented GARCH-Copula model to compute the VaR of the portfolio. Long Kang [9] modeled the joint distribution of returns of four assets using GARCH-Copula with normal, Student- $t$ , Archimedean, and mixed copulas and suggested that Student's Copula gives the highest log-likelihood.

There are different parameter estimation approaches for GARCH-Copula method [10], such as a two-step maximum likelihood method [11] where marginal distributions are fitted first and then estimate copula parameters, non-parametric Pseudo maximum likelihood [12] using empirical distributions instead of parametric models. Patton [13] summarized that a multi-stage estimator is more efficient than a one-stage estimator in most cases when estimating parameters using the Copula method .

Since VaR is an estimate of the tails of distributions, the accuracy of tail estimation becomes significant for risk measurement. Peaks over thresholds is a statistical method to investigate extreme events and a standard approach for the estimation of fat-tailed distributions. Mcneil [14] combined the GARCH model fitted to marginal distributions and EVT to estimate the tail of innovations of the GARCH model. Nyström et al. [15] applied EVT to the filter conditional residuals from ARMA-GARCH models and analyzed the tail behavior using generalized Pareto distribution. Hotta et al. [16] used GPD to give a better fit of the heavy left tail of marginal distributions.

Jondeau and Rockinger [8] argue that the dependency featured more persistence and it was higher between European stock markets than the dependency between the U.S. and European markets. This thesis aims to combine the GARCH-Copula method with the Peaks over thresholds method to find the best fit for bivariate financial times series and use it to estimate Value at Risk.

## 1.2 Structure of the Paper

The remainder of this report is organized as follows. Section 2 introduces the theory of VaR estimation and backtesting, Section 3 presents the marginal models, copula functions, and extreme value theory, Section 4 provides empirical results, and Section 5 concludes the paper with further discussion.

## 2 Theoretical background

To calculate the portfolio's Value at Risk (VaR), we adopted the GARCH-Copula model, which captures the volatility patterns of individual stock returns and their dependence structure. The copula concept is central to this approach as it enables the construction of a multivariate distribution by combining marginal distributions with a copula function. Consequently, to derive a multivariate distribution for the portfolio, we implemented the following two steps: (1) Use the ARMA-GARCH model to model the univariate marginal distributions for each stock return, and (2) Utilize the Copula function to model the dependence structure between the returns.

### 2.1 Modeling the Margins

Let  $r_t$  denote the daily return, then we have

$$r_t = \mu_t + \epsilon_t$$

where  $\mu_t$  is the mean model and  $\epsilon_t$  is the noise model.

In this thesis, the ARMA process was used to fit the mean model. General formula of ARMA(p,q) process is,

$$\mu_t = C + \sum_{i=1}^p \phi_i \mu_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j}.$$

The GARCH model is an extension to the ARMA model which assumes the time series has a constant variance  $\sigma$ , whereas the GARCH process models conditional variances based on previous values. Specifically,

$$\begin{aligned} \sigma_t^2 &= \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \\ \epsilon_t &= z_t \sigma_t, \\ z_t &\sim i.i.d, \end{aligned}$$

where innovations  $z_t$  are selected from Normal, Student- $t$ , and skewed  $t$  distribution with density functions given by.

$$\begin{aligned} f_N(x; \mu, \sigma) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2/\sigma^2}, \\ f_t(x; \nu) &= \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\Gamma(\frac{\nu}{2})} \left(1 + \frac{x^2}{\nu}\right)^{-(\nu+1)/2}, \end{aligned}$$



$$f_{skt}(x; \nu, \xi) = \frac{2\xi}{1 + \xi^2} \left[ f_t(\xi x; \nu) I(x < 0) + f_t\left(\frac{x}{\xi}; \nu\right) I(x \geq 0) \right],$$

where  $\nu$  is the degree of freedom.

To check the presence of asymmetric tails in the residuals of the marginal modeling, we fitted the Generalized Pareto distribution to the left and right tails, respectively, which is discussed below.

### 2.1.1 Ljung-Box Test

Ljung-Box is a statistical test used to check a serial correlation up to a specified lag  $k$ . It determines whether or not the residuals are independent and identically distributed. The null hypothesis is defined as,

$$H_0 : \text{The residuals are independently distributed.}$$

The test statistics is,

$$Q = n(n+2) \sum_{j=1}^k \rho_j^2 / (n-j)$$

where  $n$  is the sample size,  $k$  is the number of lags, and  $\rho_j$  represents the sample autocorrelation at lag  $j$ .  $Q$  follows a chi-squared distribution with  $k$  degree of freedom, i.e.  $Q \sim \chi^2(k)$ . We reject the null hypothesis and conclude that the residuals of the model are not independently distributed if  $Q > \chi_{1-\alpha, k}^2$  with significance level  $\alpha$ .

### 2.1.2 Goodness-of-fit Test

Adjusted Pearson goodness-of-fit test is one type of chi-squared goodness-of-fit that compares the empirical distribution of the standardized residuals with the chosen theoretical ones [17].

$$H_0 : \text{Samples follow the specified distribution.}$$

The null hypothesis will be rejected if the corresponding p-value  $\leq \alpha$ , where  $\alpha$  is the significance level and p-value is determined by the  $\chi^2$  distribution with a suitable of degree of freedom.

## 2.2 Generalized Pareto Distribution

Let  $X_1, X_2, \dots$  be a sequence of independent random variables with common distribution function  $F$ , and let

$$M_n = \max\{X_1, \dots, X_n\}.$$

If there exists sequences of constants  $\{a_n > 0\}$  and  $\{b_n\}$  such that

$$Pr\{(M_n - b_n)/a_n \leq z\} \rightarrow G(z), \quad \text{as } n \rightarrow \infty$$

where  $G$  is a non-degenerate distribution function, then  $G$  belongs to one of the Gumbel, Fréchet, and Weibull families and it could be combined into a family of the form [18]

$$G(z) = \exp \left\{ - \left[ 1 + \xi \left( \frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}.$$

Then for large enough threshold  $u$ , the Generalized Pareto distribution theorem gives the limiting distribution of  $(X - u | X > u)$  as  $u \rightarrow x^F$  where  $x^F = \sup\{x | F(x) < 1\}$  [19]. The distribution function of  $(X - u)$ , conditional on  $X > u$ , is asymptotically

$$H(y) = 1 - \left( 1 + \frac{\xi y}{\tilde{\sigma}} \right)^{-1/\xi} \quad (1)$$

defined on  $\{y : y > 0 \text{ and } (1 + \xi y / \tilde{\sigma}) > 0\}$ , where  $\tilde{\sigma} = \sigma + \xi(u - \mu)$ .

### 2.2.1 Peaks Over Threshold (POT)

Before fitting the Generalized Pareto distribution into tails, the thresholds need to be selected. A too-high threshold will lead to high variance because of the few exceedances used to estimate the model. A too-low threshold will lead to high bias due to the violation of the asymptotic model. This is called the trade-off between bias and variance. Three methods are commonly used to identify a suitable threshold: the mean residual life plot, the hill estimator, and a graphical method based on the stability of parameter estimates.

#### Mean residual life plot

Suppose a sequence  $X_1, \dots, X_n$ , the exceedances over a threshold  $u_0$  are fitted into a valid GPD given  $\xi < 1$ . Then the mean of GPD is

$$E(X - u_0 | X > u_0) = \frac{\sigma_{u_0}}{1 - \xi}.$$

For  $u > u_0$ ,

$$\begin{aligned} E(X - u | X > u) &= \frac{\sigma_u}{1 - \xi} \\ &= \frac{\sigma_{u_0} + \xi u}{1 - \xi}. \end{aligned}$$

Consequently, the conditional mean should be linear for an appropriate threshold. The mean residual life plot is defined as

$$\left\{ \left( u, \frac{1}{n_u} \sum_{i=1}^{n_u} (x_{(i)} - u) \right) : u < x_{\max} \right\}$$

where  $x_{(i)}$  denotes observation that exceed threshold  $u$ ,  $n_u$  is the number of exceedances, and  $x_{\max}$  refers to the largest value of  $x_{(i)}$ .

## Stability of GPD under change of threshold

The shape parameter  $\xi$  and scale parameter  $\sigma$  should be constant above an appropriate threshold  $u_0$ . First, given  $X \sim \text{GPD}(\sigma, \xi)$ , we have  $(X - u | X > u) \sim \text{GPD}(\sigma + \xi u, \xi)$ , indicating that GPD is stable to the change of threshold and the shape parameter stays invariant. Denote the scale parameter by  $\sigma_u$  for a threshold  $u > u_0$ , which is given by,

$$\sigma_u = \sigma_{u_0} + \xi(u - u_0). \quad (2)$$

and reparameterize it as,

$$\sigma^* = \sigma_u - \xi u.$$

Consequently,  $\sigma^*$  is constant with respect to  $u$  in light of [2].

## The Hill plot

For an ordered sequence  $X_{(1,n)} \geq X_{(2,n)} \geq \dots \geq X_{(n,n)}$  from independent and identically distributed random variables  $X_1, X_2, \dots, X_n$ , the Hill estimator using  $k + 1$  ordered statistics is given by,

$$H_{k,n} = \frac{1}{k} \sum_{i=1}^k \log \left( \frac{X_{(i,n)}}{X_{(k,n)}} \right), \quad k \leq n$$

where  $H_{k,n}$  is an estimator of the shape parameter  $\xi$  which determines the heaviness of the tails. The Hill plot shows a range of  $k$  values against the tail index  $\alpha = 1/\xi$  for different thresholds [20]. An appropriate threshold can be chosen in the region the Hill estimator tends to be stable.

### 2.2.2 Modeling the left tail with GPD

In order to show the asymmetric property tails of residuals after marginal modelling, it is necessary to fit the generalized Pareto distribution for left tails and right tails, respectively. Typically, the GPD is employed to fit exceedances, which are positive values above a specified threshold. However, when dealing with left tails, they consist of shortfalls which are negative. As a result, the inclusion of Peaks Over Threshold (POT) for negative data becomes essential for accurately capturing the characteristics of the left tails. Figure 1 illustrates the utilization of POT for both positive and negative data.

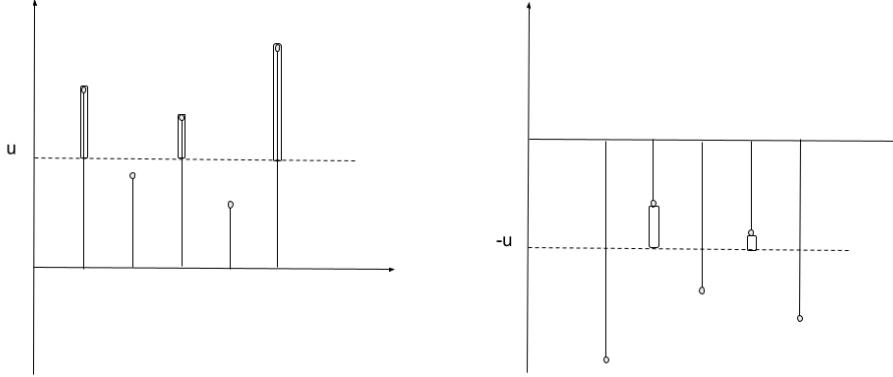


Figure 1: POT for positive(left) and negative(right) data.

According to Figure 1, Peaks over the threshold for negative data can be achieved by transforming random variables  $X_i$  to  $-X_i$  and then choosing  $-u$  as threshold after fitting generalized Pareto distribution for  $-X_i$ .

### 2.2.3 Model Diagnostics

Probability plots and quantile plots are two graphical methods used to check whether a dataset follows a given distribution. Given ordered observations,

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)},$$

the probability plot is defined by,

$$\left\{ \left( \hat{F}(x_{(i)}), \frac{i}{n+1} \right), i = 1, \dots, n \right\}$$

and the quantile plot is given by,

$$\left\{ \left( \hat{F}^{-1}\left(\frac{i}{n+1}\right), x_{(i)} \right), i = 1, \dots, n \right\}$$

where  $\hat{F}$  is a estimated distribution fitted into observations and  $\hat{F}^{-1}$  is the corresponding inverse function of  $\hat{F}$ , i.e. quantile function. The points of probability plots and quantile plots should be located near the unit diagonal if  $\hat{F}$  is a reasonable estimate of the true unknown distribution  $F$  of observations.

## 2.3 Copulas

Risk management is a necessary part of finance. For example, portfolio transactions face the problem of losing money in the future, and credit risk could cause the borrowers to fail to repay. As a result, accurate quantitative methods are essential to measuring financial risks. Copula theory is a powerful tool to measure risk applied to multivariate distribution. A copula is a function that connects univariate marginals to their multivariate distribution. It was first introduced by Sklar in 1959 [3]. In this section, we will discuss some of its fundamental properties and introduce some commonly used copula models.

### 2.3.1 Definition and properties

Let us start with Sklar's theorem, which is the core of the copula theory.

**Sklar's theorem** Let  $H$  be a joint distribution with margins  $F_1$  and  $F_2$ . Then there exists a copula  $C$  such that for all  $x_1, x_2 \in \mathbf{R}$ ,

$$H(x_1, x_2) = C(F_1(x_1), F_2(x_2)). \quad (3)$$

If  $F_1$  and  $F_2$  are continuous, then  $C$  is unique; otherwise,  $C$  is uniquely determined on  $\text{Ran}(F_1) \times \text{Ran}(F_2)$ . Conversely, if  $C$  is a copula and  $F_1, F_2$  are distribution functions, then the function  $H$  defined by (3) is a joint distribution function with margins  $F_1$  and  $F_2$  [21].

A copula function  $C(u, v)$  from  $\mathbf{I}^2$  to  $\mathbf{I}$ , where  $\mathbf{I} = [0, 1]$ , has the following properties:

1. for every  $u, v$  in  $\mathbf{I}$ ,

$$C(u, 0) = 0 = C(0, v)$$

$$C(u, 1) = u, \text{ and } C(1, v) = v$$

2. For every  $u_1, u_2, v_1, v_2$  in  $\mathbf{I}$  such that  $u_1 \leq u_2$  and  $v_1 \leq v_2$ ,

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0.$$

**Theorem** Fréchet-Hoeffding bounds for copula are defined by,

$$W(u, v) \leq C(u, v) \leq M(u, v),$$

$$W(u, v) = \max(u + v - 1, 0) \text{ and } M(u, v) = \min(u, v).$$

where  $W(u, v)$  is called as *Fréchet-Hoeffding lower bound* and  $M(u, v)$  is *Fréchet-Hoeffding upper bound* [21].

Sklar's theorem can also be generalized to the n-dimensional case in a simple manner [10]. The N-dimensional copula function is defined as,

$$C(u_1, \dots, u_d) := F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)), \quad (u_1, \dots, u_d) \in [0, 1]^d,$$

where  $F_k^{-1}$  are the univariate quantile function,  $k = 1, \dots, d$ .

### 2.3.2 Elliptical Copula

Two popular elliptical distributions are multivariate normal  $\mathcal{N}_d(\mu, \Sigma)$  and multivariate Student's t-distribution  $t_d(\mu, \Sigma, \nu)$ , which is defined below.

Let  $X \sim \mathcal{N}_d(\mu, \Sigma)$ , the density of  $X$  is given by,

$$f(x) = \frac{e^{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)}}{(2\pi)^{d/2} \det(\Sigma)^{1/2}}, \quad x \in \mathbb{R}.$$

The normal or Gaussian copula  $C_P^{Gauss}$  is the copula of  $X \sim \mathcal{N}_d(0, P)$ , where  $P$  is a correlation matrix. The copula form is obtained by,

$$C_P^{Gauss}(u_1, \dots, u_d) := \Phi_d(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d))$$

where  $\Phi_d$  is the joint distribution function and  $\phi^{-1}$  denotes the inverse function of standard normal distribution.

Let  $X \sim t_d(\mu, \Sigma, \nu)$ , then the density of  $X$  is defined as,

$$f(x) = \frac{\Gamma(\frac{1}{2}(\nu + d)) (1 + \frac{1}{\nu}(x - \mu)'\Sigma^{-1}(x - \mu))^{-(\nu+d)/2}}{\Gamma(\frac{\nu}{2})(\pi\nu)^{d/2} \det(\Sigma)^{1/2}}.$$

The form of t-copula is obtained by,

$$C_{\nu,d}^t(u_1, \dots, u_d) := t_d(t_\nu^{-1}(u_1), \dots, t_\nu^{-1}(u_d))$$

where  $t_d$  is the joint distribution and  $t_\nu$  is the univariate standard t-distribution with degree of freedom  $\nu$ .

### 2.3.3 Archimedean Copula

Let  $\phi$  be a continuous function from  $[0, 1]$  to  $[0, \infty]$  which satisfies the following properties:

- $\phi(1) = 0$
- $\phi(0) = \infty$

- For all  $t \in (0, 1)$ ,  $\phi'(t) < 0$ , i.e.  $\phi$  is decreasing,
- For all  $t \in (0, 1)$ ,  $\phi'' \geq 0$ , i.e.  $\phi$  is convex.

Then  $\phi$  is called the generator function and a Copulas is said to be an Archimedean copula if it can be written as follows,

$$C(u, v) = \phi^{[-1]}(\phi(u) + \phi(v))$$

where  $\phi^{[-1]}$  is the pseudo-inverse of  $\phi$  defined by,

$$\phi^{[-1]}(x) = \begin{cases} \phi^{-1}(x), & 0 \leq x \leq \phi(0) \\ 0, & \phi(0) \leq x \leq \infty. \end{cases}$$

Three kinds of one-parameter Archimedean copulas were implemented in this thesis: Gumbel, Frank, and Clayton with different generator functions.

### Gumbel Copula

The generator of Gumbel family is given by,  $\phi(u) = (-\ln(u))^\alpha$ , where  $\alpha \geq 1$ , and the generator inverse is  $\phi^{-1}(t) = \exp(-t^{1/\alpha})$ . Consequently, the  $d$ -dimensional Gumbel copula is defined by,

$$C(u_1, \dots, u_d) = e^{-([\sum_{i=1}^d -\ln(u_i)]^{1/\alpha})}, \quad \alpha > 1.$$

### Frank Copula

The generator of the Gumbel family is given by,

$$\phi(u) = -\ln\left(\frac{e^{-\alpha u} - 1}{e^{-\alpha} - 1}\right)$$

which its inverse is

$$\phi^{-1}(t) = -\frac{1}{\alpha} \ln(1 + e^t(e^{-\alpha} - 1)).$$

Consequently, the  $d$ -dimensional Frank copula is obtained as,

$$C(u_1, \dots, u_d) = -\frac{1}{\alpha} \ln\left(1 + \frac{\prod_{i=1}^d (e^{-\alpha u_i} - 1)}{(e^{-\alpha} - 1)^{d-1}}\right), \quad \alpha > 0.$$

### Clayton Copula

The generator of Clayton family is given by,  $\phi(u) = u^{-\alpha} - 1$ , and the generator inverse is  $\phi^{-1}(t) = (t + 1)^{-1/\alpha}$ . Consequently, the  $d$ -dimensional Clayton copula is obtained by,

$$C(u_1, \dots, u_d) = \left[ \sum_{i=1}^d u_i^{-\alpha} - d + 1 \right]^{-1/\alpha}, \quad \alpha > 0.$$

### 2.3.4 Dependence Measures

Copulas can capture the dependence of marginal distributions and are invariant under almost surely strictly increasing transformations [22], which means that the dependence of random variables remains unchanged even after strictly increasing transformation of them. This implies that any dependence measure which can be defined as function of copula will be invariant under strictly increasing monotonic transformations. In this section, we introduce three kinds of dependence measures named Kendall's tau, Spearman's rho, and Pearson's rho, where Kendall's tau and Spearman's rho defined only based on copulas measure a form of dependence known as a concordance.

#### Concordance

Let  $(x_i, y_i)$  and  $(x_j, y_j)$  denote two observations from a vector  $(X, Y)$  of continuous random variables. We say that  $(x_i, y_i)$  and  $(x_j, y_j)$  are concordant if  $(x_i - x_j)(y_i - y_j) > 0$  and they are discordant if  $(x_i - x_j)(y_i - y_j) < 0$ . Introduced by Scarsini [23], concordance measures are functions from copulas to the interval  $[-1, 1]$ . The value 1 indicates strongest positive dependence and -1 implies the opposite. Kendall's tau and Spearman's rho are two prominent examples of concordance measures for bivariate copula.

#### Kendall's tau

Assume  $(X_1, Y_1)$  and  $(X_2, Y_2)$  are two independent and identically distributed random vectors, each with joint distribution function  $H$ . Then Kendall's tau is defined as:

$$\tau = P((X_1 - X_2)(Y_1 - Y_2) > 0) - P((X_1 - X_2)(Y_1 - Y_2) < 0).$$

It equals the difference between concordance probability and discordance probability. In the context of bivariate copula, Kendall's tau for two uniform random variables  $U, V$  whose copula is denoted by  $C$  is given by

$$\tau = 4E(C(U, V)) - 1.$$

#### Spearman's rho

Assume  $(X_1, Y_1), (X_2, Y_2)$ , and  $(X_3, Y_3)$  are three independent random vectors with a common joint distribution function  $H$ . The Spearman's rho is defined as:

$$\rho = 3(P((X_1 - X_2)(Y_1 - Y_3) > 0) - P((X_1 - X_2)(Y_1 - Y_3) < 0))$$

Considering a bivariate case where  $(U, V) \sim C$ , the Spearman's rho of copula  $C$  is expressed as,

$$\rho = Cor(U, V).$$



## Pearson's rho

Pearson correlation rho is a measure of linear dependence between two variables and is the most popular tool for measuring bivariate relationships. It is defined by

$$\rho \stackrel{\text{def}}{=} \text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}}.$$

Properties include:

- (1)  $-1 \leq \text{Corr}(X, Y) \leq 1$ .
- (2) If  $X, Y$  are independent, then  $\text{Corr}(X, Y) = 0$ .
- (3)  $\text{Corr}(X, Y) = 0$  does not imply independence.

## 2.4 Parameter Estimation

Let  $X_1, X_2, \dots, X_n$  be a sequence of random variables with cumulative distribution function (cdf)  $F_k(x_k; \alpha_k)$  and probability density function (pdf)  $f_k(x_k; \alpha_k)$  with parameters  $\alpha_k$ ,  $k = 1, 2, \dots, n$ , and let  $X = (X_1, X_2, \dots, X_n)^T$ . According to Sklar's theorem, the joint distribution of  $X$  can be expressed as,

$$H(x_1, \dots, x_n; \theta) = C(F_1(x_1; \alpha_1), \dots, F_n(x_n; \alpha_n); \beta)$$

where  $C : \mathbf{I}^n \rightarrow \mathbf{I}$  is a copula function with parameter vector  $\beta$ . The corresponding pdf of  $X$  is defined by,

$$h(x_1, \dots, x_n; \theta) = c(F_1(x_1; \alpha_1), \dots, F_n(x_n; \alpha_n); \beta) \prod_{i=1}^n f_i(x_i; \alpha_i)$$
$$c(F_1(x_1; \alpha_1), \dots, F_n(x_n; \alpha_n); \beta) = \frac{\partial^n (C(F_1(x_1; \alpha_1), \dots, F_n(x_n; \alpha_n); \beta))}{\partial F_1(x_1; \alpha_1) \dots \partial F_n(x_n; \alpha_n)}$$

### 2.4.1 Full Maximum Likelihood (FML)

The full maximum likelihood estimation requires optimizing the log-likelihood function for all parameters  $\theta = (\alpha_1, \dots, \alpha_n, \beta)$  simultaneously. The log-likelihood function of  $\{x_{1t}, \dots, x_{nt}\}_{t=1}^T$  is

$$l(\theta) = \sum_{t=1}^T \log c(F_1(x_{1t}; \alpha_1), \dots, F_n(x_{nt}; \alpha_n); \beta) + \sum_{t=1}^T \sum_{i=1}^n \log f_i(x_{it}; \alpha_i).$$

and the optimized parameter  $\hat{\theta}_{MLE}$  which maximizes  $l(\theta)$  is

$$\hat{\theta}_{MLE} = \arg \max_{\theta} l(\theta).$$

### 2.4.2 Inference for Margins (IFM)

IFM is a two-step estimation method and was introduced in [11]. Since it is computationally complex to estimate all parameters in one step, the IFM method allows us to first optimize parameters of marginal distributions and then estimate the dependence parameter of the copula. Specifically, we decompose the full log-likelihood function as,

$$l(\theta) = l_m(\alpha) + l_c(\alpha, \beta)$$

where  $l_m(\alpha)$  is the log-likelihood function related to marginal distribution and  $l_c(\alpha, \beta)$  is the one related to copula function.

Step 1: Estimate marginal parameters  $\{\alpha_i\}_{i=1}^n$  by

$$\hat{\alpha}_i = \arg \max_{\alpha_i} l_m(\alpha_i) = \arg \max_{\alpha} \sum_{t=1}^T \log f_i(x_{it}; \alpha_i), \quad i = 1, \dots, n$$

Step 2: Estimate copula parameters  $\beta$  by

$$\hat{\beta} = \arg \max_{\beta} l_c(\hat{\alpha}, \beta) = \sum_{t=1}^T \log c(F_1(x_{1t}; \hat{\alpha}_1), \dots, F_n(x_{nt}; \hat{\alpha}_n); \beta).$$

### 2.4.3 Pseudo Maximum Likelihood (PML)

Both the FML and IFM are parametric estimation methods, which means the performance of algorithms is dependent on the accuracy of estimated parameters. To address the potential problem of the misspecification of marginal distributions, Pseudo Maximum Likelihood, a semiparametric estimation, was proposed in [12]. It is similar to the IFM, including two steps to carry out, but instead of estimating marginal distributions parametrically, it utilizes the empirical distribution of the samples.

More specifically, let  $\tilde{f}_k, k = 1, \dots, n$  denotes the empirical density of  $X_k$ . At step 1 in the IFM, the marginal distributions are estimated by  $\tilde{f}_k$ , followed by step 2:

$$\hat{\beta}_i^{PML} = \arg \max_{\beta_i} l_m(\beta_i) = \arg \max_{\beta} \sum_{t=1}^T \log \tilde{f}_i(x_{it}).$$

Then perform step 2 as the IFM.

In conclusion, these estimators are all consistent and asymptotically normal [12][13]. Moreover, Kim et al. [24] verified the PML method outperforms than other two estimators when the marginal distributions are misspecified or unknown.

## 2.5 Value at Risk

### 2.5.1 Introduction

Value at Risk is a popular risk management tool in the finance industry and it gives us the measured risk of a specific portfolio in monetary terms, namely, how much money we might lose in the future. It is a widely used measure of financial risk, including credit, liquidity, market price risk, and other risks. It is defined by the maximum expected loss of a portfolio or assets over a given time horizon at a given confidence level. The basic concept described by [25] shows one important characteristic that VaR takes dependence between different risk factors into consideration.

**Definition: (Value at Risk)** Given a portfolio or asset return changing from time  $t - \Delta t$  to  $t$ , the Value at Risk at time  $t$  with confidence level  $1 - \alpha$  is defined by:

$$VaR_t(\alpha) = \inf\{s : F_t(s) \geq \alpha\}$$

where  $F_t(x)$  is the distribution of portfolio or asset return  $X_{p,t}$  at time  $t$ . We also have  $P(x_{p,t} \leq VaR_t(\alpha) | \Omega_{t-1}) = \alpha$ ,  $\Omega_{t-1}$  contains all information at time  $t - 1$ , which means that the loss of portfolio or asset returns will not exceed the VaR value with confidence level  $100(1 - \alpha)\%$ .

We introduce three widely used approaches to estimate VaR: Historical Simulation, Variance-Covariance Simulation, and Monte Carlo Simulation.

### 2.5.2 Historical Simulation

The historical simulation uses actual historical data to simulate potential future losses. It is the simplest method to implement since it does not require any assumptions about the distribution of returns. However, the disadvantages of historical simulations are limitations by the limited available historical data, sensitivity to extreme events, and computational cost.

### 2.5.3 Variance-Covariance Method

The Variance-Covariance method is a parametric approach for VaR estimation and it assumes asset returns are normally distributed and the portfolio is a linear function of them. Suppose a portfolio consisting of two index returns  $[x_{1t}, x_{2t}]$  following a two-dimensional normal distribution with mean  $r_{t+1|t}$  and variance-covariance matrix  $\Sigma_t$ , the VaR over time period of 1 day can be expressed as,

$$VaR_{t+1} = r_{t+1|t} + \sigma_{t+1|t}^2 \Phi^{-1}(\alpha)$$

where  $\Phi^{-1}(\alpha)$  is the inverse function for the cumulative distribution function of standard normal and  $\sigma_{t+1|t}$  denotes standard deviation, which is formulated by,

$$\sigma_{t+1|t}^2 = W' \Sigma W = [w_1 w_2] \begin{bmatrix} \sigma_{11,t+1} & \sigma_{12,t+1} \\ \sigma_{21,t+1} & \sigma_{22,t+1} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$

Variance-Covariance method is rather simple to implement and can be applied to a wide range of financial risk estimations. However, one of the disadvantages is the assumptions about returns following a normal distribution is not always the case in the financial market. As a result, the estimation of VaR may not be accurate.

#### 2.5.4 Monte Carlo Simulation

The key idea of the Monte Carlo method as a parametric technique to compute the VaR of a portfolio is to simulate random processes based on the empirical distribution of stock returns multiple times to make it converge to the unknown true distribution of returns. Therefore we could use the simulated data to compute the VaR of the portfolio.

There are several procedures to implement Monte Carlo simulation: Firstly, choose appropriate distributions describing the behavior of stock returns and estimate related parameters. Secondly, generate a large number of random data based on estimated distribution. In the end, we could use simulated data to compute the portfolio's VaR of interest.

The Monte Carlo method could capture some complex relationships between asset returns which are difficult to find when applying the historical simulation and variance-covariance method. As it is not totally dependent on historical data which might never show again in the future and the simulation is based on the estimated model without any normal assumption.

On the other hand, the Monte Carlo method also has some drawbacks: it is computationally intensive and sensitive to the chosen distributions of returns.

#### 2.5.5 Backtesting

Backtesting is a technique used to assess the accuracy of VaR estimation. Backtesting a model involves the procedure of fitting a parametric model to observations, using estimated models to make predictions, and then comparing them to historical data. We applied Kupiec's test [26] to compare the expected losses against the actual losses realized in a given period.

#### Kupiec's Test

Kupiec's Test, also known as the unconditional Coverage Test, is based on the assumption that the number of observed exceptions, which are observations in excess of VaR,

follows a binomial distribution. Specifically, let  $N$  denote the total number of observations and  $n$  represent the number of exceptions. Assuming the model is accurate, the failure probability should converge to  $1 - c$  given the confidence level  $c$  (99%, 95% or 90%). The number of exceptions follows a binomial distribution,

$$f(n) = \binom{N}{n} p^n (1 - p)^{N-n}$$

where  $p$  is the probability of exceptions under the null hypothesis defined as

$$H_0 : p = \hat{p} = \frac{n}{N}.$$

Under the null hypothesis that there is a correct amount of exceptions, the test statistics is formulated by,

$$LR_K = -2 \ln \left( \frac{(1 - p)^{N-n} p^n}{(1 - \frac{n}{N})^{N-n} (\frac{n}{N})^n} \right).$$

Furthermore,  $LR_K$  is asymptotically  $\chi^2(1)$  distributed. If the test statistics is greater than  $\chi^2(1)$  critical value or the statistical p-value is below the significance level (1%, 5% or 10%), the null hypothesis will be rejected, and the model is seen as not good.

### 3 Empirical Results

#### 3.1 Data

In this section, we investigated the dependence structure of stock markets in Europe and Sweden. Specifically, we will analyze the data STOXX 50 and OMXS 30, which include daily equity indices from January 3, 2000, to February 28, 2023. The total number of observations is 5859.

The daily return of the index is defined as,

$$r_t = \ln \left( \frac{P_t}{P_{t-1}} \right) = \ln P_t - \ln P_{t-1}.$$

where  $P_t$  is the closing price at time  $t$ .

Daily returns for both datasets are shown in Figure 2.

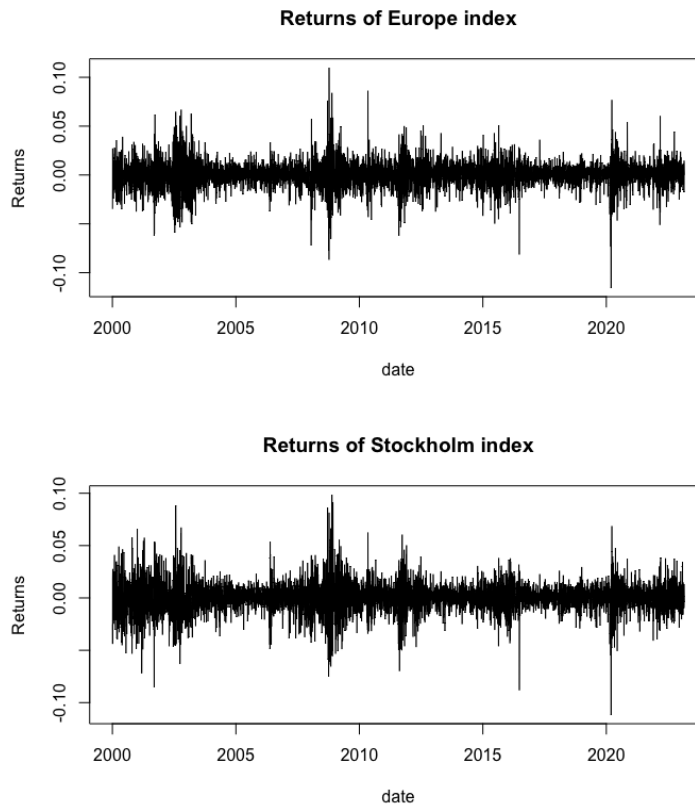


Figure 2: Daily returns of STOXX 50 and OMXS 30.

The analysis shown in Table 1 reveals that both index returns exhibit heavy tails, as indicated by kurtosis values exceeding 3, and they display left skewness due to negative skewness values. The Jarque-Bera test [27] is used to determine whether the data follows a normal distribution. Specifically, the test statistics of Jarque-Bera test is always positive and the data tend to follow a normal distribution if the value is close to zero. The results indicate that neither index return conforms to a normal distribution, given the substantial deviation of their Jarque-Bera values from zero.

Table 1: Descriptive statistics of index returns

	STOXX 50	OMXS 30
Number of Observations	5859	5859
Mean	0.0003507	0.0004422
Standard Deviation	0.01361	0.01447
Skewness	-0.1333	-0.08731
Kurtosis	5.219	4.277
Jarque-Bera Test	6674 < 2.2e-16	4479 < 2.2e-16

Figures 3 and 4 demonstrates that index returns share the features of volatility clustering, commonly referred to as autoregressive conditional heteroskedasticity in econometrics, meaning that volatility tends to persist, i.e. large movements followed by large movements and small changes followed by small changes. In light of these findings, we opted to adopt ARMA-GARCH models to fit the marginal distributions of both index returns. considering three innovation distributions for the ARMA-GARCH models: normal distribution, Student's t distribution, and skewed Student's t distribution.

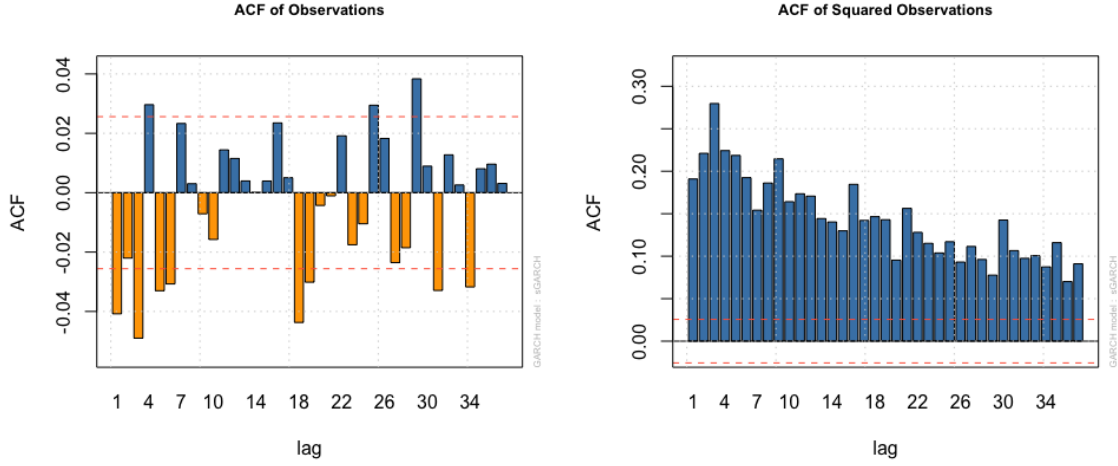


Figure 3: Autocorrelation of observations(left) and squared observations(right) for data STOXX 50.

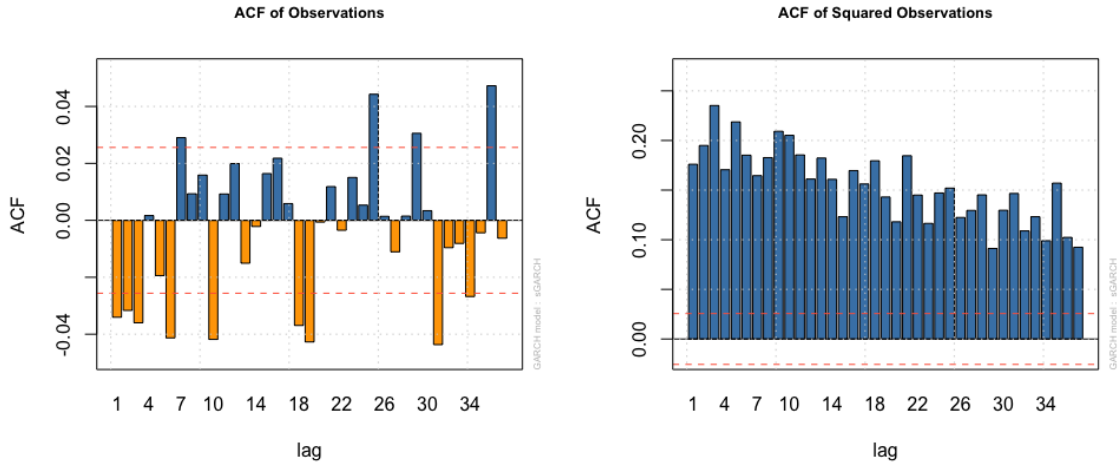


Figure 4: Autocorrelation of observations(left) and squared observations(right) for data OMXS 30.

### 3.2 Marginal Modeling

The selection of ARMA(1,1)-GARCH(1,1) models for the two return series is based on their ability to effectively capture the volatility patterns commonly observed in financial time series. Subsequently, the parameters of the GARCH models, with three different



innovation distributions, were estimated using the Maximum Likelihood method in R. A summary of the results is presented in Tables 2 and 3.

Table 2: Marginal Modelling results of European daily returns.

ARMA(1,1)-GARCH(1,1)							
Parameters		Normal		$t$		Skew $t$	
		Estimate	P-value	Estimate	P-value	Estimate	P-value
	C	0.0005	0.0000	0.0006	0.0000	0.0004	0.0001
	$\phi_1$	0.7094	0.0000	0.6840	0.0000	0.7294	0.0000
	$\theta_1$	-0.7531	0.0000	-0.7302	0.0000	-0.7835	0.0000
	$\omega$	0.0000	0.0014	0.0000	0.2305	0.0000	0.1553
	$\alpha_1$	0.0927	0.0000	0.0886	0.0001	0.0860	0.0000
	$\beta_1$	0.8885	0.0000	0.9012	0.0000	0.9032	0.0000
	$\nu$			7.3651	0.0000	7.7510	0.0000
	$\xi$			0.9049	0.0000		
Diagnostics	LLF	17756.18		17862.96		17876.23	
	AIC	-6.0591		-6.0952		-6.0994	
	BIC	-6.0523		-6.0872		-6.0903	
Ljung-Box	Lag[1]	0.4773		0.2464		0.1257	
	Lag[5]	0.4268		0.0125		6.27e-09	
	Lag[9]	0.6334		0.0083		5.61e-03	
Adjusted Pearson	Group[30]	1.28e-13		1.97e-04		0.0877	
	Group[40]	4.13e-14		1.17e-04		0.2846	
Gof Test	Group[50]	8.44e-12		8.13e-05		0.0784	

Table 3: Marginal Modelling results of daily returns in Stockholm.

ARMA(1,1)-GARCH(1,1)							
Parameters		Normal		$t$		Skew $t$	
		Estimate	P-value	Estimate	P-value	Estimate	P-value
	C	0.0006	0.0000	0.0007	0.0000	0.0005	0.0000
	$\phi_1$	0.8450	0.0000	0.8041	0.0004	0.7942	0.0000
	$\theta_1$	-0.8751	0.0000	-0.8386	0.0000	-0.8423	0.0000
	$\omega$	0.0000	0.1191	0.0000	0.0306	0.0000	0.0225
	$\alpha_1$	0.0838	0.0000	0.0776	0.0000	0.0748	0.0000
	$\beta_1$	0.9076	0.0000	0.9169	0.0000	0.9192	0.0000
	$\nu$			8.3333	0.0000	9.0657	0.0000
	$\xi$					0.8895	0.0000
Diagnostics	LLF	17493.83		17574.54		17592.69	
	AIC	-5.9696		-5.9968		-6.0026	
	BIC	-5.9627		-5.9888		-5.9935	
Ljung-Box	Lag[1]	0.8366		0.6538		1.28e-01	

	Lag[5]	0.9961	0.9110	1.12e-09
	Lag[9]	0.7230	0.6229	9.42e-04
Adjusted Pearson	Group[30]	7.64e-14	5.22e-05	0.1978
	Group[40]	5.14e-13	3.57e-06	0.2165
Gof Test	Group[50]	1.04e-13	2.87e-04	0.1644

Based on the results from Tables 2 and 3, the models with  $t$  and skewed  $t$  distributions provide better fits than the one with normal distribution due to larger likelihoods and smaller AIC and BIC values. The Ljung-Box test suggests that normal and  $t$  distribution are better but the adjusted Pearson goodness-of-fit test indicates the residuals follow a skewed  $t$  distribution.

In order to determine which distribution fits the residuals best, we draw plots comparing the empirical distribution of the residuals of GARCH models with three theoretical distributions incorporating estimated parameters. Additionally, QQ plots were shown for further comparison. Figures 5, 6 and 7 show results for data STOXX 50, and Figure 8, 9 and 10 are for data OMXS 30.

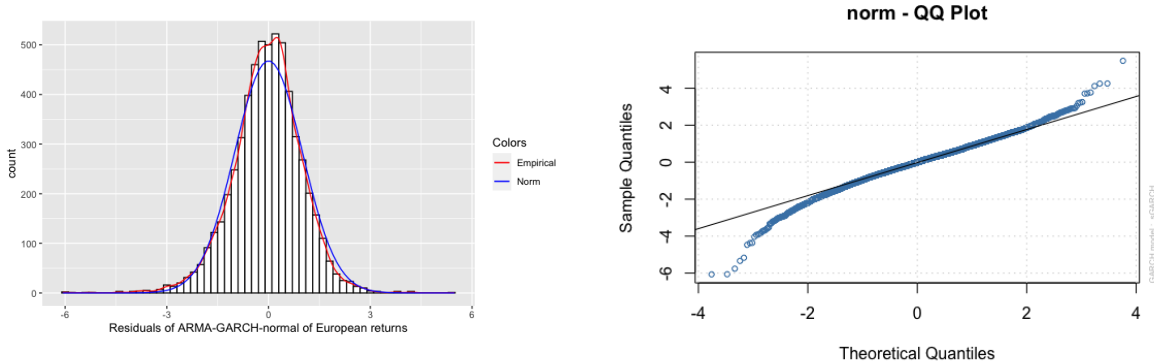


Figure 5: The left plot compares empirical residuals of data STOXX 50 with estimated normal distribution. QQ plot(right) shows the validity of the fitted normal distribution.

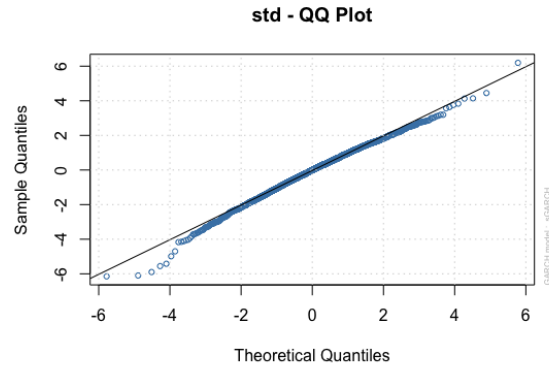
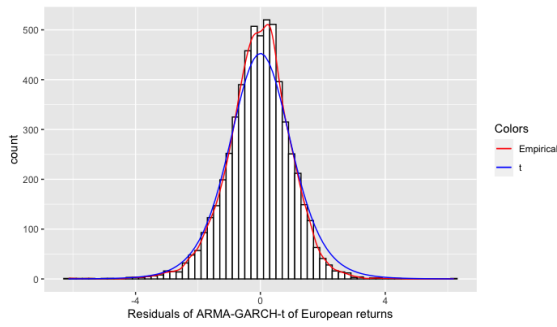


Figure 6: The left plot compares empirical residuals of data STOXX 50 with the estimated  $t$  distribution. QQ plot(right) shows the validity of the fitted  $t$  distribution.

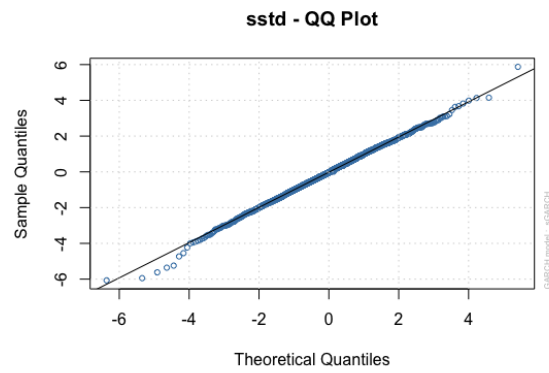
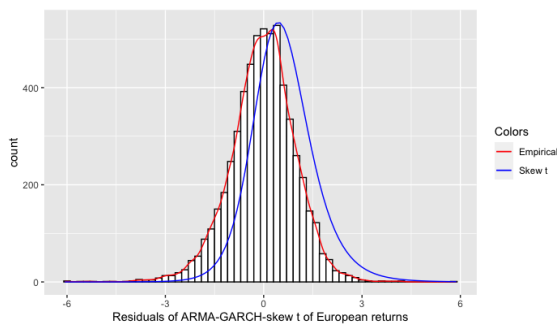


Figure 7: The left plot compares empirical residuals of data STOXX 50 with the estimated skewed  $t$  distribution. QQ plot(right) shows the validity of fitted skewed  $t$  distribution.

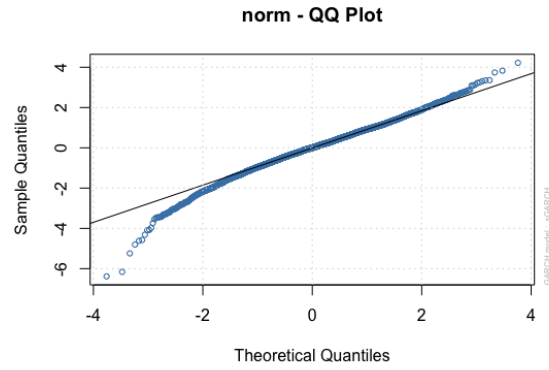
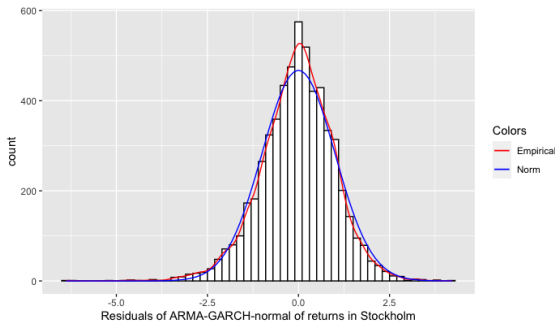


Figure 8: The left plot compares empirical residuals of data OMXS 30 with estimated normal distribution. QQ plot(right) shows the validity of the fitted normal distribution.

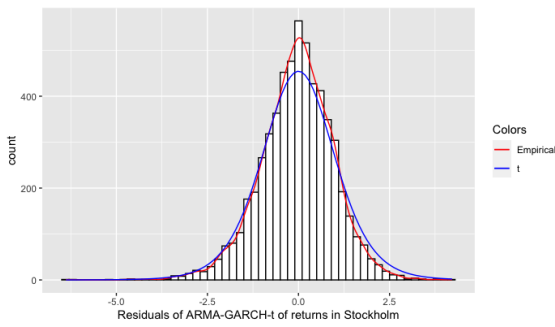


Figure 9: The left plot compares empirical residuals of data OMXS 30 with the estimated  $t$  distribution. QQ plot(right) shows the validity of the fitted  $t$  distribution.

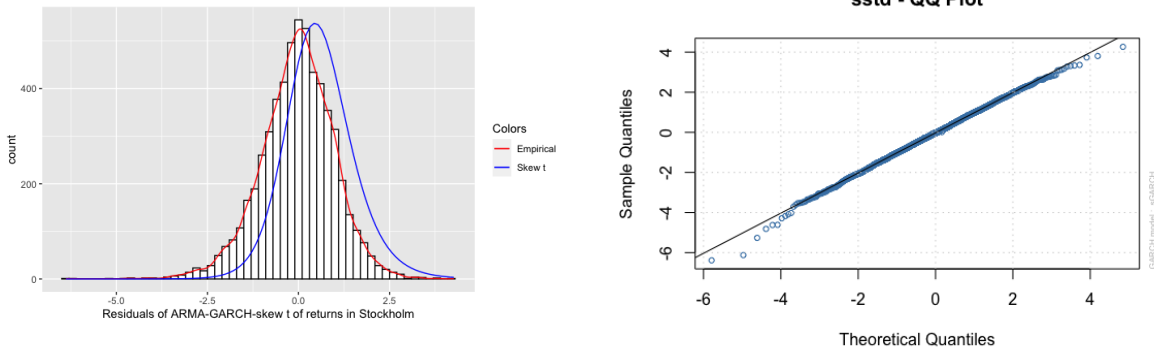


Figure 10: The left plot compares empirical residuals of data OMXS 30 with the estimated skewed  $t$  distribution. QQ plot(right) shows the validity of fitted skewed  $t$  distribution.

In conclusion, the  $t$  distribution is identified as the most suitable model for fitting the residuals. On one hand, QQ plots exhibit that  $t$  and skewed  $t$  distributions fit empirical residuals pretty well since most points in the plot lie on the line. On the other hand, the  $t$  distribution fits the tails clearly better than the skewed  $t$  distribution, which is significant for Value at risk estimation since it is calculated based on the quantiles.

### 3.3 Fitting the tails of residuals

In this section, our attention is directed towards an important observation derived from Figure 11, namely, the presence of asymmetric left and right tails in the empirical residuals. Notably, the right tail displays a significantly lighter distribution, which the  $t$  distribution fails to adequately capture.

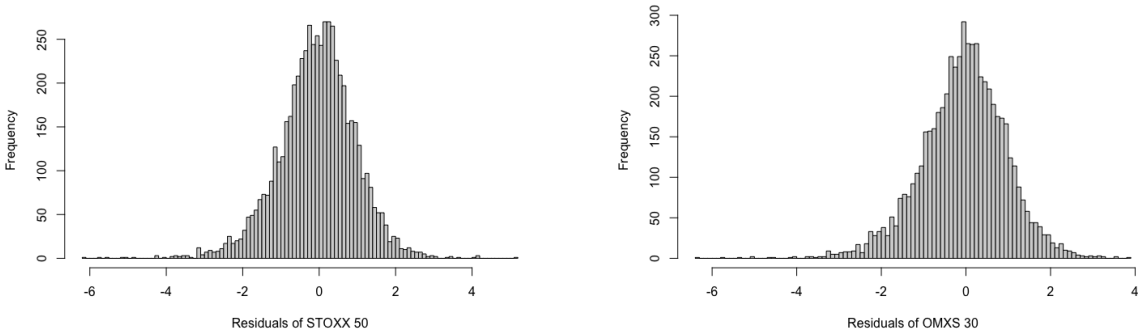


Figure 11: Histograms of residuals of marginal modeling for both data sets. The left one refers to residuals of data STOX 50 and the right one represents data OMXS 30.

## Simulation

To verify the hypothetical distributions of the residuals, we use quantile plots. To verify whether the deviations from the straight line are abnormal, we generated 1000 samples with the same sample size as our dataset and plotted the quantiles plots as shown in Figures 12 and 13. The idea is that if the underlying distribution is correct, only about 5% of points would deviate from the simulated quantile plots. These simulated samples were then compared with the empirical residuals.

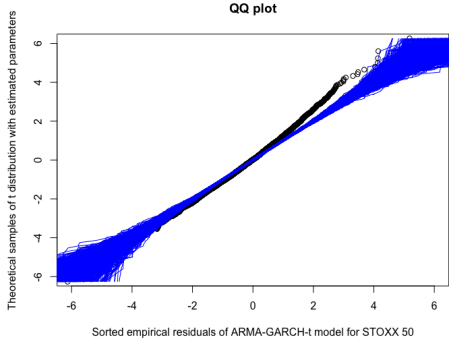


Figure 12: Qauntile plot for sorted residuals of ARMA(1,1)-GARCH(1,1)- $t$  model of data STOXX 50. 1000 lines refer to simulated samples from the  $t$  distribution with estimated parameters.

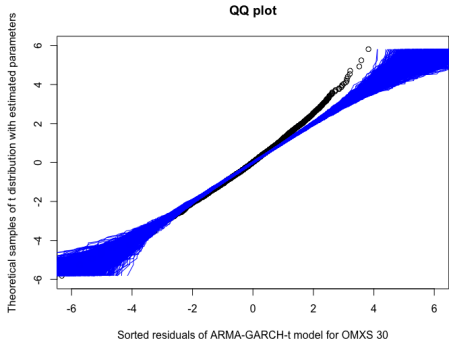


Figure 13: Qauntile plot for sorted residuals of ARMA(1,1)-GARCH(1,1)- $t$  model of data OMXS 30. 1000 lines refer to simulated samples from the  $t$  distribution with estimated parameters.

The simulation results reveal that the residuals for both index returns exhibit left skewness, indicating the presence of heavier left tails. Additionally, the simulations confirmed the asymmetry in our empirical residuals, with  $t$  distributions fitting the left tail more effectively than the right tail. Given that  $t$  distributions are symmetric, we proceeded to fit the Generalized Pareto Distribution (GPD) separately to the left and right tails of both index returns.

## GPD fitting

By looking at mean residual life plots in Figure [B.1](#), [B.2](#), and threshold range plots Figure [B.3](#)-[B.6](#), we determined that threshold 2 is appropriate for fitting the right tails of both return residuals. To address the presence of heavier left tails, we employed the Hill estimator to identify thresholds. As a result, the optimal order statistics determined by the hill estimator are 122 for data STOXX 50 and 170 for data OMXS 30, which correspond to thresholds of -2.3 and -2.1, respectively. The fitting results of the Generalized Pareto Distribution (GPD) are summarized in Table [4](#) and Table [5](#).

Table 4: GPD fitting to residuals of ARMA(1,1)-GARCH(1,1)- $t$  model for data STOXX 50.

	Left tail		Right tail	
	MLE	95% CI	MLE	95% CI
Scale parameter	0.5455	[0.3918, 0.6991]	0.4440	[0.3113, 0.5767]
Shape parameter	0.1544	[-0.0673, 0.3760]	0.1116	[-0.1142, 0.3374]
AIC	139.97		63.32	
BIC	145.61		68.52	
No. of observations	124		99	

Table 5: GPD fitting to residuals of ARMA(1,1)-GARCH(1,1)- $t$  model for data OMXS 30.

	Left tail		Right tail	
	MLE	95% CI	MLE	95% CI
Scale parameter	0.5127	[0.3949, 0.6306]	0.4545	[0.3265, 0.5825]
Shape parameter	0.1458	[-0.0306, 0.3221]	-0.1041	[-0.3056, 0.0974]
AIC	169.31		25.02	
BIC	175.61		30.19	
No. of observations	173		98	

The left tail of data OMXS 30 exhibits polynomially decreasing in contrast to the finite right tail. The related results are less clear for data STOXX 50 due to two positive shape parameter values but they still indicate asymmetric tail distributions.

## Quantile plots for GPD

In this section, we adopt an alternative approach to estimate the shape parameter  $\gamma$  in the Generalized Pareto Distribution (GPD) models. The shape parameter, also referred to as the extreme value index, plays a crucial role in characterizing the distribution of the tail. If  $\gamma < 0$  the distribution has a finite upper bound, while  $\gamma = 0$  leads to an exponentially decreasing distribution, and an exponentially decreasing tail occurs when

$\gamma > 0$ . The majority of research has focused on heavy-tailed distributions, where  $\gamma > 0$ . One prominent method for estimating  $\gamma$  is the Hill estimator, proposed by Hill in 1975 [28]. This estimator is defined as follows,

$$H_{k,n} = \frac{1}{k} \sum_{j=1}^k \log X_{n-j+1,n} - \log X_{n-k,n}$$

where  $X_{m,n}$  refers to the  $m$ th ordered observation out of  $n$  data in total. Assume  $X$  follows a Pareto distribution, the Pareto quantile plot is formulated by [29],

$$\left( \log \left( \frac{n+1}{j} \right), \log X_{n-j+1,n} \right), \quad j = 1, \dots, n$$

If the estimated GPD is a good fit to return residuals, the ordered observations served as empirical quantiles are anticipated to be linear with theoretical values, which means the Pareto quantile plot will show a line of which the slope is given by the extreme value index. Therefore, The core idea to estimate the extreme value index of heavier left tails is to fit a straight line to the Pareto quantile plot and find its slope.

Furthermore, to ensure the stability of the estimated  $\gamma$  parameter, we conducted multiple estimations using various values of  $k$ , ranging from 5300 to 5600 with a step size of 20. Figure 14 and 15 illustrate three examples of quantile plots for the left tails of the return residuals. The estimated extreme value indices are summarized in Table 6.

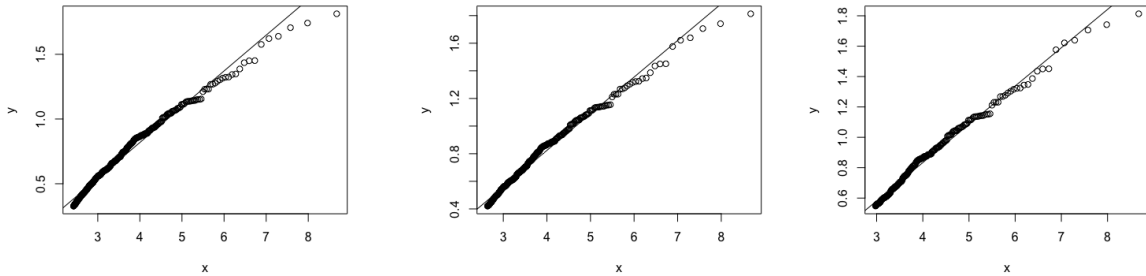


Figure 14: Pareto Quantile-Quantile plots with ordered data starting with 5340(left), 5440(middle) and 5540(right). The dots refer to empirical margins filtered by GARCH- $t$  models for data STOXX 50. The straight line is a regression line fitted to dots.



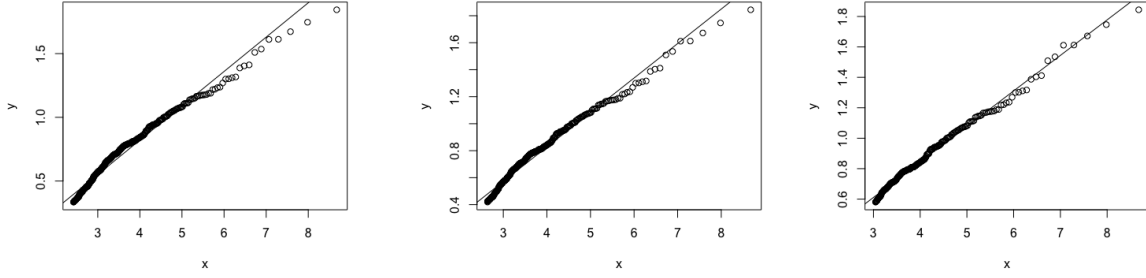


Figure 15: Pareto Quantile-Quantile plots with ordered data starting with 5340(left), 5440(middle) and 5540(right). The dots refer to empirical margins filtered by GARCH- $t$  models for data OMXS 30. The straight line is a regression line fitted to dots.

Table 6: Estimated extreme value index with respect to different ordered observations

Starting ordered data	Estimated slope	
	STOXX 50	OMXS 30
5300	0.2813885	0.2743794
5320	0.2788361	0.2719707
5340	0.2764416	0.2695846
5360	0.2741536	0.2670354
5380	0.2719008	0.2642499
5400	0.2696917	0.2612712
5420	0.2675284	0.2584317
5440	0.2653725	0.2554969
5460	0.2631412	0.2525685
5480	0.2608068	0.2493760
5500	0.2585585	0.2457505
5520	0.2564094	0.2422614
5540	0.2540690	0.2389333
5560	0.2520349	0.2361775
5580	0.2500530	0.2333648
5600	0.2476369	0.2303134

The results of estimated shape parameters are shown to be stable for both index returns and the extreme index values for left tails are clearly larger than those for right tails specified in Tables [4](#) and [5](#).

### 3.4 Copula Modelling

Regarding the elliptical copula, we used an exchangeable dispersion matrix to determine the dependence structure. In this scenario, the Gaussian copula requires the

estimation of a single parameter, while the  $t$ -copula necessitates the estimation of two parameters. Additionally, the Gumbel, Frank, and Clayton copulas belong to the class of one-parameter Archimedean copulas.

Prior to modeling the dependence structure between margins filtered by ARMA(1,1)-GARCH(1,1)- $t$  models, we transform the standardized residuals to uniformly distributed data. Subsequently, we estimated the dependence parameters using the maximum likelihood method. The results are summarized in Table 7.

Table 7: Dependence structure modelling for filtered return index by ARMA(1,1)-GARCH(1,1)- $t$ .

	Gaussian copula	$t$ copula	Gumbel copula	Frank copula	Clayton copula
Estimated parameter	0.7805	0.7889(df = 10.35)	2.2968	7.7626	2.6616
AIC	-5512.68	-5667.51	-5292.93	-5241.41	-2898.12
BIC	-5506.00	-5654.16	-5286.26	-5234.74	-2891.44
log-likelihood	2757.34	2835.75	2647.47	2621.71	1450.06

Gaussian copula and  $t$  copula are better than the other three copulas based on larger log-likelihood and smaller AIC and BIC values. And their estimated parameters are closer to dependence measurements presented in Table 8.

Table 8: Dependence measurements for empirical residuals.

Kendall's tau	Spearman's rho	Pearson's rho
0.5710	0.7583	0.78023

Figure 16 shows empirical filtered margins and five theoretical simulations based on estimated parameters. It indicates that Gaussian and  $t$  copula fitted better graphically. Gumbel and Clayton copula are both asymmetric Archimedean. Gumbel copula exhibits greater dependence in the negative tail while Clayton copula demonstrates greater dependence in the positive tail.

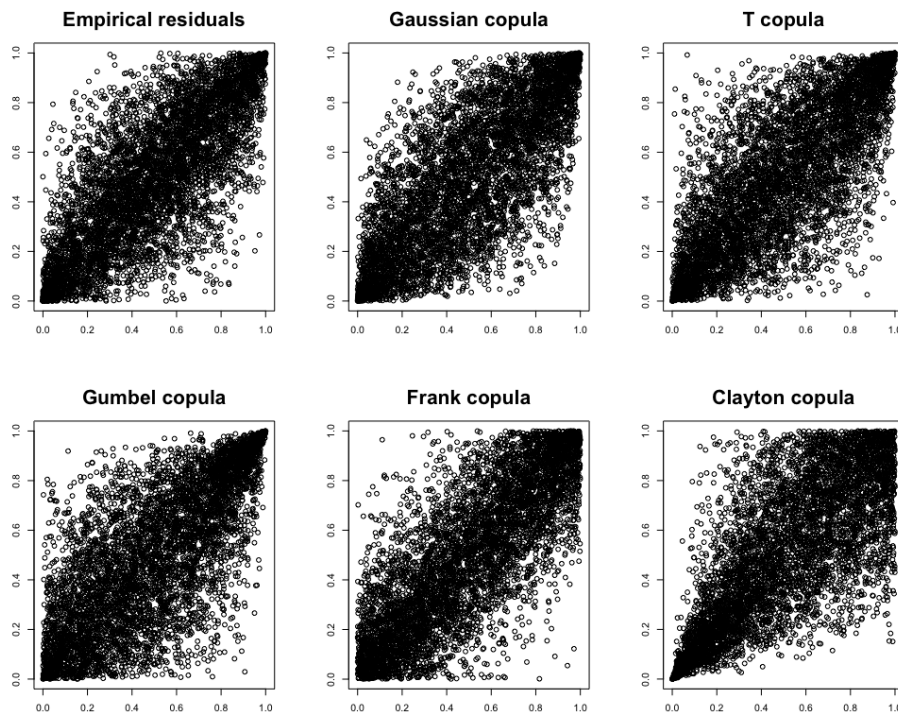


Figure 16: Scatterplots of 5859 data from Empirical residuals(upper left), and 5859 simulated samples from five kinds of copulas with estimated parameters shown in Table 7.

### 3.5 Backtesting of VaR

We implemented the GARCH-Copula approach to estimate VaR(Value-at-Risk) for the portfolio, made up of equally weighted STOXX 50 and OMXS 30, as well as index returns themselves. The detailed procedure applied to data STOXX 50 and OMX 30 is shown as followed:

1. Fit ARMA(1,1)-GARCH(1,1) with Student- $t$  distribution to both data sets. Obtain the standardized residuals.
2. Transform the marginal standardized residuals into uniformly distributed samples.
3. Estimate dependence structure of two uniformly distributed residuals using copula modeling by maximum likelihood method. Generate  $N$  random samples  $(S_{1i}, S_{2i})_{i=1}^N$  with estimated copula models.
4. Apply probability integral transform to  $(S_{1i}, S_{2i})_{i=1}^N$  and then calculate the index returns based on conditional means and conditional variances of estimated GARCH models.

5. Given that the portfolio is calculated by equally weighted two index returns, VaR could be estimated at different levels of confidence intervals.

Empirical results of five estimated copulas for portfolio, STOXX 50, and OMXS 30 are summarized in Table 9, 10, and 11.

Table 9: Number of violations of different VaR models with 99% confidence level.

$\alpha = 0.01$			
No. of observations: 5859		Expected no. of exceptions: 59	
Non-rejection region:		$39 \leq \text{exceptions} \leq 78$	
	Portfolio	STOXX 50	OMXS 30
Historical Simulation	43*	51*	49*
Variance-Covariance	93	96	138
Gaussian copula	87	73*	87
$t$ copula	49*	59*	55*
Gumbel copula	98	90	72*
Frank copula	76*	67*	64*
Clayton copula	33	55*	49*

Figures with asterisks are in the non-rejection region under the 99% confidence level of Kupiec's Test.

Table 10: Number of violations of different VaR models with 95% confidence level.

$\alpha = 0.05$			
No. of observations: 5859		Expected no. of exceptions: 293	
Non-rejection region:		$260 \leq \text{exceptions} \leq 326$	
	Portfolio	STOXX 50	OMXS 30
Historical Simulation	252	265*	259
Variance-Covariance	245	253	301*
Gaussian copula	385	357	405
$t$ copula	314*	320*	320*
Gumbel copula	376	345	385
Frank copula	368	364	336
Clayton copula	325*	304*	308*

Figures with asterisks are in the non-rejection region under 95% confidence level of Kupiec's Test.

Table 11: Number of violations of different VaR models with 90% confidence level.

$\alpha = 0.1$			
No. of observations: 5859		Expected no. of exceptions: 586	
Non-rejection region:		$548 \leq \text{exceptions} \leq 632$	
	Portfolio	STOXX 50	OMXS 30
Historical Simulation	552*	552*	539
Variance-Covariance	424	443	502
Gaussian copula	737	696	730
$t$ copula	659	624*	662
Gumbel copula	713	643	727
Frank copula	710	712	714
Clayton copula	587*	633	617*

Figures with asterisks are in the non-rejection region under 90% confidence level of Kupiec's Test.

In conclusion, the traditional approach to estimating VaR, historical simulation, tends to underestimate market risk in financial industries under the condition that a finite number of extreme events happened in the past. Some copulas, such as Gaussian, Gumbel, and Frank copula overestimate financial risk based on their relatively large number of expected exceptions. Finally,  $t$  copula and Clayton copula perform the best among elliptical and Archimedean copulas, respectively.

### 3.6 VaR prediction

Finally, rolling forecast methods are employed to generate Value at Risk (VaR) predictions for both datasets. In particular, the ARMA-GARCH- $t$  model is estimated using 4,359 observations, leaving 1,500 observations that are utilized for comparison with the forecasted values. These 1,500 samples are referred to as the 'out of sample' data, and the VaR predictions based on them are denoted as 'out of sample VaR'.

One-step prediction for ARMA-GARCH- $t$  model is obtained by,

$$\begin{aligned} r_{t+1|t} &= \hat{C} + \hat{\phi}_1 r_t + \hat{\theta}_1 \hat{\epsilon}_t \\ \sigma_{t+1|t}^2 &= \hat{\omega} + \hat{\alpha} \hat{\epsilon}_t^2 + \hat{\beta} \sigma_t^2 \\ \hat{\epsilon}_t &= r_t - \hat{r}_{t-1}. \end{aligned}$$

Then, the VaR prediction is given by,

$$VaR_\alpha = r_{t+1|t} + \sqrt{\sigma_{t+1|t}^2} t_\alpha^{-1}$$

where  $t_\alpha^{-1}$  is the  $\alpha$  quantile of Student's distribution.

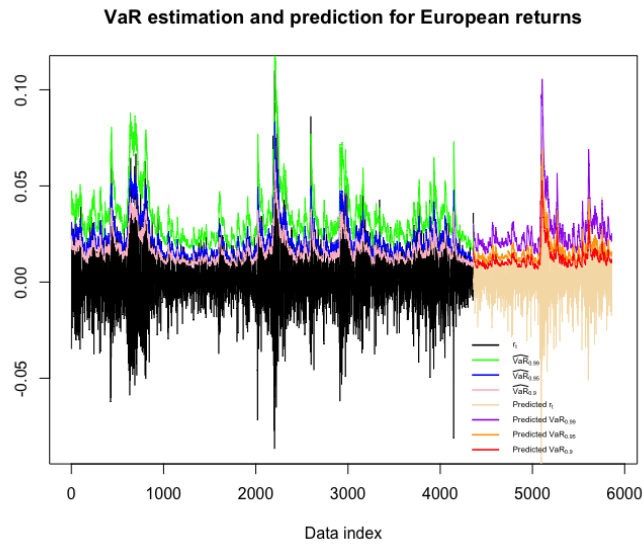


Figure 17: VaR estimation based on true observations and VaR prediction based on rolling forecast method, with confidence level 99%, 95%, and 90% for data STOXX 50.

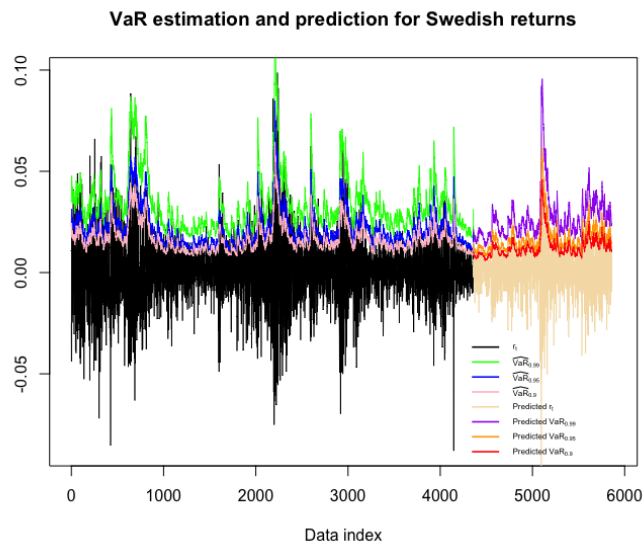


Figure 18: VaR estimation based on true observations and VaR prediction based on rolling forecast method, with confidence level 99%, 95%, and 90% for data OMXS 30.

## 4 Summary

### 4.1 Conclusions and further analysis

This thesis aims to explore the dependence structure between the European stock markets (OMXS 30) and Stockholm (STOXX 50) and utilize it to estimate Value at Risk. Accurately estimating the future losses of a portfolio requires robust multivariate joint distributions. However, the historical simulation and variance-covariance methods tend to underestimate risk across different confidence levels (99%, 95%, and 90%). The assumption of normality and linear relationship between asset returns prove inadequate in capturing the complexity of dependence structures. Monte Carlo simulation which constructs joint distribution using the combination of marginal modeling with copula function will give more reliable results. Empirical results indicate that combining the ARMA(1,1)-GARCH(1,1)- $t$  model with the  $t$  copula yields the most accurate VaR estimation. Furthermore, in the process of marginal modeling, Extreme Value theory indicates the tails of residuals are asymmetric and the left tails are heavier than the right ones for both datasets.

Future investigations could expand the analysis to include multivariate cases beyond the bivariate setting explored in this study. Furthermore, considering the asymmetric tails proved by GPD fitting and the Hill estimator simulation in the residuals from ARMA-GARCH models, incorporating adjustments based on this characteristic may improve the accuracy of VaR estimation.

## References

- [1] Kevin Dowd. *Beyond value at risk: the new science of risk management*, volume 96. John Wiley & Son Limited, 1998.
- [2] Paul Embrechts, Rüdiger Frey, and Alexander McNeil. *Quantitative risk management.*, 2011.
- [3] M Sklar. Fonctions de répartition à n dimensions et leurs marges. In *Annales de l'ISUP*, volume 8, pages 229–231, 1959.
- [4] Paul Embrechts, Alexander J. McNeil, and Daniel Straumann. *Correlation: Pitfalls and alternatives.* 1999.
- [5] Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, pages 987–1007, 1982.
- [6] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.
- [7] Jen-Jsung Huang, Kuo-Jung Lee, Hueimei Liang, and Wei-Fu Lin. Estimating value at risk of portfolio by conditional copula-garch method. *Insurance: Mathematics and economics*, 45(3):315–324, 2009.
- [8] Eric Jondeau and Michael Rockinger. The copula-garch model of conditional dependencies: An international stock market application. *Journal of international money and finance*, 25(5):827–853, 2006.
- [9] Long Kang. Modeling the dependence structure between bonds and stocks: A multidimensional copula approach. *Indiana University Bloomington*, 2007.
- [10] Jan-Frederik Mai and Matthias Scherer. *Simulating copulas: stochastic models, sampling algorithms, and applications*, volume 6. 2017.
- [11] Harry Joe and James Jianmeng Xu. The estimation method of inference functions for margins for multivariate models. 1996.
- [12] Christian Genest, Kilani Ghoudi, and L-P Rivest. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82(3):543–552, 1995.
- [13] Andrew J Patton. Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics*, 21(2):147–173, 2006.
- [14] Alexander J McNeil and Rüdiger Frey. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of empirical finance*, 7(3-4):271–300, 2000.



- [15] Kaj Nystrom and Jimmy Skoglund. Univariate extreme value theory, garch and measures of risk. *Preprint, Swedbank*, 2002.
- [16] Luiz Koodi Hotta, Edimilson C Lucas, and Helder P Palaro. Estimation of var using copula and extreme value theory. *Multinational Finance Journal*, 12(3/4):205–218, 2008.
- [17] Alexios Ghalanos. Introduction to the rugarch package.(version 1.3-1). *Manuscript*, <http://cran.r-project.org/web/packages/rugarch>. Accessed, 11, 2020.
- [18] Stuart Coles, Joanna Bawa, Lesley Trenner, and Pat Dorazio. *An introduction to statistical modeling of extreme values*, volume 208. Springer, 2001.
- [19] James Pickands III. Statistical inference using extreme order statistics. *the Annals of Statistics*, pages 119–131, 1975.
- [20] Xia Yang, Jing Zhang, and Wei-Xin Ren. Threshold selection for extreme value estimation of vehicle load effect on bridges. *International journal of distributed sensor networks*, 14(2):1550147718757698, 2018.
- [21] Roger B Nelsen. *An introduction to copulas*. Springer science & business media, 2007.
- [22] Berthold Schweizer and Edward F Wolff. On nonparametric measures of dependence for random variables. *The annals of statistics*, 9(4):879–885, 1981.
- [23] Marco Scarsini. On measures of concordance. *Stochastica*, 8(3):201–218, 1984.
- [24] Gunky Kim, Mervyn J Silvapulle, and Paramsothy Silvapulle. Comparison of semiparametric and parametric methods for estimating copulas. *Computational Statistics & Data Analysis*, 51(6):2836–2850, 2007.
- [25] Thomas J Linsmeier, Neil D Pearson, et al. Risk measurement: An introduction to value at risk. 1996.
- [26] Paul H Kupiec et al. *Techniques for verifying the accuracy of risk measurement models*, volume 95. Division of Research and Statistics, Division of Monetary Affairs, 1995.
- [27] Miodrag Lovric et al. *International encyclopedia of statistical science*. Springer Berlin Heidelberg, 2011.
- [28] Bruce M Hill. A simple general approach to inference about the tail of a distribution. *The annals of statistics*, pages 1163–1174, 1975.
- [29] Jan Beirlant, Goedele Dierckx, and Armelle Guillou. Estimation of the extreme-value index and generalized quantile plots. *Bernoulli*, 11(6):949–970, 2005.

# A ACF plots of residuals after marginal modelings

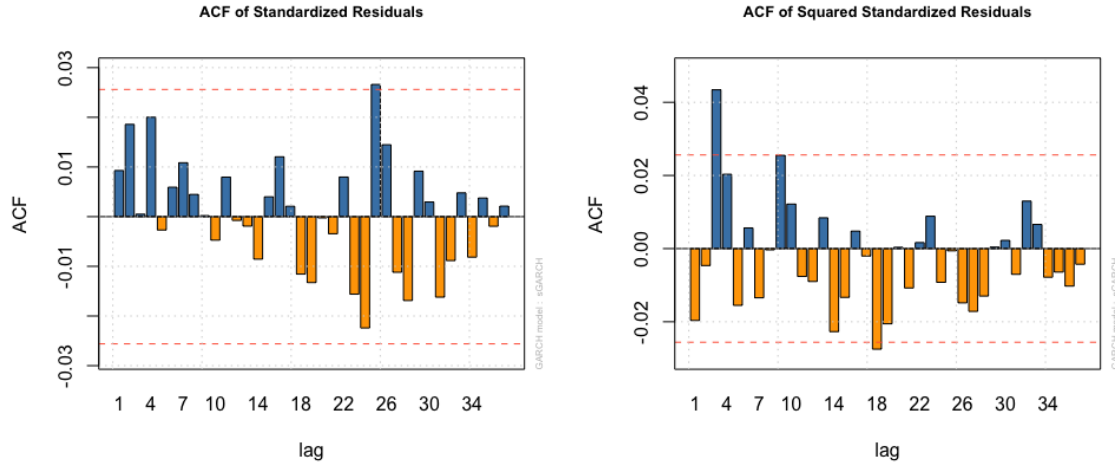


Figure A.1: Autocorrelation of standard residuals(left) and squared standard residuals(right) after ARMA(1,1)-GARCH(1,1)-norm modelling for data STOXX 50.

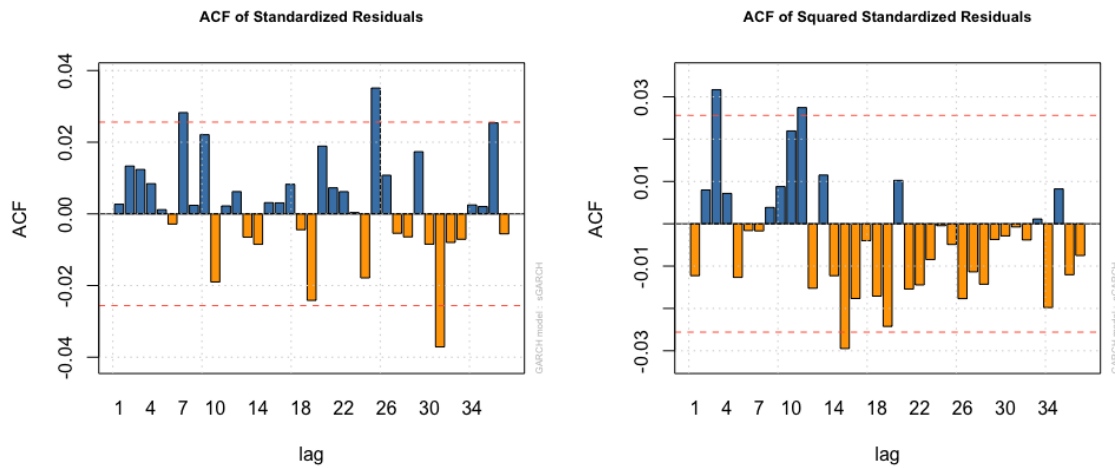


Figure A.2: Autocorrelation of standard residuals(left) and squared standard residuals(right) after ARMA(1,1)-GARCH(1,1)-norm modelling for data OMXS 30.

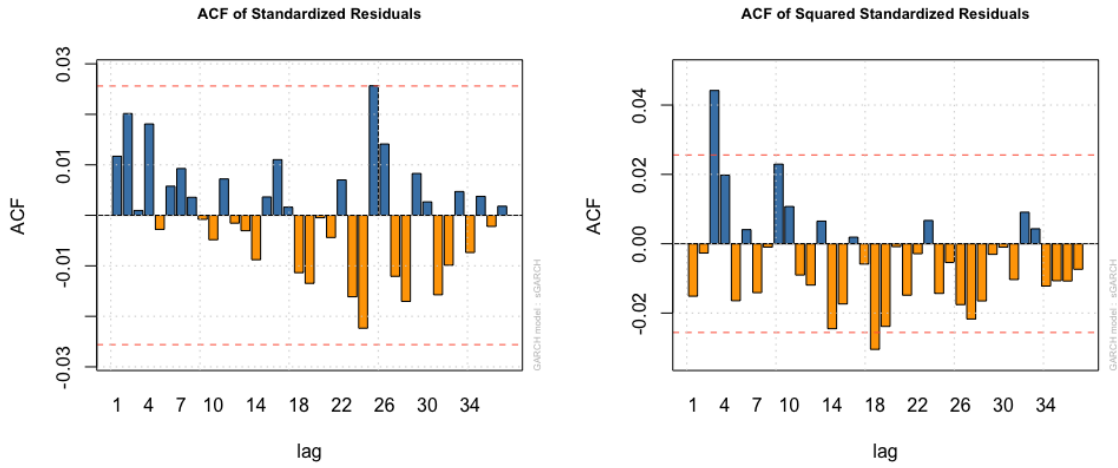


Figure A.3: Autocorrelation of standard residuals(left) and squared standard residuals(right) after ARMA(1,1)-GARCH(1,1)-t modelling for data STOXX 50.

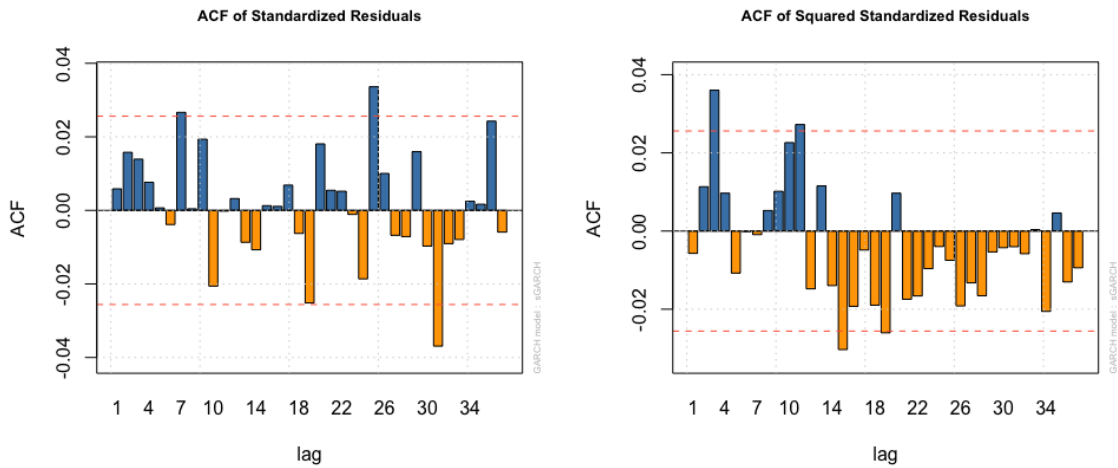


Figure A.4: Autocorrelation of standard residuals(left) and squared standard residuals(right) after ARMA(1,1)-GARCH(1,1)-t modelling for data OMXS 30.

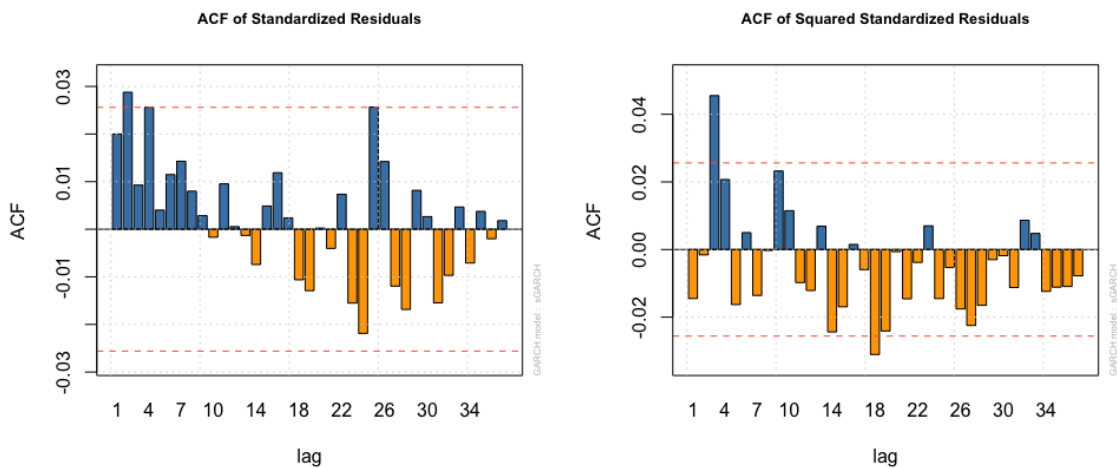


Figure A.5: Autocorrelation of standard residuals(left) and squared standard residuals(right) after ARMA(1,1)-GARCH(1,1)- skewed t modelling for data STOXX 50.

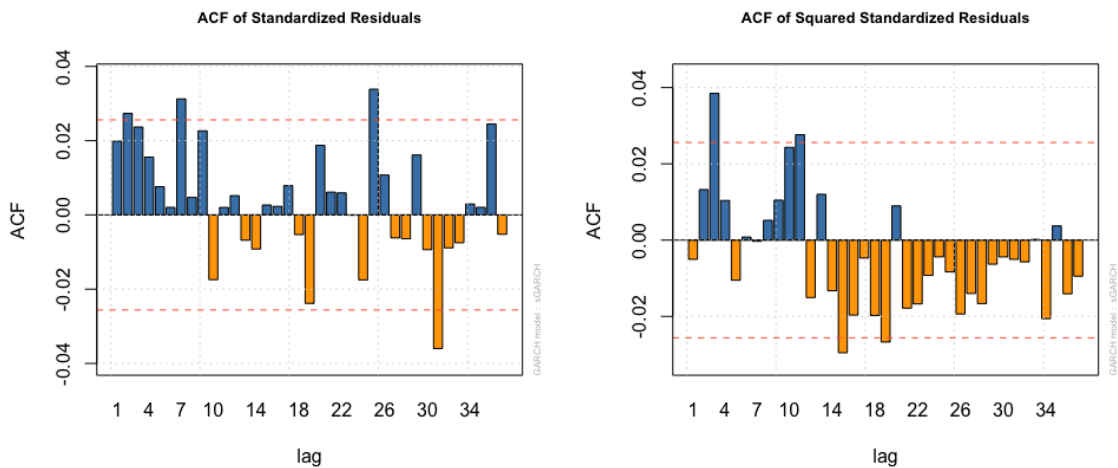


Figure A.6: Autocorrelation of standard residuals(left) and squared standard residuals(right) after ARMA(1,1)-GARCH(1,1)-skewed t modelling for data OMXS 30.

## B Mean residual life plots and threshold range plots

In this section, two kinds of threshold selection methods: mean residual life plots and plots based on the stability of parameter are shown as follows.

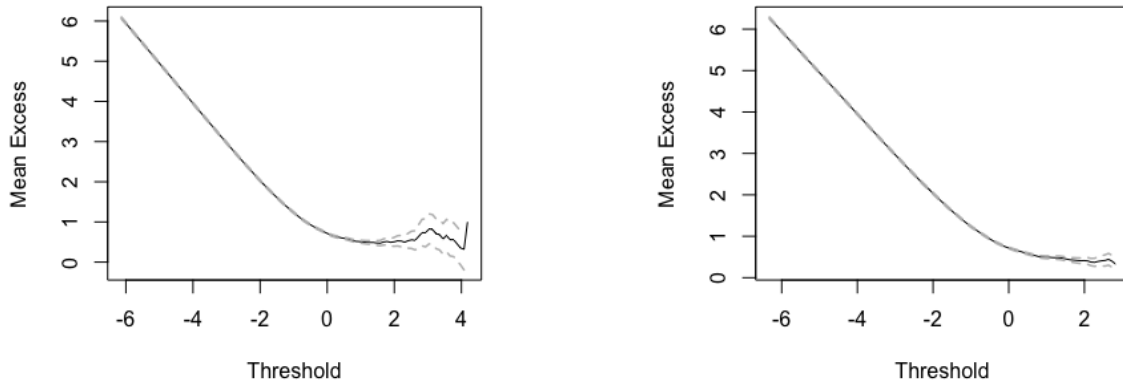


Figure B.1: Mean residual life plots of the right tail for both return residuals. The left plot is for data STOXX 50 and the right one represents OMXS 30.

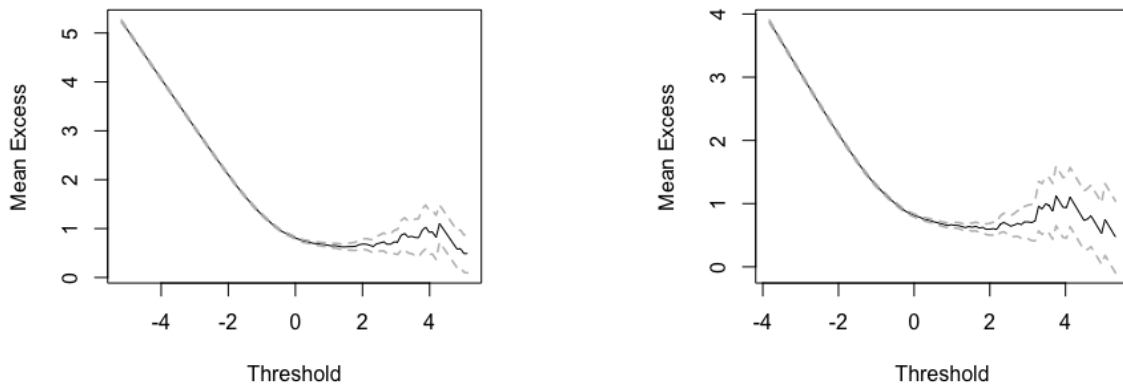


Figure B.2: Mean residual life plots of the left tail for both return residuals. The left plot is for data STOXX 50 and the right one represents OMXS 30.

```
threshrange.plot(x = stoxx_res, r = c(0, 3), type = "GP", nint = 30,  
na.action = na.fail)
```

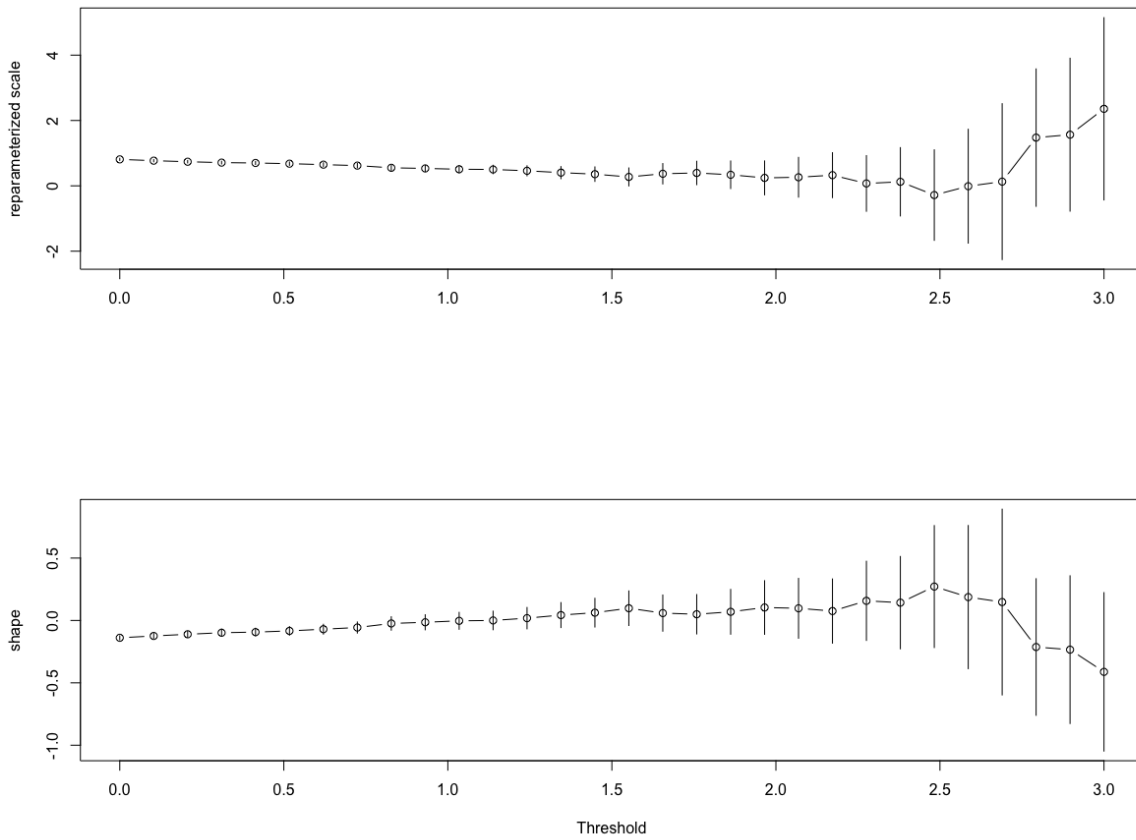


Figure B.3: Threshold range plots of the right tail for return residuals of data STOXX 50.

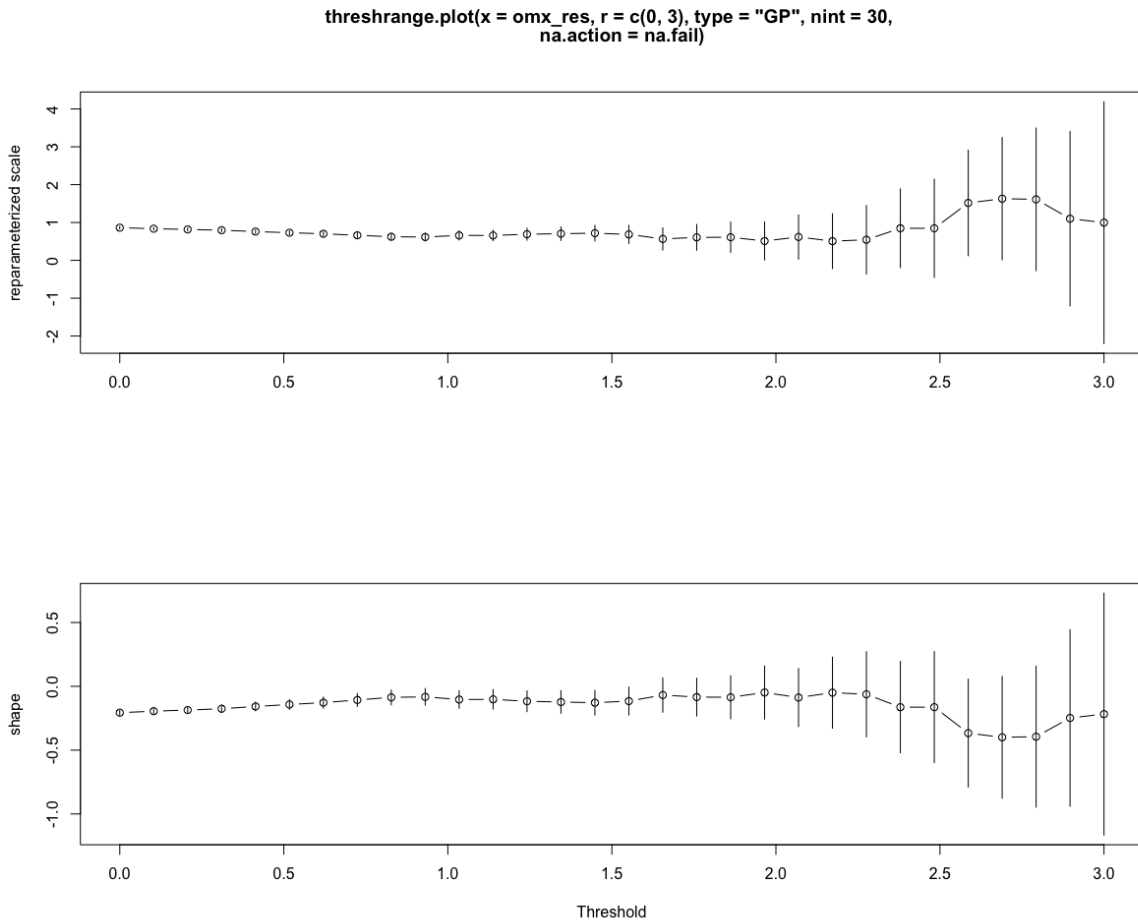


Figure B.4: Threshold range plots of the right tail for return residuals of data OMXS 30.

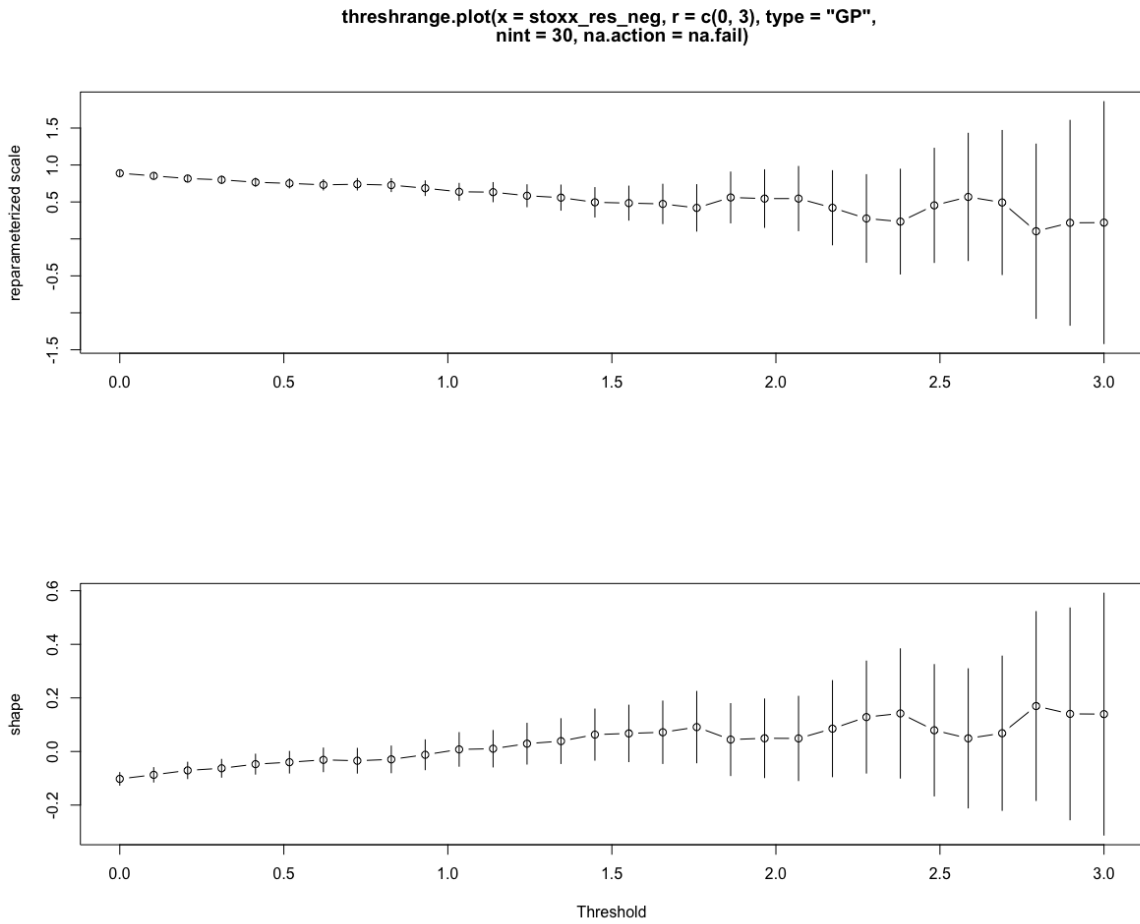


Figure B.5: Threshold range plots of the left tail for return residuals of data STOXX 50.



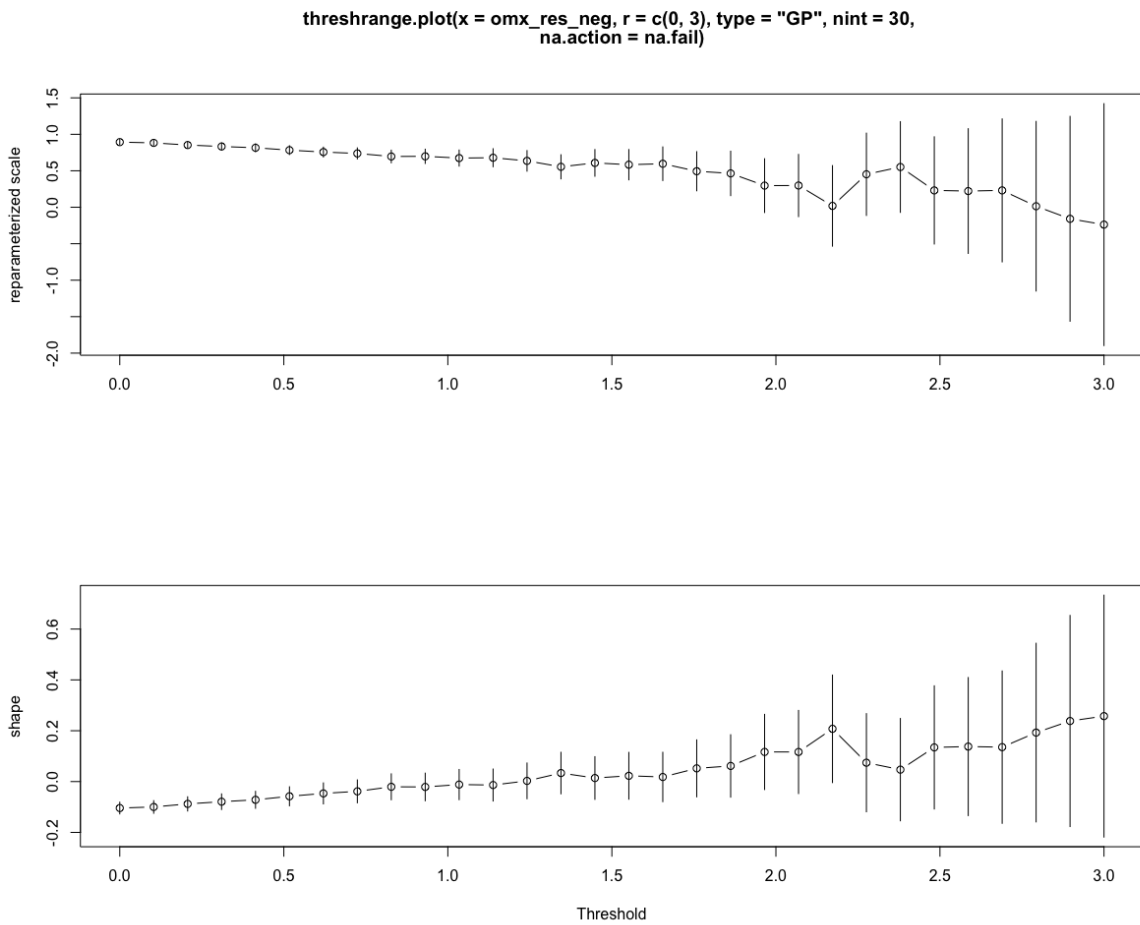
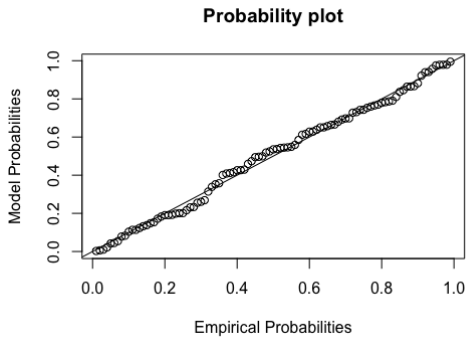


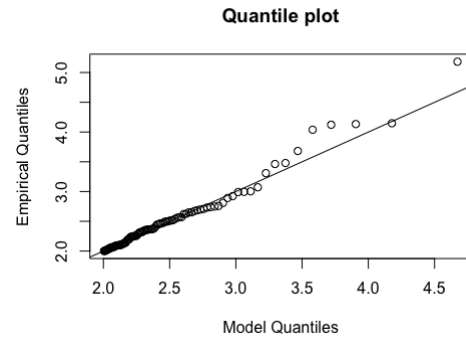
Figure B.6: Threshold range plots of the left tail for return residuals of data OMXS 30.

## C Diagnostic plots for GPD fitting

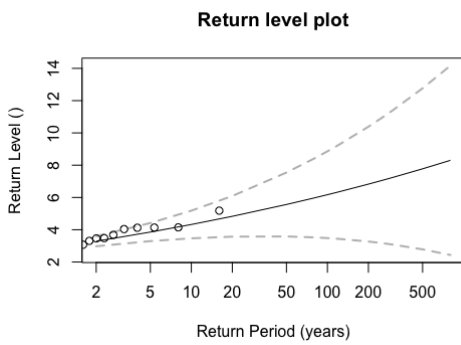
Diagnostic plots of right tail:



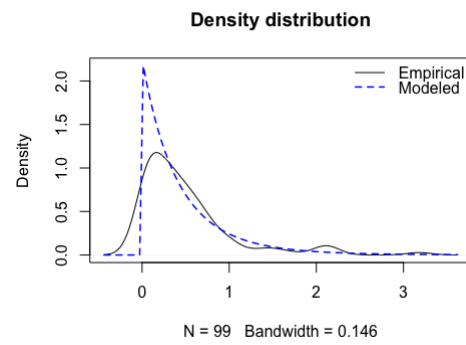
(a) Probability plot



(b) Quantile plot

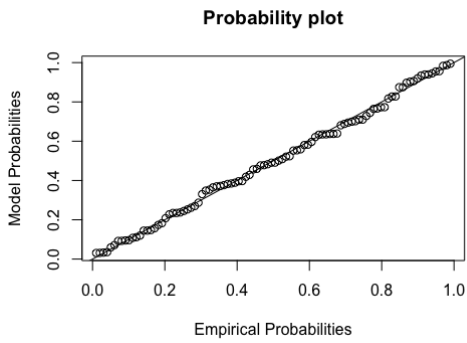


(c) Return level and confidence interval

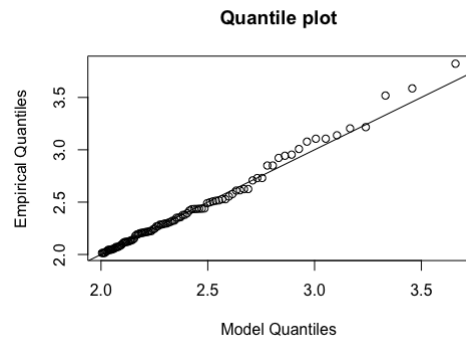


(d) Density function

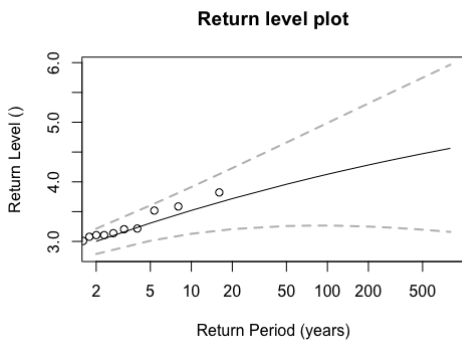
Figure C.1: Diagnostic Plots for The GPD of STOXX.



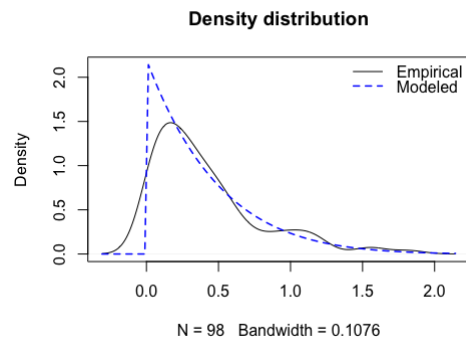
(a) Probability plot



(b) Quantile plot



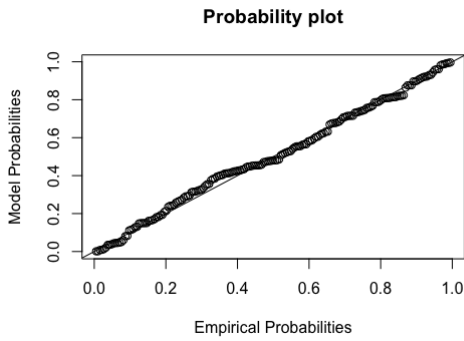
(c) Return level and confidence interval



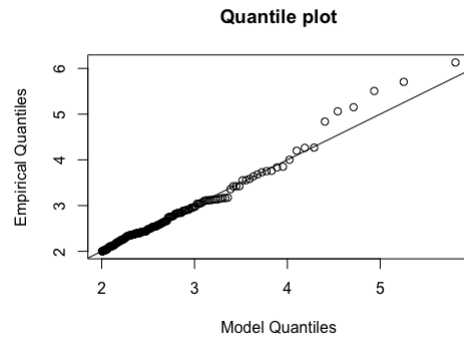
(d) Density function

Figure C.2: Diagnostic Plots for The GPD of OMX.

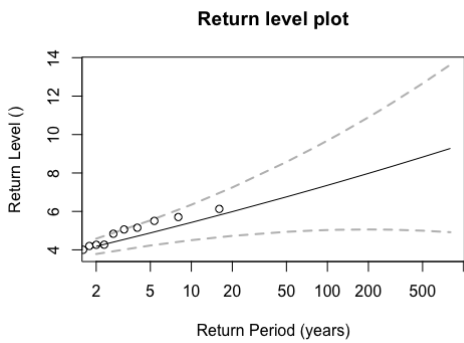
Diagnostic plots of left tail:



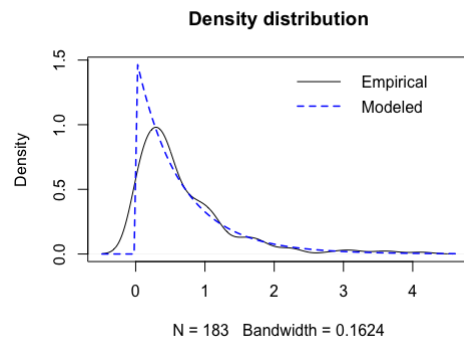
(a) Probability plot



(b) Quantile plot

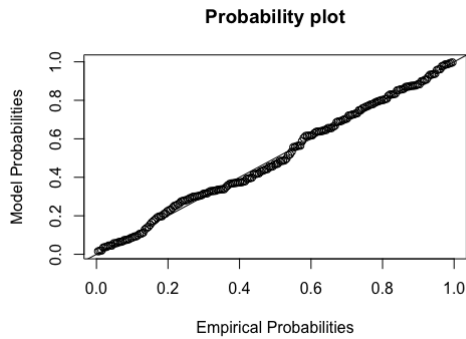


(c) Return level and confidence interval

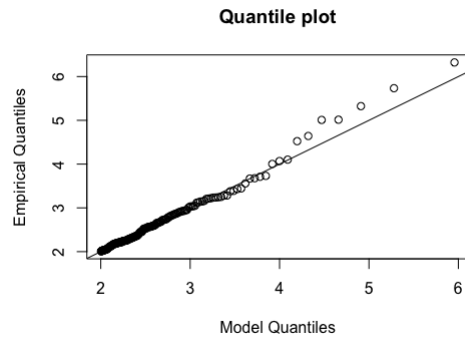


(d) Density function

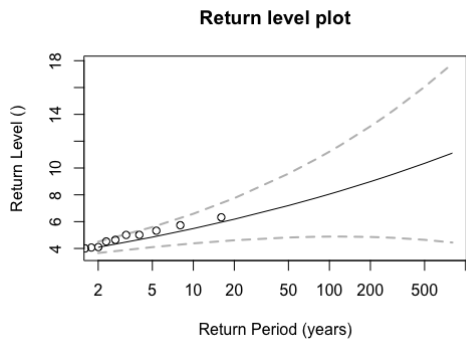
Figure C.3: Diagnostic Plots for The GPD of OMXS30.



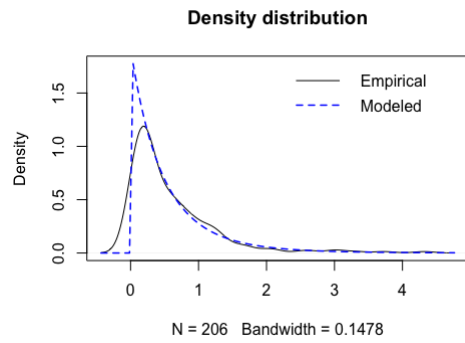
(a) Probability plot



(b) Quantile plot



(c) Return level and confidence interval



(d) Density function

Figure C.4: Diagnostic Plots for The GPD of OMXS30.

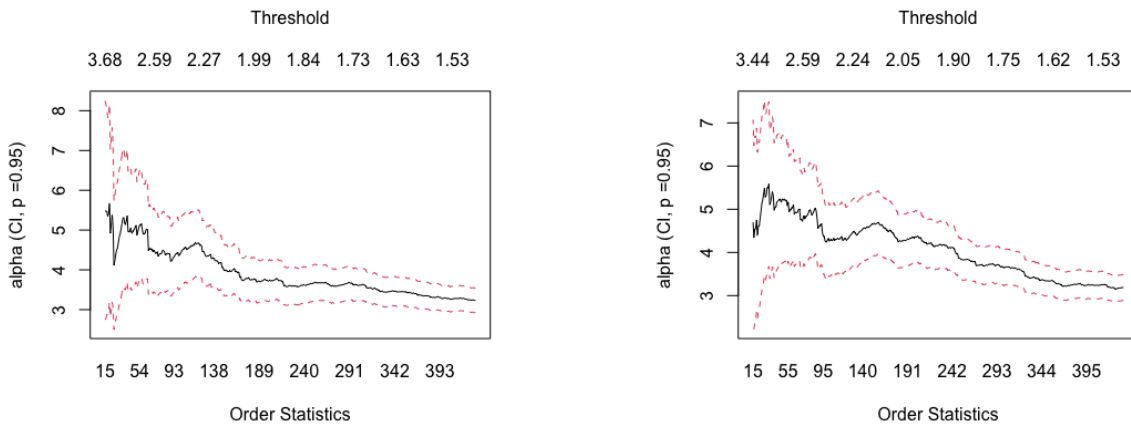


Figure C.5: Hill plots for left tails of data STOXX 50(left) and OMXS 30(right).

## D Hill Estimator simulation

The codes in R implementing the Hill estimator simulation are shown as follows:

```
Hill3 <- function(start, end, by, X){  
  
  data = sort(X)  
  N <- 5859  
  k <- seq(start, end, by)  
  num <- length(k)  
  len <- N - k  
  slope <- rep(0,num)  
  for (p in 1:num){  
    x <- rep(0, len[p])  
    y <- rep(0,len[p])  
    for (j in 1:len[p]) {  
      x[j] <- log((N+1)/j)  
      y[j] <- log(data[N-j+1])  
    }  
    plot(x,y)  
    abline(mod <- lm(y ~ x))  
    slope[p] <- mod$coefficients["x"]  
  }  
  return(slope)  
}
```

Master's Theses in Mathematical Sciences 2023:E56  
ISSN 1404-6342  
LUNFMS-3122-2023  
Mathematical Statistics  
Centre for Mathematical Sciences  
Lund University  
Box 118, SE-221 00 Lund, Sweden  
<http://www.maths.lu.se/>