DEGREE PROJECT IN MATHEMATICS,
SECOND CYCLE, 30 CREDITS
*STOCKHOLM, SWEDEN 2017*

# Copula Selection and Parameter Estimation in Market Risk Models

**CARL LJUNG**

**KTH ROYAL INSTITUTE OF TECHNOLOGY**
**SCHOOL OF ENGINEERING SCIENCES**

# Copula Selection and Parameter Estimation in Market Risk Models

**CARL LJUNG**

Degree Projects in Financial Mathematics (30 ECTS credits)
Degree Programme in Applied and Computational Mathematics (120 credits)
KTH Royal Institute of Technology year 2017
Supervisors at If Skadeförsäkring AB: Daniel Rufelt, Jerker Tapper
Supervisor at KTH: Henrik Hult
Examiner at KTH: Henrik Hult

*TRITA-MAT-E 2017:08*
*ISRN-KTH/MAT/E--17/08--SE*

Royal Institute of Technology
*School of Engineering Sciences*
**KTH** SCI
SE-100 44 Stockholm, Sweden
URL: www.kth.se/sci

# Abstract

In this thesis, literature is reviewed for theory regarding elliptical copulas (Gaussian, Student's $t$, and Grouped $t$) and methods for calibrating parametric copulas to sets of observations. Theory regarding model diagnostics is also summarized in the thesis. Historical data of equity indices and government bond rates from several geographical regions along with U.S. corporate bond indices are used as proxies of the most significant stochastic variables in the investment portfolio of If P&C. These historical observations are transformed into pseudo-uniform observations, pseudo-observations, using parametric and non-parametric univariate models. The parametric models are fitted using both maximum likelihood and least squares of the quantile function. Elliptical copulas are then calibrated to the pseudo-observations using the well known methods *Inference Function for Margins* (IFM) and *Semi-Parametric* (SP) as well as compositions of these methods and a non-parametric estimator of Kendall's tau.

The goodness-of-fit of the calibrated multivariate models is assessed in aspect of general dependence, tail dependence, mean squared error as well as by using universal measures such as Akaike and Bayesian Information Criterion, *AIC* and *BIC*. The mean squared error is computed both using the empirical joint distribution and the empirical Kendall distribution function. General dependence is measured using the scale-invariant measures Kendall's tau, Spearman's rho, and Blomqvist's beta, while tail dependence is assessed using Krupskii's tail-weighted measures of dependence (see [16]). Monte Carlo simulation is used to estimate these measures for copulas where analytical calculation is not feasible.

Gaussian copulas scored lower than Student's $t$ and Grouped $t$ copulas in every test conducted. However, not all test produced conclusive results. Further, the obtained values of the tail-weighted measures of dependence imply a systematically lower tail dependence of Gaussian copulas compared to historical observations.

# Sammanfattning

## Copulas och parameterskattning i marknadsriskmodeller

I den här uppsatsen granskas teori angående elliptiska copulas (Gaussisk, Students $t$ och s.k. Grupperad $t$) och metoder för att kalibrera parametriska copulas till stickprov av observationer. Uppsatsen summerar även teori kring olika metoder för att analysera och jämföra copula-modeller. Historisk data av aktieindex och statsobligationer från flera olika geografiska områden samt Amerikanska index för företagsobligationer används för att modellera de huvudsakliga stokastiskt drivande variablerna i investeringsportföljen hos If P&C. Dessa historiska observationer transformeras med parametriska och icke-parametriska univariata modeller till pseudolikformiga observationer, pseudo-observationer. De parametriska modellerna passas till data med både maximum likelihood och med minsta-kvadratpassning av kvantilfunktionen. Därefter kalibreras elliptiska copulas till pseudo-observationerna med de välkända metoderna *Inference Function for Margins* (IFM) och *Semi-Parametric* (SP) samt med blandningar av dessa två metoder och den icke-parametriska estimatorn av Kendalls tau.

Hur väl de kalibrerade multivariata modellerna passar de historiska data utvärderas med avseende på generellt beroende, svansberoende, rotmedelavvikelse samt genom att använda mer allmäna mått som Akaike och Bayesianskt informationskriterium, *AIC* och *BIC*. Rotmedelavvikelsen räknas ut både genom att använda den empiriska gemensamma fördelningen och den empiriska Kendall fördelningsfunktionen. Generellt beroende mäts med de skalinvarianta måtten Kendalls tau, Spearmans rho och Blomqvists beta, medan svansberoende utvärderas med Krupskiis svansviktade beroendemått (se [16]). I de fall där analytiska beräkningsmetoder inte är möjliga för copulas används Monte Carlo-simulering för att skatta dessa mått.

De Gaussiska copulas gav sämre resultat än Students $t$ och Grupperad $t$ copulas i varje enskilt test som utfördes. Dock så kan ej alla testresultat anses vara absolut definitiva. Vidare så antyder de erhållna värdena från de svansviktade beroendemåtten att modellering med Gaussisk copula resulterar i systematiskt lägre svansberoende hos modellen än hos de historiska observationerna.

# Acknowledgements

Foremost, I would like to thank my supervisors Daniel Rufelt and Jerker Tapper of Capital Management at If Skadeförskäkring for their continuous and patient support throughout the course of my Master's thesis. Their immense knowledge and insightful comments helped form the scope of this thesis and assisted me at all time during both research and writing. Without their guidance, this thesis would not have come to be.

Besides my supervisors, I would like to express my sincere gratitude to my examiner and academic advisor Professor Henrik Hult of the Department of Mathematical Statistics of the Royal Institute of Technology. Our vivid discussions provided the foundation from which the idea of this thesis grew into existence. I am grateful for the time he set aside to think of wise answers to all my questions.

I would also like to thank the entire Capital Management group at If Skadeförsäkring for their warm welcoming of me into the team. Their enthusiasm while introducing me to the work of the group has been a source of motivation.

Two persons within If Skadeförsäkring that deserve a special thanks are Solutions Developer Magnus Abrante and IT Solution Responsible Kimmo Taskinen who have been consistently helping me by providing essential computational power in the form of hardware.

My gratitude also goes to my fellow classmates and brothers in arms at the Royal Institute of Technology. The countless hours we spent in the office, both studying and not studying, over the past five years have surely paid off.

Finally, I would like to thank my family: my parents, brother and sister for their company and continuous support through all these years. Without doubt, I would not be where I am now without them. Thank you.

Stockholm, March 21, 2017

Carl Ljung

# Contents

# List of Figures

# List of Tables

# 1 INTRODUCTION

In financial mathematics, multivariate distributions and dependency of assets has been of great interest. Modern history has however highlighted the importance of proper modeling of joint extreme events and distinguishing between linear correlation and asymptotic dependence, also referred to as tail dependence.

Over the last two decades, copula modeling has found many successful applications and has seen great use in especially mathematical finance, where it has been used for just precisely dependence modeling of multivariate distributions. The finesse of copulas is that they separate the specification of univariate marginal distributions and multivariate joint distribution. Hence, rather than directly specifying a multivariate joint distribution, a model can be constructed by specifying marginal distributions and a dependence structure, i.e. a copula, for the marginals. The copula and marginals then imply a multivariate joint distribution.

If Group and its subentities (If Skadeförsäkring, If Finland, If Estonia, If Liv) are based within countries who applies the Solvency II directive of the European Union. Put shortly, without going into too much detail, the Solvency II directive means that each legal entity has a capital requirement equal to the 0.5 % quantile of the net result distribution one year into the future. This is one main reason to the interest of modelling multivariate distributions and extreme events.

Furthermore, If Group and its subentities operate throughout the entire Scandinavian region, Finland and the Baltic region. Likewise, the entities investments are mainly in the same regions. Consequently, If's financial assets are diversified over several countries and currencies. Naturally, assets in different countries and currencies are not linearly correlated. However, history has proven western economies to be greatly dependent in the case of extreme events, i.e. asymptotically dependent. Therefore, the modeling of tail dependence and the use of copulas have earned great interest.

## 1.1 If's Portfolio and the ESG Model

If's investment portfolio consists of 10–15 % equity, such as e.g. stock shares, and 85–90 % fixed income instruments, such as e.g. bonds and floating-rate notes. The majority of the assets are allocated in the Nordic region and hence the state of the Swedish, Norwegian, and Danish economies pose as significant drivers of market risk for If's portfolio. Additionally, the state of the economy of the European monetary region and the US economy should be considered to be relevant risk drivers. For convenience, these risk drivers are abbreviated as the SEK, NOK, DKK, EUR, and USD economy.

Traditionally, the Capital Management unit within If Risk Management has modeled the multivariate distribution of the portfolio assets using a Economic Scenario Generator (ESG). This model is designed to generate outcomes of the return or change in value of asset categories for several relevant monetary regions. For example, the model is used to generate outcomes of the one-year return on equity in the Swedish market.

In general, the ESG model employs the traditionally appreciated approach of using a Gaussian copula (explained in Section 2.2.3.1) as a model for the dependence structure of the multivariate distribution of the portfolio assets' driving Brownian motions. However, the Gaussian copula family implies zero asymptotic tail dependence (see Appendix C.1) which may result in an inadequate model in the aspect of dependence.

### 1.1.1 Assumptions of Portfolio Components

The stochastic behaviour of the fixed income possession of If's portfolio can in a simple model be described by the risk-free rate and credit spreads, where credit spread is the additional yield implied by a risky bond compared to a risk-free bond. Further, a reasonable model for the stochastic price change of equity in separate economies is to use stock indices consisting of stock shares that has great influence on the equity market of each separate economy.

## 1.2 Objective

This thesis aims to partly review and summarize theory regarding copula model calibration and to partly apply the theoretical methods to If's investment portfolio. The thesis objectives can thus be formulated as below

- The first objective of this thesis is to investigate methods for calibrating copula models to the driving stochastic factors of If's investment portfolio and methods to assess the adequacy- and goodness-of-fit of the models.
- The second objective of this thesis is to analyze and compare a set of copula candidates for modelling the multivariate distribution and dependence structure of the macroeconomic variables that pose as the major risk factors for If's investment portfolio. Additionally, the dependence structure and goodness-of-fit of non-Gaussian copulas should be compared with the Gaussian copula, since this is the copula model currently used in the ESG.

The considered macroeconomic variables are equity, risk-free rate, and credit spread for each of the economies SEK, NOK, DKK, EUR, and USD. Furthermore, the general idea is to attain the first objective by reviewing existing publications on the subject of copula modeling and summarize the relevant material along with a mathematical background. The second objective will be achieved by calibrating copula models to authentic market data representative for If's investment portfolio and then analyzing the goodness-of-fit each model.

The ultimate purpose of use of the ESG is Solvency II capital requirement calculations, i.e. the multivariate outcomes generated by the ESG are intended to be used for Monte Carlo approximation of a 99.5 % quantile. Thus, a copula calibration method that is weighted towards the tails can seem like a sensible approach. However, for the purpose of the first objective, this thesis will focus on more widely used calibration methods and not delve deeper into special purposed calibration criteria. No portfolio function will be specified based on the copula and thus no risk measures, such as the 99.5 % quantile, will be computed. The author acknowledges that this can be considered to be the ultimate purpose of a copula model in the given context, but this thesis limits itself to merely study the calibration methodology, resulting model dependence structure and overall model fit.

### 1.2.1 Limitations

The copulas considered within this thesis is limited to the Gaussian, Student's $t$ and Grouped $t$ copula due to their applicability in multivariate contexts, see Section 2.2.3.1. Parametric versions of copula families such as the Archimedean copulas have in general only one or sometimes two parameters which limits the possibilities of calibrating a high dimensional parametric model to have an adequate dependence structure. Archimedean copulas has therefore been excluded from the studies in this thesis, but the curious reader can find more about them in [17].

Univariate models for each of the marginal variables are limited to three parametric models and one non-parametric model. The chosen parametric models are the Normal, Student's $t$ and Polynomial normal distribution, see Section 2.2.1, and the non-parametric model is the empirical distribution, see Section 2.2.4.

This thesis does not put a great emphasis on the process of transforming observations of the marginal variables into supposedly i.i.d outcomes.

**Frequency of Data used for Modelling** As any statistician knows, having a large sample of observations provides multiple benefits almost regardless of type of study. However, the limited supply of relevant data is a common obstacle faced when proceeding from the drawing board to the real world.

When studying financial variables as is done in this thesis, there is plenty of freedom of choice regarding what time period be considered to be one single observation. In detail, one can choose whether to model based on daily observations, monthly observations or yearly observations, etc. When determining what is preferable there are two main concerns:

- **Sample size** Clearly, when modelling based on daily observations, history will provide larger samples of observations than when modelling based on e.g. yearly observations.
- **Purpose of Model** Taking into account the actual purpose of use of the model is a sensible way to avoid over-modelling. For example, one could imagine the price process of a certain asset to exhibit a certain trend or large changes when making frequent observations, but when observing over longer time periods these effects are netted out.

Evidently this is a trade-off problem with no absolute solution. One would want to set the length of the observation periods to be relevant with respect to the intended use of the model while at the same time maximizing the number of usable historical data points.

As stated earlier, the ultimate purpose of the copula models examined is to approximate a quantile value of the stochastic net result one year ahead. Consequently, it would be sensible for the modelling to be based on yearly observations. However, this would yield very small samples of observations, as will be made evident in the methodology section of this thesis. Furthermore, the quantile approximation part is not included in this thesis. Therefore, this thesis takes the approach to study daily observations in order to maximize the amount of usable historical data points.

### 1.3 Thesis Outline

The first subsequent section, Section 2, presents a mathematical background to the topics of copulas and distribution parameter estimation. It also presents the relevant concepts of copula and dependence modeling. All in all, Section 2 constitutes the foundation of this thesis as it presents both background theory and applied theory regarding copula modeling methods that are used. Thus, the advanced reader on copula theory might find it appropriate to go directly to Section 2.4 and the modeling methods.

Furthermore, Section 3 presents data set used for copula modeling and illustrates the circumstantial details of the application of the theory in Section 2. The results are presented in Section 4 and Section 5 concludes.

Regarding the objectives of this thesis, the results of the investigation and summarizing of copula calibration methods as well as adequacy- and goodness-of-fit tests are presented throughout Section 2 and are briefly discussed in Section 5. Section 3 attends mainly to the second objective by presenting the circumstantial application of the copula modeling theory in Section 2. Additionally, the results obtained through the methods in Section 3 are presented in Section 4 and discussed in Section 5.

# 2 Mathematical Theory

In the first subsection below, Section 2.1, stochastic processes for the macroeconomic quantities equity, interest rate, and credit spreads are presented. These mathematical models are used to demonstrate how time series of the quantities can be transformed into time-independent series in theory.

Section 2.2 gives a rigorous theoretical background on copulas from a probability theory perspective while Section 2.3 presents tools for fitting copula models to a set of observations and for analyzing the adequacy- and goodness-of-fit. The somewhat more important Section 2.4 then illustrates how to apply the fitting and analysis tools of Section 2.3 to model dependence and multivariate distributions using the copula theory of Section 2.2.

## 2.1 Models for Macroeconomic Quantities

### 2.1.1 Equity

As an illustrative example of the equity model, take $S_t$ to be the share price of a stock at time $t$. By implementing the lognormal model, the random process $S_t$ is assumed to have the dynamics of a geometric Brownian motion, i.e.

$$\begin{aligned} dS_t &= \left(\mu + \frac{\sigma^2}{2}\right) S_t dt + \sigma S_t dW_t \\ S_0 &= s_0 \end{aligned} \tag{2.1}$$

where $W_t$ is a Brownian motion (see [1]). Further, by proposition 5.2 in [1], this implies that

$$S_t = S_{t-1} e^{\mu + \sigma(W_t - W_{t-1})}. \tag{2.2}$$

Since $W_t$ is a Brownian motion it follows that $W_t - W_{t-1} \in N(0,1)$, or $Z_t = \mu + \sigma(W_t - W_{t-1}) \in N(\mu, \sigma)$. Consequently, for the one time period return on the stock it holds that

$$R_t = \frac{S_t}{S_{t-1}} = e^{Z_t} \quad \Leftrightarrow \quad \ln R_t = Z_t. \tag{2.3}$$

In other words, the log-returns are normally distributed provided that the stock price process is a geometric Brownian motion.

Moreover, if the current time is denoted by 0 and $S_{-n}, S_{-n+1}, \ldots, S_{-1}, S_0$ is assumed to be a time series of historical prices from the $n$ previous equally spaced points in time, then (2.2) indicates a strong dependence between the data points. The returns, on the other hand, are independent and equally distributed if $\mu$ and $\sigma$ are time-invariant and if the non-overlapping increments of the process $W_t$ are independent (the latter condition follows if $W_t$ is a Brownian motion). In terms of $Z_t$ this translates to $Z_t$ having a time-independent distribution. These model assumptions are supported by the claim in [12] that historical return series usually only are weakly dependent and can be assumed to be independent and identically distributed.

In reality however, the normality of $Z_t$ proves to be a crude assumption. Hence, the distribution of $Z_t$ is not limited to the normal distribution in this thesis.

Conclusively, by this model the future returns have the same distribution as the historical returns and the observed returns are then representative for the outcome of $R_1$, the return for the next time period.

### 2.1.2 Interest Rates

Let $p_t(T)$ denote the price of a zero-coupon bond at time $t$ with maturity at time $T$. If $r_t^c(T)$ denotes the continuously compounded risk-free rate at time $t$ with maturity at time $T$ and $r_t(T)$ denotes the annually compounded risk-free rate at time $t$ with maturity at time $T$, then

$$p_t(T) = e^{-\int_t^T f(t,s)ds} = e^{-r_t^c(T)(T-t)} = (1 + r_t(T))^{-(T-t)} \tag{2.4}$$

where $t$ and $T$ are measured in years. Further, let $T_M = T - t$ denote the time to maturity of a zero-coupon bond at time $t$, then the zero-coupon bond price at time $t$ can be expressed with $T_M$ as a parameter, i.e. $p_t(T_M) = e^{-r_t^c(T_M)T_M}$. Now, consider the price of a bond at time $t$ with maturity at time $T$, i.e. $p_t(T_M)$, and the price of a bond at time $t-1$ with maturity at time $T-1$, i.e. $p_t(T_M)$. Then, the change in market price of a $T_M$-bond is

$$R_t = \frac{p_t(T_M)}{p_{t-1}(T_M)} = e^{-\left(r_t^c(T_M) - r_{t-1}^c(T_M)\right)T_M} \quad \Rightarrow \quad Z_t := \frac{-1}{T_M}\ln R_t = r_t^c(T_M) - r_{t-1}^c(T_M).$$

Thus, studying the distribution of the log-returns is equivalent to studying the distribution of the change in the continuously compounded risk-free rate, when considering zero-coupon bonds. Further, this is approximately true also if the annually compounded risk-free rate is used. By employing Taylor expansion, one obtains

$$R_t = \frac{p_t(T_M)}{p_{t-1}(T_M)} = \left(\frac{1 + r_t(T_M)}{1 + r_{t-1}(T_M)}\right)^{-T_M} \quad \Rightarrow \quad Z_t := \frac{-1}{T_M}\ln R_t = \ln\left(\frac{1 + r_t(T_M)}{1 + r_{t-1}(T_M)}\right) \approx r_t(T_M) - r_{t-1}(T_M). \tag{2.5}$$

The validity of this linear approximation clearly depends on the length of the time periods. For example, if $t$ is measured in days and $R_t$ is the daily change in price of a zero-coupon bond which matures in $T_M$ days, then the change in price can be assumed to be small and consequently the linear approximation above is justified.

As in the case of equity, the time series of the process $Z_t$ will be assumed to be independent and identically distributed. This means that the observed changes in the risk-free rate are assumed to be representative for the future changes in the risk-free rate.

### 2.1.3 Credit Spreads

Consider a corporate bond with time to maturity $T_M$ and price process $c_t(T_M)$. A corporate bond carries a certain credit risk in the form of counterparty default risk but does in turn generally give a yield to maturity higher than the risk-free rate. The amount of which the yield of a corporate bond exceeds the risk-free rate is referred to as the credit spread and will be denoted $s_t(T_M)$. The price process $c_t(T_M)$ can now be formulated as

$$c_t(T_M) = e^{-\left(r_t^c(T_M) + s_t^c(T_M)\right)T_M} = \left(1 + r_t(T_M) + s_t(T_M)\right)^{-T_M}$$

where $s_t^c$ is the continuously compounded credit spread and $s_t$ is the annually compounded credit spread. By operations similar to the preceding sections, one can derive

$$R_t = \frac{c_t(T_M)}{c_{t-1}(T_M)} = e^{-\left(r_t^c(T_M) + s_t^c(T_M) - r_{t-1}^c(T_M) - s_{t-1}^c(T_M)\right)T_M} \tag{2.6}$$

$$\Rightarrow \quad Z_t := \frac{-1}{T_M}\ln R_t + \left(r_t^c(T_M) - r_{t-1}^c(T_M)\right) = s_t^c(T_M) - s_{t-1}^c(T_M). \tag{2.7}$$

Let $\mathscr{F}_t^c = \sigma\left(\ldots, r_{t-2}^c(T_M), r_{t-1}^c(T_M), r_t^c(T_M)\right)$, i.e. $\mathscr{F}_t^c$ is the sigma-algebra generated by the random variables $\ldots, r_{t-2}^c(T_M), r_{t-1}^c(T_M), r_t^c(T_M)$ (the reader is referred to [1] for more information about sigma-algebras). Then, the process of $Z_t|\mathscr{F}_t^c$ is a process driven only by the stochastics of the changes in the credit spread. In other words, provided that the risk rate is known up to the considered time, the distribution of $Z_t$ depends solely on the stochastics of $s_t^c(T_M) - s_{t-1}^c(T_M)$.

Furthermore, using a similar argument to the one in the preceding section, an expression for $Z_t$ using the annually compounded credit spread can be obtained. By linear approximation, $Z_t$ can be expressed as the negative change in the credit spread. Hence, under the assumption that the credit spread changes are independent for non-overlapping time periods and constitute a process with stationary distribution, the observations of $Z_t$ will be outcomes of independent and identically distributed variables. Thus, the historical outcomes of $Z_t$ will be representative for future changes in the credit spread.

## 2.2 Probability Distributions

The use of copulas allows for marginal distributions and joint distribution to be specified separately. More specifically, when modeling a multivariate joint distribution using copulas, the marginals are specified individually and a copula is used to set the dependence between the marginals. Thereby, this section presents theory regarding univariate and multivariate distributions as well as copulas.

An important family of distribution in mathematical finance and quantitative risk analysis is the family of elliptical distributions. These distributions pose good applicability in problems in high dimensions. Consequently, this section focuses on the family of elliptical distributions and two its more famous members, namely the normal distribution and Student's $t$ distribution.

### 2.2.1 Univariate Distributions

Generally, financial returns distributions are heavy-tailed and slightly skewed, i.e. asymmetric. Consequently, these are the features that a parametric marginal model should capture.

Three univariate distributions and their characteristics are presented in this section: the normal, Student's $t$, and polynomial normal distribution. For the purpose of providing a thorough theoretical background, the distribution and density functions of the normal distribution and Student's t- distribution is defined in this section. The more exotic polynomial normal distribution, which allows both heavy tails and skewness, is presented with a shorter derivation.

**2.2.1.1 Normal Distribution and Student's t Distribution** A random variable $Z$ is said to be standard normally distributed, $N(0,1)$, if its density function, $f_Z$ is of the form

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} =: \phi(z). \tag{2.8}$$

The corresponding distribution function is denoted $\Phi$. Moreover, if $X = \mu + \sigma Z$, then $X$ has a general normal distribution with mean $\mu$ and variance $\sigma^2$, which is denoted $N(\mu, \sigma)$. The density function of $X$ is then

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{2.9}$$

The distribution function of $X$ can be expressed using the standard normal distribution function as $F_X(x) = \Phi((x-\mu)/\sigma)$. The normal distribution has neither heavy tails or asymmetric shape and is thereby a poor candidate in theory, as mentioned in the preceding section. However, it is included to pose as a benchmark.

Furthermore, if $Y = \mu + \sigma\sqrt{\nu/S_\nu}Z$, where the random variable $S_\nu$ has a $\chi^2$-distribution with mean $\mu$, scale parameter $\sigma$ and $\nu$ degrees of freedom, then $Y$ has a Student's $t$ distribution with $\nu$ degrees of freedom. The density function of $Y$ is then

$$f_Y(y) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sigma\sqrt{\nu\pi}} \left(1 + \frac{(y-\mu)^2}{\nu\sigma^2}\right)^{-\frac{\nu+1}{2}}. \tag{2.10}$$

General Student's $t$ distribution is denoted $t(\mu, \sigma, \nu)$. When $\mu = 0$ and $\sigma = 1$, the distribution is referred to as standard Student's $t$ distribution and the corresponding distribution function is denoted as $t_\nu$. It follows that if $Y \in t(\mu, \sigma, \nu)$, then $\text{Var}[Y] = \sigma^2\nu/(\nu-2)$ and consequently $(Y-\mu)/\sigma$ has standard Student's $t$ distribution with variance $\nu/(\nu-2)$.

**2.2.1.2 Polynomial Normal Distribution** In this thesis, the polynomial normal (PN) distribution model is limited in several ways in order to fit the intended application. The construction of our polynomial normal model relies heavily on Proposition 6.3 in [12], cited below.

**Proposition 2.1.** *If $g: \mathbb{R} \to \mathbb{R}$ is a non-decreasing and left continuous function, then, for any random variable $Z$, it holds that $F^{-1}_{g(Z)}(p) = g\left(F^{-1}_Z(p)\right)$, for all $p \in (0,1)$.*

The reader is referred to [12] for a proof of the proposition. Now, the approach – evidently inspired by [12] – is to model the random variable under study as a polynomial of a standard normal random variable, i.e. $Z = g(Y)$ where $Y$ is standard normally distributed. Then, if $g$ is a non-decreasing and left continuous function, it follows from Proposition 2.1 that $F_Z^{-1}(p) = g\left(F_Y^{-1}(p)\right) = g\left(\Phi^{-1}(p)\right)$. In other words, the quantile function of $Z$ is a function of the standard normal quantile function.

Furthermore, if $g$ is chosen as a polynomial of odd degree, then, provided a sufficient coefficient constraint, $g$ is both non-decreasing and continuous. Hence, $g$ is set as

$$g\left(y;\boldsymbol{\theta}\right) = \theta_0 + \theta_1 y + \theta_2 y^2 + \theta_3 y^3, \tag{2.11}$$

where $\boldsymbol{\theta} = [\theta_0, \theta_1, \theta_2, \theta_3]^T$. Here, it is possible to have higher order terms in the model, however, it would increase the complexity of the model since more parameters would be included. A polynomial degree of three thus provides a sufficient model while preserving simplicity (a polynomial degree of one would imply that $Z$ is normally distributed with mean $\theta_0$ and standard deviation $\theta_1$ and is thereby not considered).

Provided the constraints $\theta_3 \geq 0$ and $g'\left(y;\boldsymbol{\theta}\right) \geq 0$, $g$ is non-decreasing and by Proposition 2.1 it holds that

$$F_Z^{-1}(p) = F_{g(Y;\boldsymbol{\theta})}^{-1}(p) = \theta_0 + \theta_1 \Phi^{-1}(p) + \theta_2 \Phi^{-1}(p)^2 + \theta_3 \Phi^{-1}(p)^3. \tag{2.12}$$

Moreover, by noticing that $g'\left(y;\boldsymbol{\theta}\right) = \theta_1 + 2\theta_2 y + 3\theta_3 y^2$ has a global minimum if $\theta_3 \geq 0$, the constraint $g'\left(y;\boldsymbol{\theta}\right) \geq 0$ can be equivalently formulated as $g'\left(y^\star;\boldsymbol{\theta}\right) \geq 0$ and $\theta_3 \geq 0$, where $y^\star = \operatorname{argmin}\left\{g'\left(y;\boldsymbol{\theta}\right)\right\}$. Due to the constraint $\theta_3 \geq 0$, $g'\left(y;\boldsymbol{\theta}\right)$ is convex and it follows that $y^\star$ is the solution to $g''\left(y;\boldsymbol{\theta}\right) = 0$, i.e. $y^\star = -\theta_2/(3\theta_3)$. Thus, $g'\left(y^\star;\boldsymbol{\theta}\right) \geq 0$ is equivalent to $3\theta_1\theta_3 - \theta_2^2 \geq 0$.

Further, $g\left(y;\boldsymbol{\theta}\right)$ is strictly increasing if $g'\left(y;\boldsymbol{\theta}\right) > 0$, which translates into $3\theta_1\theta_3 - \theta_2^2 > 0$. If $g\left(y;\boldsymbol{\theta}\right)$ is strictly increasing, then

$$F_Z(x) = \mathrm{P}\left(g(Y;\boldsymbol{\theta}) \leq x\right) = \mathrm{P}\left(Y \leq g^{-1}(x;\boldsymbol{\theta})\right) = \Phi\left(g^{-1}(x;\boldsymbol{\theta})\right),$$

since $Y$ has standard normal distribution. By the chain rule and the inverse function theorem, the density function of $Z$ is

$$f_Z(x) = \frac{d}{dx}\left(\Phi\left(g^{-1}(x;\boldsymbol{\theta})\right)\right) = \phi\left(g^{-1}(x;\boldsymbol{\theta})\right)\frac{d}{dx}\left(g^{-1}(x;\boldsymbol{\theta})\right) = \frac{\phi\left(g^{-1}(x;\boldsymbol{\theta})\right)}{g'\left(g^{-1}(x;\boldsymbol{\theta})\right)}.$$

Here $g^{-1}(x;\boldsymbol{\theta})$ is defined as the real root of the polynomial $g\left(y;\boldsymbol{\theta}\right) - x$, which, since $g\left(y;\boldsymbol{\theta}\right)$ is strictly increasing, will be unique.

Furthermore, due to the symmetry of the standard normal distribution it follows that

$$\mathbb{E}[Z] = \mathbb{E}\left[g(Y;\boldsymbol{\theta})\right] = \theta_0 + \theta_2. \tag{2.13}$$

Moreover, this yields that $\mathrm{Var}[Z] = \mathbb{E}\left[Z^2\right] - \mathbb{E}[Z]^2 = \mathbb{E}\left[Z^2\right] - (\theta_0 - \theta_2)^2$, where $\mathbb{E}\left[Z^2\right] = \theta_0^2 + \left(2\theta_0\theta_2 + \theta_1^2\right)\mathrm{E}\left[Y^2\right] + \left(2\theta_1\theta_3 + \theta_2^2\right)\mathbb{E}\left[Y^4\right] + \theta_3^2\mathbb{E}\left[Y^6\right]$. Thereby, since $\mathbb{E}\left[Y^2\right] = 1$, $\mathbb{E}\left[Y^4\right] = 3$, and $\mathbb{E}\left[y^6\right] = 15$, the variance of $Z$ is obtained as

$$\mathrm{Var}[Z] = \theta_1^2 + 2\theta_2^2 + 6\theta_1\theta_3 + 15\theta_3^2 \tag{2.14}$$

### 2.2.2 Spherical and Elliptical Distributions

A random vector $\mathbf{Z}$ has a spherical distribution if its distribution is invariant under rotatios and reflections, i.e. its distribution is spherically symmetric. Mathematically, this implies that $\mathbf{Z}$ has a spherical distribution if

$$\mathbf{OZ} \stackrel{d}{=} \mathbf{Z}, \text{ for every orthogonal matrix } \mathbf{O}. \tag{2.15}$$

Two examples of spherical distributions are $d$-dimensional standard normal distribution ($\mathbf{Z} \in \mathrm{N}_d(\mathbf{0}, \mathbf{I})$, and independent components) and $d$-dimensional standard normal variance mixture ($W\mathbf{Z}$ where $\mathbf{Z}$ is standard

normal and independent of $W$ which is a non-negative random variable).
A random vector $\mathbf{Y}$ has an elliptical distribution if it holds that

$$\mathbf{Y} \stackrel{d}{=} \boldsymbol{\mu} + \mathbf{A}\mathbf{Z} \tag{2.16}$$

where $\mathbf{Z}$ is spherically distributed. Then, $\boldsymbol{\mu}$ is the location vector and $\mathbf{A}\mathbf{A}^T = \boldsymbol{\Sigma}$ is called the dispersion matrix. Further, the correlation (see Section 2.3.4.1 for correlation) matrix, $\mathbf{R}$, can be derived from the dispersion matrix through

$$R_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}} \tag{2.17}$$

Moreover, if $\mathbb{E}\left[|\mathbf{Y}|^2\right] < \infty$, then the covariance of $\mathbf{Y}$ exists finitely and it follows that the dispersion matrix is equal to a positive real constant times $\mathrm{Cov}[\mathbf{Y}]$. Additionally, the correlations in (2.17) are equivalent to (2.53).

**2.2.2.1 Multivariate Normal Distribution** If $\mathbf{Z} \in N_d(\mathbf{0}, \mathbf{I})$, then $Y$ in (2.16) has multivariate normal distribution, i.e. $\mathbf{Y} \in N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. It follows that the density of $\mathbf{Y}$ is

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^d |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right) \tag{2.18}$$

where $|\boldsymbol{\Sigma}|$ is the determinant of $\boldsymbol{\Sigma}$. Further, this implies that the joint distribution function of $\mathbf{Y}$ is

$$F_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^d |\boldsymbol{\Sigma}|}} \int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_d} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right) dx_1 \ldots dx_d \tag{2.19}$$

**2.2.2.2 Multivariate Student's t Distribution** An important variant of (2.16) presented in [12] is the normal variance mixture distribution of random vectors with stochastic representation

$$\mathbf{Y} = \boldsymbol{\mu} + W\mathbf{A}\mathbf{Z}$$

where $\mathbf{Z} \in N_d(\mathbf{0}, \mathbf{I})$ and $W$ is a non-negative random variable independent of $\mathbf{Z}$. By setting $W \stackrel{d}{=} \sqrt{\nu/S_\nu}$, where $S_\nu \in \chi^2(\nu)$, the resulting distribution of $\mathbf{Y}$ is a multivariate Student's $t$ distribution with $\nu$ degrees of freedom, denoted $t_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$.

In [21], the joint density function of $\mathbf{Y} \in t_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$ formulated as

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\pi\nu)^d |\boldsymbol{\Sigma}|}} \left(1 + \frac{(\mathbf{y}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})}{\nu}\right)^{\frac{-\nu+d}{2}}, \tag{2.20}$$

where $|\boldsymbol{\Sigma}|$ is the determinant of $\boldsymbol{\Sigma}$. Further, this implies that the joint distribution function of $\mathbf{Y}$ is

$$F_{\mathbf{Y}}(\mathbf{y}) = \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\pi\nu)^d |\boldsymbol{\Sigma}|}} \int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_d} \left(1 + \frac{(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{\nu}\right)^{-\frac{\nu+d}{2}} dx_1 \ldots dx_d \tag{2.21}$$

### 2.2.3 Copula Families and Properties

Copulas provide a method of introducing a dependence structure to the components of a random vector with known univariate marginal distribution. As the name implies, the marginal distribution functions are "coupled" to the multivariate joint distribution function using the copula function. The introduction of copulas do, however, require understanding of two transformations called the probability and the quantile transform and, hence, Proposition 6.1 in [12] will be cited here.

**Proposition 2.2.** *Let $F$ be a distribution function on $\mathbb{R}$. Then:*

*(i)* $u \leq F(x)$ *if and only if* $F^{-1}(u) \leq x$.

*(ii) If $F$ is continuous, then* $F\left(F^{-1}(u)\right) = u$.

*(iii) (Quantile transform) If $U$ is $U(0,1)$-distributed, then* $\mathrm{P}\left(F^{-1}(U) \leq x\right) = F(x)$.

*(iv) (Probability transform) If $X$ has distribution function $F$, then $F(X)$ is $U(0,1)$-distributed if and only if $F$ is continuous.* □

The reader is encouraged to read section 6.2 in [12] for a complete proof. The resulting conclusion of the quantile and probability transform, however, can be equivalently formulated as $F^{-1}(U) \stackrel{d}{=} X$ and $F(X) \stackrel{d}{=} U$ respectively.

Consider a random vector $\mathbf{X} = [X_1, ..., X_d]^T$ with known marginal distribution functions $F_1, ..., F_d$, then its joint distribution function is

$$F_{\mathbf{X}}(\mathbf{x}) = \mathrm{P}(X_1 \leq x_1, ..., X_d \leq x_d). \tag{2.22}$$

However, by the quantile transform and property $(i)$ of Proposition 2.2 it holds that $\mathrm{P}(X_1 \leq x_1, ..., X_d \leq x_d) = \mathrm{P}\left(F_1^{-1}(U_1) \leq x_1, ..., F_d^{-1}(U_d) \leq x_d\right) = \mathrm{P}(U_1 \leq F_1(x_1), ..., U_d \leq F_d(x_d))$, where $U_1, ..., U_d$ are uniformly distributed on $(0,1)$. Consequently, (2.22) can be formulated as a joint distribution function of uniform random variables, i.e.

$$F_{\mathbf{X}}(\mathbf{x}) = \mathrm{P}(U_1 \leq F_1(x_1), ..., U_d \leq F_d(x_d)) := C(F_1(x_1), ..., F_d(x_d)). \tag{2.23}$$

The joint distribution function $C(u_1, ..., u_d)$ of a random vector $\mathbf{U}$ whose components are uniformly distributed on $(0,1)$ is called a copula function. Conversely, and more practically, we can express the vector $\mathbf{X}$ as $\mathbf{X} = \left[F_1^{-1}(U_1), ..., F_d^{-1}(U_d)\right]^T$, where $\mathbf{U} = [U_1, ..., U_d]^T$ has the joint distribution function $C(u_1, ..., u_d)$. Evidently, $\mathbf{X}$ will inherit the dependence structure of $\mathbf{U}$ and the choice of copula for $\mathbf{X}$ is equivalent with the choice of how to model $\mathbf{U}$.

A more formal description of the nature of copulas is in [17] given by Sklar's theorem, which is presented below.

**Theorem 2.3. Sklar's theorem.** *Let $\mathbf{X} = [X_1, ..., X_d]$ be a random vector with joint distribution function $F_{\mathbf{X}}$ and marginals $F_1, ..., F_d$. Then there exists a $d$-dimensional copula $C$ such that for all $\mathbf{x} = [x_1, ..., x_d]$ in $[-\infty, \infty]^d$,*

$$F_{\mathbf{X}}(\mathbf{x}) = C(F_1(x_1), ..., F_d(x_d)). \tag{2.24}$$

*If $F_1, ..., F_d$ are all continuous, then $C$ is unique; otherwise, $C$ is uniquely determined on the image room of $F_1, ..., F_d$. Conversely, if $C$ is a $d$-dimensional copula and $F_1, ..., F_d$ are univariate distribution functions, then the function $F_{\mathbf{X}}$, defined as above, is a $d$-dimensional multivariate distribution function with marginals $F_1, ..., F_d$.*

Commonly, if (2.24) holds true, then $\mathbf{X}$ is said to have the copula $C$. Furthermore, if $F_1, ..., F_d$ are continuous, then, by the probability transform and Sklar's theorem, it holds that

$$C(\mathbf{u}) = F_{\mathbf{X}}\left(F_1^{-1}(u_1), ..., F_d^{-1}(u_d)\right) \tag{2.25}$$

**2.2.3.1 Elliptical Copulas** In the case of elliptical copulas, the vector $\mathbf{U}$ is modeled using a direct application of the probability transform. Note that, for any random vector $\mathbf{Y} = [Y_1, ..., Y_d]^T$ with continuous marginals $H_1, ..., H_d$, it holds that

$$\mathbf{U} = [U_1, ...U_d]^T \stackrel{d}{=} [H_1(Y_1), ..., H_d(Y_d)]^T. \tag{2.26}$$

Inserting (2.26) into (2.23) and applying property $(i)$ of Proposition 2.2 yields that

$$F_{\mathbf{X}}(\mathbf{x}) = \mathrm{P}\left(Y_1 \le H_1^{-1}(F_1(x_1)), ..., Y_d \le H_d^{-1}(F_d(x_d))\right) = F_{\mathbf{Y}}\left(H_1^{-1}(F_1(x_1)), ..., H_d^{-1}(F_d(x_d))\right), \tag{2.27}$$

where $F_{\mathbf{Y}}$ is the joint distribution function of the vector $\mathbf{Y}$. This result implies that the joint distribution of $\mathbf{X}$ can be specified by the choice of the marginals $H_1, ..., H_d$ of $\mathbf{Y}$. Further, formulated using the copula function, this relation takes the form

$$C(\mathbf{u}) = \mathrm{P}(H_1(Y_1) \le u_1, ..., H_d(Y_d) \le u_d) = F_{\mathbf{Y}}\left(H_1^{-1}(u_1), ..., H_d^{-1}(u_d)\right) \tag{2.28}$$

**Gaussian copula** One important elliptical copula is the Gaussian copula and is obtained by setting $\mathbf{Y}$ to have $d$-dimensional normal distribution, with standard normal components and linear correlation matrix $\mathbf{R}$. Then $H_i(x) = \Phi(x)$, for $i = 1, ..., d$, and (2.28) simplifies into

$$C_{\mathbf{R}}^{Ga}(\mathbf{u}) = \Phi_{\mathbf{R}}^{d}\left(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)\right), \tag{2.29}$$

where $\Phi_{\mathbf{R}}^{d}$ is the joint distribution function of $\mathbf{Y}$. The function $C_{\mathbf{R}}^{Ga}$ is called a Gaussian copula. Furthermore, by (2.19), the above equation can be formulated as

$$C_{\mathbf{R}}^{Ga}(\mathbf{u}) = \frac{1}{\sqrt{(2\pi)^d |\mathbf{R}|}} \int_{-\infty}^{\Phi^{-1}(u_1)} \dots \int_{-\infty}^{\Phi^{-1}(u_d)} \exp\left(-\frac{1}{2}\mathbf{s}^T \mathbf{R}^{-1} \mathbf{s}\right) ds_1 \dots ds_d.$$

As outlined in [14], the copula density can be obtained through application of the inverse function theorem as

$$c_{\mathbf{R}}^{Ga}(\mathbf{u}) = \frac{\phi_{\mathbf{R}}^{d}\left(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)\right)}{\prod_{j=1}^{d} \phi\left(\Phi^{-1}(u_j)\right)}.$$

By using (2.8) and (2.18), the Gaussian copula density can be expressed as

$$c_{\mathbf{R}}^{Ga}(\mathbf{u}) = \frac{1}{\sqrt{|\mathbf{R}|}} \exp\left(-\frac{1}{2}\mathbf{y}^T \left(\mathbf{R}^{-1} - \mathbf{I}\right) \mathbf{y}\right) \tag{2.30}$$

where $\mathbf{y} = \left[\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)\right]$ and $\mathbf{I}$ is the identity matrix. As explained in Appendix C.1, the Gaussian copula implies no asymptotic tail dependence.

**Student's t-copula** The Student's $t$ copula can be derived in a similar manner of that of the Gaussian copula. Let $\mathbf{Y}$ have $d$-dimensional t-distribution with $\nu$ degrees of freedom, standard t-distributed components and linear correlation matrix $\mathbf{R}$. Then $H_i(x) = t_\nu(x)$, for $i = 1, ..., d$, and (2.28) simplifies into

$$C_{\nu,\mathbf{R}}^{t}(\mathbf{u}) = t_{\nu,\mathbf{R}}^{d}\left(t_\nu^{-1}(u_1), ..., t_\nu^{-1}(u_d)\right), \tag{2.31}$$

where $t_{\nu,\mathbf{R}}^{d}$ is the joint distribution function of $\mathbf{Y}$. The function $C_{\nu,\mathbf{R}}^{t}$ is called a t-copula. Moreover, by (2.21), the above equation can be formulated as

$$C_{\nu,\mathbf{R}}^{t}(\mathbf{u}) = \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\pi\nu)^d |\mathbf{R}|}} \int_{-\infty}^{t_\nu^{-1}(u_1)} \dots \int_{-\infty}^{t_\nu^{-1}(u_d)} \left(1 + \frac{\mathbf{s}^T \mathbf{R}^{-1} \mathbf{s}}{\nu}\right)^{-\frac{\nu+d}{2}} ds_1 \dots ds_d,$$

and as is stated in [5], the corresponding copula density function is

$$c^t_{\nu,\mathbf{R}}(\mathbf{u}) = \frac{\Gamma\left(\frac{\nu+d}{2}\right)\Gamma\left(\frac{\nu}{2}\right)^{d-1}}{\Gamma\left(\frac{\nu+1}{2}\right)^{d}\sqrt{|\mathbf{R}|}} \frac{\prod_{k=1}^{d}\left(1+\frac{y_k^2}{\nu}\right)^{\frac{\nu+1}{2}}}{\left(1+\frac{\mathbf{y}^T\mathbf{R}^{-1}\mathbf{y}}{\nu}\right)^{\frac{\nu+d}{2}}} \tag{2.32}$$

where $y_k = t_\nu^{-1}(u_k)$ and $\mathbf{y} = \left[t_\nu^{-1}(u_1), ..., t_\nu^{-1}(u_d)\right]$. Unlike the Gaussian copula, the Student's $t$ copula do imply asymptotic tail dependence, with coefficients of tail dependence given by (C.3).

**2.2.3.2 Grouped t copula** The grouped $t$ copula is a generalization, introduced in [5], of which the Student's $t$ copula is a special case. The derivation of the grouped $t$ copula presented in this thesis should be considered a short summary and the reader is directed to [5] for a more thorough presentation. The grouped $t$ copula is not a perfectly elliptical copula, although it has approximately elliptical properties in some contexts.

Let $\mathbf{Z} \in N_J(\mathbf{0}, \mathbf{R})$, where $\mathbf{R}$ is an arbitrary linear correlation matrix, be independent of $U$, which is uniformly distributed on $(0,1)$. Further, let $W_\nu \in \chi^2_\nu$ and let $G_\nu$ be the distribution function of $\sqrt{\nu/W_\nu}$, i.e. $G_\nu(x) = \mathrm{P}\left(\sqrt{\nu/W_\nu} \le x\right)$. Next, partition the sequence $S = \{1, ..., J\}$ into $1 \le m \le J$ subsets of sizes $s_1, ..., s_m$ and denote the distinct subsets as $S_1, ..., S_m$. Introduce the functions $S_k(n_k)$, for $n_k = 1, ..., s_k$ and $k = 1, ..., m$, such that

$$S_k(n_k) = \text{ The } n_k\text{:th member of subset } S_k.$$

Moreover, let $R_k = G^{-1}_{\nu_k}(U)$ – then, by the quantile transform, $R_k \stackrel{d}{=} \sqrt{\nu_k/W_{\nu_k}}$ – for $k = 1, ..., m$. Thereby, if $\mathbf{Y}$ is defined as

$$\mathbf{Y} = \left[R_1 Z_{S_1(1)}, ..., R_1 Z_{S_1(s_1)}, R_2 Z_{S_2(1)}, ..., R_{k-1} Z_{S_{k-1}(s_{k-1})}, R_k Z_{S_k(1)}, ..., R_k Z_{S_k(s_k)}, ..., R_m Z_{S_m(s_m)}\right]^T, \tag{2.33}$$

then the random vector $\left[Y_1, ..., Y_{s_1}\right]^T = \left[R_1 Z_{S_1(1)}, ..., R_1 Z_{S_1(s_1)}\right]^T$ has an $s_1$-dimensional t-distribution with $\nu_1$ degrees of freedom. Further, the vector $\left[Y_{s_1+...+s_k+1}, ..., Y_{s_1+...s_k+s_{k+1}}\right]^T = \left[R_{k+1} Z_{S_{k+1}(1)}, ..., R_{k+1} Z_{S_{k+1}(s_{k+1})}\right]^T$ has an $s_{k+1}$-dimensional t-distribution with $\nu_{k+1}$ degrees of freedom, for $k = 1, ..., m-1$. Finally, if $F_k$ is the distribution function of $Y_k$, for $k = 1, ..., J$, and $H_1, ..., H_J$ are some arbitrary continuous strictly increasing distribution functions, then

$$\begin{aligned}\mathbf{X} &= \left[H_1^{-1}(F_1(Y_1)), ..., H_J^{-1}(F_J(Y_J))\right]^T \\ &= \left[H_1^{-1}\left(t_{\nu_1}(Y_1)\right), ..., H_{s_1}^{-1}\left(t_{\nu_1}\left(Y_{s_1}\right)\right), H_{s_1+1}^{-1}\left(t_{\nu_2}\left(Y_{s_1+1}\right)\right), ..., H_{s_1+s_2}^{-1}\left(t_{\nu_2}\left(Y_{s_1+s_2}\right)\right), ..., H_J^{-1}\left(t_{\nu_m}(Y_J)\right)\right]^T\end{aligned}$$

is a generalization of the model using a Student's $t$ copula. Here, the $m$ different subsets of the components are allowed to have different degrees of freedom parameter. The copula of $\mathbf{X}$ is called a grouped $t$ copula. With $m = 1$, the sequence $\{1, ..., J\}$ is partitioned into only one subset, i.e. there is no real partitioning, and the Student's $t$ copula is obtained. Lastly, the grouped $t$ copula function is

$$C^{Gt}_{\nu,\mathbf{R}}(\mathbf{u}) = \mathrm{P}\left(t_{\nu_1}(Y_1) \le H_1(u_1), ..., t_{\nu_1}\left(Y_{s_1}\right) \le H_{s_1}\left(u_{s_1}\right), t_{\nu_2}\left(Y_{s_1+1}\right) \le H_{s_1+1}\left(u_{s_1+1}\right), ..., t_{\nu_m}(Y_J) \le H_J^{-1}(u_J)\right) \tag{2.34}$$

Note that this implies that each group within the grouped $t$ copula has a regular Student's $t$ copula.

**The Group Specific Random Variable** $R$ As visible in (2.33), each group share a common $R_k$ variable and since $R_k = G^{-1}_{\nu_k}(U)$ this introduces a dependency between all elements of the random vector $\mathbf{Y}$. Unlike the Student's $t$ copula, the $R$ variable is scaled using a different $\nu$ for each group. Further, by its definition the random variable $R_k$ has distribution function $G_{\nu_k}(x) = \mathrm{P}\left(\sqrt{\nu/W_\nu} \le x\right)$. However, given this, it is not an easy task to compute $R_k = G^{-1}_{\nu_k}(U)$. Proposition 6.4 in [12] useful for this task and is therefore cited below.

**Proposition 2.4.** *For any random variable X*

$$F_{-X}^{-1}(p) = -F_X^{-1}((1-p)+)$$

*for all* $p \in (0,1)$*, where* $F_X^{-1}((1-p)+) = \lim_{\epsilon\downarrow 0} F_X^{-1}(1-p+\epsilon)$*. In particular, if* $F_X$ *is continuous and strictly increasing, then*

$$F_{-X}^{-1}(p) = -F_X^{-1}(1-p).$$

□

For a proof of this proposition, the reader is directed to [12]. Moreover, set $g(x;\nu) = \sqrt{\nu/x}$, then $g$ is a strictly decreasing and strictly positive of $x$. Conversely, $-g$ is a strictly increasing function and it is thus possible to reformulate $G_\nu^{-1}$ using Proposition 2.1 and 2.4. First, note that $\sqrt{\nu_k/W_{\nu_k}} = g(W_{\nu_k};\nu_k)$ and therefore

$$G_{\nu_k}^{-1}(u) = F_{g\left(W_{\nu_k};\nu_k\right)}^{-1}(u).$$

Due to $g$ being a continuous strictly positive function and $W_{\nu_k}$ being a continuous random variable, it follows that $F$ is continuous and strictly increasing. Then, for all $u \in (0,1)$ , Proposition 2.4 yields that

$$G_{\nu_k}^{-1}(u) = F_{g\left(W_{\nu_k};\nu_k\right)}^{-1}(u) = -F_{-g\left(W_{\nu_k};\nu_k\right)}^{-1}(1-u)$$

Further, since $-g$ is strictly increasing, Proposition 2.1 yields that

$$G_{\nu_k}^{-1}(u) = -F_{-g\left(W_{\nu_k};\nu_k\right)}^{-1}(1-u) = g\left(F_{W_{\nu_k}}^{-1}(1-u);\nu_k\right) = \sqrt{\frac{\nu_k}{F_{W_{\nu_k}}^{-1}(1-u)}}$$

Finally, since the random variable $R_k$ is defined as $R_k = G_{\nu_k}^{-1}(U)$ where $U$ is uniformly distributed on $(0,1)$, it follows that

$$R_k = G_{\nu_k}^{-1}(U) = \sqrt{\frac{\nu_k}{F_{W_{\nu_k}}^{-1}(1-U)}} \tag{2.35}$$

where $F_{W_{\nu_k}}$ is the distribution function of $W_{\nu_k} \in \chi^2_{\nu_k}$.

**2.2.3.3 Kendall Distribution Function** Consider a multivariate uniform $(0,1)$ random vector $[U_1,...,U_d]$ with joint distribution function $C$ and let $V = C(U_1,...,U_d)$. Then, the *Kendall distribution function* is defined as

$$K(v) = \mathrm{P}(V \le v). \tag{2.36}$$

If the vector $[U_1,...,U_d]$ had been one-dimensional – i.e. $d = 1$ –, then it would follow from the probability transform that $V$ has uniform $(0,1)$ distribution, under the assumption that $C$ is continuous. This result does however not hold in higher dimensions. Nevertheless, the Kendall distribution function provides useful in multidimensional copula analysis, since it describes the distribution of $V$.

### 2.2.4 Empirical Distributions

When modeling using historical data, a convenient approach is to fit a parametric model. This however introduces the risk of model misspecification. A robust alternative approach, presented in [12], is to use non-parametric estimators, so called empirical distributions. In this section, the theory in [12] is outlined for uni- and multivariate distributions and also expanded to comprise copulas and the Kendall distribution function.

Consider a set of $n$ independent and identically distributed $d$-dimensional random vectors $\mathbf{X}_1,...,\mathbf{X}_n$ with a common unknown distribution function $F(\mathbf{x}) = \mathrm{P}(X_1 \le x_1,...,X_d \le x_d)$. Then, the empirical distribution of the random sample $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ is

$$F_{n,\mathbf{X}}(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(\mathbf{X}_i \le \mathbf{x}), \tag{2.37}$$

where $\mathbb{1}$ is an indicator function (equal to unity if the logical statement is true and otherwise equal to zero). The function (2.37) is a discontinuous step-function with step-length $1/n$ and range $[0,1]$ Here $F_{n,\mathbf{X}}$ is random due to the underlying sample being random. Thus, provided a sample of observations $\mathbf{x}_1, ..., \mathbf{x}_n$ of $\mathbf{X}_i$, $i = 1, ..., n$, the empirical distribution function is obtained as

$$F_n(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(\mathbf{x}_i \le \mathbf{x}) \tag{2.38}$$

and is an outcome of $F_{n,\mathbf{X}}$. Further, $F_n$ is a distribution function in the sense that it is not random.

In the case when $d = 1$, i.e. when $\mathbf{X}$ is univariate, the quantile function corresponding to $F_n$ is the empirical quantile function and is given as

$$F_n^{-1}(p) = \min\{x : F_n(x) \ge p\}$$

In [12], it is shown that if the sample $\{x_1, ..., x_n\}$ is ordered such that $x_{1,n} \le \cdots \le x_{n,n}$, then the empirical quantile function is

$$F_n^{-1}(p) = x_{\lceil np \rceil, n}$$

where $\lceil x \rceil$ is the smallest integer larger than $x$.

**2.2.4.1 Empirical Copulas** Consider the empirical distribution function (2.38) and note that the steps will occur exactly at the points where $[x_1, ..., x_d] = [x_{1,i}, ..., x_{d,i}]$ for $i = 1, ..., n$. Thereby, the range of (2.38) is $\mathrm{Ran}(F_n) = \{0, F_{n,1}, ..., F_{n,n}\}$ where

$$F_{n,j} = F_n(x_{1,j}, ..., x_{d,j}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(x_{1,i} \le x_{1,j}, ..., x_{d,i} \le x_{d,j}). \tag{2.39}$$

Under the assumption that the components of $\mathbf{X}$ have continuous univariate distribution functions $F_1, ..., F_d$, the empirical copula function can be determined by implementing (2.25) as

$$C_n(\mathbf{u}) = F_n\left(F_1^{-1}(u_1), ..., F_d^{-1}(u_d)\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\left(x_{1,i} \le F_1^{-1}(u_1), ..., x_{d,i} \le F_d^{-1}(u_d)\right) \tag{2.40}$$

where $u_k = F_k(x_k)$, $k = 1, ..., d$, i.e. the sample of observations transformed by the probability transform in Proposition 2.1. These transformed observations are commonly called *pseudo-observations*. Further, $C_n$ and $F_n$ evidently have identical range since $C_{n,i} = F_{n,i}$, $i = 1, ..., n$, where

$$C_{n,j} = C_n(u_{1,j}, ..., u_{d,j}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(u_{1,i} \le u_{1,j}, ..., u_{d,i} \le u_{d,j}), \tag{2.41}$$

and $[u_{1,i}, ..., u_{d,i}] = [F_1(x_{1,i}), ..., F_d(x_{d,i})]$ for $i = 1, ..., n$. The set of points where the steps occurs for $F_n$ and $C_n$ respectively, say $S_{F_n}$ and $S_{C_n}$, are not identical, instead, they are related through the non-decreasing transform $T(y_1, ..., y_d) = [F_1(y_1), ..., F_d(y_d)]$ such that $S_{C_n} = \{T(\mathbf{x}) \mid \mathbf{x} \in S_{F_n}\}$. In other words, $S_{F_n} = \{\mathbf{x}_i\}_{i=1}^{n}$ and $S_{C_n} = \{\mathbf{u}_i\}_{i=1}^{n}$, where $\mathbf{u}_i = [u_{1,i}, ..., u_{d,i}]$.

Moreover, since $C_n$ and $F_n$ have identical range, estimating $C$ by $C_n$ on the set $S_{C_n}$ is equivalent to estimating $F_{\mathbf{X}}$ by $F_n$ on the set $S_{F_n}$.

**2.2.4.2 Empirical Kendall Distribution Function** Assume that $\mathbf{u}_i = \left[u_{1,i},...,u_{d,i}\right]$, $i = 1,...,n$, is a sample of observations from a known copula $C$. Then, a sample $v_1,...,v_n$ of observations of the random variable $V = C(U_1,...,U_d)$ can be computed in two different ways depending on the premises. If the copula $C$ and have a simple explicit form, then the observations $v_1,...,v_n$ can be computed analytically as

$$v_i = C(\mathbf{u}_i), \quad i = 1,...,n,$$

or numerically by approximating the simple explicit form of $C$ in case $C$ cannot be computed analytically. On the other hand, if the copula $C$ is known but do not have a simple explicit form, then $v_1,...,v_n$ can be approximated by Monte Carlo simulation and the empirical copula function (2.40). Then,

$$v_i = C_m(\mathbf{u}_i) = \frac{1}{m}\sum_{j=1}^{m} \mathbb{1}\left(u^\star_{1,j} \leq u_{1,i},...,u^\star_{d,j} \leq u_{d,i}\right)$$

where $\mathbf{u}^\star_1,...,\mathbf{u}^\star_m$ is a simulated sample from the copula $C$ and preferably $m >> n$. This yields, as in shown in [7], a non-parametric estimator of the Kendall distribution function given by

$$K_n(v) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(v_i \leq v). \tag{2.42}$$

Similarly to the empirical functions discussed in the preceding section, this is a discontinuous step-function with step-length $1/n$ and range $[0,1]$. Further, the steps occur exactly when $v = v_i$, $i = 1,...,n$, and consequently the range of (2.42) is $\operatorname{Ran}(K_n) = \left\{0, K_{n,1},...,K_{n,n}\right\}$ where

$$K_{n,j} = K_n\left(C_{n,j}\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\left(C_{n,i} \leq C_{n,j}\right). \tag{2.43}$$

In the case where the copula $C$ is known, a sample of the variable $V$ can be constructed from the observed sample without estimation. An example is when outcomes are simulated from a copula whose Kendall distribution function does not have an explicit form.

If $\mathbf{x}_i$, $i = 1,...,n$ is the observed data and $\mathbf{u}_i = T(\mathbf{x}_i)$ is the transformation by marginal distribution functions as defined in the preceding section, then the corresponding sample of the random variable $V$ is $v_i = C(\mathbf{u}_i)$, $i = 1,...,n$. The Kendall distribution function can then be estimated on the sample $v_i$, $i = 1,...,m$, as

$$K^C_n(v) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(v_i \leq v), \tag{2.44}$$

and, using notation similar to (2.43), the non-zero points in the range of $K^C_n$ are denoted

$$K^C_{n,j} = K^C_n\left(v_j\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\left(v_i \leq v_j\right). \tag{2.45}$$

## 2.3 Parametric Estimators and Model Diagnostics

This section presents different methods of finding the best fit of parametric models to a given set of data. Obviously, there is no universal definition of "best fit" which justifies the use of different estimators. Two categories of methods of fitting parametric models by estimating parameters are considered in this thesis, and consequently presented in this section, namely: Maximum Likelihood and Least-Squares. Moreover, non-parametric estimators of certain dependence measures are presented in Section 2.3.4. These estimators are mainly used to assist numerical optimization and asses the obtained models. Section 2.3.6 presents methods for construction of approximate confidence intervals of estimated parameters.

### 2.3.1 Maximum Likelihood Estimator (MLE)

If $f$ is the joint density of a parametric model of $\mathbf{X} = [X_1, ..., X_d]$, with parameters $\boldsymbol{\theta}$, then, provided a sample $\mathbf{x}_i = \left[x_{i,1}, ..., x_{i,d}\right]$, $i = 1, ..., n$, of independent observations of $\mathbf{X}$, the likelihood function is defined as

$$\mathscr{L}(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(\mathbf{x}_i; \boldsymbol{\theta}). \tag{2.46}$$

Moreover, the maximum likelihood estimate of $\boldsymbol{\theta}$, here denoted as $\widehat{\boldsymbol{\theta}}_{MLE}$, is the value for which $\mathscr{L}(\boldsymbol{\theta})$ attains its maximum value. More concisely, $\widehat{\boldsymbol{\theta}}_{MLE}$ can be expressed as

$$\widehat{\boldsymbol{\theta}}_{MLE} = \underset{\boldsymbol{\theta}}{\arg\max} \left\{\mathscr{L}(\boldsymbol{\theta})\right\}. \tag{2.47}$$

However, it is easier to use the logarithm of (2.46) since the product of densities then turns into a sum. The maximum of the likelihood and the log-likelihood function will coincide since the logarithm is a monotone increasing function. In explicit form, the log-likelihood function is

$$\ln \mathscr{L}(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ln f(\mathbf{x}_i; \boldsymbol{\theta}). \tag{2.48}$$

The asymptotic normality of $\sqrt{n}\left(\widehat{\boldsymbol{\theta}}_{MLE} - \boldsymbol{\theta}\right)$ is established under sufficient regularity conditions in [20]. Specifically, $\sqrt{n}\left(\widehat{\boldsymbol{\theta}}_{MLE} - \boldsymbol{\theta}\right)$ is normally distributed with zero mean and finite variance as $n$ tends to infinity.

**2.3.1.1 Log-Likelihood of Univariate Distributions** For the univariate distributions in Section 2.2.1, the log-likelihood functions can easily be derived from the distribution density functions. Additionally, an analytic expression of the maximum likelihood estimator can be obtained when considering the normal distribution.

Let $x_1, ..., x_n$ be observations from a $N(\mu, \sigma)$-distribution, then (2.9) yields that (2.48) takes the form

$$\ln \mathscr{L}(\mu, \sigma) = -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 - n \ln\left(\sigma\sqrt{2\pi}\right).$$

Since $0 \leq f(x; \mu, \sigma) < \infty, \forall x$ and $\lim_{x \to \pm\infty} f(x; \mu, \sigma) = 0$, the log-likelihood attains its maximum value at an inner extreme point. Setting the gradient of the log-likelihood function to zero yields

$$\begin{aligned} \widehat{\mu}_{MLE} &= \bar{x} \\ \widehat{\sigma}_{MLE} &= \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2} \end{aligned}$$

where $\bar{x} = n^{-1} \sum_{i=1}^{n} x_i$ is the arithmetic mean of the observations. When considering the Student's $t$ distribution, with density function given by (2.10), the log-likelihood function (2.48) is obtained as

$$\ln \mathscr{L}(\mu, \sigma, \nu) = n\left(\ln \Gamma\left(\frac{\nu+1}{2}\right) - \ln \Gamma\left(\frac{\nu}{2}\right) - \ln\left(\sigma\sqrt{\nu\pi}\right)\right) - \frac{\nu+1}{2} \sum_{i=1}^{n} \ln\left(1 + \frac{(x_i - \mu)^2}{\nu\sigma^2}\right).$$

Consequently, the partial derivatives of the log-likelihood function are

$$\frac{\partial \ln \mathscr{L}(\mu,\sigma,\nu)}{\partial \mu} = (\nu+1)\sum_{i=1}^{n} \frac{x_i-\mu}{\nu\sigma^2+(x_i-\mu)^2}$$

$$\frac{\partial \ln \mathscr{L}(\mu,\sigma,\nu)}{\partial \sigma} = \frac{1}{\sigma}\left((\nu+1)\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{\nu\sigma^2+(x_i-\mu)^2} - n\right)$$

$$\frac{\partial \ln \mathscr{L}(\mu,\sigma,\nu)}{\partial \nu} = \frac{n}{2}\left(\psi\left(\frac{\nu+1}{2}\right)-\psi\left(\frac{\nu}{2}\right)-\frac{1}{\nu}\right)+\frac{1}{2}\sum_{i=1}^{n}\left[\frac{(\nu+1)(x_i-\mu)^2}{\nu^2\sigma^2+\nu(x_i-\mu)^2}-\ln\left(1+\frac{(x_i-\mu)^2}{\nu\sigma^2}\right)\right]$$

where $\psi(x)$ is the *digamma function*, defined as $\psi(x) := \frac{d}{dx}\ln\Gamma(x) = \Gamma'(x)/\Gamma(x)$. Finding the roots of the partial derivatives above is a non-trivial task and hence the maximum likelihood estimators are best found using numerical optimization methods.

**2.3.1.2 Copula Log-Likelihood** This section presents a summary of the log-likelihood functions of the copulas considered in Section 2.2.3. In Table 2.1, the logarithm of the copula density function are presented. As visible in (2.48), the log-likelihood is obtained by summation of the logarithmic densities. Note that the density function of the grouped $t$ copula is not explicitly stated. Due to the complex structure of the grouped $t$ copula, the density function is considered beyond the scope of this thesis. Instead, the fact that the copula margins have regular Student's $t$ copulas will be utilized when estimating the parameters of the grouped $t$ copula, as recommended in [5].

Table 2.1: Explicit forms of the logarithms of copula densities.

| Copula | $\ln c(\mathbf{u}_i;\boldsymbol{\theta})$ | Comment |
|---|---|---|
| Gaussian | $-\frac{1}{2}\ln\lvert\mathbf{R}\rvert - \frac{1}{2}\mathbf{y}_i^T\left(\mathbf{R}^{-1}-\mathbf{I}\right)\mathbf{y}_i$ | $y_{i,k} = \Phi^{-1}\left(u_{i,k}\right)$ |
| Student's t | $\ln\left(\frac{\Gamma\left(\frac{\nu+d}{2}\right)\Gamma\left(\frac{\nu}{2}\right)^{d-1}}{\Gamma\left(\frac{\nu+1}{2}\right)^d\sqrt{\lvert\mathbf{R}\rvert}}\right)+\frac{\nu+1}{2}\sum_{k=1}^{d}\ln\left(1+\frac{y_{i,k}^2}{\nu}\right)-\frac{\nu+d}{2}\ln\left(1+\frac{\mathbf{y}_i^T\mathbf{R}^{-1}\mathbf{y}_i}{\nu}\right)$ | $y_{i,k} = t_\nu^{-1}\left(u_{i,k}\right)$ |

As a consequence of the extensiveness of the copula log-likelihood functions, analytic expressions of the maximum likelihood estimator (2.47) or the gradient are not derived.

### 2.3.2 Least-Squares Estimator (LSE)

Consider a sample of independent observations, $x_1,...,x_n$, of the random variable $X$. If $F(\cdot;\boldsymbol{\theta})$ is a parametric model of the distribution function of $X$ with parameter $\boldsymbol{\theta}$, then the quadratic quantile sum function $\mathscr{Q}^q$ is implicitly defined in [12] as

$$\mathscr{Q}^q(\boldsymbol{\theta}) = \sum_{i=1}^{n}\left(x_i^{(n)} - F^{-1}\left(\frac{n-i+1}{n+1};\boldsymbol{\theta}\right)\right)^2, \tag{2.49}$$

where $x_1^{(n)},...,x_n^{(n)}$ is the ordered sample of observations, i.e. the empirical quantiles. The least-squares estimate of $\boldsymbol{\theta}$, here denoted as $\widehat{\boldsymbol{\theta}}_{LSE}^q$, is the value for which $\mathscr{Q}^q(\boldsymbol{\theta})$ attains its minimum value. More concisely, $\widehat{\boldsymbol{\theta}}_{LSE}$ can be expressed as

$$\widehat{\boldsymbol{\theta}}_{LSE}^q = \underset{\boldsymbol{\theta}}{\arg\min}\left\{\mathscr{Q}^q(\boldsymbol{\theta})\right\} \tag{2.50}$$

Problem arises when the multivariate case is considered, since no inverse exist for multivariate distribution functions. However, [21] uses a similar quadratic sum function which is applicable in higher dimensions. It is identical to (2.49) in all aspects except that it implements the empirical and parametric joint distribution function instead of the empirical and parametric quantile function. Provided a sample of independent observations $\mathbf{x}_i = \left[x_{i,1},...,x_{i,d}\right]$, $i = 1,...,n$, of the random vector $\mathbf{X} = [X_1,...,X_d]$, the quadratic distribution sum function $\mathscr{Q}^{cdf}$ is defined as

$$\mathscr{Q}^{cdf}(\boldsymbol{\alpha}) = \sum_{i=1}^{n}\left(F_n(\mathbf{x}_i) - F(\mathbf{x}_i;\boldsymbol{\alpha})\right)^2 = \sum_{i=1}^{n}\left(C_{n,i} - F(\mathbf{x_i};\boldsymbol{\alpha})\right)^2$$

where the last equality follows from the conclusions of Section 2.2.4.1. Further, by Sklar's Theorem in Section 2.2.3, it follows that $F(\mathbf{x}_i;\boldsymbol{\alpha}) = C\left(F_1\left(x_{i,1};\boldsymbol{\eta}_1\right),...,F_d\left(x_{i,d};\boldsymbol{\eta}_d\right);\boldsymbol{\theta}\right)$, where $F_1\left(\cdot;\eta_1\right),...,F_d\left(\cdot;\eta_d\right)$ are parametric marginals of $\mathbf{X}$ and $\boldsymbol{\theta}$ is the parameter of the copula model. The equation above can then be rephrased as

$$\mathscr{Q}^{cdf}\left(\boldsymbol{\eta}_1,...,\boldsymbol{\eta}_d,\boldsymbol{\theta}\right) = \sum_{i=1}^{n}\left(C_{n,i} - C\left(F_1\left(x_{i,1};\boldsymbol{\eta}_1\right),...,F_d\left(x_{i,d};\boldsymbol{\eta}_d\right);\boldsymbol{\theta}\right)\right)^2. \tag{2.51}$$

With $\boldsymbol{\zeta} = \left[\boldsymbol{\eta}_1^T,...,\boldsymbol{\eta}_d^T,\boldsymbol{\theta}^T\right]^T$, the least-squares estimate $\widehat{\boldsymbol{\zeta}}_{LSE}^{cdf}$ of $\boldsymbol{\zeta}$ using (2.51) can more concisely be expressed as

$$\widehat{\boldsymbol{\zeta}}_{LSE}^{cdf} = \arg\min_{\boldsymbol{\zeta}}\left\{\mathscr{Q}^{cdf}(\boldsymbol{\zeta})\right\} \tag{2.52}$$

Additionally, it should be mentioned that [21] investigated several minimum-distance estimators for parametric copula models and found little difference in the qualities of the estimators. Clearly, there are more than one way to measure distance in a multidimensional vector space. Using the $L_2$-norm is one way, and leads to the minimum-distance estimator being equal to the least-squares estimator above. In the light of the results of [21], and in order to preserve simplicity, only the $L_2$-norm will be considered in this thesis.

### 2.3.3 Strengths and Shortcomings of the MLE and the LSE

In [21] it is shown that the MLE outmatches every considered minimum-distance estimator (of which the LSE is one) in aspect of bias, efficiency and computational complexity when estimating parametric copulas. This study was however carried out for bivariate copulas and consequently in a low-dimensional context.

As stated by [13], the surface of the log-likelihood relatively flat in the neighbourhood of its maximum. Thus, if in a context with several varying parameters, then a numerical optimizer with a finite tolerance might fail to find a good parametrization since several parameter combination yield a close-to-optimal log-likelihood value. One consequence highlighted in [12] is that the ML estimates are sensitive to variations in the data. Consider for example the Student's $t$ distribution, where both the scale parameter $\sigma$ and the degrees of freedom parameter $v$ affects the heaviness of the tails. Hence, parameter combinations may exist that yield log-likelihoods values close to that of the ML estimates.

### 2.3.4 Dependence Measures and Non-Parametric Estimators

In the context of copula analysis, the property of invariance to scaling through strictly increasing functions possessed by several dependence measures should be highlighted. This property, here referred to as scale-invariance, is defined in Definition 2.5 in a way similar to as in [14].

**Definition 2.5.** A measure $\delta$ is scale-invariant if it is invariant under strictly increasing transforms, i.e. if $\delta(Y_1,Y_2) = \delta(h_1(Y_1),h_2(Y_2))$, where $h_1$ and $h_2$ are strictly increasing functions.

Consider a random vector $\mathbf{X} = [X_1,...,X_d]$ with copula $C$ and a random vector $\mathbf{U} = [U_1,...,U_d]$ with joint distribution function $C$. It then follows from Proposition 2.2 and Sklar's theorem that $\delta\left(X_i,X_j\right) = \delta\left(U_i,U_j\right)$, $i,j = 1,...,d$, if the marginal distribution functions of $\mathbf{X}$ are strictly increasing. Concisely, a scale-invariant measure $\delta$ is a direct property of the dependence structure of a random vector, i.e. it is independent of the marginal distributions of the vector.

Furthermore, if $\mathbf{Y}$ is elliptically distributed with continuous strictly increasing marginal distribution functions, then (2.26) holds and it follows that $\delta\left(X_i,X_j\right) = \delta\left(U_i,U_j\right) = \delta\left(Y_i,Y_j\right)$, for $i,j = 1,...,d$.

**2.3.4.1 Pearson's Product-Moment Correlation Coefficient** Pearson's product-moment correlation coefficient, often referred to as simply the coefficient of linear correlation, is a measure of linear correlation between two random variables. For a random vector $\mathbf{X} = [X_1, ..., X_d]$, it is defined as

$$\rho\left(X_i, X_j\right) = \frac{\operatorname{Cov}\left[X_j, X_j\right]}{\sqrt{\operatorname{Var}\left[X_j\right]\operatorname{Var}\left[X_j\right]}}. \tag{2.53}$$

provided that $\mathbf{X}$ has finite covariance, i.e. $\mathbb{E}\left[|X|^2\right] < \infty$. In case the covariance does not exist finitely for an elliptical distribution, the linear correlation can be defined through (2.17). Further, the correlation matrix $\mathbf{R}$ of the random vectors $\mathbf{X}$ is then such that $(\mathbf{R})_{ij} = \rho_{i,j} = \rho\left(X_i, X_j\right)$. The coefficient of linear correlation is not a scale-invariant measure. Moreover, given a sample of identically distributed random vectors $\mathbf{X}_k = \left[X_{1,k}, X_{2,k}\right]$, $k = 1, ..., n$, the coefficient of linear correlation can be estimated using the estimator

$$\widehat{\rho} = \frac{\sum_{k=1}^{n}\left(X_{1,k} - \bar{X}_1\right)\left(X_{2,k} - \bar{X}_2\right)}{\sqrt{\sum_{k=1}^{n}\left(X_{1,k} - \bar{X}_1\right)^2 \sum_{k=1}^{n}\left(X_{2,k} - \bar{X}_2\right)^2}} \tag{2.54}$$

where $\bar{X}_i$, $i = 1, 2$, is the sample mean.

**2.3.4.2 Kendall's Tau** Kendall's tau is a coefficient measuring forms of dependence known as bivariate concordance and discordance. Concordance refers to a monotone, not necessarily linear, relationship, where an increase in one variable can be associated with an increase in the other variable. Obviously, discordance refers to the inverse relationship. In [11], Kendall's tau is defined as in the following definition

**Definition 2.6.** For a pair of random variables $X_1, X_2$, Kendall's tau is defined as

$$\tau\left(X_1, X_2\right) = \mathrm{P}\left(\left(X_1 - X_1'\right)\left(X_2 - X_2'\right) > 0\right) - \mathrm{P}\left(\left(X_1 - X_1'\right)\left(X_2 - X_2'\right) < 0\right), \tag{2.55}$$

where $\left[X_1', X_2'\right]$ is an independent copy of $[X_1, X_2]$.

Note that this is the probability of concordant pairs minus the probability of discordant pairs. Further, Kendall's tau is a scale-invariant measure and is therefore useful when studying copulas as is made evident by the following two propositions.

**Proposition 2.7.** *Let* $\mathbf{U}$ *be a random vector whose components are uniformly distributed on* $[0, 1]$*. If* $\mathbf{U}$ *has joint distribution* $C$ *and the joint marginal distribution of the components* $U_i$ *and* $U_j$ *is denoted* $C_{i,j}$ *for* $i, j \in \{1, ..., d\}$*, then*

$$\tau\left(U_i, U_j\right) = 4\mathbb{E}\left[C_{i,j}\left(U_i, U_j\right)\right] - 1 \tag{2.56}$$

□

For a proof, the reader is referred to [14]. Note that since Kendall's tau is scale-invariant the proposition above is applicable whenever a random vector with copula $C$ is studied.

**Proposition 2.8.** *If* $\mathbf{X}$ *has* $d$*-dimensional elliptical distribution such that* $X_i$ *and* $X_j$ *have linear correlation coefficient* $\rho_{i,j}$ *for* $i, j \in \{1, ..., d\}$*, then*

$$\tau\left(X_i, X_i\right) = \frac{2}{\pi}\arcsin\left(\rho_{i,j}\right). \tag{2.57}$$

□

A proof of the proposition can be found in [12]. The above proposition together with the property of scale-invariance of Kendall's tau means that for any random vector $\mathbf{X}$ with elliptical copula $C$, there is a unique relationship between Kendall's tau of the components of $\mathbf{X}$ and the correlation matrix of the copula $C$. Furthermore, it is shown in [5] that equation (2.57) is approximately true for a random vector $\mathbf{X}$ with a grouped $t$ copula $C$. This relation is formulated mathematically in the lemma below.

**Lemma 2.9.** *If* $\mathbf{X}$ *is a* $d$*-dimensional random vector with a copula* $C$ *such that* $C$ *is a grouped* $t$ *copula with correlation coefficients* $\rho_{i,j}$ *for* $i, j \in \{1, ..., d\}$*, then*

$$\tau\left(X_i, X_j\right) \approx \frac{2}{\pi} \arcsin\left(\rho_{i,j}\right). \tag{2.58}$$

□

Note that in the lemma above, $\rho_{i,j}$ is taken to be the correlation coefficients of the copula and not necessarily the correlations coefficients of the random vector $\mathbf{X}$.

Provided a sample of identically distributed random vectors $\mathbf{X}_k = \left[X_{1,k}, X_{2,k}\right]$, $k = 1, ..., n$, Kendall's tau can be estimated through

$$\hat{\tau} = \binom{n}{2}^{-1} \sum_{j<k} \text{sign}\left(\left(X_{1,j} - X_{1,k}\right)\left(X_{2,j} - X_{2,k}\right)\right). \tag{2.59}$$

This estimator can in several special cases provide useful when estimating other coefficients. As an example, by (2.57), an estimator of the linear coefficient of correlation of an elliptically distributed bivariate random vector is

$$\hat{\rho} = \sin\left(\frac{\pi}{2}\hat{\tau}\right). \tag{2.60}$$

Moreover, as is outlined in [14], the process $n^{-1}(\hat{\tau} - \tau)$ is asymptotically normal as $n \to \infty$ with zero mean and variance given through

$$\lim_{n\to\infty} n\text{Var}\left[\hat{\tau}\right] = 16 \int_{[0,1]^2} \left(C(u_1, u_2) + \bar{C}(u_1, u_2)\right)^2 dC(u_1, u_2) - 4(\tau + 1)^2 \tag{2.61}$$

where $\bar{C}$ denotes a survival copula defined as $\bar{C}(u_1, ..., u_d) = \text{P}(U_1 > u_1, ..., U_d > u_2)$, where $U_1, ..., U_d$ are uniformly distributed on $[0, 1]$ and have joint distribution function $C$. Note that since Kendall's tau is a function of two random variables, the copulas in (2.61) are bivariate. Moreover, it is generally true that the copula of $\left[X_i, X_j\right]$, $i, j = 1, ...d$ belong to the same family as the copula of $[X_1, ..., X_d]$.

A large-sample estimator of the variance of $n^{-1}(\hat{\tau} - \tau)$ can be constructed using the empirical copula function in Section 2.2.4.1. With $C_{n,j}$ defined as in (2.41) and $\bar{C}_{n,j}$ defined analogous for the survival copula, the integral in (2.61) can for large $n$ be approximated as

$$\int_{[0,1]^2} \left(C(u_1, u_2) + \bar{C}(u_1, u_2)\right)^2 dC(u_1, u_2) \approx \frac{1}{n} \sum_{j=1}^{n} \left(C_{n,j} + \bar{C}_{n,j}\right)^2.$$

Now, let $\bar{V}_n$ denote the arithmetic mean of $C_{n,j}$, $j = 1, .., n$, i.e. $\bar{V}_n = n^{-1}\sum_{j=1}^{n} C_{n,j}$. It then follows from (2.56) that

$$\hat{\tau} = 4\bar{V}_n - 1 \quad \Rightarrow \quad 4(\tau + 1)^2 \approx 64\bar{V}_n^2.$$

Hence, if the asymptotic variance of $n^{-1}(\hat{\tau} - \tau)$ is denoted $v^2$, then the large-sample estimator, $\hat{v}^2$, of $v^2$ is

$$\hat{v}^2 = \frac{16}{n^4} \sum_{j=1}^{n} \left[\left(C_{n,j} + \bar{C}_{n,j}\right)^2 - 4\bar{V}_n^2\right] \tag{2.62}$$

As is shown in [7], the arithmetic mean of $\bar{C}_{n,j}$, $j = 1, .., n$, is also $\bar{V}_n$, which can be used to show that (2.62) is equivalent to the large-sample estimator in Proposition 3.1 in [7]. Further, (2.62) yields that

$$\frac{n\left(\hat{\tau}-\tau\right)}{4\sqrt{\sum_{j=1}^{n}\left[\left(C_{n,j}+\bar{C}_{n,j}\right)^{2}-4\bar{V}_{n}^{2}\right]}}\sim AsN\left(0,1\right) \tag{2.63}$$

where $AsN$ abbreviates asymptotically normal.

**2.3.4.3 Spearman's Rho** The dependence measure known as Spearman's rho – or Spearman's rank correlation – is, just like Kendall's tau, based on concordance and discordance. In [17], Spearman's rho is defined as in the following definition.

**Definition 2.10.** Consider a random vector $[X_1, X_2]$. Spearman's rho is then defined as

$$\rho_S\left(X_1,X_2\right)=3\left(\mathrm{P}\left(\left(X_1-X_1'\right)\left(X_2-X_2''\right)>0\right)-\mathrm{P}\left(\left(X_1-X_1'\right)\left(X_2-X_2''\right)<0\right)\right), \tag{2.64}$$

where $\left[X_1', X_2'\right]$ and $\left[X_1'', X_2''\right]$ are independent copies of $[X_1, X_2]$.

In words, Spearman's rho is the difference in probability of concordance and discordance for the two vectors $[X_1, X_2]$ and $\left[X_1', X_2''\right]$, i.e., two vectors with identical marginal distributions but where one has independent components. Additionally, Spearman's rho is a scale-invariant measure (see [14]). By using the probability transform in proposition 2.2 Spearman's rho can be expressed as in the proposition below.

**Proposition 2.11.** *If $[X_1, X_2]$ is a continuous random vector with marginal distributions $F_1$ and $F_2$, then*

$$\rho_S\left(X_1,X_2\right)=12\mathbb{E}\left[U_1U_2\right]-3, \tag{2.65}$$

*where $U_1 = F_1(X_1)$ and $U_2 = F_2(X_2)$.*

□

A derivation of (2.65) can be found in [17]. Further, the random variables $U_1$ and $U_2$ in (2.65) are sometimes referred to as the distribution (or population) grades. The more commonly known concept of sample ranks, explained below, is the sample analogy of grades.

Unlike Kendall's tau, Spearman's rho is not invariant in the class of elliptical distributions, as is shown in [11]. However, under the more specific assumption that $\mathbf{X}$ is normally distributed, an explicit expression exists. The expression is given by the following proposition, found and proved in [11].

**Proposition 2.12.** *Let $\mathbf{X} \in N_d\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right)$, where $\Sigma_{i,i}, \Sigma_{j,j} > 0$ for $i, j \in \{1, ..., d\}$. Then*

$$\rho_S\left(X_i,X_j\right)=\frac{6}{\pi}\arcsin\left(\frac{\rho_{i,j}}{2}\right) \tag{2.66}$$

*where $\rho_{i,j}$ is the linear correlation coefficient.*

□

In [14], a time efficient sample estimator is introduced using the concept of ranks. Provided a sample of identically distributed random vectors $\mathbf{X}_k = \left[X_{1,k}, X_{2,k}\right]$, $k = 1, ..., n$, the sample ranks $\left[R_{1,k}, R_{2,k}\right]$ are such that $R_{1,k} = i$ if $X_{1,k}$ is the $i$th smallest element among $X_{1,1}, ..., X_{1,n}$ and $R_{2,k} = j$ if $X_{2,k}$ is the $j$th smallest element among $X_{2,1}, ..., X_{2,n}$. Under the assumption that there are no ties in the sample, the estimator of (2.64) is

$$\hat{\rho}_S=\frac{\sum_{k=1}^{n}R_{k,1}R_{k,2}-n\left(\frac{n+1}{2}\right)^2}{n\left(n^2+1\right)/12}. \tag{2.67}$$

When studying continuous random variables, such as e.g. financial log-returns, the probability of two observations being identical is in theory equal to zero. Hence, the assumption of no ties within the sample ranks is justified. Furthermore, the process $n^{-1}\left(\widehat{\rho}^S - \rho^S\right)$ is asymptotically normal with zero mean as $n$ tends to infinity. In [14], the asymptotic variance is given by

$$\lim_{n\to\infty} n\text{Var}\left[\widehat{\rho}_S\right] = 144\int_{[0,1]^2}\left(u_1u_2 + g_1(u_1) + g_2(u_2)\right)^2 dC(u_1,u_2) - 9\left(\widehat{\rho}_S + 3\right)^2, \tag{2.68}$$
$$g_1(u_1) = \int_0^1 \bar{C}\left(u_1, y\right) dy,$$
$$g_2(u_2) = \int_0^1 \bar{C}(x, u_2)\, dx,$$

where $\bar{C}$ denotes the survival copula, mentioned in the preceding section. In analogy to the approach in the previous section, a large-sample estimator of the variance of $n^{-1}\left(\widehat{\rho}_S - \rho_S\right)$ can be constructed using the empirical copula function in Section 2.2.4.1. Consider a sample $\left[u_{1,k}, u_{2,k}\right]$, $k = 1,...,n$, of observations from the unknown copula $C$. Then, let $\widehat{g}_{1,k}$ and $\widehat{g}_{2,k}$ be defined as

$$\widehat{g}_{1,k} = \frac{1}{n}\sum_{j=1}^{n}\bar{C}_n\left(u_{1,k}, u_{2,j}\right) = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n}\mathbb{1}\left(u_{1,i} > u_{1,k}, u_{2,i} > u_{2,j}\right)$$
$$\widehat{g}_{2,k} = \frac{1}{n}\sum_{j=1}^{n}\bar{C}_n\left(u_{1,j}, u_{2,k}\right) = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n}\mathbb{1}\left(u_{1,i} > u_{1,j}, u_{2,i} > u_{2,k}\right).$$

For large $n$, it holds that $g_1\left(u_{1,k}\right) \approx \widehat{g}_{1,k}$ and $g_2\left(u_{2,k}\right) \approx \widehat{g}_{2,k}$. The procedure of computing $\widehat{g}_{1,k}$ and $\widehat{g}_{2,k}$ for every $k = 1,...,n$ results in a triple sum over the sample size which may cause computational inconvenience.The numerical computations can be simplified by noting that

$$\begin{aligned}\widehat{g}_{1,k} &= \frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n}\mathbb{1}\left(u_{1,i} > u_{1,k}, u_{2,i} > u_{2,j}\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}\left(u_{1,i} > u_{1,k}\right)\frac{1}{n}\sum_{j=1}^{n}\mathbb{1}\left(u_{2,i} > u_{2,j}\right)\\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}\left(u_{1,i} > u_{1,k}\right)F_{n,u_2}\left(u_{2,i}\right),\end{aligned}$$

where $F_{n,u_2}(u) = n^{-1}\sum_{j=1}^{n}\mathbb{1}\left(u > u_{2,j}\right)$. Note here that $F_{n,u_2}$ does not depend on the value of $k$ and consequently this coefficient only needs to be computed once for each $\widehat{g}_{1,1},...,\widehat{g}_{1,n}$. The terms $\widehat{g}_{2,k}$, $k = 1,...,n$ can be simplified in a similar manner. Furthermore, taking the empirical estimator of (2.65) yields

$$\widehat{\rho}_S = \frac{12}{n}\sum_{k=1}^{n}u_{1,k}u_{2,k} - 3 \quad\Rightarrow\quad \left(\widehat{\rho}_S + 3\right)^2 = \frac{12^2}{n^2}\left(\sum_{k=1}^{n}u_{1,k}u_{2,k}\right)^2.$$

Hence, if the asymptotic variance of $n^{-1}\left(\widehat{\rho}_S - \rho_S\right)$ is denoted $w^2$, then its large-sample estimator $\widehat{w}^2$ is obtained as

$$\widehat{w}^2 = \frac{12^2}{n^4}\left(\sum_{k=1}^{n}\left(u_{1,k}u_{2,k} + \widehat{g}_{1,k} + \widehat{g}_{2,k}\right)^2 - \frac{9}{n}\left(\sum_{k=1}^{n}u_{1,k}u_{2,k}\right)^2\right). \tag{2.69}$$

A standardized approximately asymptotic normal expression can then be derived from $n^{-1}\left(\widehat{\rho}_S - \rho_S\right)$ as

$$\frac{n\left(\widehat{\rho}_S - \rho\right)}{12\sqrt{\sum_{k=1}^{n}\left(u_{1,k}u_{2,k} + \widehat{g}_{1,k} + \widehat{g}_{2,k}\right)^2 - \frac{9}{n}\left(\sum_{k=1}^{n}u_{1,k}u_{2,k}\right)^2}} \sim AsN(0,1). \tag{2.70}$$

**2.3.4.4 Blomqvist's Beta** The dependence measure known as Blomqvist's Beta – or the medial correlation coefficient – is defined as the difference in probability of concordance and discordance and is a scale-invariant measure (see [14]). In contrast to Kendall's tau and Spearman's rho though, Blomqvist's beta is defined using a bivariate random vector and the component medians of the vector. The definition is given below.

**Definition 2.13.** If $[X_1, X_2]$ is a random vector, then Blomqvist's beta is defined as

$$\beta(X_1, X_2) = \mathrm{P}((X_1 - \tilde{x}_1)(X_2 - \tilde{x}_2) > 0) - \mathrm{P}((X_1 - \tilde{x}_1)(X_2 - \tilde{x}_2) < 0). \tag{2.71}$$

where $\tilde{x}_1$ is the median of $X_1$ and $\tilde{x}_2$ is the median of $X_2$.

In [14], the most efficient sample estimator is claimed to be

$$\widehat{\beta} = \frac{2}{n}\sum_{k=1}^{n} \mathbb{1}\left(\left(R_{k,1} - \frac{n+1}{2}\right)\left(R_{k,2} - \frac{n+1}{2}\right) \geq 0\right) - 1 \tag{2.72}$$

where the sample ranks $R_{k,1}, R_{k,2}$ of $\left[X_{k,1}, X_{k,2}\right]$, $k = 1, \ldots, n$, are defined as in the preceding section.

For elliptical distributions, the median equals the mean due to symmetry which results in Blomqvist's beta being equal to Kendall's tau. Furthermore, the process $\sqrt{n}\left(\widehat{\beta} - \beta\right)$ is asymptotically normal with zero mean and asymptotic variance $1 - \beta^2$ as $n \to \infty$. This yields that

$$\frac{\sqrt{n}\left(\widehat{\beta} - \beta\right)}{\sqrt{1 - \widehat{\beta}^2}} \sim AsN(0, 1) \tag{2.73}$$

### 2.3.5 Diagnostics of Parametric Models

Suppose that estimates have been obtained under the maximum likelihood and least-squares criterion respectively for various parametric models. The problem then remains to decide which model bears the greatest resemblance to the data.

This section presents procedures and methods for assessing the Adequacy- and Goodness-of-fit for copula models. Section 2.3.5.1 presents numerical measures for assessing how good a model fits a certain set of data. The measures presented are however relative measures; when considering several models, the measures enable comparison of the goodness-of-fit of the models and thus provide a criterion for selecting the best model. However, the measures do not address the problem as to whether any one model fits the data set sufficiently good or not.

Section 2.3.5.2 presents methods for assessing whether or not the model adequately approximates the dependence of the data set by analyzing dependence measures. Additionally, statistics for copula hypothesis tests are introduced. The statistics are used to test the hypothesis implied by a certain model selection.

**2.3.5.1 Goodness-of-Fit Measures** This section presents three distinct criteria for measuring the goodness-of-fit of a model: Mean squared error, Akaike information criterion (AIC) and Bayesian information criterion (BIC). Mean squared error simply measures the deviations between the observed data and the model while AIC and BIC measures the likelihood of the observations under the model and penalize complexity. Given a set of plausible models, these measures can be used to internally rank the models by how good they fit a certain data set.

**Mean Squared Error** Model selection using the mean squared error as a criterion means selecting the model with the smallest sum of squared deviations from the observations. When fitting probabilistic models, there are several approaches to define the deviations of the model from the observed data. One way is to define the model deviations as the difference of the model quantiles and the empirical quantiles. Then the mean squared error is simply the function (2.49) divided by the sample size, i.e.

$$\mathrm{MSE}^q = \frac{1}{n}\mathcal{Q}^q\left(\widehat{\boldsymbol{\theta}}\right), \tag{2.74}$$

where $\widehat{\boldsymbol{\theta}}$ is a vector of the model parameters. Similarly to (2.49), this is only applicable for univariate models. Just as in Section 2.3.2, the joint distribution function can instead be considered in higher dimensions and the joint distribution can in turn be expressed using the corresponding copula. Defining the model deviations as the difference of the parametric copula function and the empirical copula function, the mean squared error of the model is

$$\mathrm{MSE}^{cdf} = \frac{1}{n}\mathscr{Q}^{cdf}\left(\widehat{\boldsymbol{\zeta}}\right), \tag{2.75}$$

where $\mathscr{Q}^{cdf}$ is as in (2.51) and $\widehat{\boldsymbol{\zeta}}$ is a vector of the model parameters. For most copula families, the copula function of the parametric copula in (2.51) cannot easily be computed analytically and is hence preferably approximated using Monte Carlo methods.

Another measure used for assessing goodness-of-fit is the mean squared error of the Kendall distribution function, used in [6]. This is however done for bivariate Archimedean copulas where the analytic form of the Kendall distribution function is well known. Deriving an explicit expression for the Kendall distribution function for copulas in general proves to be hard and a more fruitful approach is to resort to Monte Carlo approximation. A mean squared error goodness-of-fit measure can then be defined as

$$\mathrm{MSE}^{kdf} = \frac{1}{n}\sum_{i=1}^{n}\left(K_{n,i} - K_n^C\left(C_{n,i}\right)\right)^2, \tag{2.76}$$

where $K_{n,i}$ is estimated from the sample data by (2.43), $K_n^C$ is estimated as in (2.44) and is evaluated over the range of the empirical copula of the *sample* data.

In most cases, it is necessary to determine the copula function or Kendall distribution function by Monte Carlo approximation. The mean squared error measure is thus stochastic since the simulation step of Monte Carlo approximation introduces randomness. One way of producing a more robust measure is to repeat the Monte Carlo procedure to obtain a sample of (2.75) or (2.76) and compute the mean value. Mathematically, the procedure is to compute

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}\mathrm{MSE}_i. \tag{2.77}$$

This operation reduces the standard deviation by a factor of approximately $1/N$. How large $N$ is required to be for a sufficiently robust measure clearly depends on the number of simulations used for Monte Carlo approximation.

**Akaike and Bayesian Information Criteria** Both the Akaike information criterion (AIC) and Bayesian information criterion (BIC) are based on rewarding high likelihood while penalizing complexity, where complexity is taken to be the number of model parameters. If $\widehat{\boldsymbol{\theta}}$ is vector of the parameters of the model, $n_p$ is the number of parameters and $n$ is the sample size used to estimate the parameters, then AIC and BIC are defined as

$$\mathrm{AIC} = -2\ln\mathscr{L}\left(\widehat{\boldsymbol{\theta}}\right) + 2n_p, \tag{2.78}$$

$$\mathrm{BIC} = -2\ln\mathscr{L}\left(\widehat{\boldsymbol{\theta}}\right) + n_p\ln n. \tag{2.79}$$

In AIC, the sample size is not taken into consideration, whereas the complexity penalization in BIC proportional to the logarithm of the sample size. Using this definition, models with low AIC and BIC values are preferable.

**2.3.5.2 Diagnostics for Adequacy-of-Fit** When assessing the adequacy of fit of a copula model it is of particular interest to study the dependence structure of the model and and its sensitivity to the tail behavior of the data. A procedure that focuses on these characteristics of the model is suggested in [14].

The scale-invariant dependence measures of Section 2.3.4 provide a simple method to compare the general dependence structure of the model and of the data. By comparing empirical estimates with model values of the dependence measures it is possible to evaluate whether the dependence allowed within the model adequately approximates the dependence observed in the data. More specifically, provided a sufficiently large sample, the asymptotic distributions (2.63), (2.70), and (2.73) can be used to approximate confidence intervals of Kendall's tau, Spearman's rho, and Blomqvist's beta of the observed data. It is then easy to evaluate whether the corresponding model values falls within the intervals or not.

**Tail-Weighted Measures of Dependence** To assess if a copula model adequately approximates the tail dependence of a set of data requires more complex measures. One example of such measures are the tail-weighted measures of dependence presented in [16], defined as

$$\varrho_L\left(U_1,U_2;a(\cdot),p\right)=\mathrm{Cor}\left[a\left(1-\frac{U_1}{p}\right),a\left(1-\frac{U_2}{p}\right)\middle|U_1<p,U_2<p\right]$$
$$\varrho_U\left(U_1,U_2;a(\cdot),p\right)=\mathrm{Cor}\left[a\left(\frac{U_1}{p}-1\right),a\left(\frac{U_2}{p}-1\right)\middle|U_1>p,U_2>p\right],$$

where $a(\cdot):[0,1]\to(0,\infty)$ is a continuous function, $0\le p\le 1/2$ and $[u_1,U_2]$ is a random uniform $(0,1)$ vector with joint distribution function $C$. Additionally, $\mathrm{Cor}[\cdot,\cdot]$ denotes the Pearson's product-moment correlation in Section 2.3.4.1.

Further, [14] suggests a choice of $a(\cdot)$ and $p$ that provides useful in the application of assessing tail dependence of a copula model. Let $C_{i,j}$ be the bivariate copula marginal distribution of a $d$-dimensional copula and set $a(x;r)=x^r$ and $p=1/2$. Then, the tail-weighted measures of dependence are obtained as

$$\varrho_L\left(U_i,U_j;r\right)=\mathrm{Cor}\left[(1-2U_i)^r,(1-2U_j)^r\middle|U_i<\frac{1}{2},U_j<\frac{1}{2}\right] \tag{2.80}$$
$$\varrho_U\left(U_i,U_j;r\right)=\mathrm{Cor}\left[(2U_i-1)^r,(2U_j-1)^r\middle|U_i>\frac{1}{2},U_j>\frac{1}{2}\right], \tag{2.81}$$

where $U_i,U_j$ are the components corresponding to the marginal $C_{i,j}$. For convenience, $\varrho_L$ and $\varrho_U$ will be referred to as the lower and upper tail-weighted dependence coefficient respectively. It is suggested in [14] that $r$ equal to 5 or 6 is a good choice that balances variability and capability to discriminate tail dependence from intermediate tail dependence.

The distribution of the linear coefficient of correlation depends on the distribution of the underlying random variables, and is non-trivial for any distribution but the normal distribution. Hence, the distributions of (2.80) and (2.81) have to be approximated through Monte Carlo methods for the purpose of constructing confidence intervals.

### 2.3.6 Confidence Intervals

Let $\theta_n$ be a random variable such that $\lim_{n\to\infty}\theta_n\in N(\theta,\sigma)$, i.e. $\theta_n$ is asymptotically $N(\theta,\sigma)$ as $n\to\infty$. Then, if $\widehat{\theta}_n$ is an observation of $\theta_n$, an approximate confidence interval for $\theta$ with significance level $\alpha$ is

$$I_\theta=\widehat{\theta}_n\pm\Phi^{-1}(1-\alpha)\widehat{\sigma} \tag{2.82}$$

where $\widehat{\sigma}$ is the large-sample estimator of the asymptotic variance of $\theta_n$.

**Monte Carlo Approximation by Bootstrap** Consider the observations $x_1,...,x_n$ of independent and identically distributed random variables $X_1,...,X_n$ with unknown univariate distribution $F$. Suppose that the sought quantity $\theta$ depends on the distribution $F$, i.e. $\theta=\theta(F)$. Let $\widehat{\theta}$ be the estimator of $\theta$ on the random variables $X_1,...,X_n$ and let $\widehat{\theta}_{obs}=\theta(F_n)$ be the point estimate based on the sample $x_1,...,x_n$. Then, the problem here is that the sample $x_1,...,x_n$ is sufficient to obtain $\widehat{\theta}_{obs}$, but the approximation of a confidence interval for $\theta$ requires knowledge of the distribution of $\widehat{\theta}$.

One approach to approximate the confidence interval is to use Monte Carlo approximation which means that the distribution of $\widehat{\theta}$ is approximated as the empirical distribution of the sample $\widehat{\theta}^\star_1,...\widehat{\theta}^\star_N$ for some large $N$. This, however, requires a method such as the Bootstrap method to create $\widehat{\theta}^\star_k$, $k=1,...,N$. Below, the Monte Carlo and Bootstrap procedure for confidence interval approximation from [12] is presented step-by-step. For notational clarity, let $\widehat{\theta}_{obs}$ denote the point estimate of $\theta$ on the original sample $x_1, ..., x_n$ and let $\widehat{\theta}^\star$ be the estimator of $\theta$ on the samples created by bootstrapping, i.e. $\widehat{\theta}^\star_k$, $k = 1, ..., N$, are outcomes of $\widehat{\theta}^\star$.

(i) For each $k \in \{1, ..., N\}$, draw uniformly with replacement $n$ elements from the sample $x_1, ..., x_n$ to obtain the sample $\{X_1^{\star(k)}, ..., X_n^{\star(k)}\}$ and denote the corresponding empirical distribution function $F_n^{\star(k)}$.

(ii) Compute the estimates $\widehat{\theta}^\star_k = \theta\left(F_n^{\star(k)}\right)$ and the residuals $R^\star_k = \widehat{\theta}_{obs} - \widehat{\theta}^\star_k$ for $k = 1, ..., N$.

(iii) Approximate a confidence interval for $\theta$ with significance level $\alpha$ by

$$I_\theta = \left[\widehat{\theta}_{obs} + R^\star_{\lceil N\alpha/2\rceil, N}; \widehat{\theta}_{obs} + R^\star_{\lceil N(1-\alpha/2), N\rceil}\right] \tag{2.83}$$

where $R^\star_{1,N} \leq \cdots \leq R^\star_{N,N}$ is the ordered sample of residuals $R^\star_k$, $k = 1, ..., N$ and $\lceil y \rceil$ is the smallest integer larger than $y$.

The assumption within this procedure is that the distributions of $\theta - \widehat{\theta}$ and $\widehat{\theta}_{obs} - \widehat{\theta}^\star$ are approximately equal. This assumption is true if the bootstrap resampling is successful, which requires a sufficiently large $n$. For a large $n$, $\widehat{\theta}_{obs} \approx \theta$ and if $N$ is also large then $\widehat{\theta}^\star \overset{d}{\approx} \widehat{\theta}$. A more thorough explanation of the validity of (2.83) can be found in Section 7.5.2 in [12].

## 2.4 Copula Modeling

The construction of a copula model, i.e. fitting a copula to a set of data, involves several steps which can be described using the theory in Section 2.2 and 2.3. In Section 2.4.1, the procedures used to calibrate the copula parameters are presented. Section 2.4.2 illustrates how to generate random variables from a certain copula. Simulation from copulas is necessary for the Monte Carlo approximation required in the implementation of several of the diagnostic methods in Section 2.3.5. Further, Section 2.4.3 presents a summary of how to compare and select models using these methods.

### 2.4.1 Calibration of Copula Model Parameters

Consider a parametric copula model of the joint distribution of a random vector $\mathbf{X} = [X_1, ..., X_d]$

$$F_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\alpha}) = C\left(F_1\left(x_1; \boldsymbol{\eta}_1\right), ..., F_d\left(x_d; \boldsymbol{\eta}_d\right); \boldsymbol{\theta}\right),$$

where $\boldsymbol{\eta}_1, ..., \boldsymbol{\eta}_d$ are the parameters of the marginal models $F_1, ..., F_d$, $\boldsymbol{\theta}$ is the copula parameter. The joint distribution function parameter $\boldsymbol{\alpha}$ is merely present to emphasize that the final purpose of the copula modeling is a parametric model of a multivariate distribution. Further, let $\mathbf{x}_i = \left[x_{1,i}, ..., x_{d,i}\right]$, $i = 1, ..., n$, be a sample of observations of $\mathbf{X}$. Then, the aim is to estimate the parameters $\boldsymbol{\eta}_1, ..., \boldsymbol{\eta}_d, \boldsymbol{\theta}$ from the observations $\mathbf{x}_i$, $i = 1, ..., n$. Evidently, this estimation procedure is likely to comprise a large number of parameters and hence several estimation methods have been discussed in literature such as e.g. [10], [8], [15], and [14]. Two of the methods are presented in the subsections below.

**2.4.1.1 Inference Function for Margins – IFM** A two-stage estimation method that has been used frequently over the last two decades is the method commonly referred to as *Inference Function for Margins* (IFM). This method separates the estimation of the marginal parameters from the estimation of the copula parameters. The method is described in a step-by-step manner below.

1. For each univariate marginal distribution, estimate the model parameters $\boldsymbol{\eta}_j$, $j = 1, ..., d$, by $\widehat{\boldsymbol{\eta}}_j$, $j = 1, ..., d$, using either maximum likelihood (2.47) or least-squares of the quantile function (2.49).
2. Obtain an estimate, $\widehat{\boldsymbol{\theta}}$, of $\boldsymbol{\theta}$ by maximizing the IFM-log-likelihood function of the copula $C$, i.e. maximize

$$\ln \mathscr{L}_C(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ln c\left(F_1\left(x_{1,i}; \widehat{\boldsymbol{\eta}}_1\right), ..., F_d\left(x_{d,i}; \widehat{\boldsymbol{\eta}}_d\right); \boldsymbol{\theta}\right)$$

where the marginal function parameters are held fix. Here $c$ denotes the density function of the copula. Note that this it not the true log-likelihood of the copula since the marginal parameters are fixed as $\widehat{\boldsymbol{\eta}}_j$, $j = 1, ..., d$, which is why the function is referred to as the IFM-log-likelihood function.

Under a reasonable set of regularity conditions, it holds that $\sqrt{n}\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)$ is asymptotically normal with zero mean; see [15].

**Composite IFM – CIFM** In the case when the model copula $C$ is elliptical, the second step described above can by executed in a different manner. If $C$ is elliptical, then some or all components of $\widehat{\boldsymbol{\theta}}$ can be obtained using (2.60). If $C$ is Gaussian, then $\widehat{\boldsymbol{\theta}}$ is the correlation matrix and all components can be obtained using (2.60). If $C$ is a Student's $t$ copula, then all components but one of $\widehat{\boldsymbol{\theta}}$ can be obtained through (2.60). The remaining component, the degrees of freedom parameter $\nu$, is then preferably obtained by maximum likelihood estimation with all the other components held fixed, as is suggested in [5]. Formally, $\widehat{\nu}$ is obtained by maximizing

$$\ln \mathscr{L}_C(\nu) = \sum_{i=1}^{n} \ln c\left(F_1\left(x_{1,i}; \widehat{\boldsymbol{\eta}}_1\right), ..., F_d\left(x_{d,i}; \widehat{\boldsymbol{\eta}}_d\right); \widehat{\mathbf{R}}, \nu\right) \tag{2.84}$$

where $\widehat{\mathbf{R}}$ is the estimate of $\mathbf{R}$ using (2.60).

This method is referred to as *Composite Inference Function for Margins* (CIFM) since it is a composition of (2.60) and the IFM method.

**2.4.1.2 Semi-Parametric Method – SP** A possible shortcoming of the IFM method is the inevitable risk of inconsistency due to miss-specification. The IFM method requires $d$ parametric marginal models and one parametric copula model to be chosen appropriately and [15] shows that miss-specification of just one marginal can have severe effects on the estimation of the copula parameters.

Another two-stage estimation method is introduced in [8] and is referred to as the *Semi-Parametric Method* (SP). Unlike the IFM method, the SP method does not require specification of the marginal distributions, instead they are estimated by their empirical distribution function (2.38). The method is described in a step-by-step manner below.

1. For each univariate marginal distribution, compute the empirical distribution function $F_{1,n}$ as (2.38).
2. Obtain an estimate, $\widehat{\boldsymbol{\theta}}$, of $\boldsymbol{\theta}$ by maximizing the SP-log-likelihood function of the copula $C$, i.e. maximize

$$\ln \mathscr{L}_C(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ln c\left(F_{1,n}\left(x_{1,i}\right), ..., F_{d,}\left(x_{d,i}\right); \boldsymbol{\theta}\right)$$

where $c$ denotes the density function of the copula. Note that this it not the true log-likelihood of the copula since the marginal distribution functions have been replaced by the empirical marginal distribution functions, which is why the function is referred to as the SP-log-likelihood function.

In [8] it is shown that $\sqrt{n}\left(\widehat{\theta}_k - \theta_k\right)$ is asymptotically $N\left(0, \upsilon^2\right)$, $\forall k$, for some $\upsilon^2$. Further, a consistent estimator of $\upsilon$ is given in Section 3 of [8]. Moreover, the results of [15] suggest that the SP method should be preferred over IFM in most practical situations due to its robustness against miss-specification of the marginal distributions.

**Composite SP – CSP** Similarly to CIFM, a composite estimation method can be constructed to replace the second step in the SP method if the model copula is elliptical. In the case when $C$ is elliptical, some or all of the components of $\boldsymbol{\theta}$ can be estimated by (2.60). If $C$ is Gaussian, then $\boldsymbol{\theta}$ is the correlation matrix and all of its components can be obtained using (2.60). If $C$ is a Student's $t$ copula, then all components but one of $\boldsymbol{\theta}$ can be obtained through (2.60). Just as for CIFM, the degrees of freedom parameter $\nu$ is then obtained by maximum likelihood with all other components of $\boldsymbol{\theta}$ held fixed. Formally, $\widehat{\nu}$ is obtained by maximizing

$$\ln \mathscr{L}_C(\nu) = \sum_{i=1}^{n} \ln c\left(F_{1,n}\left(x_{1,i}\right), ..., F_{d,n}\left(x_{d,i}\right); \widehat{\mathbf{R}}, \nu\right) \tag{2.85}$$

where $\widehat{\mathbf{R}}$ is the estimate of $\mathbf{R}$ using (2.60).

Analogous to the case of CIFM, this method is referred to as *Composite Semi-Parametric Method* (CSP) since it is a composition of (2.60) and the SP method.

**2.4.1.3 Advantages and Drawbacks of Composite Methods** If provided with pseudo-observations $\mathbf{u}_i$, $i = 1, ... n$, then there is no difference between the IFM and SP methods for parameter calibration, i.e. there is no difference when the marginal models are disregarded. Hence, the four calibration methods in Section 2.4.1 (IFM, CIFM, SP, and CSP) will be referred to as composite methods (CIFM, CSP) and non-composite methods (IFM, SP) from here on.

When fitting a Student's $t$ copula to the set of observations, then the composite methods will prevent the ambiguity of the MLE described in 2.3.3. In a context of a large number of free parameters, such as calibration of high dimensional elliptical copulas, the possible inaccuracy of MLE due to a "flat" maximum is more palpable. In CIFM and CSP, the copula correlation matrix is set correspond the dependence structure in the set of observations as estimated by (2.60). When fitting a Gaussian copula using CIFM or CSP, the procedure does not require any specification or estimation of the marginal distribution since Kendall's tau is scale-invariant.

In [5], the grouped $t$ copula is fitted using composite methods. Further, the ML estimation of the degrees of freedom parameters is done separately for each group. When considering the grouped $t$ copula, an IFM or SP approach may not be feasible due to the copula density function being non-trivial. By the same argument, applying (2.84) or (2.85) on the entire copula to simultaneously estimate the degrees of freedom parameter of every group may not be possible. Hence, the composite methods with separate optimization of (2.84) and (2.85) are the only implementable methods among the copula calibration methods illustrated above. One evident drawback of the separated ML estimation is that the group specific random variables $R_1, ..., R_m$ in (2.33) are by definition not independent.

**Remedy for Indefinite Correlation Matrices** In high-dimensional contexts, in the sense that the number of marginals is high, the estimator (2.60) may yield indefinite correlation matrices. The problem is most commonly that the obtained correlation matrix has one negative eigenvalue very close to zero. Evidently, this is a major drawback since true correlation matrices are positive semidefinite by definition. A remedy to this shortcoming is presented in [19], where a method called the eigenvalue method for modifying estimated correlation matrices to ensure positive semidefiniteness is illustrated. In this method, the negative eigenvalues of the correlation matrix is replaced by small positive numbers. On a further note, positive semidefinite correlation matrices are necessary for any application of maximum likelihood on Gaussian or Student's $t$ copulas since the likelihood function contains the inverse of the correlation matrix.

When the IFM- or SP-log-likelihood functions are maximized using a numerical optimizer, practical problems can occur as the optimizer passes an indefinite correlation matrix while iterating. Obviously, each correlation parameter should be constrained such that $\rho_{i,j} \in [-1, 1]$, $\forall i, j$. Further, constraints for the correlation matrix to be positive definite is required for the log-likelihood to be defined on the entire domain. This second constraint is however by far more complicated to implement. In practice, an easier solution to redefine the log-likelihood functions of the Gaussian and Student's $t$ copulas to yield sub-maximal function values when the inputted correlation matrix is not positive semidefinite.

### 2.4.2 Simulation from Joint Distributions Using Copulas

The main steps when simulating from a joint distribution through a copula are significantly different depending on copula family. In general, the method is to draw a sample from the copula and then transform the sample using the quantile transform to obtain a sample from the sought joint distribution.

Suppose the aim is to simulate a sample from the joint distribution $F_{\mathbf{X}}$, where $\mathbf{X}$ has marginals $F_1, ..., F_d$ and copula $C$. Then, as was shown intially in Section 2.2.3, it holds that

$$\mathbf{X} = \left[F_1^{-1}(U_1), ..., F_d^{-1}(U_d)\right]$$

where $\mathbf{U}$ has distribution function $C$ and uniform $(0, 1)$ components. Thus, if a sample from the copula $C$ can be simulated, then a sample from $F_{\mathbf{X}}$ can be obtained. However, for many applications it is sufficient to simulate a uniform sample from the copula itself. For example, if the aim is to implement Monte Carlo approximation of scale-invariant measures, then it is sufficient to work with samples from the copula function $C$.

**2.4.2.1 Simulation from Elliptical Copulas** In Section 2.2.3.1 it was shown that the copula function $C$ of elliptical copulas is (2.28) and that the random vector $\mathbf{U}$ with joint distribution function $C$ can be formulated as (2.26). Thereby, a sample of the copula $C$ can be obtained by first simulating from the elliptical distribution of $\mathbf{Y}$ and then transforming using the marginal distribution functions of $\mathbf{Y}$. More concise, the procedure is to:

1. Simulate $\mathbf{Y}$ from the elliptical distribution $F_{\mathbf{Y}}$
2. Set $U_k = H_k(Y_k)$
3. Set $X_k = F_k^{-1}(U_k)$

Above, the third step is excluded in case the aim is to simulate a uniform sample from the copula $C$.

**Gaussian Copula** Simulations from a Gaussian copula with correlation matrix $\mathbf{R}$ are obtained if the distribution of $\mathbf{Y}$ is set to be multivariate standard normal with correlation matrix $\mathbf{R}$, i.e. $\mathbf{Y} \in N_d(\mathbf{0}, \mathbf{R})$. In this case, the multivariate distribution function in step one above is $F_{\mathbf{Y}}(\mathbf{y}) = \Phi^d_{\mathbf{R}}(\mathbf{y})$ and its univariate marginal distribution functions are $H_k(y) = \Phi(y)$ for every $k$.

**Student's t-Copula** Simulations from a Student's $t$ copula with correlation matrix $\mathbf{R}$ and degrees of freedom $\nu$ are obtained if the distribution of $\mathbf{Y}$ is set to be multivariate standard Student's $t$ with correlation matrix $\mathbf{R}$ and degrees of freedom $\nu$, i.e. $\mathbf{Y} \in t_d(\mathbf{0}, \mathbf{R}, \nu)$. In this case, the multivariate distribution function in step one above is $F_{\mathbf{Y}}(\mathbf{y}) = t^d_{\nu,\mathbf{R}}(\mathbf{y})$ and its univariate marginal distribution functions are $H_k(y) = t_\nu(y)$ for every $k$.

**2.4.2.2 Simulation from the Grouped t copula** The procedure of simulating from the grouped $t$ copula is similar to that of elliptical copula but do require some extra steps due to the grouping. A concise step-by-step description of how to simulate a uniform $(0,1)$ random vector $\mathbf{V}$ from a grouped $t$ copula with correlation matrix $\mathbf{R}$ is presented in [5].

1. Simulate $\mathbf{Z}$ from $N_J(\mathbf{0}, \mathbf{R})$
2. Simulate $U$, independently from $\mathbf{Z}$, from the uniform distribution on $(0,1)$
3. Set $R_k = G^{-1}_{\nu_k}(U)$ for $k = 1, ..., m$ by using (2.35)
4. Obtain the $J$-dimensional random vector $\mathbf{Y}$ by (2.33)
5. Set $V_k = F_k(Y_k)$ where $F_k$ is the distribution function of $Y_k$, for $k = 1, ..., J$, i.e. $F_k = t_{\nu_1}$ for $k \in \{1, ..., s_1\}$, $F_k = t_{\nu_2}$ for $k \in \{s_1 + 1, ..., s_1 + s_2\}$, etc

To obtain a sample with arbitrary marginal distributions, the random variables $V_k$, $k = 1, ..., J$, are simply transformed using the quantile transform and some arbitrary continuous strictly increasing distribution functions $H_1, ..., H_J$.

### 2.4.3 Selection of Copula Model

This thesis presents a methodology for finding the best copula model for a set of observations consisting of two steps: first reject all inadequate models and then find model with the best goodness-of-fit. The first step, rejection of inadequate models, is given a more conceptual presentation since the methods necessary for significant hypothesis testing require immense computational power. Thus, this thesis will study and implement these methods in small scale and then argue that the same methodology can be repeated in larger scale to attain significant results. Both steps are executed using the dependence measures in Section 2.3.4 and the model diagnostics in Section 2.3.5.

It is perhaps enlightening to state a clear definition of what is meant by a 'copula model' before proceeding. Henceforth, a copula model of the multivariate joint distribution of a data set is defined to consist of the following components:

1. Model for marginal distributions
2. Estimator type for marginal models
3. Copula family
4. Estimator type for the copula

Here, "model for marginal distributions" refers to both parametric (normal, Student's $t$, polynomial normal) distribution models and the non-parametric model using the empirical distribution. The estimator types for marginal models considered is this thesis are MLE and LSE for parametric marginal models and the empirical distribution function for the non-parametric marginal model. The three families of copulas included are the Gaussian copula, Student's $t$ copula, and the grouped $t$ copula. Both composite and non-composite estimation methods are used for all types of copulas except the grouped $t$ copula, for which only composite methods are used. Copula models will thus be categorized based on these four components.

One could consider the two-step approach assessing adequacy- and goodness-of-fit for each of the models for the marginal distributions before proceeding to constructing multivariate models using copulas. This approach is, however, not embraced in the analysis within this thesis. Instead, the fit of the copula model – which by the definition above includes marginal models – is assessed as a whole. The motivation for this is that the relevance is in the fit of the entire model rather than the individual fit of its components.

Deducing adequacy through dependence measures and hypothesis tests. Identifying the most appropriate model by comparative measures such as MSE, AIC and BIC.

**2.4.3.1 Rejection of Inadequate Models** This section mainly aims to summarize the application of the adequacy-of-fit assessment methods of Section 2.3.5.2 that utilize dependence measures. As an act of environmental friendliness, the method applications are presented briefly in step-by-step guides.

**Assessing Adequacy of Overall Dependence**

1. For the set of observations, estimate Kendall's tau, Spearman's rho, and Blomqvist's beta through (2.59), (2.67), and (2.72) and their respective asymptotic variances by (2.62), (2.69), and (2.73). Use the point estimates and their respective estimated asymptotic variances to compute approximate confidence intervals. The precision of the large-sample estimators for the asymptotic variances clearly depends on the size of the set of observations.
2. For each considered copula model, compute Kendall's tau, Spearman's rho, and Blomqvist's beta. If no analytic expression exists for one or more of the dependence measures, simulate a large uniform sample from the copula model and estimate the sought dependence measure. Provided that the simulated sample is large, confidence intervals of the estimates can be omitted as they will be of negligible width.
3. Check whether or not the estimates from step two are within the corresponding confidence intervals from step one. If for one copula model many estimates fall outside their intervals, then the copula model should be deemed to inadequately capture the overall dependence structure of the set of observations.

If a significance level of 5 % is employed, then there is an approximately 95 % probability that the confidence intervals obtained through step one will encompass the true values of the dependence measures. Conversely, if the copula model calibration methods are unbiased, then the estimates in step three would be encompassed by the confidence intervals with a probability of approximately 95 %. Note however that this only holds true for each pair-wise estimate individually since the values of Kendall's tau, Spearman's rho and Blomqvist's beta are not independent for different pairs of marginals.

**Assessing Adequacy of Tail Dependence**

1. For each considered combination of marginal models and estimator of marginal models, transform the set of observations into pseudo-uniform observations by using the probability transform in Proposition 2.1. For each obtained set of pseudo-uniform observations, estimate the upper and lower tail-weighted dependence coefficients in (2.80) and (2.81).

2. For each set of pseudo-uniform observations, use the Monte Carlo and bootstrap method in Section 2.3.6 to approximate confidence intervals for the upper and lower tail-weighted dependence coefficients through (2.83).

3. For each considered copula model, simulate a large uniform sample from the copula model and estimate the upper and lower tail-weighted dependence coefficients in (2.80) and (2.81). Provided that the simulated sample is large, confidence intervals of the estimates can be omitted as they will be of negligible width.

4. Check whether or not the estimates from step three are within the corresponding confidence intervals. If for one copula model many estimates fall outside their intervals, then the copula model should be deemed to inadequately capture the tail dependence in the set of observations.

If a significance level of 5 % is employed, then there is an approximately 95 % probability that the confidence intervals obtained through step one will encompass the true values of the dependence measures. Conversely, if the copula model calibration methods are unbiased, then the estimates in step three would be encompassed by the confidence intervals with a probability of approximately 95 %.

As a comment on the last step in both of the assessment methods above, it should be stated that this rejection criterion has not been developed into full scientific elegance. Though this method one will only obtain a single model estimate of each dependence measure for each variable pair. A more robust approach would be to approximate the distributions of the dependence measure estimates of the parametric copula model by a double bootstrap method analogous to the ones described in [10] and then assess whether or not it is probable that the dependence measure estimates on the data set of observations are probable to be outcomes of the corresponding approximate distributions. This approach is however considered to be beyond the scope of this thesis, but its implementation in future studies is encouraged by the author.

Without a sample of model estimates of each dependence measure for each variable pair, statistical conclusions regarding model rejection using the confidence intervals cannot be made. Double bootstrap would be one way to attain this. However, even without double bootstrap the test results can be used to compare the models internally.

**2.4.3.2 Identifying Model with Best Goodness-of-Fit** Provided a set of adequate copula models, the distinct measures of goodness-of-fit are used to deduce which model provides the best fit. The goodness-of-fit measures used to assess copula models are:

- Mean-squared error using the copula function, (2.75)
- Mean-squared error using the Kendall distribution function, (2.76)
- Akaike information criterion, (2.78)
- Bayesian information criterion, (2.79)

Among the Gaussian, Student's $t$, and grouped $t$ copula, only the Gaussian copula distribution function can be approximated from its analytic formulation. For the others, Monte Carlo approximation is used to approximate the copula function when computing (2.75). In detail, a large sample is simulated from the copula and then the copula function is approximated by the empirical copula function (2.40) on the simulated sample. Further, the Kendall distribution function lack an explicit expression for all three copulas mentioned above. Therefore, the approach of (2.77) is preferable for both (2.75) and (2.76) to reduce the randomness of the Monte Carlo approximations.

These measures provide means for comparison, however, they all measure different model qualities and the model choice is clearly not necessarily unambiguous. The key is to consult the goodness-of-fit measures with the intended use of the copula model in mind.

# 3 METHODOLOGY

The subsequent four subsections concisely presents the application of the theory in Section 2, the real world data used for model calibration and assessment, and the computational details of this project.

## 3.1 Usage of Software and Hardware for Computations

All simulations and numerical computations have been executed using the software RStudio. Further, the optimization of log-likelihood functions and quadratic sum functions (for the purpose of least-squares) has been done using the *optim* numerical optimizer.

If Skadeförsäkring has provided a virtual server which has been utilized for computations during the entire course of the thesis. The server used eight cores but was temporarily scaled up to 16 cores for some of the heavier computations. The functionality of the package *parallel* for RStudio was implemented in order to enable the use of multiple processor cores.

## 3.2 Acquisition, Selection and Processing of Financial Data

In summary, the studies in this thesis will employ time series data from 16 macroeconomic variables of which five are stock indices, five are government bond implied five-year zero rates, and six are U.S. credit spread indices for various ratings. All time series data considered are with daily observation frequency, which justifies the linear approximations of asset returns in Section 2.1. The details of each variable are presented in the following subsections.

It should be mentioned that it is not an obvious choice to use daily observations. As was mentioned in Section 1, using daily observations results in historical data yielding a larger sample of observations which leads to more robust statistical methods such as parameter estimation. Just to present one example where daily observations might pose a poor choice, suppose that the volatility of daily log-returns of e.g. an equity index is high but the volatile movement tend to net out over time so that the volatility of monthly log-returns would be lower. Suppose further that one wants to estimate the return distribution of the equity index one month ahead in time. Then, a model based on constructing the monthly return distribution by estimating the daily return distribution and generating a month's worth of daily returns would infer a higher volatility than what can be historically observed. The high model volatility can then be considered a consequence of over-modelling.

On the other side, if the I.I.D. assumption holds true for daily log-returns, then it follows that it is true for log-returns of any positive integer number of days, e.g. months, quarters, or years. Alas, the I.I.D assumption is only an assumption and rarely completely true. Thus, at the end of the day the choice of observation frequency comes down to finding a frequency that maximizes the "trueness" of the I.I.D. assumption. This thesis does however not treat this optimization problem and instead uses daily observations to maximize the sample size.

### 3.2.1 Equity

One stock index from each of the monetary regions is used to model the stochastics of the return on equity. Table 3.1 presents the stock indices elected to model equity and the length of the acquired time series. The five stock indices included in this study are indices which consist of stocks with great liquidity from large companies. This choice has been made because it pose good resemblance to the composition of If's equity portfolio. Additionally, it is partly the aim of If's investment strategy to have high correlation between If's equity portfolio and the OMXS30 index.

As a comparison, one rejected alternative is to use all-share indices, which can be assumed to reflect the overall equity market stochastics, including smaller and less liquid companies.

Table 3.1: The five stock indices used to model equity within each considered monetary region.

| Index | Description | Start Date | End Date |
|---|---|---|---|
| Euro STOXX 50 (SX5E) | The Euro STOXX 50 index consists of 50 of the largest and most liquid stocks from the *Eurozone* states, i.e. the member states of the European monetary union. | 1987-01-01 | 2016-04-08 |
| Dow Jones Industrial Average (INDU) | The Dow Jones Industrial Average is a price-weighted index consisting of stocks from 30 U.S. blue-chip companies and covers all industries except transportation and utilities. | 1925-01-05 | 2016-04-08 |
| OMX Copenhagen 20 (OMXC20) | Formerly known as the KFX index, this index consists of the 20 most liquid stocks on the Copenhagen Stock Exchange (CSE). | 1989-12-04 | 2016-04-08 |
| OBX Index (OBX) | The OBX index is a stock index on the Oslo Stock Exchange (OSE) and includes the 25 most liquid stock from companies listed on the OSE. | 1996-01-02 | 2016-04-08 |
| OMX Stockholm 30 (OMXS30) | The OMX Stockholm 30 index is a capitalization-weighted stock index and includes the 30 most liquid stocks of companies listed on the Swedish Stock Exchange (SSE). Up until 2005 this index was known as simply OMX index. | 1986-09-30 | 2016-04-08 |

### 3.2.2 Five-Year Risk-Free Interest Rates

The five-year risk-free rate is modeled using the interest rate implied by government bonds with a time to maturity of five years. When collecting data, the Swedish National Bank was used as a main source, as they offer long time-series of both Swedish and some non-Swedish government bonds on a daily basis. The Swedish National Bank did not however have data for Danish or Norwegian government bond rates and consequently these data were obtained from Bloomberg and Norges Bank (the National Bank of Norway). Bloomberg was used simply because the Danish National Bank does not provide historical data of government bond rates. Table 3.2 summarizes the government bonds included in this study.

Table 3.2: Government bonds included in the thesis and the source used to obtain historical data

| Macroeconomic Quantity | Source | Start Date | End Date |
|---|---|---|---|
| U.S. 5-Year Government Bond (USGVB5Y) | Swedish National Bank | 1987-02-02 | 2016-04-22 |
| German 5-Year Government Bond (DEGVB5Y) | Swedish National Bank | 1987-02-09 | 2016-04-22 |
| Danish 5-Year Government Bond (DKGVB5Y) | Bloomberg | 1993-01-04 | 2016-04-08 |
| Norwegian 5-Year Government Bond (NOGVB5Y) | Norges Bank | 1989-12-29 | 2016-04-21 |
| Swedish 5-Year Government Bond (SEGVB5Y) | Swedish National Bank | 1985-01-02 | 2016-04-21 |

The German 5-year government bond is used to model the risk-free interest rate of the Eurozone. The German government is among the most financially stable governments in the Eurozone which is why the the interest rate implied by the German government bond consistently has been among the lowest for Eurozone countries. Hence, the German government bond provides the best approximation of the risk-free interest rate for the euro currency.

### 3.2.3 Ten-Year Credit Spreads

The ten-year credit spread is modeled using credit spread indices. The available data for credit spread indices is sparse in several aspects. First, data series of credits spread indices are in general shorter, due to more recent start dates, than data series for equity or interest rates. Second, credit spread indices for monetary regions other than the U.S. are rare and very few were encountered during the data survey for this thesis. The non-U.S. spread indices encountered were quoted as total yield (spread rate plus risk-free rate) without providing a clear definition of the index, thus making it impossible to deduce the spread rate implied by the index yield. As a result, only U.S. spread indices are included in the studies.

All credit spread time series data have been obtained from Bloomberg and their definitions have been deciphered with assistance of Bloomberg and public corporate reports from Guotai Junan Securities[1], BlackRock[2], and Narodowy Bank Polski[3].

All but one of the credit spread indices are quoted as the actual spread rate, i.e. the process $s_t$ in Section 2.1.3. The one exception is the BICLB10Y index which is quoted as an index which follows the actual spread rate. However, this should not be of concern for the application of this thesis since it is the changes of the spread rate, and not the actual value of the spread rate, that is relevant for the purpose of dependency modeling and analysis.

Table 3.3: Credit spreads indices included in the thesis along with a description of the index compositions. Note that different rating agencies uses different labels, e.g. Moody's BAA rating corresponds to S&P's BBB rating.

| Index | Description | Quoted as | Start Date | End Date |
|---|---|---|---|---|
| BASPCAAA | The BASPCAAA Index represents the spread between Moody's 10-year corporate bond yields for bonds rated AAA and the U.S. 10-year government bond. | Credit spread | 1986-01-10 | 2016-04-08 |
| CSI A | The CSI A Index represents the spread between 10-year corporate bond yields for bonds rated A and the U.S. 10-year government bond. | Credit spread | 2002-09-25 | 2016-04-08 |
| CSI BB | The CSI BB Index represents the spread between 10-year corporate bond yields for bonds rated BB and the U.S. 10-year government bond. | Credit spread | 2002-09-25 | 2016-04-08 |
| CSI BBB | The CSI BBB Index represents the spread between 10-year corporate bond yields for bonds rated BBB and the U.S. 10-year government bond. | Credit spread | 2002-09-25 | 2016-04-08 |
| CSI BARC | The CSI BARC Index represents the spread between the yield to worst of Barclays Capital U.S. corporate high yield index and the U.S. 10-year government bond. | Credit spread | 1987-01-30 | 2016-04-08 |
| BICLB10Y | The BICLB10Y Index represents the spread between Moody's 10-year corporate bond yields for bonds rated BAA and the U.S. 10-year government bond. | Index | 1986-01-02 | 2016-04-08 |

[1] `http://www.gtja.com.hk/UploadFiles/gtja_enReport/2015/12/FI_RF_Dec.pdf` (Accessed 2016-05-19)
[2] `http://www.blogg.etfsverige.se/wp-content/uploads/2012/05/etpl_industrysummary_apr2012_Global_final.pdf` (Accessed 2016-05-19)
[3] `https://www.nbp.pl/publikacje/materialy_i_studia/213_en.pdf` (Accessed 2016-05-20)

Time series data for U.S. credit spread indices were found for more indices than those presented in Table 3.3. However, some of the time series acquired were deemed to be too short and were therefore not included. Consequently, the indices in Table 3.3 does not represent a complete range of credit ratings. The negative effect of including one or more short time series in the modeling is explained in the next section.

### 3.2.4 Processing of Marginal Data

The first processing step is to transform the time series data into outcomes of supposedly independent and identically distributed random variables, following the transformations described by (2.3), (2.5), and (2.6) in Section 2.1. The credit spread index BICLB10Y has to be transformed differently due to it being quoted as an index. Since an index value describes relative change, the BICLB10Y index is transformed into log-returns, similarly to the stock indices.

In order to be able to implement the copula modeling methods in 2.4, the marginal time series data has to be fused into a multivariate time series of joint observations. This is done by inspecting all the marginal time series data for observation dates that exist for every marginal. If an certain observation date exists for every marginal, then a joint observation exists for that date. The multivariate time series is constructed of these joint observations.

The number of missing internal data points within each marginal time series is low in relation to the total number of data points and hence the acquired number of joint observations is mainly dependent on the start and end dates of the marginal time series. Clearly, the number of joint observations relies heavily on the length of the shortest marginal time series, i.e. CSI A, CSI BB, and CSI BBB. Table 3.4 presents the length of each marginal time series. Additionally, the average number of data points per year have been computed for each time series to assess the completeness of data points. The first and last year of each time series are excluded from the average as the time series rarely start on the first day of a year or end on the last day of a year. For reference, the US calendar has an average of 252 working days per year between the years 1920 and 2016. This reference value is used to compute the right-most column in Table 3.4.

Table 3.4: The total number of raw data points used for each macroeconomic variable. The fourth column shows the ratio in percent between the average data points per year of each variable and the yearly average number of working days in the US calendar from 1920 to 2016.

| **Variable** | **Data Points** | **Avg. Points per Year** | **Ratio** % |
|---|---|---|---|
| SX5E | 7542 | 258 | 102 |
| INDU | 23022 | 252 | 100 |
| OMXC20 | 6592 | 250 | 99 |
| OBX | 5084 | 251 | 100 |
| OMXS30 | 7406 | 251 | 100 |
| USGVB5Y | 6971 | 251 | 100 |
| DEGVB5Y | 7072 | 251 | 100 |
| DKGVB5Y | 5532 | 237 | 94 |
| NOGVB5Y | 6440 | 261 | 104 |
| SEGVB5Y | 7828 | 251 | 100 |
| BASPCAAA | 8300 | 250 | 99 |
| CSI A | 3517 | 260 | 103 |
| CSI BB | 3523 | 260 | 103 |
| CSI BBB | 3523 | 260 | 103 |
| CSI BARC | 7615 | 261 | 104 |
| BICLB10Y | 7612 | 251 | 100 |

**3.2.4.1 Validity of IID Assumption of Marginals** It is fundamental for the application of a copula model that the considered data sample are outcomes of independent and identically distributed random variables. Evidently, the majority of the theory presented in this thesis relies on the assumption that IID data is at hand. Consequently, the transformation of the marginal time series into IID time series crucial for the reliability of the results of this thesis.

The lognormal, from which the transformations (2.3), (2.5), and (2.6) were derived, does not perfectly capture the characteristics of the financial market and hence the transformed marginal time series cannot be expected to be perfectly IID. The IID hypothesis can be tested using the *sample autocorrelation function*. How to test the IID hypothesis for a time series using the sample autocorrelation function is outlined in section 1.6 of [2]. In short, approximately 95% of the sample autocorrelations should be within the bounds $\pm 1.96/\sqrt{n}$, where $n$ is the sample size, if the considered sample is a realization of an IID sequence. However, testing using the sample autocorrelation function is not included in this thesis.

## 3.3 Estimation of Model Parameters

### 3.3.1 Modeling Marginal Distributions

Parametric distribution models of the normal, Student's $t$, and the polynomial normal (PN) distribution are fitted to the transformed marginal data using MLE (see (2.47)) and LSE (see (2.50)). Additionally, the empirical distribution function of each marginal is computed through (2.38). Note here that the entire transformed marginal time series data is used, rather than only the data points that are comprised within the multivariate time series of joint observations.

This choice results in more data points for the fitting of marginal models and hence less variability and data sensitivity. However, if the distribution characteristics of the transformed marginal time series data has a significant non-periodic time-dependency, then this choice results in the risk of the fitted marginal models being unrepresentative for the marginal data within the multivariate time series of joint observations.

Once the parametric and non-parametric marginal models have been computed, the marginal data in the set of joint observations are transformed into sets of pseudo-observations using the probability transform in Proposition 2.1. Thus, one set of joint pseudo-observations is obtained for each parametric marginal model and type of parametric estimator and one set of joint pseudo-observations is obtained for the non-parametric marginal model. All in all, seven sets of joint pseudo-observations are obtained. Mathematically, the set of joint pseudo-observations is

$$\mathbf{u}_i = \left[F_1\left(x_{1,i};\widehat{\boldsymbol{\eta}}_1\right),..,F_d\left(x_{d,i};\widehat{\boldsymbol{\eta}}_d\right)\right] \quad \text{and} \quad \mathbf{u}_i = \left[F_{1,n_1}\left(x_{1,i}\right),..,F_{d,n_d}\left(x_{d,i}\right)\right] \quad i = 1,...,m$$

for the parametric models and non-parametric model respectively, where $m$ is the number of joint observations and $n_1,...,n_d$ are the lengths of the transformed marginal time series.

### 3.3.2 Calibrating Copula Model Parameters to Pseudo-Observations

Gaussian copulas and Student's $t$ copula are fitted using both composite and non-composite methods to each of the sets of joint pseudo-observations. The resulting correlation matrix from (2.60) was not positive semidefinite and was thus modified using the eigenvalue method mentioned in Section 2.4.1.3. It was found to be sufficiently good to replace the negative eigenvalue with the value of the smallest positive eigenvalue times a factor of one half. Further, when calibrating Student's $t$ copulas using composite methods, a initial value of $\nu = 15$ was used for the numerical ML estimations.

For the numerical ML estimation of the copula parameters comprised within the non-composite methods, the obtained copula parameter estimates of the composite methods were used as initial values.

The grouped $t$ copula is only calibrated using composite methods. Further, the optimization of (2.84) and (2.85) is done for each group separately within the copula rather than for the entire copula.

The results of the model calibration, with calibration methods taken into account, are five calibrated parametric copulas for each of the seven sets of pseudo-observations, namely:

- Gaussian copula calibrated by composite methods
- Gaussian copula calibrated by non-composite methods
- Student's $t$ copula calibrated by composite methods
- Student's $t$ copula calibrated by non-composite methods
- Grouped $t$ copula calibrated by composite methods

For the grouped $t$ copula the marginals are grouped per instrument category as presented in Section 3.2.1, 3.2.2, and 3.2.3, i.e. the marginals are modeled in three groups, namely: equity, interest rates and credit spreads.

## 3.4 Computation of Model Diagnostics

Provided with the $5 \cdot 7 = 35$ calibrated copula models, all of the diagnostic methods in Section 2.3.5 can be applied. The dependence measures for assessment of adequacy-of-fit are computed as described in Section 2.4.3.1. Further, the four goodness-of-fit measures listed in Section 2.4.3.2 are computed for each of the 35 calibrated copula models. One exception however is that AIC and BIC are not computed for the grouped $t$ copula due to it not having a well-defined likelihood function.

Regarding the Monte Carlo approximation required for computation of the mean squared error measures, (2.75) and (2.76), the scale-invariant dependence measures, the tail-weighted dependence measures and the hypothesis tests, samples of 50 000 simulated joint observations were used. Due to the large number of simulations, $N = 10$ was used for the computations of (2.77).

# 4 RESULTS

This section presents the results of the calibration of the copula models as well as the results of the goodness-of-fit and adequacy-of-fit measures. However, due to the large number of considered macroeconomic variables, the obtained numerical values of pairwise measures such as linear correlation coefficients or other dependence measures are presented separately in the Appendix.

## 4.1 Adequacy-of-Fit of Copula Models

The numbers presented in this section are the most computationally heavy results in this thesis, due to the frequent use of Monte Carlo approximation and bootstrapping. In total, the results required months of computation time (computations were executed partly on an 8 core machine and partly on a 16 core machine). As an example, the goodness-of-fit mean squared errors required 24 days of computational time.

### 4.1.1 Scale-invariant Dependency Measures

The parametric copula models imply certain values of the dependency measures Kendall's tau, Spearman's rho and Blomqvist's beta. These values are computed using Monte Carlo simulation and are then compared to the confidence intervals approximated from the sample data. For some models it is possible to analytically compute the dependence measures instead of using Monte Carlo simulation.

Table 4.1: Kendall's tau, percentage inside confidence interval

| | Gaussian (Composite) | Gaussian | Student's t (Composite) | Student's t | Grouped t (Composite) |
|---|---|---|---|---|---|
| Normal & MLE | 0.992 | 0.967 | 0.992 | 0.992 | 1 |
| Normal & LSE | 0.992 | 0.942 | 0.992 | 0.967 | 1 |
| Student's t & MLE | 0.992 | 0.992 | 0.992 | 1 | 1 |
| Student's t & LSE | 0.992 | 0.967 | 0.992 | 1 | 1 |
| PNN & MLE | 0.992 | 0.992 | 0.992 | 1 | 1 |
| PNN & LSE | 0.992 | 0.967 | 1 | 0.992 | 1 |
| Empirical | 0.992 | 0.983 | 0.992 | 1 | 1 |

As can be seen in Table 4.1, nearly all model estimates are encompassed by their corresponding confidence intervals. The grouped $t$ copula model values of Kendall's tau are all inside their respective confidence interval for all marginal models.

On a further note, $1/120 = 0.008$, meaning that most of the copula models only "miss" the Kendall's tau value of one variable pair. By looking at Figure 1 through Figure 5, it appears that the credit spread indices *CSI BB* and *CSI BBB* have an extremely high Kendall's tau value. Table B.1 shows that the approximate confidence interval is $(0.953, 1.000)$. Additionally, Figure 1 through 5 show that only the models calibrated using non-composite methods capture this high Kendall's tau value. The grouped $t$ copula model capture the high Kendall's tau value despite being calibrated using composite methods however.

Table 4.2: Spearman's rho, percentage inside confidence interval

| | Gaussian (Composite) | Gaussian | Student's t (Composite) | Student's t | Grouped t (Composite) |
|---|---|---|---|---|---|
| Normal & MLE | 1 | 0.992 | 1 | 1 | 1 |
| Normal & LSE | 1 | 0.992 | 1 | 1 | 1 |
| Student's t & MLE | 1 | 1 | 1 | 1 | 1 |
| Student's t & LSE | 1 | 1 | 1 | 1 | 1 |
| PNN & MLE | 1 | 1 | 1 | 1 | 1 |
| PNN & LSE | 1 | 1 | 1 | 1 | 1 |
| Empirical | 1 | 1 | 1 | 1 | 1 |

The approximated confidence intervals for Spearman's rho on the sample data are relatively wide, as can be seen in the figures in Appendix A.1. Consequently, the adequacy-of-fit test using Spearman's rho has a high acceptance rate and as shown in Table 4.2, all copula models are assessed to be adequate.

Table 4.3: Blomqvist's beta, percentage inside confidence interval

| | Gaussian (Composite) | Gaussian | Student's t (Composite) | Student's t | Grouped t (Composite) |
|---|---|---|---|---|---|
| Normal & MLE | 0.892 | 0.783 | 0.883 | 0.875 | 0.908 |
| Normal & LSE | 0.908 | 0.775 | 0.908 | 0.883 | 0.933 |
| Student's t & MLE | 0.892 | 0.908 | 0.883 | 0.925 | 0.875 |
| Student's t & LSE | 0.908 | 0.883 | 0.892 | 0.927 | 0.917 |
| PNN & MLE | 0.908 | 0.917 | 0.892 | 0.925 | 0.908 |
| PNN & LSE | 0.900 | 0.867 | 0.908 | 0.917 | 0.900 |
| Empirical | 0.892 | 0.908 | 0.892 | 0.933 | 0.908 |

Table 4.3 shows a lower acceptance rate than both Table 4.1 and 4.2. Looking at Table 4.3, there are actually several marginal models for which the Gaussian copulas score higher than the Student's $t$ copulas. The highest results are obtained for the non-composite Student's $t$ copula with the empirical distributions as marginal models, i.e. the Student's $t$ copula calibrated using the SP method. Equally high results are obtained for the grouped $t$ copula model with normal marginals fitted using LSE.

Inspection of Figure 11 through 15 lead to the impression that there are variable pairs whose Blomqvist's beta are not captured by any model. More specifically, the Blomqvist's beta value of *DE GVB 5Y* and *BICLB 10Y*, *DK GVB 5Y* and *CSI A*, *OMXS30* and *DK GVB 5Y*, as well as *US GVB 5Y* and *BICLB 10Y*. Additionally, there are another three pairs whose Blomqvist's beta are not captured in the majority of the copula models. This absence of randomness in the copula models' shortcomings gives reason to suspect that the process of converting the market data into series of outcomes of supposedly IID random variables was inadequate for certain time series.

Worth mentioning is that the pair *NO GVB 5Y* and *BICLB 10Y* has a Blomqvist's beta value of 1.

#### 4.1.2 Tail-Weighted Measures of Dependence

The implied tail dependence of the parametric copula models are compared to the observed tail dependence in the sample data, using the tail-weighted measures of dependence. The tables below present ratios, for each copula model, of the number of model values that are within the corresponding confidence intervals approximated from the sample data. The model values of the tail-weighted measures of dependence are approximated using Monte Carlo simulation.

Table 4.4: Lower tail-weighted dependence measure, percentage inside confidence interval

| | Gaussian (Composite) | Gaussian | Student's t (Composite) | Student's t | Grouped t (Composite) |
|---|---|---|---|---|---|
| Normal & MLE | 0.258 | 0.200 | 0.408 | 0.442 | 0.683 |
| Normal & LSE | 0.233 | 0.175 | 0.358 | 0.292 | 0.642 |
| Student's t & MLE | 0.200 | 0.217 | **0.742** | 0.717 | 0.258 |
| Student's t & LSE | 0.183 | 0.158 | 0.667 | 0.675 | 0.308 |
| PNN & MLE | 0.175 | 0.200 | 0.650 | 0.625 | 0.267 |
| PNN & LSE | 0.175 | 0.192 | 0.708 | 0.733 | 0.158 |
| Empirical | 0.217 | 0.192 | 0.592 | 0.550 | 0.333 |

Table 4.4 contains the ratios of the number of model values of the lower tail-weighted dependence measure inside the approximated confidence intervals. Evidently, the number of model values within the confidence intervals are in every way lower if comparing with the scale-invariant dependence measures. These results are in no way strange or unexpected since none of the calibration methods have been weighted towards the tail. The calibration methods weight all observations equally.

The copula model with the highest ratio is the Student's t copula calibrated with composite methods and Student's t marginals estimated using MLE. The ratio of this model has been highlighted, despite not being deemed adequate by this test, simply because it is the model which provides the closest to adequate tail dependence.

By inspecting Figure 16 through 22 and comparing the estimates of the lower tail-weighted dependence measures on the copula models with the confidence intervals estimated on the set of observations, it becomes evident that the estimates on the grouped $t$ copula generally are above the confidence interval when not inside, and the estimates on the other copulas are generally below the confidence interval when not inside. This gives reason to believe that this grouped $t$ copula (with the chosen grouping considered) implies a stronger tail dependence structure than what can be observed in the set of observations.

Table 4.5: Upper tail-weighted dependence measure, percentage inside confidence interval

| | Gaussian (Composite) | Gaussian | Student's t (Composite) | Student's t | Grouped t (Composite) |
|---|---|---|---|---|---|
| Normal & MLE | 0.258 | 0.233 | 0.433 | 0.408 | 0.558 |
| Normal & LSE | 0.208 | 0.200 | 0.375 | 0.350 | 0.525 |
| Student's t & MLE | 0.142 | 0.150 | **0.725** | 0.708 | 0.267 |
| Student's t & LSE | 0.175 | 0.133 | 0.667 | 0.650 | 0.358 |
| PNN & MLE | 0.167 | 0.117 | 0.717 | 0.683 | 0.275 |
| PNN & LSE | 0.141 | 0.117 | **0.725** | 0.708 | 0.208 |
| Empirical | 0.158 | 0.125 | 0.675 | 0.658 | 0.317 |

In Table 4.5, the ratios obtained using the upper tail-weighted dependence measure are presented. Similarly to the case with the lower tail-weighted dependence measure, the ratios are all lower than those obtained using the scale-invariant dependence measures for adequacy-of-fit assessment.

By inspecting Figure 23 through 29 and comparing the estimates of the upper tail-weighted dependence measures on the copula models with the confidence intervals estimated on the set of observations, similar conclusions can be made as to that of with the lower tail-weighted dependence measure and Figure 16 through 22. However, when considering the upper tail-weighted dependence measure, the copula models with normal marginal models form an exception. For neither normal marginals fitted with MLE or LSE the grouped $t$ copula displays systematically higher estimates than those of the set of observations.

## 4.2 Goodness-of-Fit of Copula Models

### 4.2.1 Mean Squared Error

The goodness-of-fit of the copula models are assessed in two distinct ways using mean squared error, using the copula function and using the Kendall distribution function. In both cases, the mean squared error values are computed using the robust procedure (2.77). Further, Table 4.6 presents the mean squared error values computed using the copula function, see (2.75), and Table 4.7 presents the mean squared error values computed using the Kendall distribution function, see (2.76).

Table 4.6: The mean squared error of the copula function of the all the considered copula models. Values have been computed as the mean of 10 mean squared error values.

| | Gaussian (Composite) | Gaussian | Student's t (Composite) | Student's t | Grouped t (Composite) |
|---|---|---|---|---|---|
| Normal & MLE | 0.001708 | 0.001967 | **0.001682** | 0.002002 | 0.001909 |
| Normal & LSE | 0.002068 | 0.002326 | 0.002092 | 0.002543 | 0.002289 |
| Student's t & MLE | 0.003338 | 0.003683 | 0.003537 | 0.003698 | 0.003876 |
| Student's t & LSE | 0.003539 | 0.004023 | 0.003823 | 0.003979 | 0.004227 |
| PNN & MLE | 0.003706 | 0.004064 | 0.003897 | 0.004012 | 0.004292 |
| PNN & LSE | 0.004534 | 0.004972 | 0.005080 | 0.005190 | 0.005701 |
| Empirical | 0.004049 | 0.004580 | 0.004302 | 0.004521 | 0.004874 |

Table 4.7: The mean squared error of the Kendall distribution function of the all the considered copula models. Values have been computed as the mean of 10 mean squared error values.

| | Gaussian (Composite) | Gaussian | Student's t (Composite) | Student's t | Grouped t (Composite) |
|---|---|---|---|---|---|
| Normal & MLE | 232.4 | 223.2 | 229.8 | 219.7 | 223.9 |
| Normal & LSE | 232.5 | 224.3 | 229.8 | 214.7 | 220.4 |
| Student's t & MLE | 232.2 | 222.5 | 227.6 | 224.3 | 221.6 |
| Student's t & LSE | 232.5 | 219.9 | 225.2 | 224.9 | 221.1 |
| PNN & MLE | 232.3 | 222.4 | 226.7 | 224.3 | 215.2 |
| PNN & LSE | 232.9 | 220.0 | 228.4 | 224.0 | **214.4** |
| Empirical | 232.3 | 221.4 | 226.0 | 227.4 | 216.1 |

As can be seen from the results in Table 4.7, the copula model that poses the best fit to the data, when goodness-of-fit is measured using the mean squared error of the Kendall distribution function, is the Student's $t$ copula calibrated with non-composite methods and with normal marginal models fitted using least-squares. Note further that for any marginal model other than normal, the Gaussian copula with non-composite calibration is preferable.

### 4.2.2 Akaike and Bayesian Information Criterion

Below the Akaike and Bayesian information criterions are presented in Table 4.8 and 4.9. The values have been computed by applying (2.78) and (2.79) to the calibrated copula models.

Table 4.8: Akaike Information Criterion of all considered copula models, except those including the grouped $t$ copula.

| | Gaussian (Composite) | Gaussian | Student's t (Composite) | Student's t |
|---|---|---|---|---|
| Normal & MLE | −20413 | −42803 | −32639 | −49718 |
| Normal & LSE | −20034 | −43913 | −33727 | −**50168** |
| Student's t & MLE | −29303 | −45284 | −33828 | −49520 |
| Student's t & LSE | −28976 | −45426 | −33387 | −49315 |
| PNN & MLE | −28828 | −45192 | −33622 | −49368 |
| PNN & LSE | −28928 | −45284 | −33758 | −49455 |
| Empirical | −29355 | −45539 | −33511 | −44835 |

By inspecting Table 4.8 it is evident that from an AIC perspective, the most preferable copula model is the Student's $t$ copula calibrated with non-composite methods and with normal marginal models fitted using least-squares. Note additionally that the non-composite Student's $t$ copula is preferable over the the other copula models for every considered marginal model. With the exception of the case with empirical marginals, the non-composite Student's $t$ copula models are preferable to all other models, regardless of marginal models.

Table 4.9: Bayesian Information Criterion of all considered copula models, except those including the grouped $t$-copula.

| | Gaussian (Composite) | Gaussian | Student's t (Composite) | Student's t |
|---|---|---|---|---|
| Normal & MLE | −19522 | −41912 | −31741 | −48821 |
| Normal & LSE | −19143 | −43022 | −32829 | −**49270** |
| Student's t & MLE | −28318 | −44398 | −32837 | −38529 |
| Student's t & LSE | −27991 | −44441 | −32396 | −48324 |
| PNN & MLE | −27749 | −44113 | −32538 | −48283 |
| PNN & LSE | −27849 | −44205 | −32673 | −48370 |
| Empirical | −28651 | −44835 | −32801 | −48476 |

The results in Table 4.9 are similar to those in Table 4.8. However, the non-composite Student's $t$ copula models are not equally superior when the number of model parameters is penalized as it is when considering BIC. The most preferable copula model is the Student's $t$ copula calibrated with non-composite methods and with normal marginal models fitted using least-squares, exactly as when AIC is considered. One difference in the results relative to 4.8 is that the model with Student's $t$ marginals fitted using MLE and non-composite Student's $t$ copula scores notably worse compared to other models; the non-composite Gaussian copula is in fact preferable to the non-composite Student's $t$ copula, if only that marginal model is considered. Another notable difference is that the model with empirical marginals and non-composite Student's $t$ copula scores better and is the third most preferable model.

# 5 DISCUSSION

Regarding the first objective, several approaches to calibrating parametric copula models have been outlined. In literature, copula parameter estimation using non-parametric marginal models have been greatly cherished, mainly due to its obvious robustness to miss-specification. This choice is favorable when aiming to isolate the copula properties and achieve results with a more general application. However, the objective of this thesis concerns the driving stochastic factors of If's investment portfolio and hence the marginals are neither arbitrary or completely unknown. Although the best choice of model is not obvious, many of the marginal characteristics are known. Hence, the risk and mainly the severity of miss-specification of the marginals is lower in this setting. Thus, copula calibration methods capable of handling parametric marginal model are relevant.

In literature such as [15], the effect of marginal miss-specification is analyzed and found to be significant enough for the SP method to be preferred in general. However, the study is limited to bivariate Archimedean copulas and the marginal miss-specifications are limited to the hypothesized model being normal while the true distribution is more complicated (Student's $t$ distribution, skew-$t$ distribution or $\chi^2$-distribution). For example, the normal distribution is a special case of the Student's $t$ distribution ($\nu \to \infty$) and, as opposed to the normal distribution, the $\chi^2$-distribution is non-negative. In other words, the type of miss-specification investigated is one-sided and not necessarily representative to the miss-specifications one can encounter after having performed marginal analysis. For the case of this thesis, where both quantitative and qualitative information exist for the portfolio components, the degree of marginal miss-specification is likely to be less severe. The IFM and SP methods for estimation of copula parameters are the most widely used and researched methods in publications to this date. However, the performance of these methods have mainly been evaluated for bivariate Archimedean copulas, e.g. in [10] and [15].

Summarily, calibration of copula models to If's investment portfolio would preferably be done using high dimensional elliptical copulas on a set of marginals with similar and predictable characteristics, which opens up for parametric marginal models. This is a not thoroughly explored subcategory of the complex field of applied copula theory. It should also be mentioned that [14] presents the vine copula as a good model for high dimensional contexts. This copula family is beyond the scope of this thesis however.

Worth mentioning is that the methods for transforming the marginal observations to i.i.d. outcomes can have a far larger impact than the choice of whether to implement IFM or SP. This is due to the fundamental dilemma of financial mathematics that methods for estimating distribution parameters rely on the availability of i.i.d. data while financial time series tend to not be i.i.d. The choice of IFM versus SP or composite versus non-composite loses significance if the underlying data is not of sufficient quality for the methods to be reliable.

The second thesis objective have been met, although there is lot of room for improvement as has been pointed out. The ability of the copula models and calibration methods to fit the dependence structure of a set of historical observations have been analyzed and compared. However, a more robust implementation of the adequacy-of-fit tests would have yielded results with more of a statistical significance. Although this point does not completely invalidate the achieved results in this thesis, it should be kept in mind when analyzing the results; conclusions can still be made from the results.

The endeavour to assess copula model plausibility in this thesis should be put into perspective considering the limited published research within the specific field of application. Over the last two and a half decades, plenty of research has been published on the topic of copula theory and copula parameter estimation. However, the majority of the research done have chosen focus on the bivariate case, leaving the high dimensional multivariate case unmentioned or arguing that the findings are applicable in higher dimensions as well. That being true, higher dimensional applications can give rise to problems not encountered in the bivariate case. One example is the problem mentioned in [5] with the non-parametric estimation of elliptical copulas' correlation matrices using Kendall's tau, where the estimator may yield indefinite correlations matrices in high dimensional contexts. The inevitable increase in simulation and computation time that follows when moving to a high dimensional model should not be neglected. Sophisticated real world applications tend to require a high number of marginals, especially in market risk analysis.

The copula model assessments done in this thesis should be considered as an exploratory step in the environment of applied high dimensional copula modeling of macroeconomic variables. Some conclusions can be made regarding the copula models however. The results of the tail-weighted measures of dependence show something interesting. The systematic deviations of the copula models from the sample results imply a non-stochastic difference in the dependence structure between the models as well as with the sample of observations. Further, the results of the adequacy of fit tests using scale-invariant dependence measures are inconclusive as there is only one sample estimate for each variable pair that is being checked against a confidence interval. These results would carry greater significance had the method been extended with the double bootstrap method suggested in Section 2.4.3.1. This extension would however had increase the requirements regarding computational power of this thesis tremendously. The results attained can however be used to compare the models under study.

Regarding the tests for goodness-of-fit, the results are not unambiguous but there are some coherence. For example the Gaussian copula is found inferior by every test performed, regardless of the marginal models used. Additionally, the tests based on the density functions both prefer the Student's $t$ copula with LSE normal marginals. The Student's $t$ copula was one one best performers in adequacy-of-fit test as well, but the test results were inconclusive regarding the adequacy of all copulas tested due to insufficient methodology (no double bootstrap or equivalent method was implemented). For future studies it would be of interest to include the grouped $t$ copula in the AIC and BIC tests and perhaps include several plausible groupings of the marginals for the grouped $t$ copula.

Lastly, it should also be mentioned that whenever assessing how good a certain model fits data, the results are highly dependent of how "good" is defined, especially for multivariate models. A natural approach is to define "good" based on the intended purpose of the model. In this thesis, the purpose of the multivariate copula models was to capture the dependence structure of the observed data, without specifying whether focus should be on tail dependence or overall dependence. As was initially mentioned, another possible purpose of a copula model could be to estimate the 99.5% quantile of the one-year return of an aggregate portfolio where the components dependence structure is modelled using the copula. Then "good" should preferably be defined based on the estimation bias and estimation error of the return quantile.

A natural extension to the studies conducted in this thesis would be to define a portfolio value process that maps the macroeconomic variables into a single monetary value and then compute the distribution of the one-year return of the portfolio using a copula. One could then estimate the 0.5 % quantile (99.5 % quantile of the loss distribution) of the portfolio return for the purpose of risk calculation. Moreover, this would give reason to revisit the initial choice in this thesis to base the copula modelling on daily observations. It would be idyllic to further investigate the applicability of the I.I.D. assumption depending on the observation frequency.

The subsequent subsection briefly summarizes the author's ideas of how to proceed and further investigate relevant topics related to this thesis.

## 5.1 Suggestion of Topics for Future Research

1. Investigate the accuracy of estimation of elliptical copula correlation matrix through non-parametric estimation (the estimator (2.60)). Also, analyze of the estimator's accuracy depends on the size of the sample of observations. **Suggestion:** A simulation study where (2.63) is further investigated. A good general approach would be to simulate several samples from elliptical copulas with known correlation matrices and then estimate using (2.60). With estimates from several samples, and a varying sample size, then the distributional characteristics of (2.60) can be studied as a function of the sample size.

2. Investigate the capability of goodness-of-fit tests to identify the most preferable copula model. **Suggestion:** A simulation study where multiple samples are simulated from known copulas, then several copula models are calibrated to the simulated samples and compared using the goodness-of-fit tests presented in this thesis. With multiple simulated samples, once can statistically analyze the rate at which each goodness-of-fit test identifies the correct copula as the most preferable model. Additionally, the effect of the size of the simulated samples on the results can be included in the study.

3. Investigate the accuracy of confidence intervals for Kendall's tau, Spearman's rho and Blomqvist's beta approximated by using the large-sample estimators of the asymptotic variances. This study should be seen as an extension to the topic in 1. above. **Suggestion:** A large simulation study where several samples

are simulated from copulas where Kendall's tau, Spearman's rho and Blomqvist's beta are known. Then confidence intervals can be approximated on each sample by assuming that the estimators have their respective asymptotic distributions. The true confidence level of the approximated intervals can then be assessed. Further, the simulated samples can be used to perform a deeper analysis of the asymptotic distributions of the estimators. The theoretical distributions can be compared to the observed outcomes of the estimators and the large-sample estimators of the asymptotic variances for Kendall's tau and Spearman's rho can be assessed. Finally, studying the speed of convergence to the asymptotic distributions as function of sample size would be of great interest.

4. Investigate the capability of adequacy-of-fit tests to reject invalid copula models by assessing confidence intervals of scale-invariant dependence measures approximated using large-sample estimators of asymptotic variance. **Suggestion:** A simulation study where samples are simulated from copulas with known Kendall's tau, Spearman's rho and Blomqvist's beta. For each copula used for simulation, take the simulated sample to be the observed outcomes, estimate dependence measures and approximate confidence intervals. Then, list a set of possible copula models and assess their adequacy to describe the dependence structure of the observed outcomes by simulating a large sample from each model and estimating dependence measures. It would at this point be of great value to have significant results from the study described in 1. and 3. to use as grounds for determining what sample size to use for the samples simulated from the models. The dependency between the power of the adequacy-of-fit tests and the sample size of the observed outcomes constitutes another dimension of interest for this study.

Both step three and four can in theory also be repeated using the tail-weighted measures of dependence. However, since there are no analytic expressions for the distributions of the estimators, the tasks will be far heavier from a computational perspective and for example require bootstrapping.

# A FIGURES

## A.1 Scale-invariant Dependence Measures

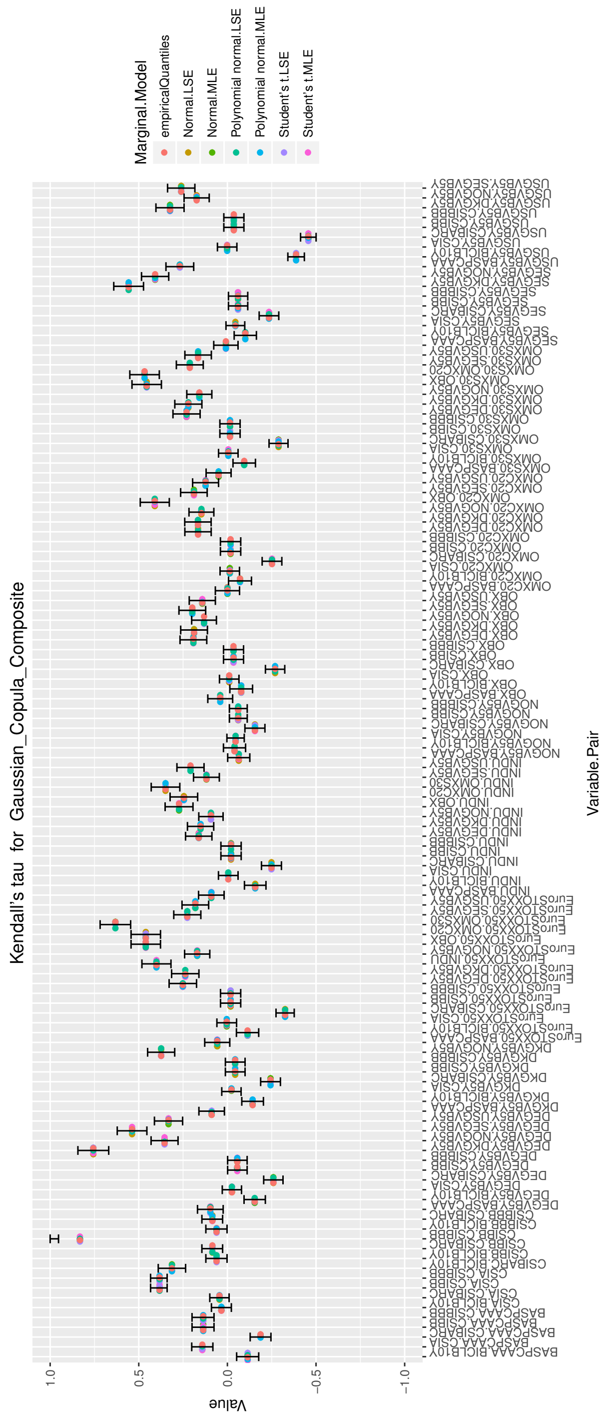


Figure 1: Kendall's tau for Gaussian copula with composite methods.

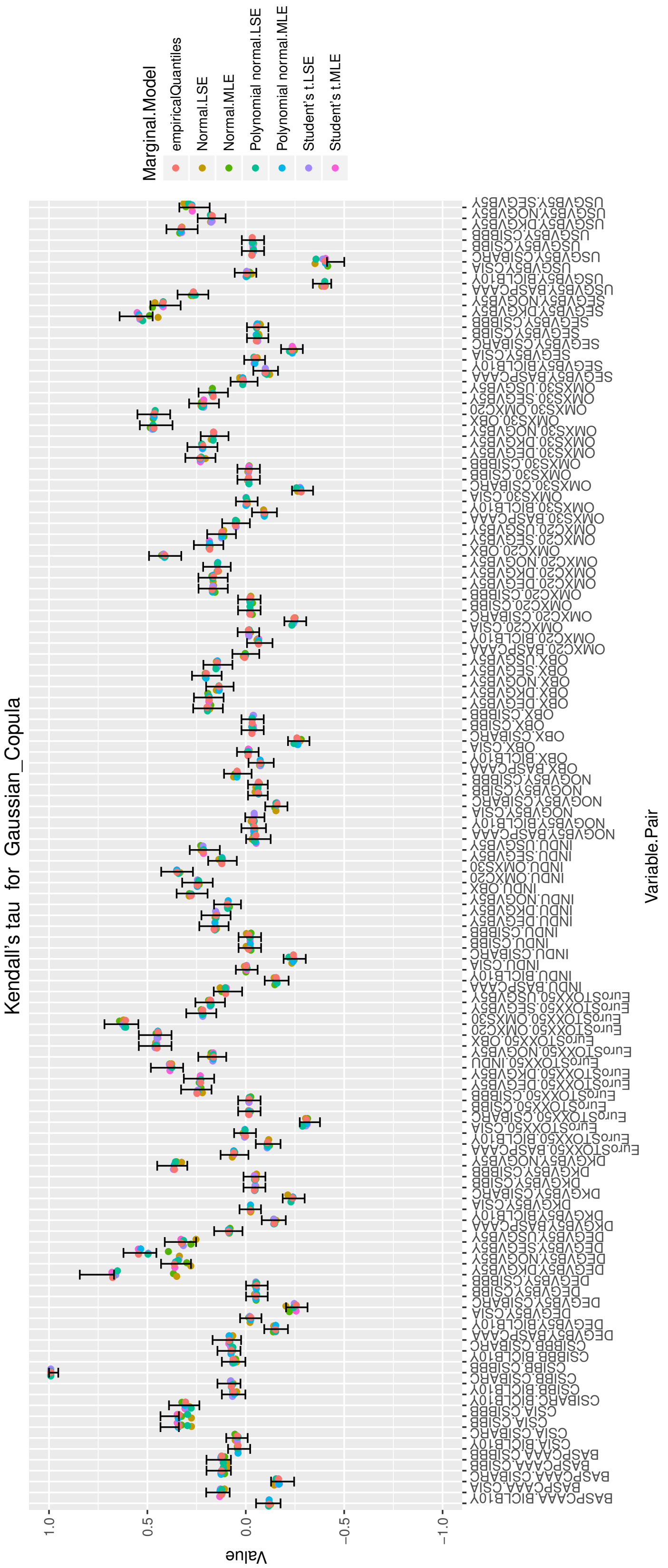


Figure 2: Kendall's tau for Gaussian copula with non-composite methods.

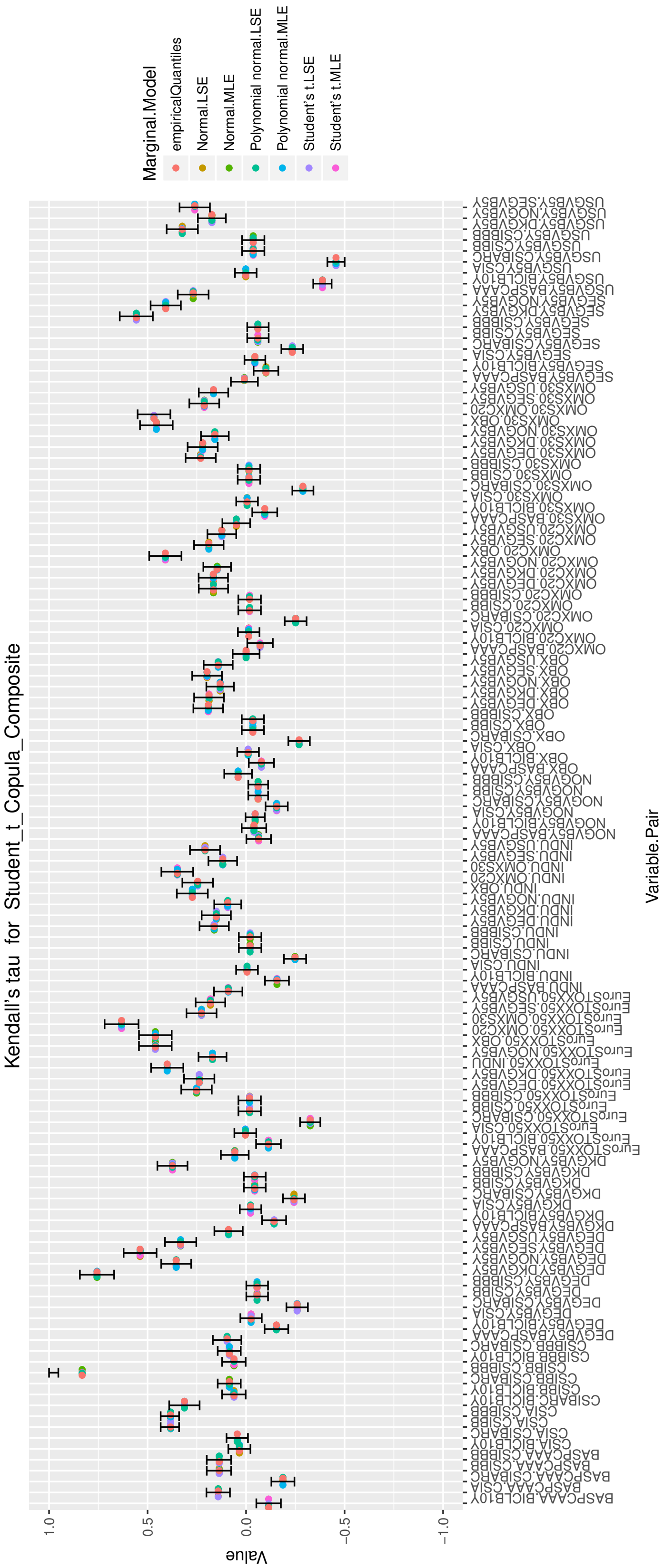


Figure 3: Kendall's tau for Student's t copula with composite methods.

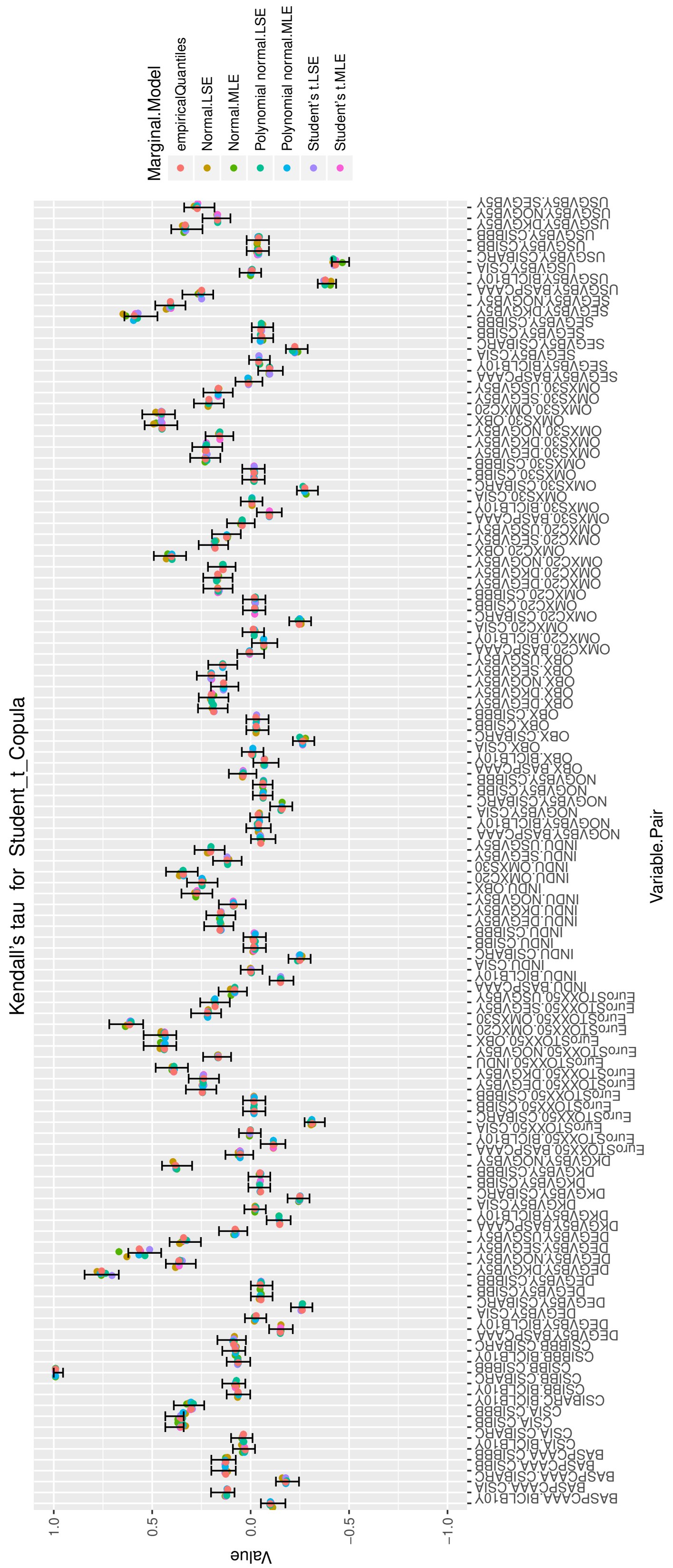


Figure 4: Kendall's tau for Student's t copula with non-composite methods.

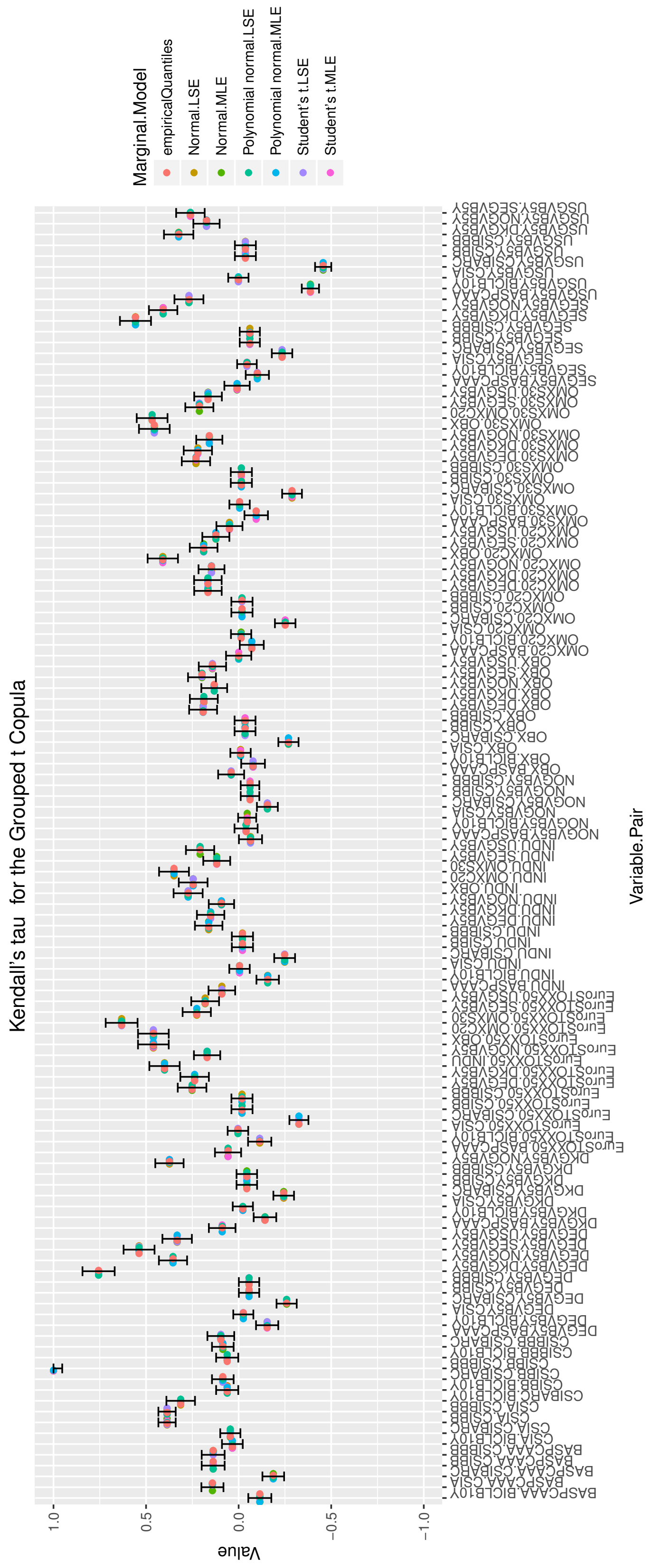


Figure 5: Kendall's tau for Grouped t copula with non-composite methods.

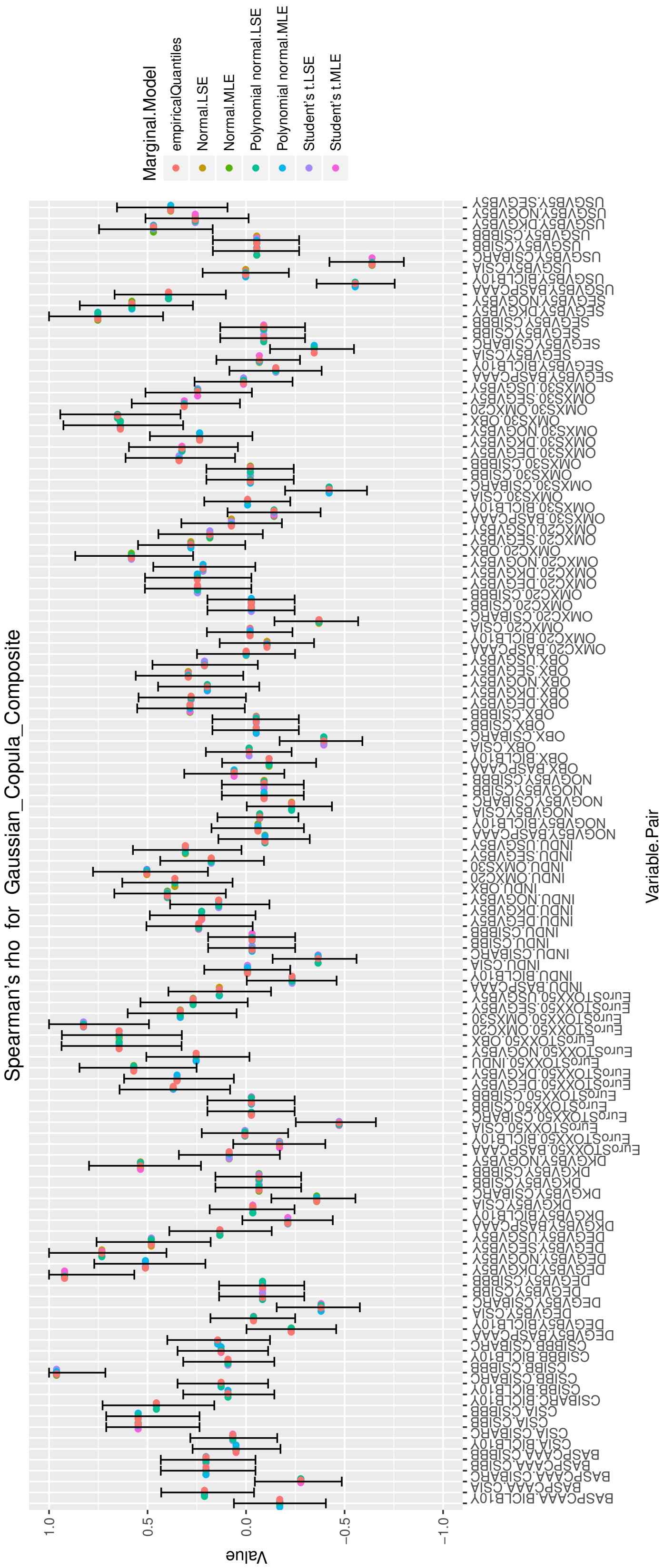


Figure 6: Spearman's rho for Gaussian copula with composite methods.

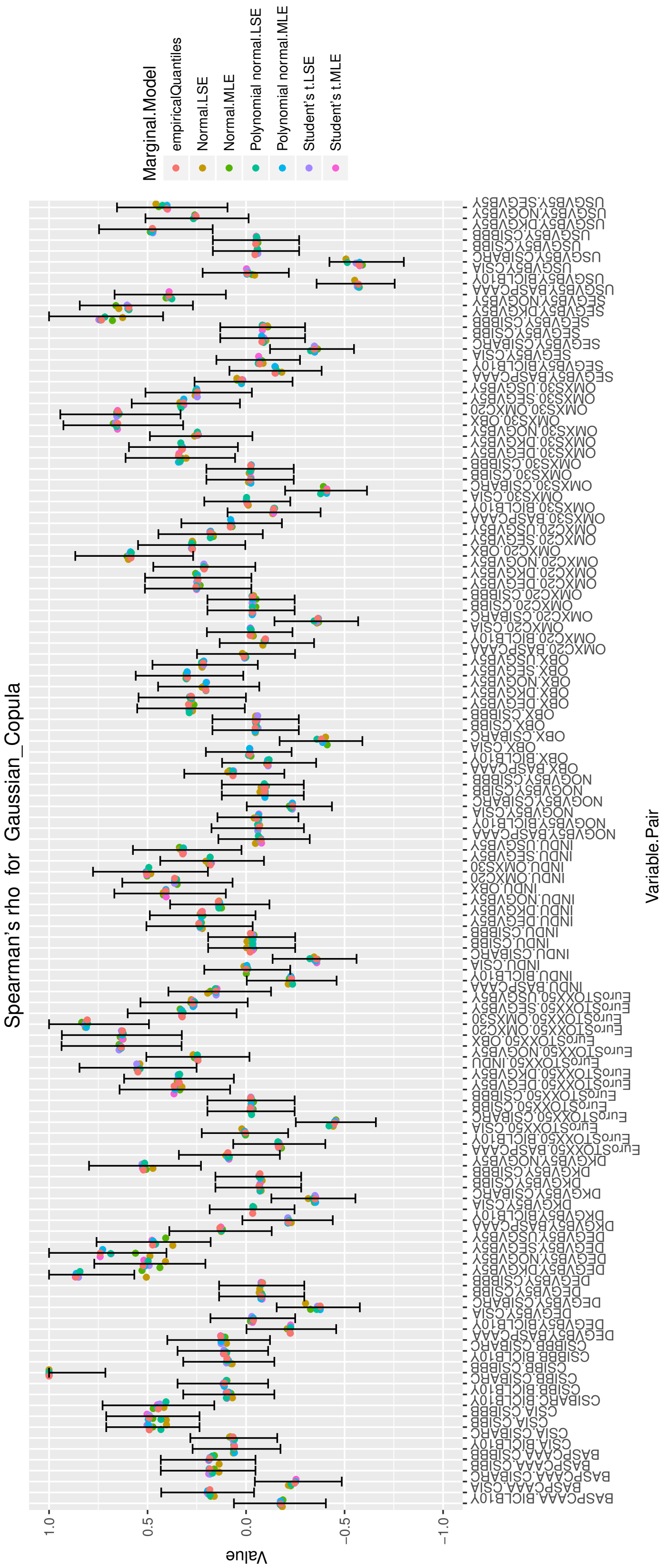


Figure 7: Spearman's rho for Gaussian copula with non-composite methods.

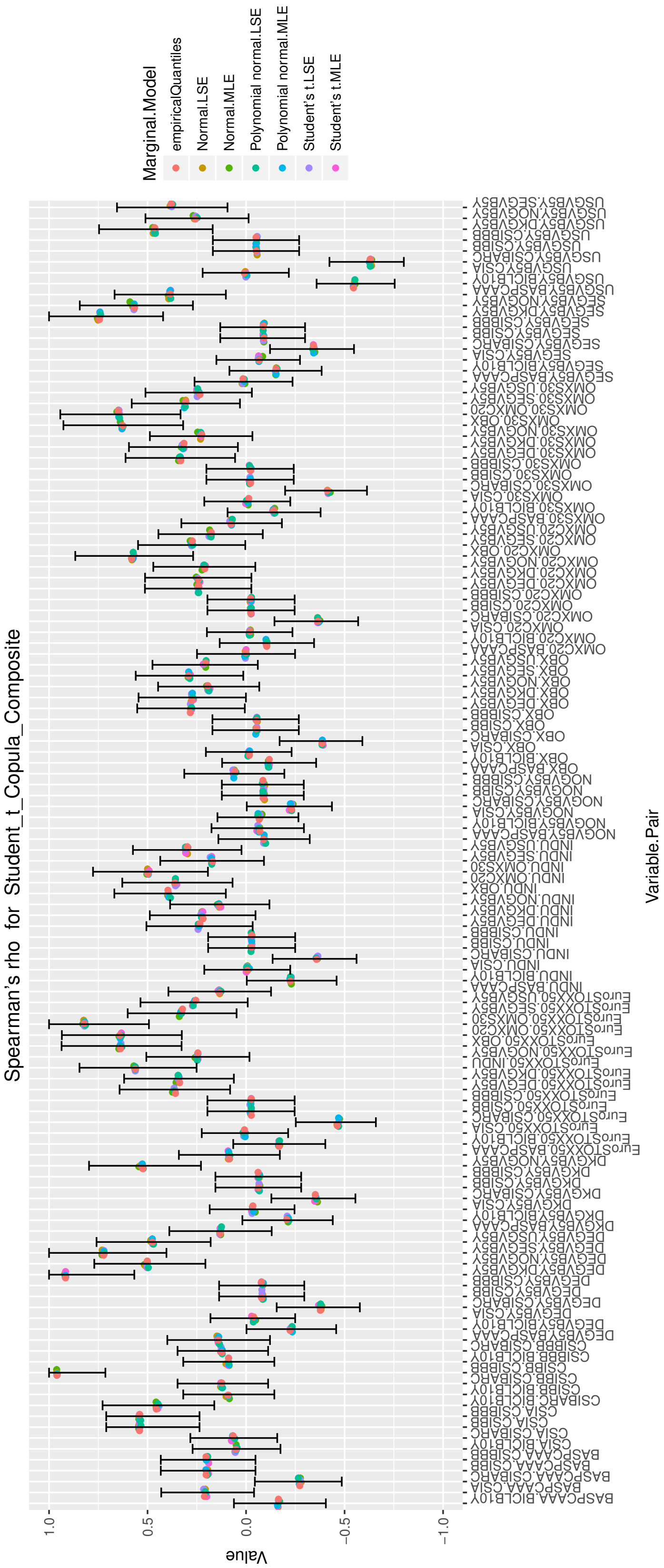


Figure 8: Spearman's rho for Student's t copula with composite methods.

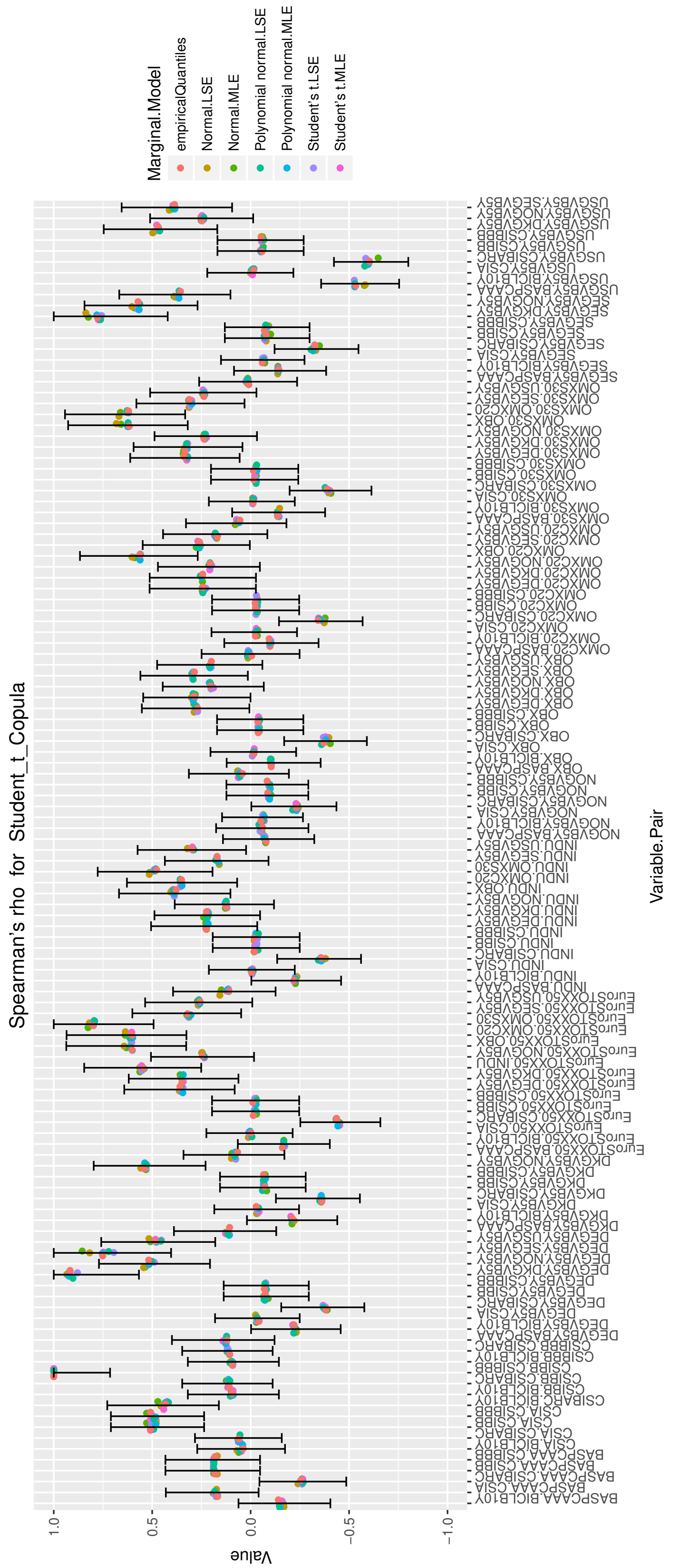


Figure 9: Spearman's rho for Student's $t$-copula with non-composite methods.

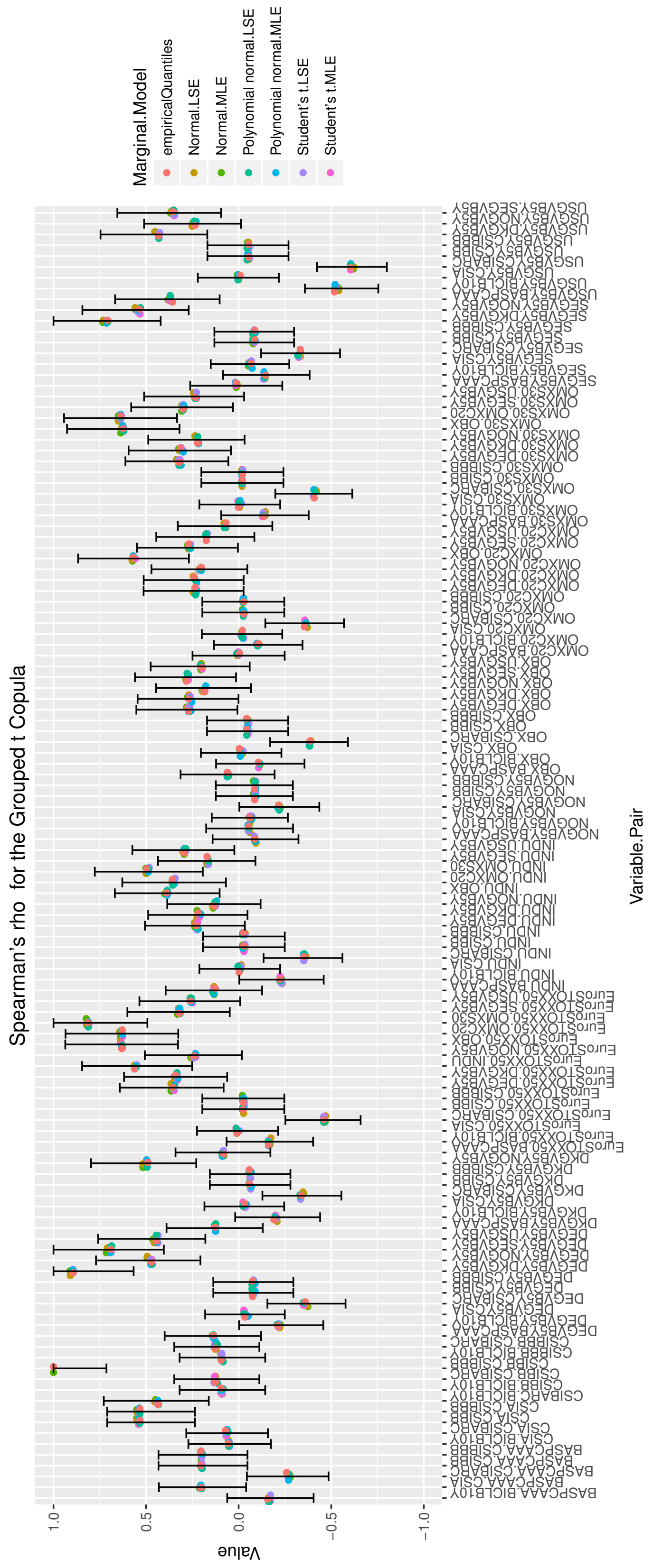


Figure 10: Spearman's rho for Grouped t copula with non-composite methods.

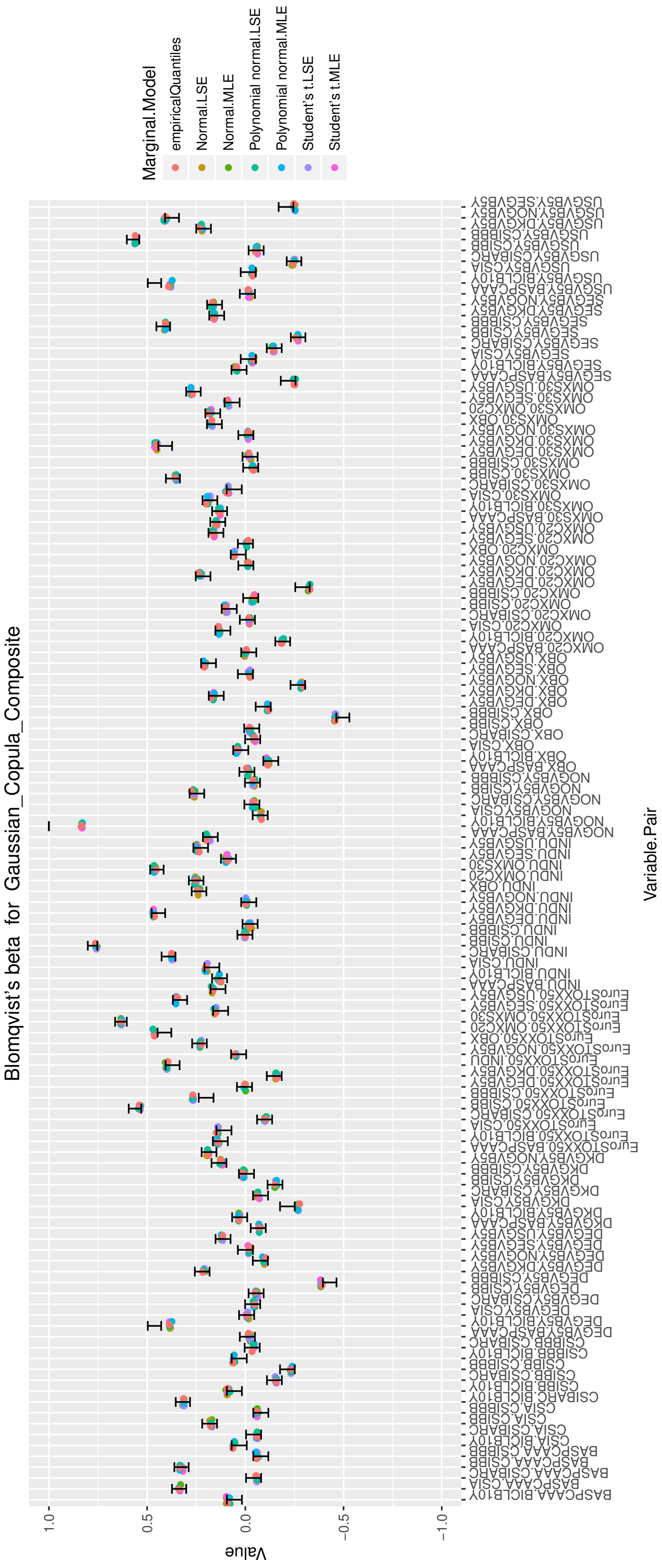


Figure 11: Blomqvist's beta for Gaussian copula with composite methods.

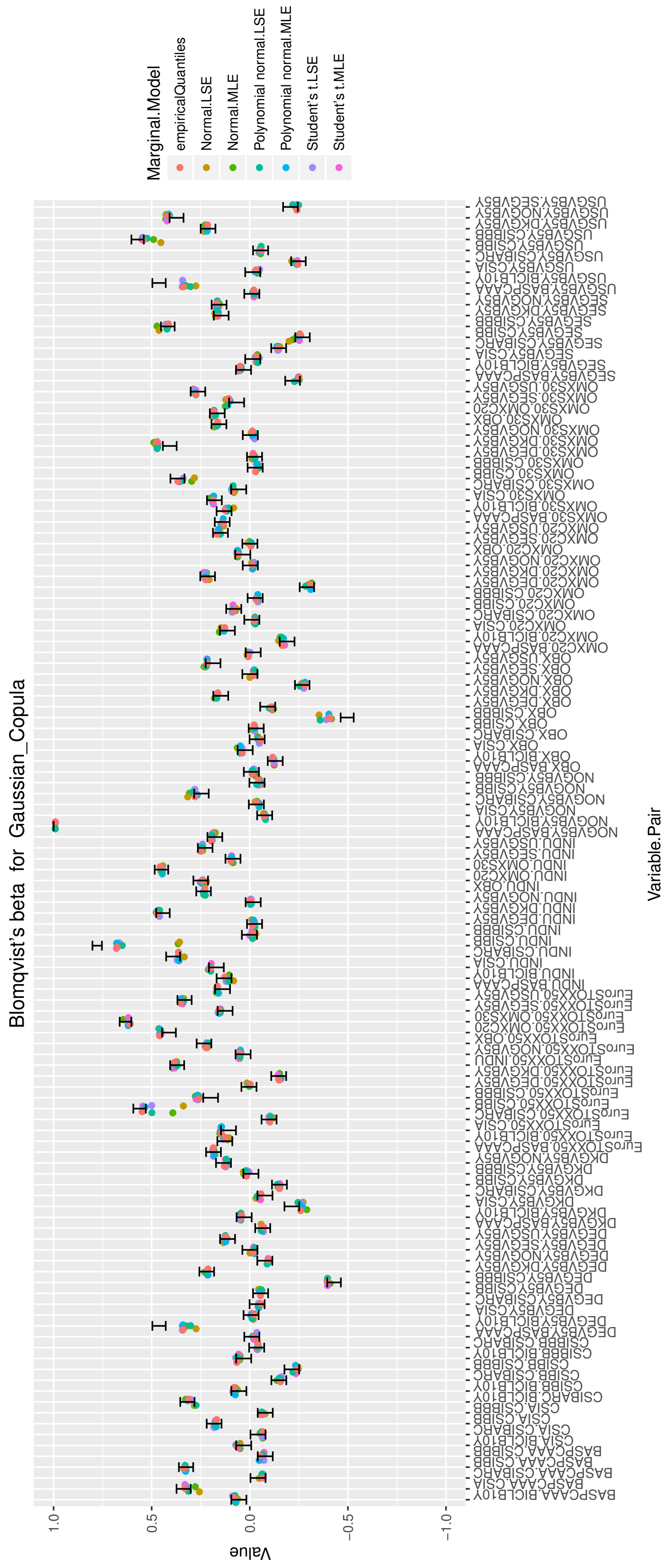


Figure 12: Blomqvist's beta for Gaussian copula with non-composite methods.

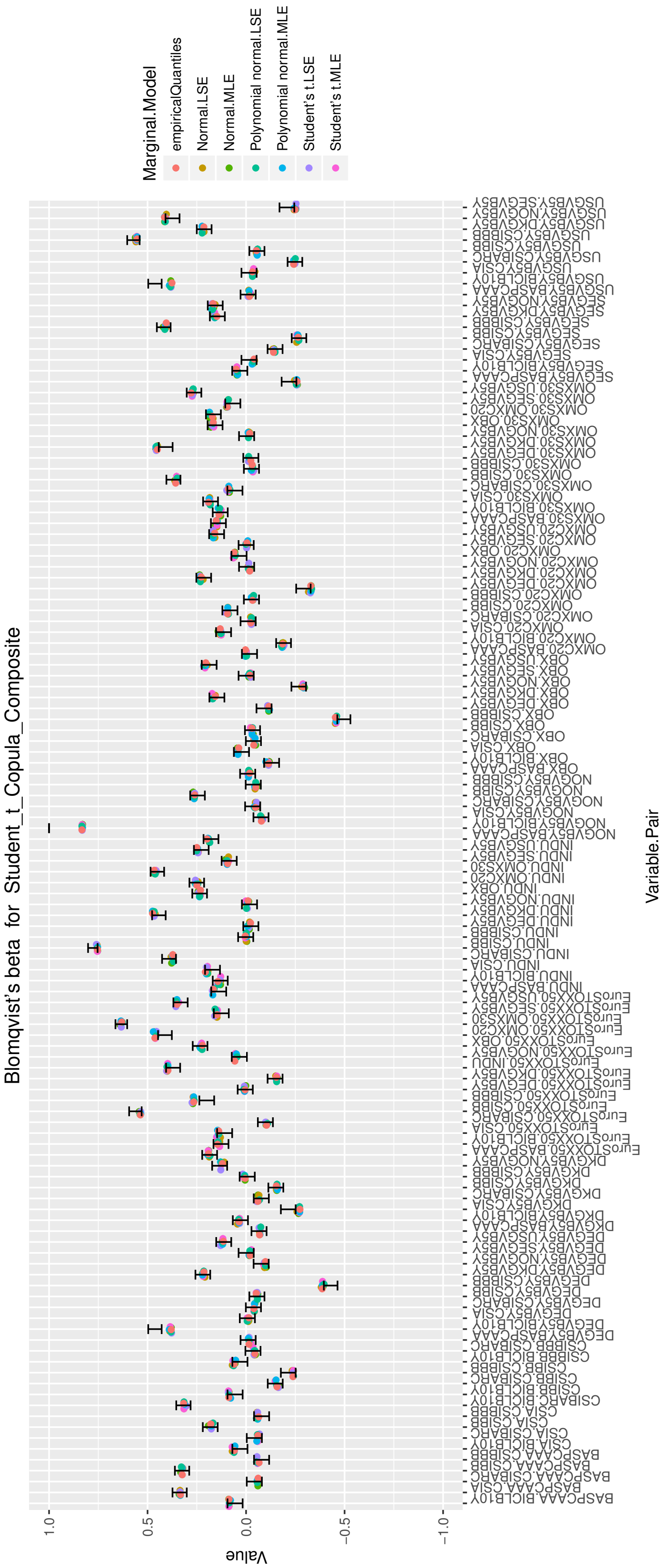


Figure 13: Blomqvist's beta for Student's $t$-copula with composite methods.

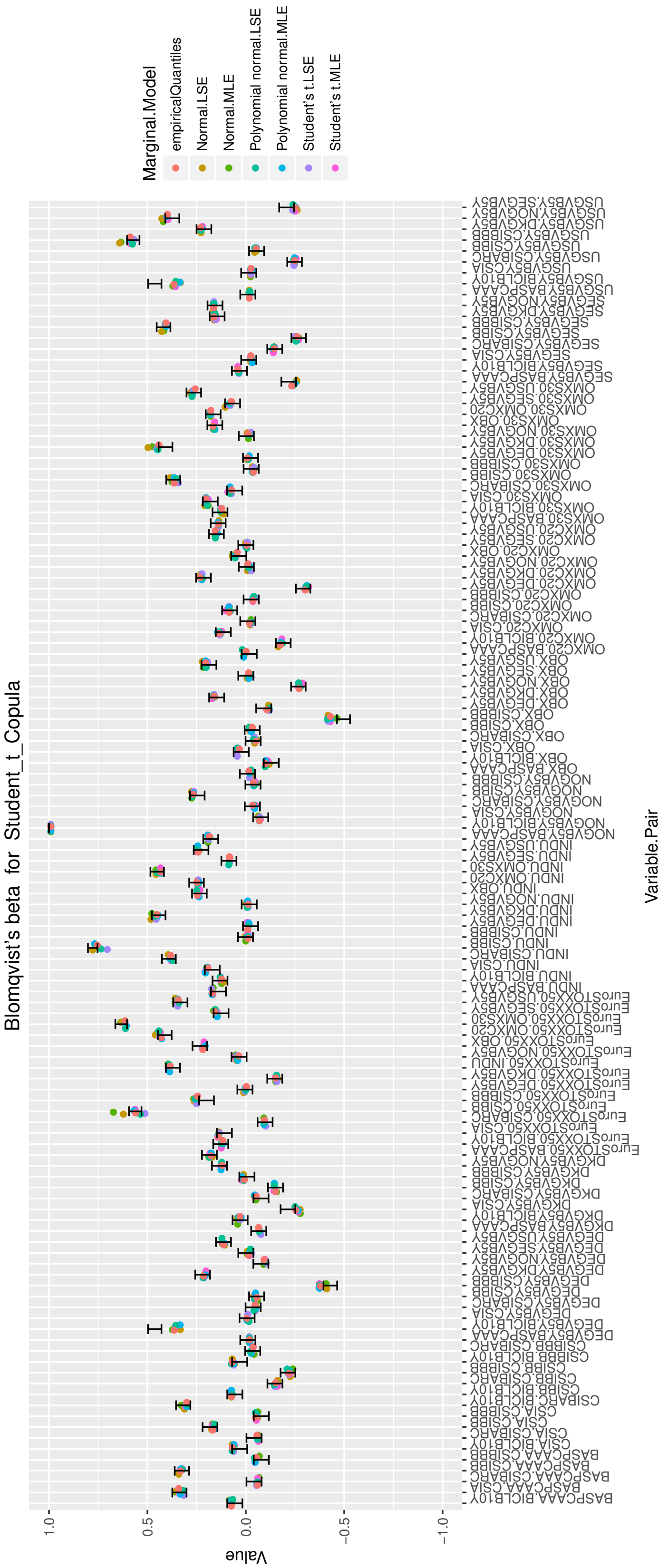


Figure 14: Blomqvist's beta for Student's $t$-copula with non-composite methods.

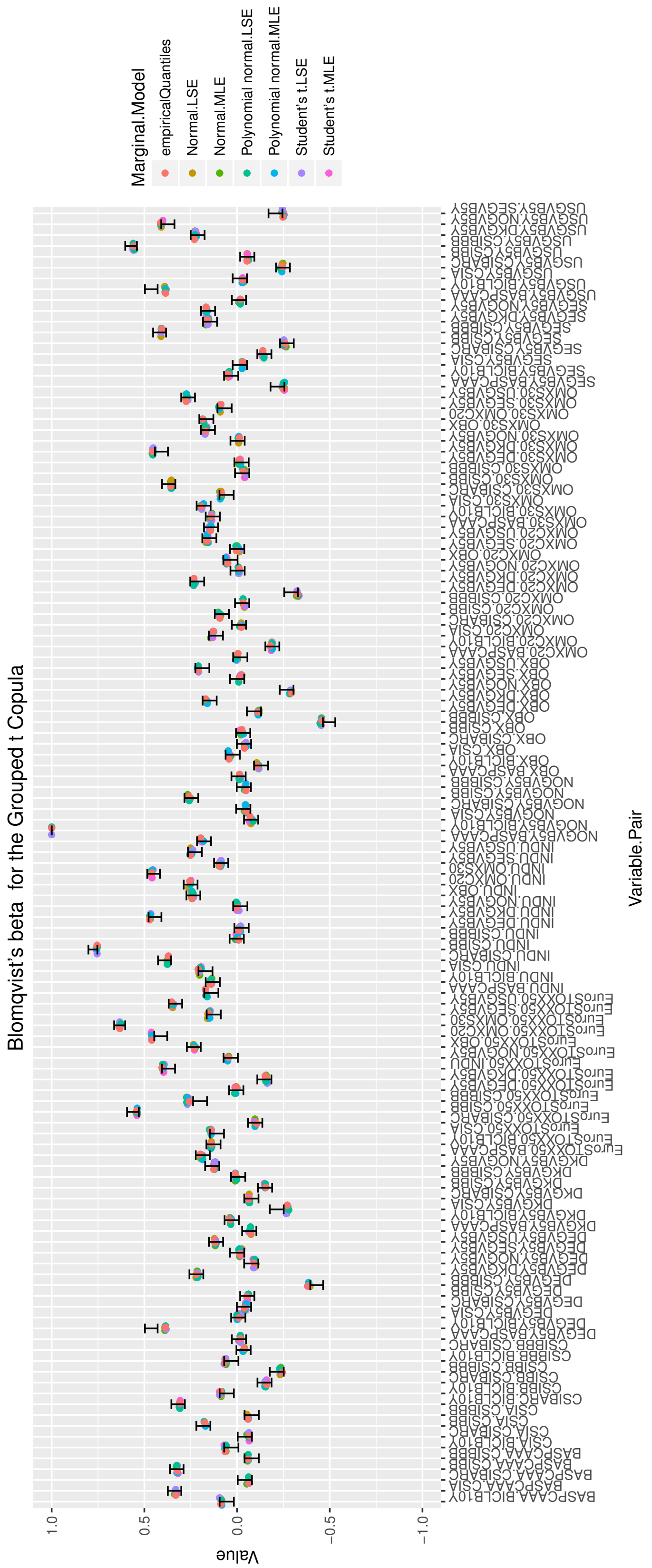


Figure 15: Blomqvist's beta for Grouped t copula with non-composite methods.

## A.2 Tail-weighted Dependence Coefficients

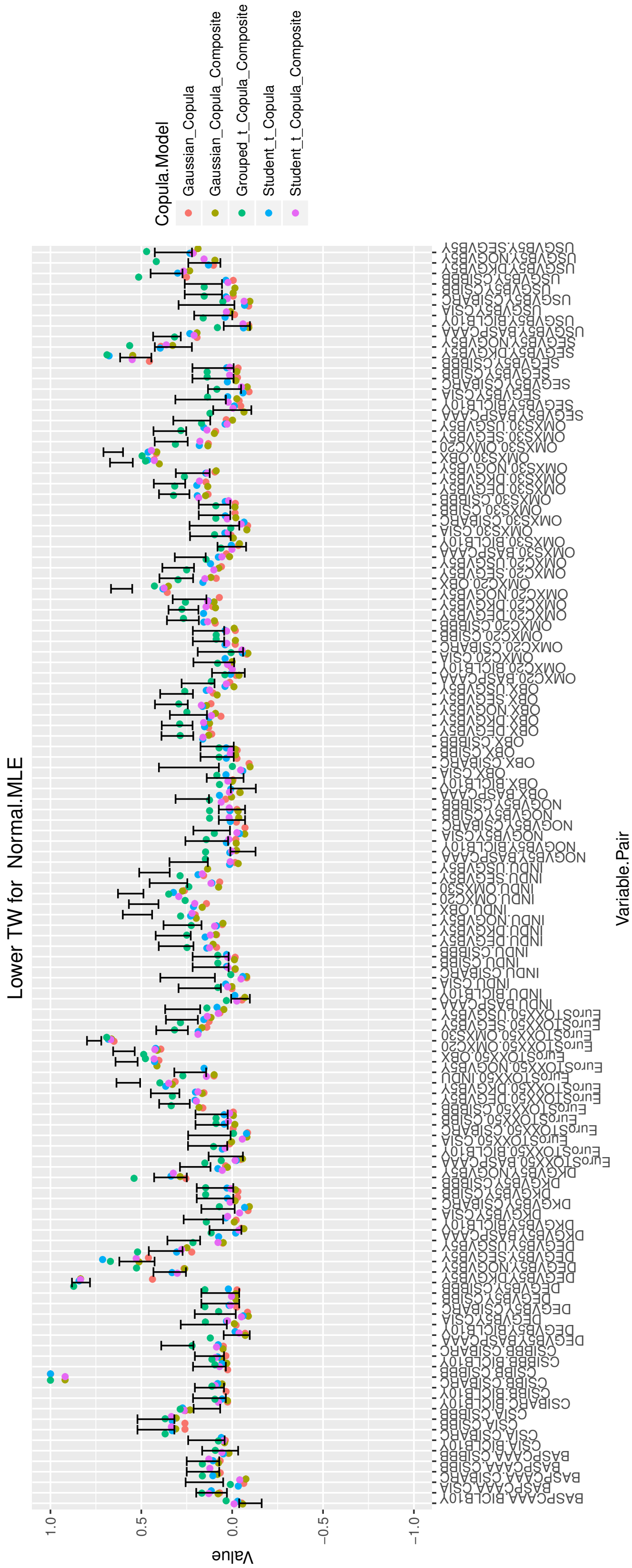


Figure 16: Lower tail-weighted dependence measure for copula models with normal MLE marginals.

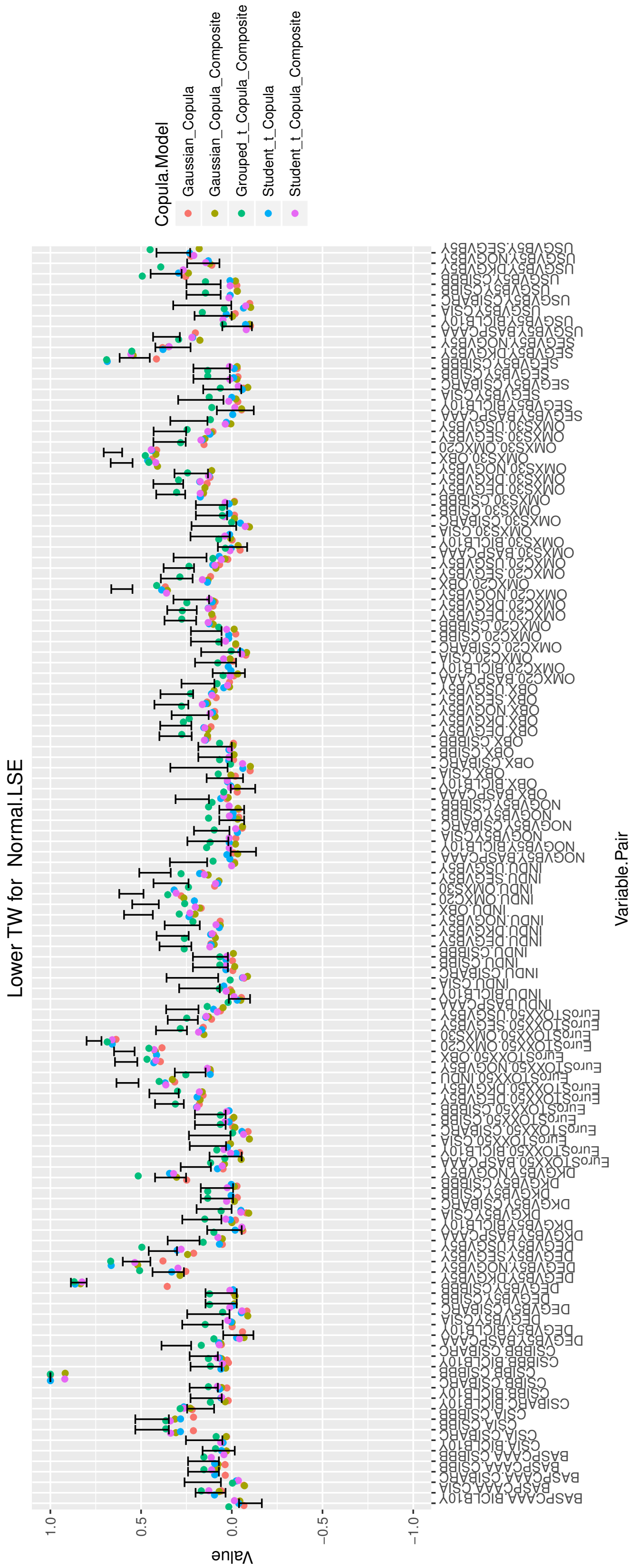


Figure 17: Lower tail-weighted dependence measure for copula models with normal LSE marginals.

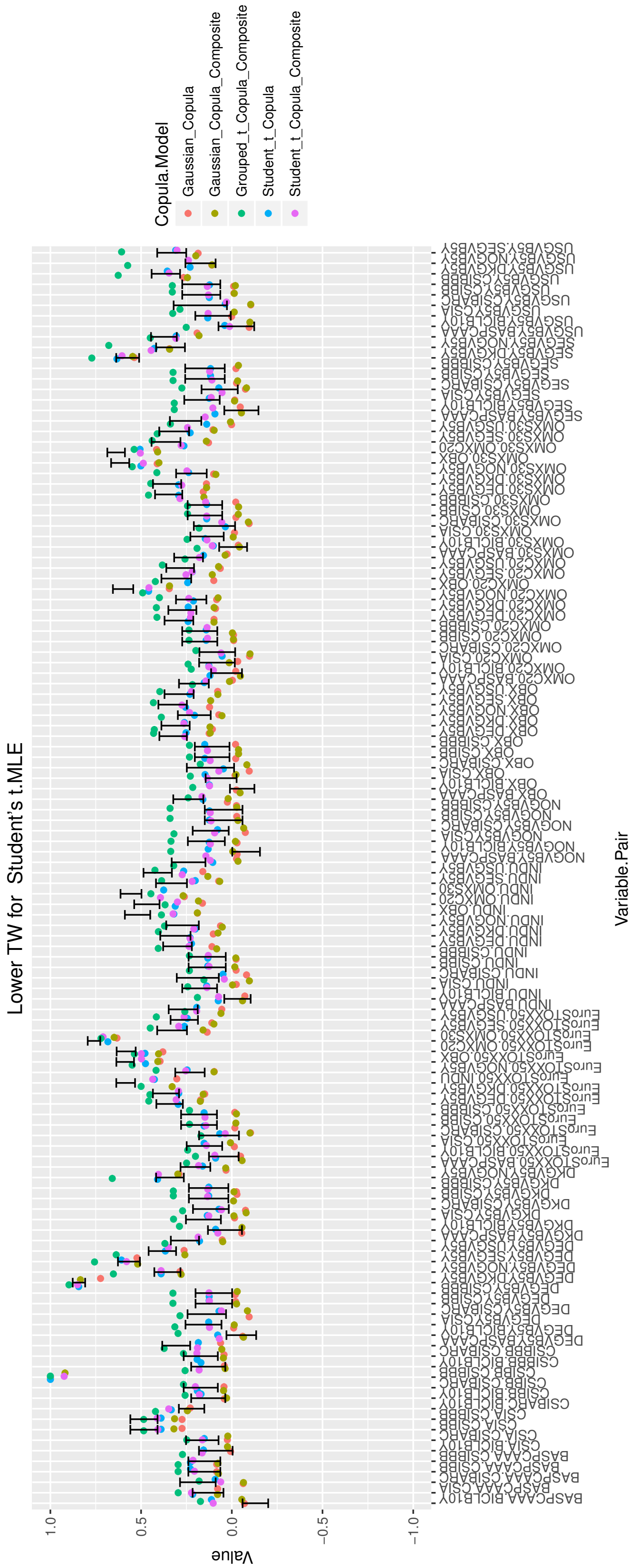


Figure 18: Lower tail-weighted dependence measure for copula models with Student's t MLE marginals.

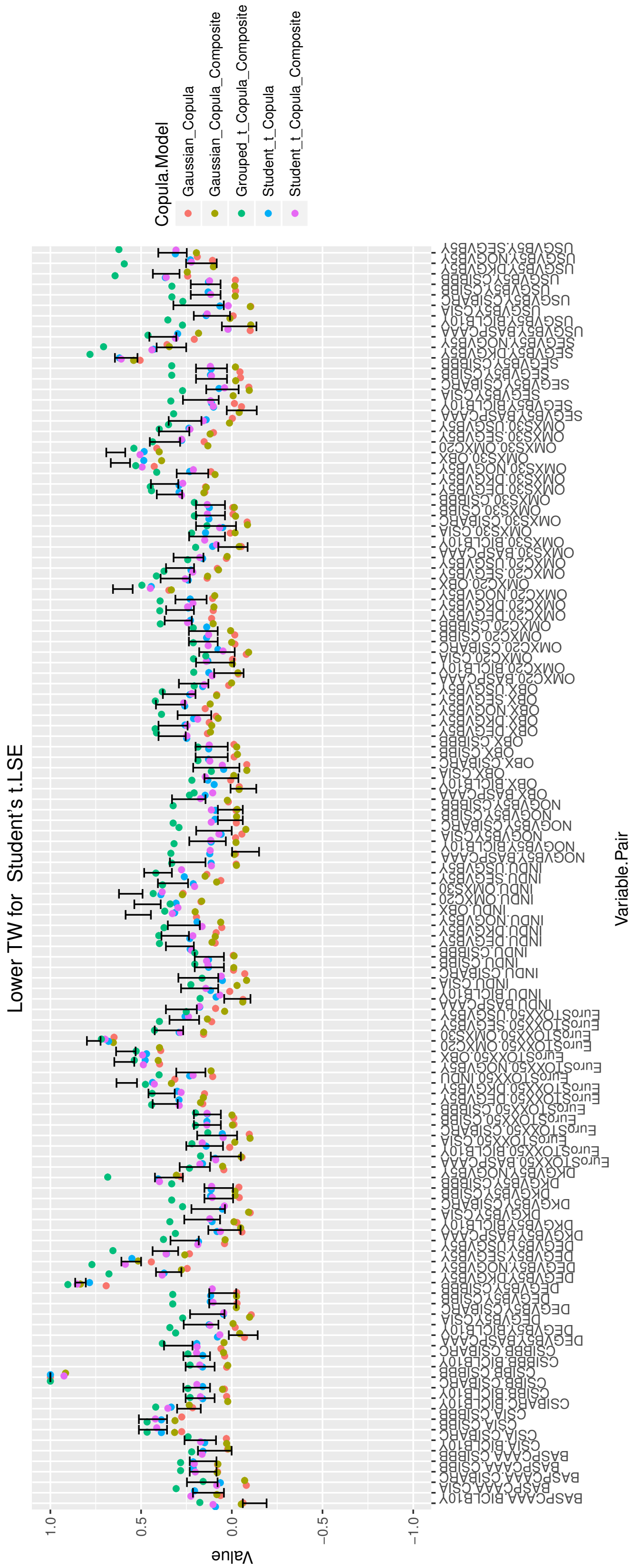


Figure 19: Lower tail-weighted dependence measure for copula models with Student's t LSE marginals.

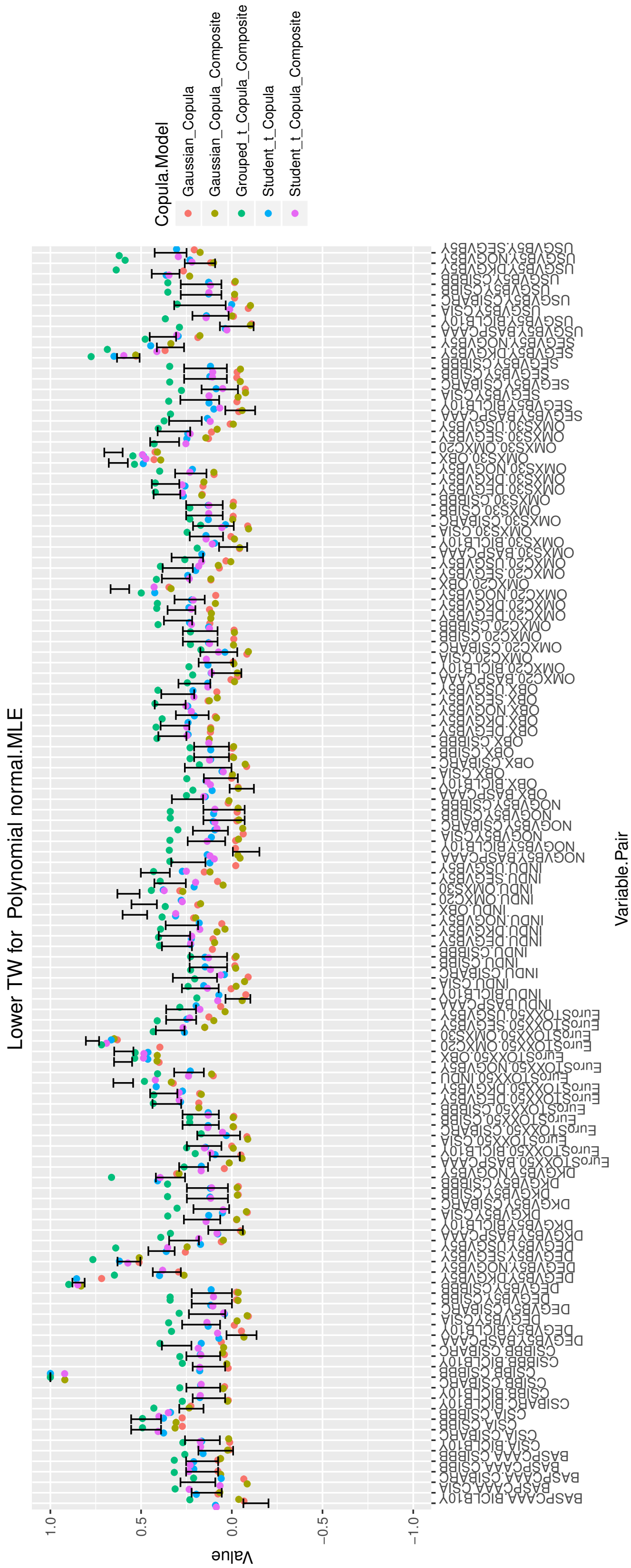


Figure 20: Lower tail-weighted dependence measure for copula models with polynomial normal MLE marginals.

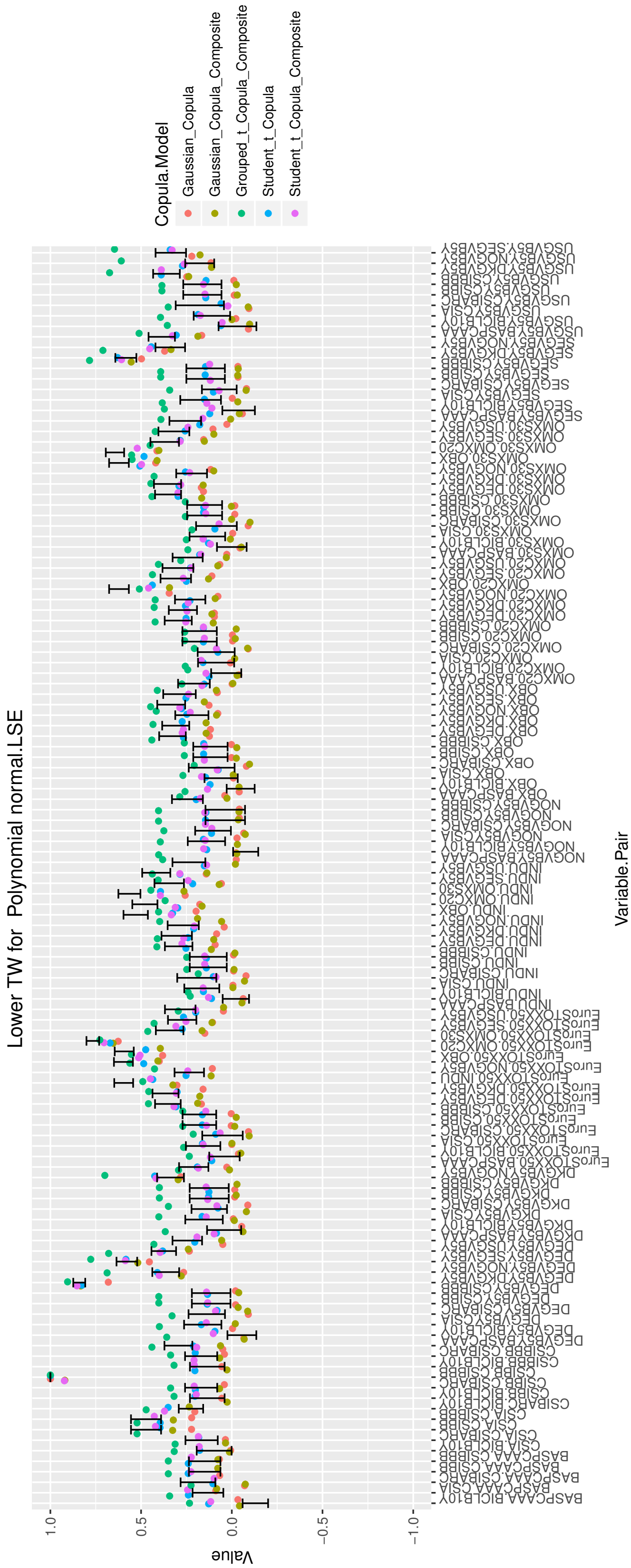


Figure 21: Lower tail-weighted dependence measure for copula models with polynomial normal LSE marginals.

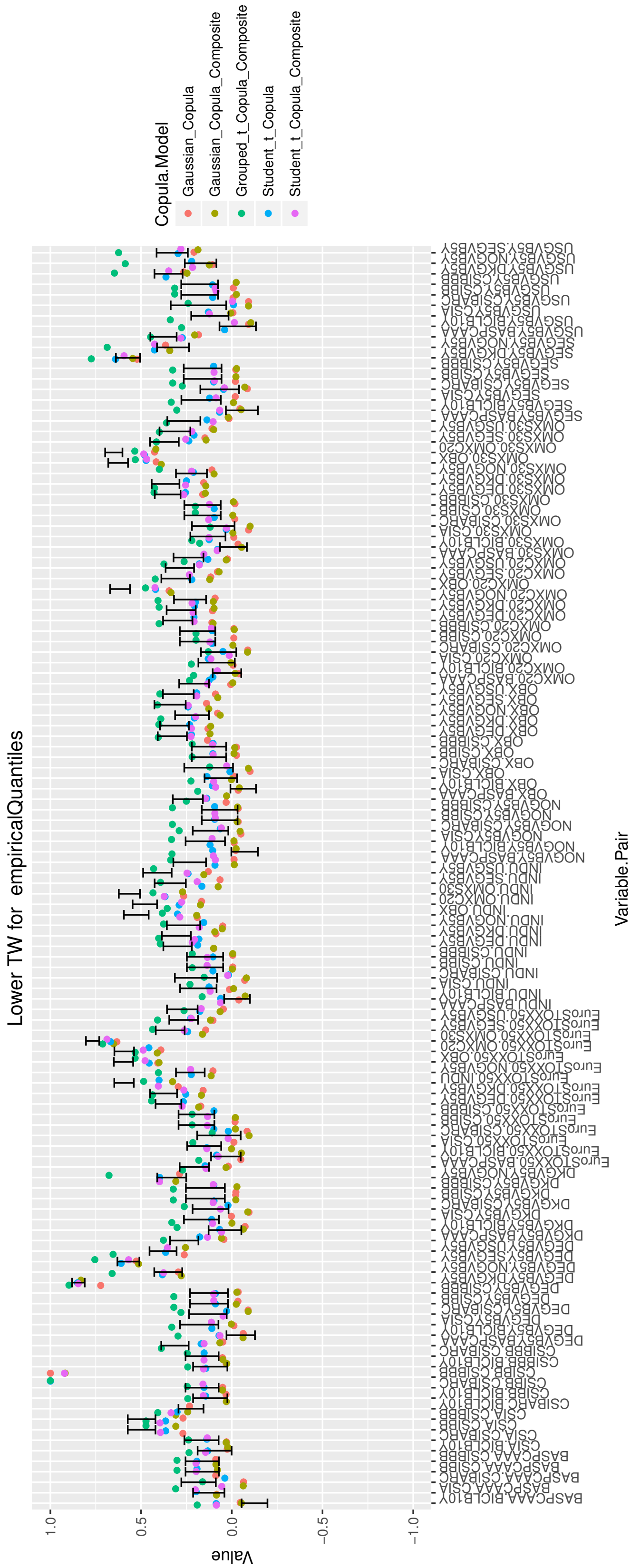

Figure 22: Lower tail-weighted dependence measure for copula models with empirical marginals.

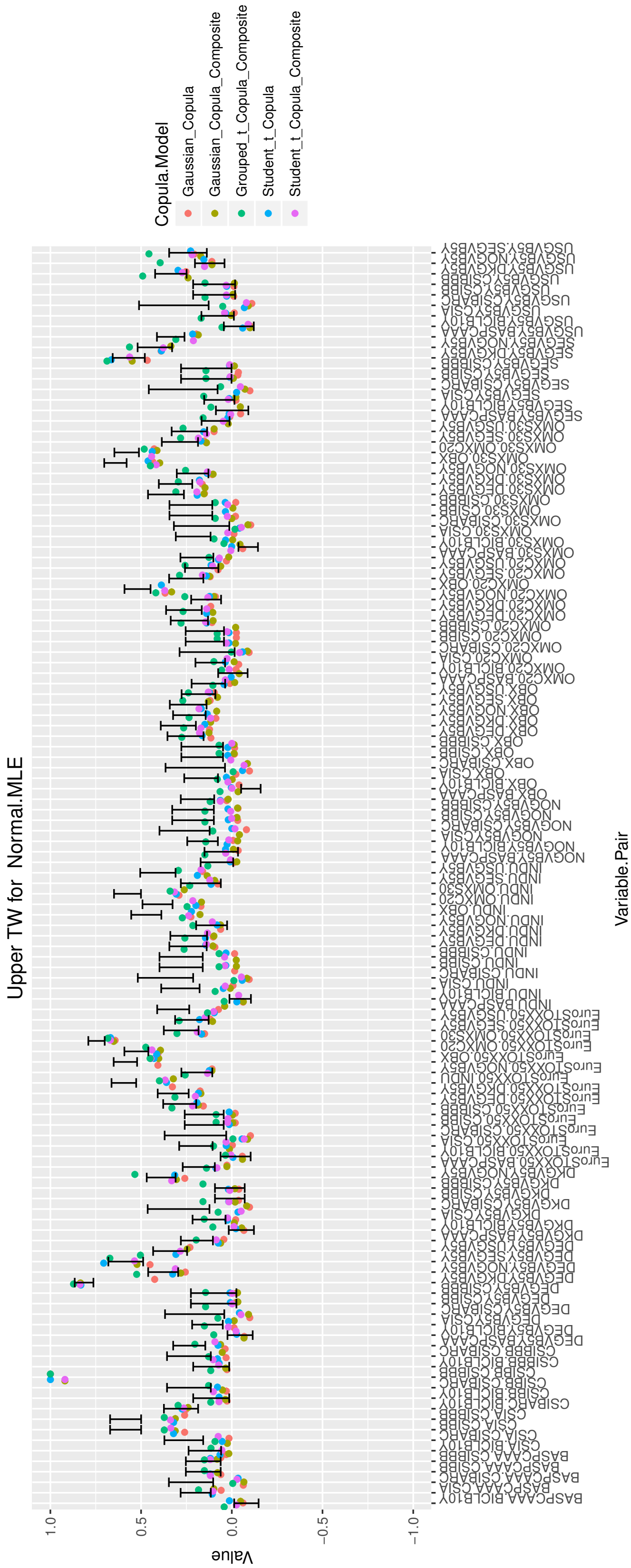


Figure 23: Upper tail-weighted dependence measure for copula models with normal MLE marginals.

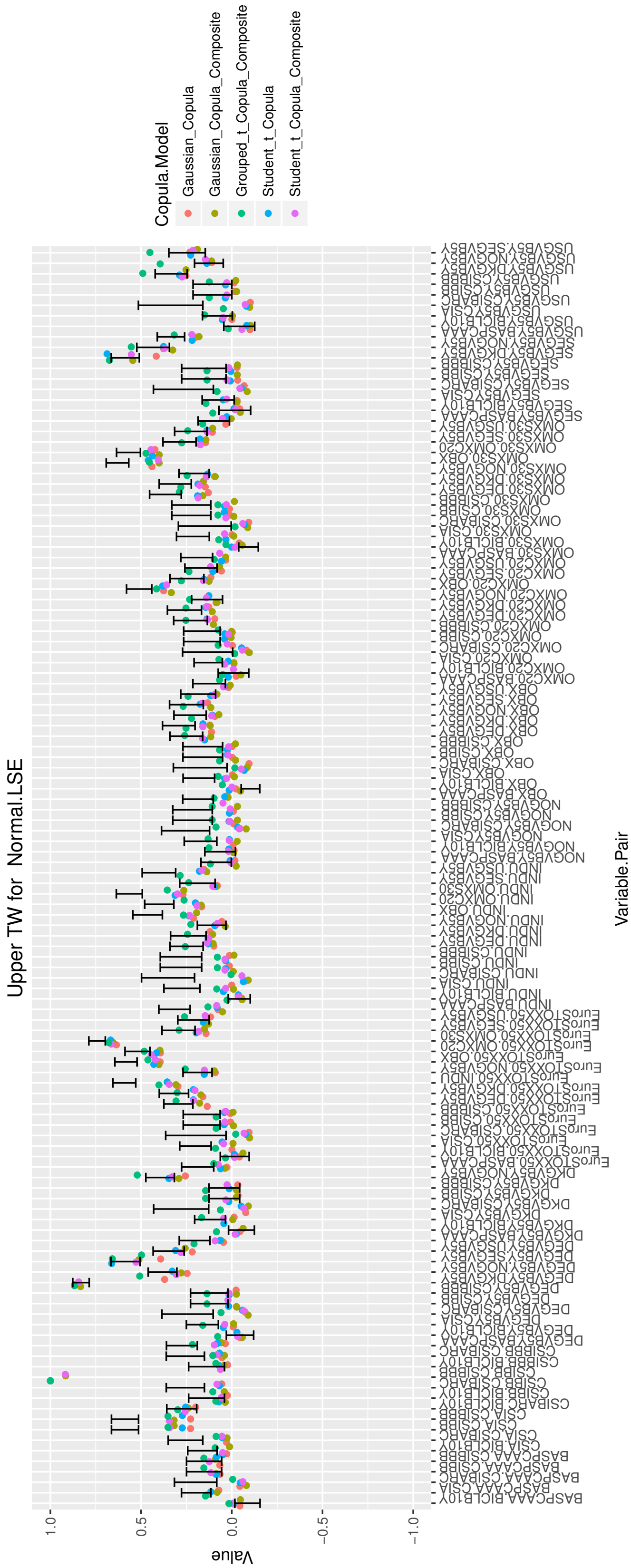


Figure 24: Upper tail-weighted dependence measure for copula models with normal LSE marginals.

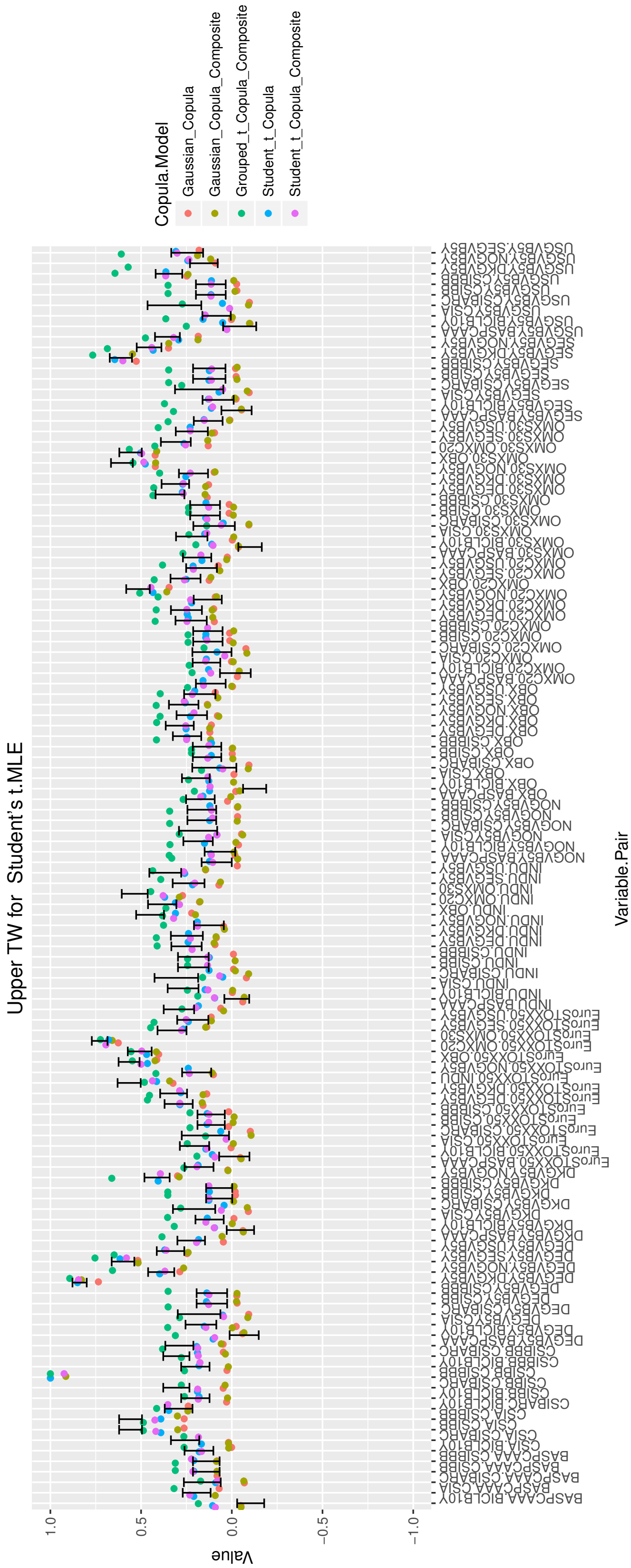


Figure 25: Upper tail-weighted dependence measure for copula models with Student's t MLE marginals.

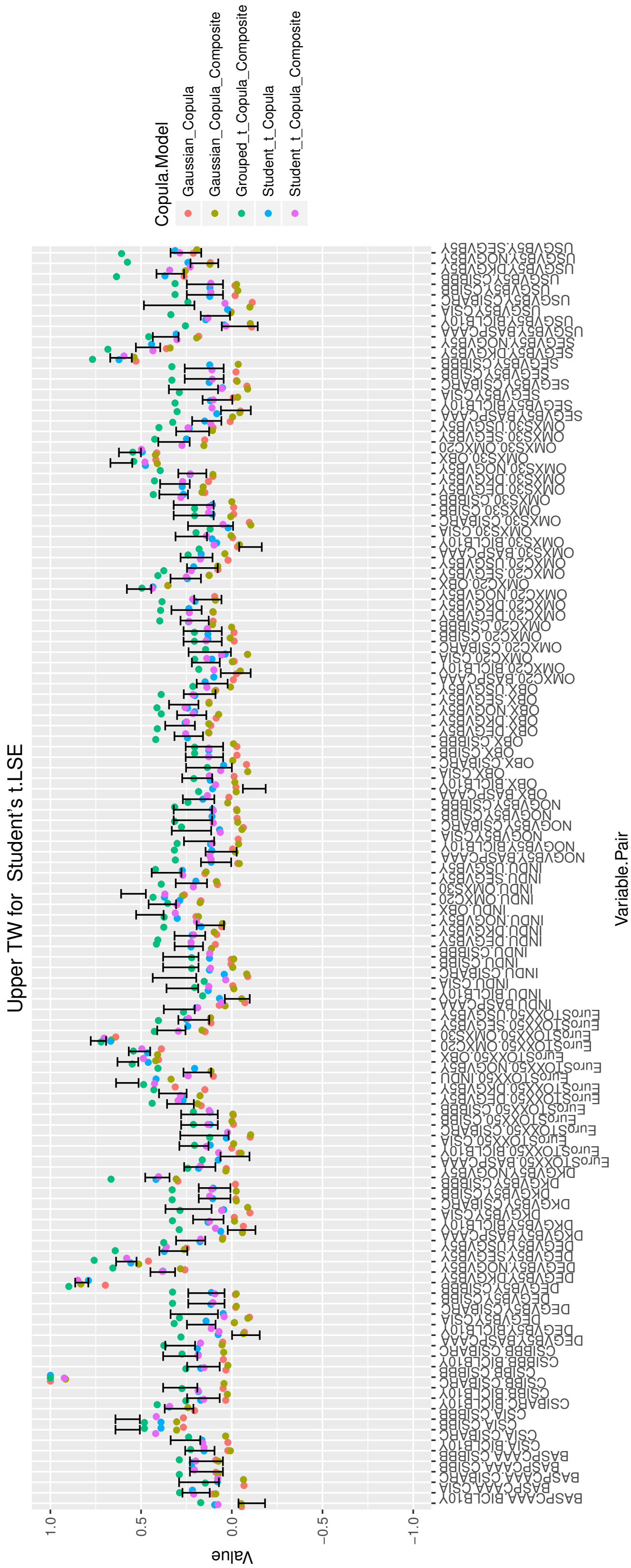


Figure 26: Upper tail-weighted dependence measure for copula models with Student's t LSE marginals.

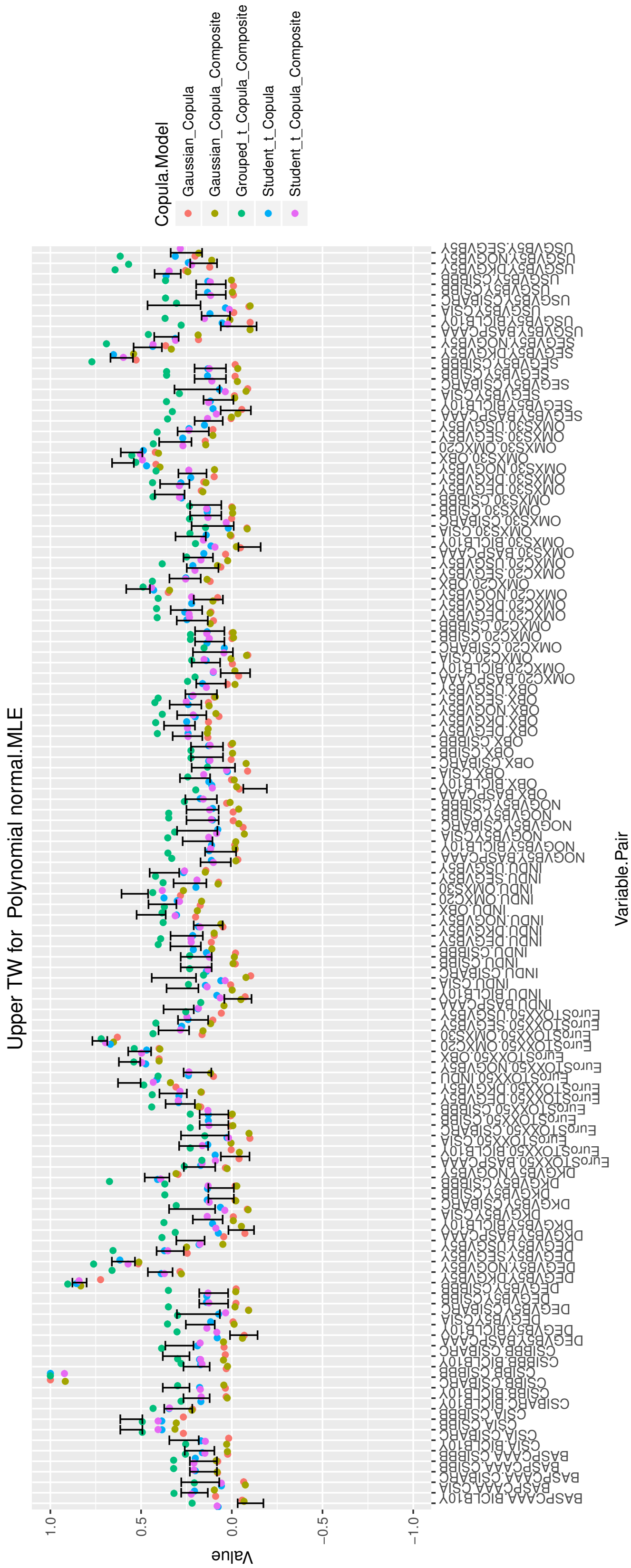


Figure 27: Upper tail-weighted dependence measure for copula models with polynomial normal MLE marginals.

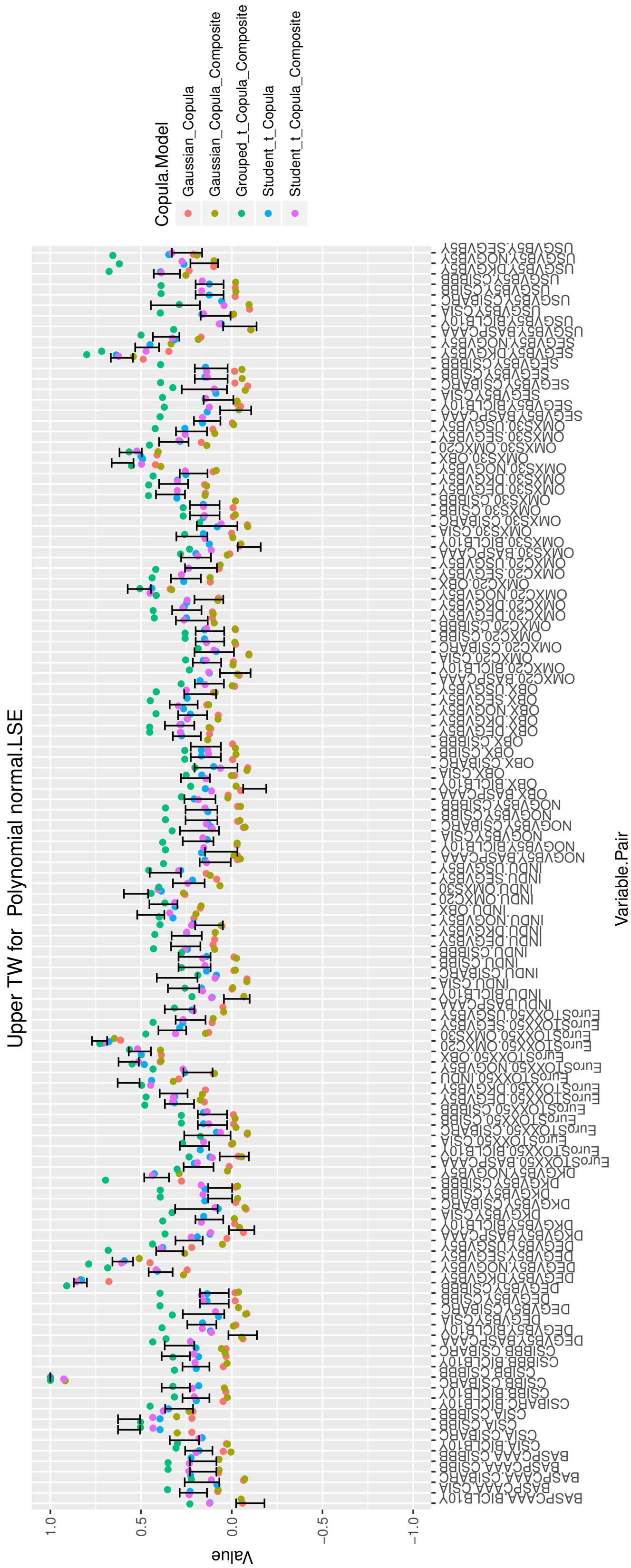


Figure 28: Upper tail-weighted dependence measure for copula models with polynomial normal LSE marginals.

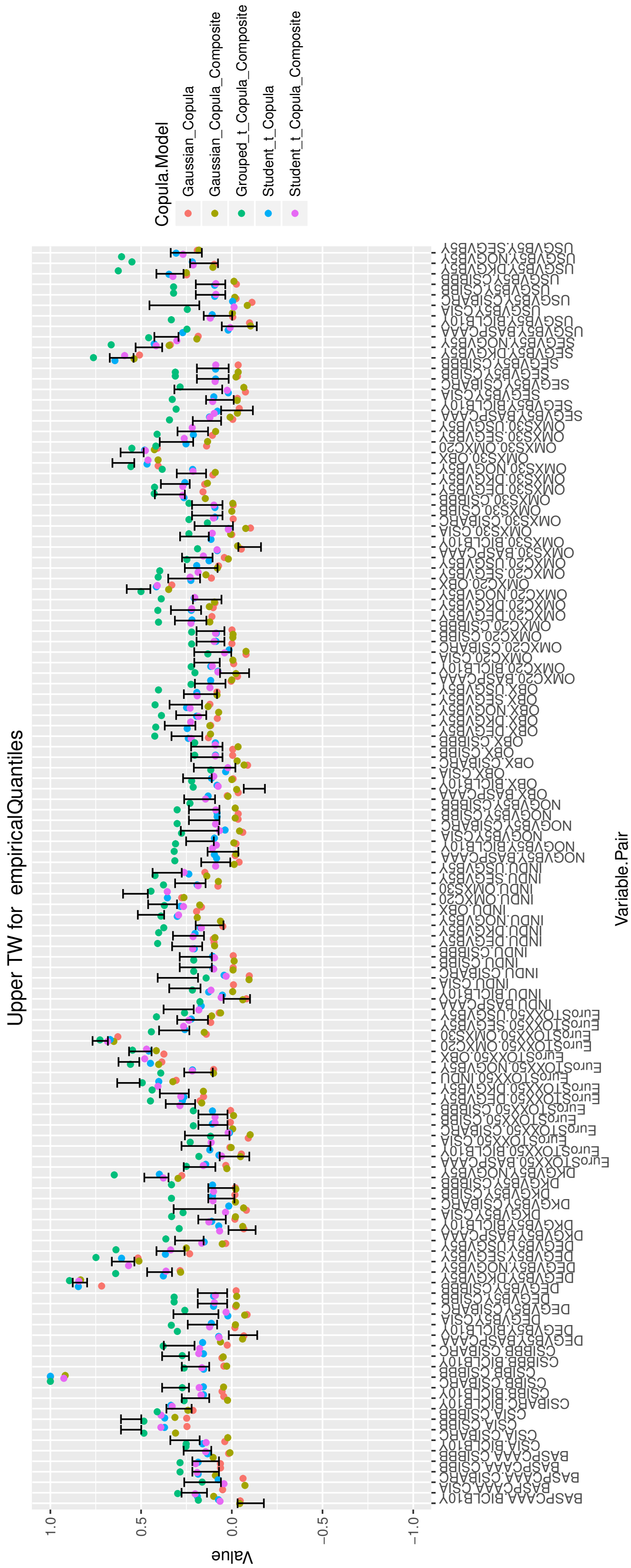


Figure 29: Upper tail-weighted dependence measure for copula models with empirical marginals.

# B Dependence Coefficients

Table B.1: Kendall's tau for the data sample of macroeconomic variables.

| **Variable Pair** | **Lower Bound** | **Upper Bound** | **Variable Pair** | **Lower Bound** | **Upper Bound** | **Variable Pair** | **Lower Bound** | **Upper Bound** |
|---|---|---|---|---|---|---|---|---|
| EuroSTOXX50.INDU | 0.318 | 0.482 | OBX.NOGVB5Y | 0.062 | 0.202 | OMXS30.CSIBBB | -0.0717 | 0.0425 |
| EuroSTOXX50.OMXS30 | 0.546 | 0.718 | DEGVB5Y.NOGVB5Y | 0.278 | 0.431 | OMXC20.CSIBBB | -0.0752 | 0.0394 |
| INDU.OMXS30 | 0.268 | 0.43 | USGVB5Y.NOGVB5Y | 0.103 | 0.245 | OBX.CSIBBB | -0.0911 | 0.0219 |
| EuroSTOXX50.OMXC20 | 0.377 | 0.544 | SEGVB5Y.NOGVB5Y | 0.331 | 0.485 | DEGVB5Y.CSIBBB | -0.111 | -0.00122 |
| INDU.OMXC20 | 0.168 | 0.323 | DKGVB5Y.NOGVB5Y | 0.297 | 0.45 | USGVB5Y.CSIBBB | -0.0931 | 0.0202 |
| OMXS30.OMXC20 | 0.384 | 0.551 | EuroSTOXX50.BASPCAAA | -0.0134 | 0.128 | SEGVB5Y.CSIBBB | -0.115 | -0.00589 |
| EuroSTOXX50.OBX | 0.378 | 0.544 | INDU.BASPCAAA | 0.0186 | 0.163 | DKGVB5Y.CSIBBB | -0.0996 | 0.0114 |
| INDU.OBX | 0.194 | 0.351 | OMXS30.BASPCAAA | -0.0208 | 0.12 | NOGVB5Y.CSIBBB | -0.112 | -0.0113 |
| OMXS30.OBX | 0.372 | 0.538 | OMXC20.BASPCAAA | -0.0685 | 0.0681 | BASPCAAA.CSIBBB | 0.0756 | 0.2 |
| OMXC20.OBX | 0.328 | 0.492 | OBX.BASPCAAA | -0.0296 | 0.111 | CSIA.CSIBBB | 0.34 | 0.433 |
| EuroSTOXX50.DEGVB5Y | 0.174 | 0.329 | DEGVB5Y.BASPCAAA | 0.0239 | 0.169 | CSIBB.CSIBBB | 0.953 | 1 |
| INDU.DEGVB5Y | 0.0874 | 0.236 | USGVB5Y.BASPCAAA | 0.19 | 0.347 | EuroSTOXX50.CSIBARC | -0.377 | -0.274 |
| OMXS30.DEGVB5Y | 0.154 | 0.307 | SEGVB5Y.BASPCAAA | -0.0599 | 0.077 | INDU.CSIBARC | -0.305 | -0.192 |
| OMXC20.DEGVB5Y | 0.0914 | 0.24 | DKGVB5Y.BASPCAAA | 0.0168 | 0.161 | OMXS30.CSIBARC | -0.342 | -0.235 |
| OBX.DEGVB5Y | 0.116 | 0.267 | NOGVB5Y.BASPCAAA | -0.126 | -0.00156 | OMXC20.CSIBARC | -0.307 | -0.196 |
| EuroSTOXX50.USGVB5Y | 0.106 | 0.256 | EuroSTOXX50.CSIA | -0.0518 | 0.0593 | OBX.CSIBARC | -0.324 | -0.215 |
| INDU.USGVB5Y | 0.132 | 0.285 | INDU.CSIA | -0.0601 | 0.0505 | DEGVB5Y.CSIBARC | -0.314 | -0.205 |
| OMXS30.USGVB5Y | 0.0906 | 0.24 | OMXS30.CSIA | -0.0602 | 0.05 | USGVB5Y.CSIBARC | -0.5 | -0.413 |
| OMXC20.USGVB5Y | 0.05 | 0.196 | OMXC20.CSIA | -0.0682 | 0.0414 | SEGVB5Y.CSIBARC | -0.29 | -0.179 |
| OBX.USGVB5Y | 0.0678 | 0.215 | OBX.CSIA | -0.0652 | 0.0448 | DKGVB5Y.CSIBARC | -0.299 | -0.188 |
| DEGVB5Y.USGVB5Y | 0.253 | 0.412 | DEGVB5Y.CSIA | -0.0796 | 0.029 | NOGVB5Y.CSIBARC | -0.212 | -0.0988 |
| EuroSTOXX50.SEGVB5Y | 0.15 | 0.303 | USGVB5Y.CSIA | -0.0535 | 0.0562 | BASPCAAA.CSIBARC | -0.246 | -0.128 |
| INDU.SEGVB5Y | 0.0456 | 0.191 | SEGVB5Y.CSIA | -0.0979 | 0.00805 | CSIA.CSIBARC | -0.00955 | 0.0994 |
| OMXS30.SEGVB5Y | 0.136 | 0.288 | DKGVB5Y.CSIA | -0.0772 | 0.0318 | CSIBB.CSIBARC | 0.027 | 0.144 |
| OMXC20.SEGVB5Y | 0.114 | 0.264 | NOGVB5Y.CSIA | -0.095 | 0.00276 | CSIBBB.CSIBARC | 0.027 | 0.144 |
| OBX.SEGVB5Y | 0.123 | 0.273 | BASPCAAA.CSIA | 0.0819 | 0.201 | EuroSTOXX50.BICLB10Y | -0.177 | -0.0509 |
| DEGVB5Y.SEGVB5Y | 0.454 | 0.621 | EuroSTOXX50.CSIBB | -0.0751 | 0.0389 | INDU.BICLB10Y | -0.218 | -0.0962 |
| USGVB5Y.SEGVB5Y | 0.183 | 0.338 | INDU.CSIBB | -0.0777 | 0.0367 | OMXS30.BICLB10Y | -0.158 | -0.0315 |
| EuroSTOXX50.DKGVB5Y | 0.161 | 0.315 | OMXS30.CSIBB | -0.0717 | 0.0425 | OMXC20.BICLB10Y | -0.136 | -0.00619 |
| INDU.DKGVB5Y | 0.0772 | 0.225 | OMXC20.CSIBB | -0.0752 | 0.0394 | OBX.BICLB10Y | -0.142 | -0.0138 |
| OMXS30.DKGVB5Y | 0.144 | 0.297 | OBX.CSIBB | -0.0911 | 0.0219 | DEGVB5Y.BICLB10Y | -0.214 | -0.0939 |
| OMXC20.DKGVB5Y | 0.0918 | 0.241 | DEGVB5Y.CSIBB | -0.111 | -0.00122 | USGVB5Y.BICLB10Y | -0.434 | -0.341 |
| OBX.DKGVB5Y | 0.113 | 0.263 | USGVB5Y.CSIBB | -0.0931 | 0.0202 | SEGVB5Y.BICLB10Y | -0.164 | -0.0382 |
| DEGVB5Y.DKGVB5Y | 0.67 | 0.843 | SEGVB5Y.CSIBB | -0.115 | -0.00589 | DKGVB5Y.BICLB10Y | -0.203 | -0.0812 |
| USGVB5Y.DKGVB5Y | 0.245 | 0.403 | DKGVB5Y.CSIBB | -0.0996 | 0.0114 | NOGVB5Y.BICLB10Y | -0.103 | 0.0223 |
| SEGVB5Y.DKGVB5Y | 0.473 | 0.641 | NOGVB5Y.CSIBB | -0.112 | -0.0113 | BASPCAAA.BICLB10Y | -0.177 | -0.0522 |
| EuroSTOXX50.NOGVB5Y | 0.0994 | 0.241 | BASPCAAA.CSIBB | 0.0756 | 0.2 | CSIA.BICLB10Y | -0.0223 | 0.0901 |
| INDU.NOGVB5Y | 0.0244 | 0.161 | CSIA.CSIBB | 0.34 | 0.433 | CSIBB.BICLB10Y | 0.00203 | 0.122 |
| OMXS30.NOGVB5Y | 0.0878 | 0.229 | EuroSTOXX50.CSIBBB | -0.0751 | 0.0389 | CSIBBB.BICLB10Y | 0.00203 | 0.122 |
| OMXC20.NOGVB5Y | 0.0765 | 0.217 | INDU.CSIBBB | -0.0777 | 0.0367 | CSIBARC.BICLB10Y | 0.236 | 0.39 |

Table B.2: Spearman's rho for the data sample of macroeconomic variables.

| Variable Pair | Lower Bound | Upper Bound | Variable Pair | Lower Bound | Upper Bound | Variable Pair | Lower Bound | Upper Bound |
|---|---|---|---|---|---|---|---|---|
| EuroSTOXX50.INDU | 0.251 | 0.845 | OBX.NOGVB5Y | -0.0669 | 0.447 | OMXS30.CSIBBB | -0.242 | 0.201 |
| EuroSTOXX50.OMXS30 | 0.493 | 1 | DEGVB5Y.NOGVB5Y | 0.206 | 0.77 | OMXC20.CSIBBB | -0.247 | 0.196 |
| INDU.OMXS30 | 0.194 | 0.777 | USGVB5Y.NOGVB5Y | -0.0129 | 0.511 | OBX.CSIBBB | -0.268 | 0.171 |
| EuroSTOXX50.OMXC20 | 0.326 | 0.935 | SEGVB5Y.NOGVB5Y | 0.27 | 0.844 | DEGVB5Y.CSIBBB | -0.295 | 0.137 |
| INDU.OMXC20 | 0.0685 | 0.628 | DKGVB5Y.NOGVB5Y | 0.229 | 0.797 | USGVB5Y.CSIBBB | -0.269 | 0.169 |
| OMXS30.OMXC20 | 0.333 | 0.943 | EuroSTOXX50.BASPCAAA | -0.172 | 0.341 | SEGVB5Y.CSIBBB | -0.299 | 0.131 |
| EuroSTOXX50.OBX | 0.327 | 0.936 | INDU.BASPCAAA | -0.127 | 0.394 | DKGVB5Y.CSIBBB | -0.28 | 0.156 |
| INDU.OBX | 0.103 | 0.669 | OMXS30.BASPCAAA | -0.182 | 0.328 | NOGVB5Y.CSIBBB | -0.293 | 0.123 |
| OMXS30.OBX | 0.319 | 0.927 | OMXC20.BASPCAAA | -0.249 | 0.249 | BASPCAAA.CSIBBB | -0.0478 | 0.433 |
| OMXC20.OBX | 0.269 | 0.867 | OBX.BASPCAAA | -0.194 | 0.314 | CSIA.CSIBBB | 0.236 | 0.71 |
| EuroSTOXX50.DEGVB5Y | 0.0811 | 0.642 | DEGVB5Y.BASPCAAA | -0.121 | 0.4 | CSIBB.CSIBBB | 0.714 | 1 |
| INDU.DEGVB5Y | -0.0335 | 0.506 | USGVB5Y.BASPCAAA | 0.103 | 0.667 | EuroSTOXX50.CSIBARC | -0.659 | -0.252 |
| OMXS30.DEGVB5Y | 0.0558 | 0.612 | SEGVB5Y.BASPCAAA | -0.236 | 0.262 | INDU.CSIBARC | -0.561 | -0.135 |
| OMXC20.DEGVB5Y | -0.0268 | 0.513 | DKGVB5Y.BASPCAAA | -0.13 | 0.39 | OMXS30.CSIBARC | -0.614 | -0.198 |
| OBX.DEGVB5Y | 0.00628 | 0.552 | NOGVB5Y.BASPCAAA | -0.323 | 0.141 | OMXC20.CSIBARC | -0.569 | -0.144 |
| EuroSTOXX50.USGVB5Y | -0.00842 | 0.536 | EuroSTOXX50.CSIA | -0.214 | 0.225 | OBX.CSIBARC | -0.59 | -0.17 |
| INDU.USGVB5Y | 0.0232 | 0.574 | INDU.CSIA | -0.224 | 0.212 | DEGVB5Y.CSIBARC | -0.577 | -0.155 |
| OMXS30.USGVB5Y | -0.0295 | 0.511 | OMXS30.CSIA | -0.224 | 0.212 | USGVB5Y.CSIBARC | -0.801 | -0.423 |
| OMXC20.USGVB5Y | -0.085 | 0.445 | OMXC20.CSIA | -0.236 | 0.199 | SEGVB5Y.CSIBARC | -0.548 | -0.122 |
| OBX.USGVB5Y | -0.0596 | 0.475 | OBX.CSIA | -0.231 | 0.204 | DKGVB5Y.CSIBARC | -0.555 | -0.128 |
| DEGVB5Y.USGVB5Y | 0.18 | 0.759 | DEGVB5Y.CSIA | -0.249 | 0.181 | NOGVB5Y.CSIBARC | -0.436 | -0.00357 |
| EuroSTOXX50.SEGVB5Y | 0.0481 | 0.601 | USGVB5Y.CSIA | -0.217 | 0.22 | BASPCAAA.CSIBARC | -0.486 | -0.0446 |
| INDU.SEGVB5Y | -0.0908 | 0.436 | SEGVB5Y.CSIA | -0.274 | 0.151 | CSIA.CSIBARC | -0.158 | 0.283 |
| OMXS30.SEGVB5Y | 0.031 | 0.58 | DKGVB5Y.CSIA | -0.246 | 0.185 | CSIBB.CSIBARC | -0.112 | 0.348 |
| OMXC20.SEGVB5Y | 0.00394 | 0.548 | NOGVB5Y.CSIA | -0.266 | 0.145 | CSIBBB.CSIBARC | -0.112 | 0.348 |
| OBX.SEGVB5Y | 0.0142 | 0.56 | BASPCAAA.CSIA | -0.0404 | 0.431 | EuroSTOXX50.BICLB10Y | -0.402 | 0.0652 |
| DEGVB5Y.SEGVB5Y | 0.404 | 1 | EuroSTOXX50.CSIBB | -0.247 | 0.196 | INDU.BICLB10Y | -0.459 | -0.00354 |
| USGVB5Y.SEGVB5Y | 0.0942 | 0.655 | INDU.CSIBB | -0.249 | 0.192 | OMXS30.BICLB10Y | -0.378 | 0.094 |
| EuroSTOXX50.DKGVB5Y | 0.0619 | 0.619 | OMXS30.CSIBB | -0.242 | 0.201 | OMXC20.BICLB10Y | -0.345 | 0.134 |
| INDU.DKGVB5Y | -0.0479 | 0.489 | OMXC20.CSIBB | -0.247 | 0.196 | OBX.BICLB10Y | -0.356 | 0.122 |
| OMXS30.DKGVB5Y | 0.0414 | 0.594 | OBX.CSIBB | -0.268 | 0.171 | DEGVB5Y.BICLB10Y | -0.457 | -0.00218 |
| OMXC20.DKGVB5Y | -0.0271 | 0.513 | DEGVB5Y.CSIBB | -0.295 | 0.137 | USGVB5Y.BICLB10Y | -0.754 | -0.358 |
| OBX.DKGVB5Y | 0.000657 | 0.546 | USGVB5Y.CSIBB | -0.269 | 0.169 | SEGVB5Y.BICLB10Y | -0.384 | 0.0844 |
| DEGVB5Y.DKGVB5Y | 0.567 | 1 | SEGVB5Y.CSIBB | -0.299 | 0.131 | DKGVB5Y.BICLB10Y | -0.44 | 0.0188 |
| USGVB5Y.DKGVB5Y | 0.169 | 0.746 | DKGVB5Y.CSIBB | -0.28 | 0.156 | NOGVB5Y.BICLB10Y | -0.294 | 0.175 |
| SEGVB5Y.DKGVB5Y | 0.421 | 1 | NOGVB5Y.CSIBB | -0.293 | 0.123 | BASPCAAA.BICLB10Y | -0.405 | 0.0617 |
| EuroSTOXX50.NOGVB5Y | -0.0172 | 0.506 | BASPCAAA.CSIBB | -0.0478 | 0.433 | CSIA.BICLB10Y | -0.175 | 0.271 |
| INDU.NOGVB5Y | -0.119 | 0.386 | CSIA.CSIBB | 0.236 | 0.71 | CSIBB.BICLB10Y | -0.144 | 0.319 |
| OMXS30.NOGVB5Y | -0.0323 | 0.488 | EuroSTOXX50.CSIBBB | -0.247 | 0.196 | CSIBBB.BICLB10Y | -0.144 | 0.319 |
| OMXC20.NOGVB5Y | -0.047 | 0.471 | INDU.CSIBBB | -0.249 | 0.192 | CSIBARC.BICLB10Y | 0.161 | 0.728 |

Table B.3: Blomqvist's beta for the data sample of macroeconomic variables.

| Variable Pair | Lower Bound | Upper Bound | Variable Pair | Lower Bound | Upper Bound | Variable Pair | Lower Bound | Upper Bound |
|---|---|---|---|---|---|---|---|---|
| EuroSTOXX50.INDU | 0.335 | 0.406 | OBX.NOGVB5Y | -0.304 | -0.23 | OMXS30.CSIBBB | -0.0657 | 0.0111 |
| EuroSTOXX50.OMXS30 | 0.603 | 0.663 | DEGVB5Y.NOGVB5Y | -0.115 | -0.0382 | OMXC20.CSIBBB | -0.0657 | 0.0111 |
| INDU.OMXS30 | 0.416 | 0.485 | USGVB5Y.NOGVB5Y | 0.338 | 0.409 | OBX.CSIBBB | -0.53 | -0.463 |
| EuroSTOXX50.OMXC20 | 0.377 | 0.447 | SEGVB5Y.NOGVB5Y | 0.119 | 0.194 | DEGVB5Y.CSIBBB | -0.464 | -0.395 |
| INDU.OMXC20 | 0.213 | 0.287 | DKGVB5Y.NOGVB5Y | 0.0961 | 0.172 | USGVB5Y.CSIBBB | 0.54 | 0.603 |
| OMXS30.OMXC20 | 0.128 | 0.204 | EuroSTOXX50.BASPCAAA | 0.147 | 0.223 | SEGVB5Y.CSIBBB | 0.383 | 0.453 |
| EuroSTOXX50.OBX | 0.196 | 0.271 | INDU.BASPCAAA | 0.101 | 0.177 | DKGVB5Y.CSIBBB | -0.0442 | 0.0327 |
| INDU.OBX | 0.198 | 0.273 | OMXS30.BASPCAAA | 0.102 | 0.178 | NOGVB5Y.CSIBBB | -0.0749 | 0.00188 |
| OMXS30.OBX | 0.119 | 0.195 | OMXC20.BASPCAAA | -0.0557 | 0.0211 | BASPCAAA.CSIBBB | -0.117 | -0.0405 |
| OMXC20.OBX | -0.00265 | 0.0742 | OBX.BASPCAAA | -0.0465 | 0.0304 | CSIA.CSIBBB | -0.117 | -0.0405 |
| EuroSTOXX50.DEGVB5Y | -0.0342 | 0.0427 | DEGVB5Y.BASPCAAA | -0.0488 | 0.028 | CSIBB.CSIBBB | -0.252 | -0.177 |
| INDU.DEGVB5Y | -0.0626 | 0.0142 | USGVB5Y.BASPCAAA | -0.0488 | 0.028 | EuroSTOXX50.CSIBARC | -0.136 | -0.0598 |
| OMXS30.DEGVB5Y | -0.0626 | 0.0142 | SEGVB5Y.BASPCAAA | -0.255 | -0.18 | INDU.CSIBARC | 0.356 | 0.426 |
| OMXC20.DEGVB5Y | -0.328 | -0.254 | DKGVB5Y.BASPCAAA | -0.104 | -0.0274 | OMXS30.CSIBARC | 0.0181 | 0.0949 |
| OBX.DEGVB5Y | -0.129 | -0.0528 | NOGVB5Y.BASPCAAA | 0.14 | 0.216 | OMXC20.CSIBARC | -0.0488 | 0.028 |
| EuroSTOXX50.USGVB5Y | 0.296 | 0.369 | EuroSTOXX50.CSIA | 0.0706 | 0.147 | OBX.CSIBARC | -0.0757 | 0.00111 |
| INDU.USGVB5Y | 0.19 | 0.265 | INDU.CSIA | 0.132 | 0.208 | DEGVB5Y.CSIBARC | -0.0757 | 0.00111 |
| OMXS30.USGVB5Y | 0.227 | 0.301 | OMXS30.CSIA | 0.143 | 0.218 | USGVB5Y.CSIBARC | -0.285 | -0.211 |
| OMXC20.USGVB5Y | 0.112 | 0.188 | OMXC20.CSIA | 0.076 | 0.152 | SEGVB5Y.CSIBARC | -0.185 | -0.109 |
| OBX.USGVB5Y | 0.149 | 0.225 | OBX.CSIA | -0.015 | 0.0619 | DKGVB5Y.CSIBARC | -0.116 | -0.039 |
| DEGVB5Y.USGVB5Y | 0.0752 | 0.152 | DEGVB5Y.CSIA | -0.045 | 0.0319 | NOGVB5Y.CSIBARC | -0.0719 | 0.00496 |
| EuroSTOXX50.SEGVB5Y | 0.0876 | 0.164 | USGVB5Y.CSIA | -0.0534 | 0.0234 | BASPCAAA.CSIBARC | -0.0803 | -0.00351 |
| INDU.SEGVB5Y | 0.0474 | 0.124 | SEGVB5Y.CSIA | -0.0534 | 0.0234 | CSIA.CSIBARC | -0.0803 | -0.00351 |
| OMXS30.SEGVB5Y | 0.0289 | 0.106 | DKGVB5Y.CSIA | -0.252 | -0.177 | CSIBB.CSIBARC | -0.186 | -0.11 |
| OMXC20.SEGVB5Y | -0.0388 | 0.038 | NOGVB5Y.CSIA | -0.114 | -0.0374 | CSIBBB.CSIBARC | -0.0734 | 0.00342 |
| OBX.SEGVB5Y | -0.0388 | 0.038 | BASPCAAA.CSIA | 0.301 | 0.373 | EuroSTOXX50.BICLB10Y | 0.0891 | 0.165 |
| DEGVB5Y.SEGVB5Y | -0.0388 | 0.038 | EuroSTOXX50.CSIBB | 0.53 | 0.594 | INDU.BICLB10Y | 0.093 | 0.169 |
| USGVB5Y.SEGVB5Y | -0.245 | -0.17 | INDU.CSIBB | 0.754 | 0.802 | OMXS30.BICLB10Y | 0.093 | 0.169 |
| EuroSTOXX50.DKGVB5Y | -0.185 | -0.109 | OMXS30.CSIBB | 0.333 | 0.404 | OMXC20.BICLB10Y | -0.228 | -0.153 |
| INDU.DKGVB5Y | 0.408 | 0.477 | OMXC20.CSIBB | 0.0436 | 0.12 | OBX.BICLB10Y | -0.168 | -0.0915 |
| OMXS30.DKGVB5Y | 0.373 | 0.443 | OBX.CSIBB | -0.0711 | 0.00573 | DEGVB5Y.BICLB10Y | 0.428 | 0.497 |
| OMXC20.DKGVB5Y | 0.177 | 0.252 | DEGVB5Y.CSIBB | -0.0934 | -0.0166 | USGVB5Y.BICLB10Y | 0.428 | 0.497 |
| OBX.DKGVB5Y | 0.11 | 0.186 | USGVB5Y.CSIBB | -0.0934 | -0.0166 | SEGVB5Y.BICLB10Y | -0.0065 | 0.0703 |
| DEGVB5Y.DKGVB5Y | 0.182 | 0.257 | SEGVB5Y.CSIBB | -0.306 | -0.232 | DKGVB5Y.BICLB10Y | -0.00958 | 0.0672 |
| USGVB5Y.DKGVB5Y | 0.175 | 0.25 | DKGVB5Y.CSIBB | -0.189 | -0.113 | NOGVB5Y.BICLB10Y | 1 | 1 |
| SEGVB5Y.DKGVB5Y | 0.108 | 0.184 | NOGVB5Y.CSIBB | 0.209 | 0.284 | BASPCAAA.BICLB10Y | 0.0174 | 0.0941 |
| EuroSTOXX50.NOGVB5Y | -0.00419 | 0.0726 | BASPCAAA.CSIBB | 0.289 | 0.361 | CSIA.BICLB10Y | -0.00727 | 0.0696 |
| INDU.NOGVB5Y | -0.0557 | 0.0211 | CSIA.CSIBB | 0.144 | 0.22 | CSIBB.BICLB10Y | 0.0174 | 0.0941 |
| OMXS30.NOGVB5Y | -0.0411 | 0.0357 | EuroSTOXX50.CSIBBB | 0.162 | 0.237 | CSIBBB.BICLB10Y | -0.00727 | 0.0696 |
| OMXC20.NOGVB5Y | -0.0411 | 0.0357 | INDU.CSIBBB | -0.0365 | 0.0404 | CSIBARC.BICLB10Y | 0.282 | 0.354 |

# C Supplemental Theory

## C.1 Coefficients of Tail Dependence

For many applications, the dependence of extreme events is of particular interest. This type of dependence is referred to as tail dependence or asymptotic dependence. Consider a random vector $\mathbf{X} = (X_1, X_2)$ with marginals $F_1$ and $F_2$, the coefficient of lower tail dependence is then defined as the limit of the conditional probability that $X_2$ is less than or equal to the quantile $F_2^{-1}(q)$ provided that $X_1$ is less than or equal to $F_1^{-1}(q)$

as $q$ approaches 0, i.e.

$$\lambda_L = \lim_{q \to 0^+} \mathrm{P}\left(X_2 \le F_2^{-1}\left(q\right) | X_1 \le F_1^{-1}\left(q\right)\right). \tag{C.1}$$

Conversely, the coefficient of upper tail dependence defined as the limit of the conditional probability that $X_2$ is greater than the quantile $F_2^{-1}\left(q\right)$ provided that $X_1$ is greater than $F_1^{-1}\left(q\right)$ as $q$ approaches 1, i.e

$$\lambda_U = \lim_{q \to 1^-} \mathrm{P}\left(X_2 > F_2^{-1}\left(q\right) | X_1 > F_1^{-1}\left(q\right)\right) \tag{C.2}$$

If $\lambda_L = 0$, $X_1$ and $X_2$ have independent tails; if $\lambda \in (0, 1]$, $X_1$ and $X_2$ have dependent lower tails; and similar for $\lambda_U$.
When **X** has an elliptical distribution, $\lambda_L$ and $\lambda_U$ takes a special form. In order to address this case, however, the notion of regularly varying tails must be introduced.

**Definition C.1.** A distribution function $F$ is said to have a regularly varying left tail with tail index $k$ if there exists a number $k$ such that

$$\lim_{t \to -\infty} \frac{F\left(tx\right)}{F\left(t\right)} = x^k \text{ for every } x > 0.$$

Similarly, the condition for $F$ to have a regularly varying right tail with tail index $k$ is that there exists a number $k$ such that

$$\lim_{t \to \infty} \frac{1 - F\left(tx\right)}{1 - F\left(t\right)} = x^k \text{ for every } x > 0.$$ □

Now, if **X** is assumed to be elliptically distributed with linear correlation coefficient $\rho$, it holds that

$$\lambda_U = \frac{\int_{(\pi/2 - \arcsin \rho)/2}^{\pi/2} (\cos t)^\alpha \, dt}{\int_0^{\pi/2} (\cos t)^\alpha \, dt} \tag{C.3}$$

provided that $X_1$ and $X_2$ are equally distributed and $F_1\left(x\right)$ has a regularly varying left tail with index $-\alpha$. For more information about a proof, the reader is directed to [12]. Due to the symmetry of elliptical distribution it follows that $\lambda_U = \lambda_L$. Additionally, a special case worth mentioning is when **X** has a bivariate standard normal distribution, which leads to the tails of $X_1$ and $X_2$ being independent, i.e. $\lambda_L = \lambda_U = 0$.

# REFERENCES


[1] Bjork, T., *Arbitrage Theory in Continuous Time*, third edition, New York, Oxford University Press, 2009.

[2] Brockwell, P. and Davis, R., *Introduction to Time Series and Forecasting*, second edition, New York, Springer, 2002.

[3] Chollete, L. and Ning, C., 'The Dependence Structure of Macroeconomic Variables in the US', *Working Paper Series 0204*, Ryerson University, Department of Economics, November 2009.

[4] Chollete, L. and Ning, C., 'Asymmetric Dependece in the US Economy: Application to Money and the Phillips Curve', *Working Paper Series 046*, Ryerson University, Department of Economics, October 2012.

[5] Daul, S. et al., 'The Grouped t-Copula with an Application to Credit Risk', *RISK*, November 2003, p. 73-76.

[6] Dowd, K., 'Copulas in Macroeconomics', *Journal of International and global Economic Studies*, vol. 1, no. 1, June 2008, p. 1-16.

[7] Genest, C. and Rivest, L.-P., 'Statistical Inference Procedures for Bivariate Archimedean Copulas', *Journal of the American Statistical Association*, vol. 88, no. 423, September 1993, p. 1034-1043.

[8] Genest, C., Ghoudi, K., and Rivest, L.-P., 'A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions', *Biometrika*, vol. 82, no. 3, September 1995, p. 543-552.

[9] Genest, C., and Rivest, L.-P., 'On the Multivariate Pobability Integral Transformation', *Statistics & Probability Letters*, vol. 53, no. 4, July 2001, p. 391-399.

[10] Genest, C., Rémillard, B., and Beaudoin, D., 'Goodness-of-Fit Tests for Copulas: A Review and a Power Study', *Insurance: Mathematics and Economics*, vol. 44, no. 2, April 2009, p. 199-213.

[11] Hult, H. and Lindskog, F., 'Multivariate Extremes, Aggregation and Dependence in Elliptical Distributions', *Advances in Applied Probability*, vol. 34, no. 3, September 2002, p. 587-608

[12] Hult, H. et al., *Risk and Portfolio Analyis: Principles and Methods*, London, Springer, 2012.

[13] Hult, Henrik; Professor at the department of mathematical statistics, Royal Institute of Technology. Tutoring 2016-03-23 and 2016-01-19.

[14] Joe, H., *Dependence Modeling with Copulas*, Boca Raton, FL, CRC Press, 2014.

[15] Kim, G., Silvapulle, M., and Silvapulle, P., 'Comparison of Semiparametric and Parametric Methods for Estimating Copulas', *Computational Statistics & Data Analysis*, vol. 51, no. 6, March 2007, p. 2836-2859.

[16] Krupskii, P. and Joe, H., 'Tail-Weighted Measures of Dependence', *Journal of Applied Statistics*, vol. 42, no. 3, 2015, p. 614-629.

[17] Nelsen, R.B., *An Introduction to Copulas*, second edition, New York, NY, Springer, 2006.

[18] Patton, A., 'A Review of Copula Models for Economic Time Series', *Journal of Multivariate Analysis*, vol. 110, 2012, p. 4-18.

[19] Rousseeuw, P., and Molenberghs, G., 'Transformation of Non Positive Semidefinite Correlation Matrices', *Communications in Statistics – Theory and Methods*, vol. 22, no. 4, 1993, p. 965-984.

[20] Silvey, S., *Statistical Inference*, Chapman Hall/CRC Press, 1975.

[21] Weiß, G., 'Copula Parameter Estimation by Maximum-Likelihood and Minimum-Distance Estimators: A Simulation Study', *Computational Statistics*, vol. 26, no. 1, March 2011, p. 31-54.

TRITA -MAT-E 2017:08
ISRN -KTH/MAT/E--17/08--SE