Technische Universität München

Zentrum Mathematik

Selecting pair-copula families for regular vines with application to the multivariate analysis of European stock market indices

Diplomarbeit

von

Natalia Belgorodski

Themensteller:	Prof. Dr. Claudia Czado
Themensteller:	Christian Libor (Deutsche Börse)
Betreuer:	Dr. Mathias Hofmann
Abgabetermin:	29.07.2010

Hiermit erkläre ich, dass ich diese Diplomarbeit selbstständig angefertigt und nur die angegebenen Quellen verwendet habe.

Garching, den July 29, 2010

Danksagung

Allen, die mich bei der Anfertigung meiner Diplomarbeit auf unterschiedlichste Weise unterstützt haben, möchte ich an dieser Stellen meinen herzlichen Dank aussprechen.

Ganz besonders bedanke ich mich bei Frau Prof. Dr. Claudia Czado für ein interessantes und meinen Wünschen entsprechendes Thema dieser Diplomarbeit, für ihre hervorragende fachliche Unterstützung, interessante Ideen und die sorgfältige Korrektur meiner Arbeit.

Als nächstes möchte ich mich bei meinem Betreuer, Herrn Dr. Mathias Hofmann, für seine hilfreiche Tipps bedanken. Ein besonderer Dank geht an Ulf Schepsmeier für seine umfassende Betreuung in allen möglichen Fragen zum R Paket `CDVineMLE`. Weiterhin möchte ich mich bei Herrn Christian Libor, meinem Betreuer seitens der Deutchen Börse, bedanken, dass er mich durch den Wald der Börsenbegriffe geführt hat. Einen ganz herzlichen Dank richte ich an Frau Katharina Schüller, Geschäftsführerin der Firma STAT-UP, die mir den Kontakt zu der Deutschen Börse ermittelt hat.

Meinen Eltern danke ich dafür, dass sie mich in meinem Vorhaben, ein Auslandsstudium zu machen, bestärkt haben und in schwierigen Zeiten für mich immer da waren. Ich richte einen besonderen Dank an meinen Mann für seine Liebe, Verständniss und unendliche Geduld in den letzten acht Monaten. Ohne Unterstützung seitens meiner Familie wäre mir mein Studium viel schwerer gefallen.

Zusammenfassung

Das Thema dieser Diplomarbeit wurde von der Deutschen Börse initiiert, Technische Universität München stellte notwendige theoretische Grundlagen und die Software dazu bereit. Das Ziel dieser Arbeit ist, Renditen von zehn internationalen Europäischen Aktienindizes multivariat zu modellieren.

Zur Modellierung von multivariaten Daten wird ein zweistufiges Verfahren eingesetzt, das Zeitreihenmodelle mit Pair-Copula Zerlegungen kombiniert. Copula-Modelle benötigen unabhängig identisch verteilte Daten als Input. In der Praxis erfüllen nur wenige von beobachteten Daten diese Bedingung. Insbesondere werden diese Voraussetzungen von Finanzzeitreihen, zu denen auch die Renditen gezählt werden, verletzt. Finanzzeitreihen sind meistens nicht normalverteilt und zeichnen sich durch starke Wölbung, ausgeprägte Schiefe und nicht konstante Varianz aus. Im ersten Schritt des Verfahrens werden marginale Verteilungen eines multivariaten Datensatzen mittels (ARMA-) GARCH Zeitreihenmodellen modelliert. Dabei lassen sich besondere Eigenschaften von Finanzzeitreihen durch eine einschlägige Wahl von Fehlertermenverteilung komplett beschreiben. Sind die Modellvoraussetzungen richtig formuliert, so erwartet man, dass standardisierte Residuen unabhängig identisch verteilt sind. Durch eine passende Transformation lassen sie sich dann auf das Einheitsintervall projizieren und können nun als Input für Copula-Modelle verwendet werden. Im zweiten Schritt werden die Abhängigkeiten zwischen den univariaten Daten mittels Copulas modelliert. Es gibt zahlreiche parametrische Copula-Familien, die zur Modellierung von unterschiedlichen Abhängigkeitsstrukturen geeignet sind. Allerdings funktioniert es sehr gut nur auf der bivariaten Ebene. Die Erweiterung von Copulas auf höhere Dimensionen ist mit einigen Hindernissen verbunden und ist nicht immer möglich. Die so genannten Pair-Copula Zerlegungen, die bei Aas, Czado, Frigessi, and Bakken (2009) präsentiert wurden, lösen dieses Problem. Die Idee dieser Methode besteht darin, die gemeinsame Dichte von Zufallsvariablen in deren Randdichten und bivariaten Copuladichten zu zerlegen. Da solche Zerlegungen nicht eindeutig sind, werden sie durch von Kurowicka and Cooke (2006) eingeführten Vine-Modellen bestimmt.

Das bedeutende Problem in der Pair-Copula Theorie stellt die Auswahl einer passenden Copula-Familie für bestimmte Variablenpärchen dar. Ein grosser Teil dieser Arbeit beschreibt die Entwicklung eines Verfahrens für Goodness-Of-Fit Testing für Copulas. Zuerst werden zwei nicht parametrische Verfahren, der Vuong (Vuong (1989)) und der Clarke Test (Clarke (2007)) betrachtet, die zum Vergleich von nicht eingeschachtelten Modellen einge-

führt wurden. Beide Tests basieren auf dem Kullback-Leibler Informationskriterium, das den Abstand zwischen zwei Modellen misst. Weiter werden die Ergebnisse einer umfangreichen Simulationsstudie dargestellt und diskutiert. Das Ziel dieser Studie ist, herauszufinden, wie gut beide Tests zum bivariaten Vergleich von Copulas geeignet sind. Die Berechnungen wurden für Daten aus verschiedenen Copula-Klassen und in Abhängigkeit von Stichprobenlänge und Abhängigkeitsstärke durchgeführt. Als Nächstes werden diese zwei Tests erweitert, dass sie zum Vergleich von mehr als Copulas einsetzbar sind. Dazu wurde ebenfalls eine Simulationsstudie durchgeführt, deren Ergebnisse detailliert diskutiert werden. Letztendlich werden bivariate Vuong und Clarke Tests mit den gängigen Goodness-Of-Fit Tests von Genest, Rémillard, and Beaudoin (2009) verglichen.

Contents

List of figures vii

List of tables viii

Introduction 1

1 Data description 4

2 Common time series models 8

2.1 Basic concepts . . . 8
2.2 ARMA models . . . 12
2.3 ARCH and GARCH models . . . 14
2.4 ARMA-GARCH models . . . 16
2.4.1 Forecasting future mean and volatility . . . 17
2.5 Partial autocorrelation function . . . 18
2.6 Specification of ARMA-GARCH models . . . 19
2.6.1 Order of ARMA models . . . 19
2.6.2 Order of GARCH models . . . 20
2.7 Model diagnostics for time series models . . . 20
2.8 An illustrative example . . . 22

3 Modeling multivariate dependence 27

3.1 Moments and their estimation . . . 27
3.2 Correlation coefficients . . . 29
3.3 Tail dependence . . . 31
3.4 Graphical tools for detecting (tail) dependence . . . 33
3.4.1 Chi-plots . . . 33
3.4.2 Kendall's process plots . . . 34
3.4.3 Examples of Chi- and K-plots . . . 35
3.4.4 Chi-plots for detecting tail dependences . . . 37
3.5 Copulas . . . 37

3.5.1 Joint distributions . . . 39
3.5.2 Definition and basic copula properties . . . 40
3.5.3 Copula and dependence measures . . . 42
3.5.4 Most common parametric copula families . . . 43
3.6 Pair-copula constructions and vines . . . 49
3.6.1 Pair-copula constructions . . . 54
3.6.2 D-vines . . . 58
3.6.3 Inference for D-vine . . . 60
3.6.4 Simulation from D-vine . . . 60
3.7 Time series pair-copula approach . . . 63

4 Goodness-of-fit tests for copulas **66**
4.1 Vuong Test . . . 66
4.2 Clarke Test . . . 69
4.3 Simulation study for the Vuong and Clarke tests . . . 72
4.4 Distances between different copula families . . . 73
4.5 Relationship between accuracy of the Vuong/Clarke test and KLICs . . . 75
4.6 Comparison to Genest study . . . 76
4.7 Comparison more than two models . . . 77

5 Application: Financial Market Indices **91**
5.1 Univariate modeling of margins . . . 92
5.2 Analysing dependence structure . . . 93
5.3 D-vine identification . . . 94
5.4 D-vine estimation . . . 95
5.4.1 Sequential estimation . . . 95
5.4.2 Joint MLE estimation . . . 96
5.5 Validation results . . . 96
5.6 Simulation of future returns . . . 97

Conclusion **116**

Bibliography **118**

A Notations **123**

B R packages **125**

C Statistical distributions used **126**
C.1 Univariate distributions . . . 126

C.1.1 Normal distribution 126
C.1.2 Skewed normal distribution 126
C.1.3 Student's t distribution 127
C.1.4 Skewed Student's t distribution 127
C.2 Multivariate distributions 128
C.2.1 Multivariate normal distribution 129
C.2.2 Multivariate Student's t Distribution 130

D R code of Vuong and Clarke tests **132**
D.1 Vuong Test 132
D.2 Clarke Test 134

E Data for D-vine trees **137**

F D-vine sequential estimation **145**

List of Figures

1.1 Time series plots of data set FINANCIAL MARKET INDICES 6
1.2 Scatter plots and Kendall's $\hat{\tau}$ of FINANCIAL MARKET INDICES 7
2.1 Time series plots of AR(1), MA(1) and ARMA(1,1) processes 14
2.2 Time series with constant and variable volatility 15
2.3 ACF and PACF of AR(1), MA(1) and ARMA(1,1) processes 19
2.4 Time series plots, ACF and PACF plots of DAX returns 23
2.5 Diagnostics of ARMA-GARCH models fitted to the DAX data 25
3.1 Drawbacks of correlation coefficients 31
3.2 Lower and upper tails of different dependence forms 33
3.3 Examples of Chi- and K-plots 36
3.4 Detecting tail dependency using Chi-plots 38
3.5 Most common copula families, univariate margins 50
3.6 Most common copula families, standard normal margins (1) 51
3.7 Most common copula families, standard normal margins (2) 52
3.8 Most common copula families, standard normal margins (3) 53
3.9 Representation of 4-dimensional D-vine 59
4.1 KLIC distances between different copula families 83
5.1 Sample ACF and PACF of data set FMI 99
5.2 Residual analysis of the GARCH models (1) 103
5.3 Residual analysis of the GARCH models (1) 104
5.4 Scatter and contour plots of transformed GARCH residuals 105
5.5 Chi- and K-plots of transformed GARCH residuals 106
5.6 Chi-plot of transformed residuals for detecting tail dependences 107
5.7 Visualization of empirical Kendall's τ correlation matrix 108
5.8 D-vine for transformed standardized residuals 109
5.11 Simulated returns 113
C.1 Density plots of several distributions 128
C.2 Bivariate standard normal distribution 131

List of Tables

1.1 Descriptive statistics of data set FMI 5
2.1 General behavior of ACF and PACF in AR, MA and ARMA processes . . 19
2.2 Summaries of estimated ARMA-GARCH models fitted to the DAX data . 24
3.1 Different copula families 43
4.1 Calculation of *better*, *equivalent* and *worse* entries 78
4.2 Comparison efficiency of the bivariate Vuong and Clarke tests, $\tau = 0.25$ . . 80
4.3 Comparison efficiency of the bivariate Vuong and Clarke tests, $\tau = 0.5$ . . . 81
4.4 Comparison efficiency of the bivariate Vuong and Clarke tests, $\tau = 0.75$ . . 82
4.5 Extended Genest study for $\tau = 0.25$ 84
4.6 Extended Genest study for $\tau = 0.5$ 85
4.7 Extended Genest study for $\tau = 0.75$ 86
4.8 Example of the score method based on the Vuong test 87
4.9 Example of the score method based on the Clarke test 88
4.10 Simulation study: goodness-of-fit testing via score approach 89
4.11 Simulation study: goodness-of-fit testing via score approach 90
5.1 Testing the null hypothesis about equivalence of several D-vines 97
5.2 Summaries of estimated time series models for dataset FMI 100
5.3 Estimated parameters for 10-dimensional D-vine 111
5.4 Validation the D-vine model by simulation 112
5.5 Validation time series properties of simulated returns 114
5.6 Validation simulated returns 115

Introduction

The procedure of modeling multivariate return series of financial indices has to take into account two factors. The first one is making allowance for time-varying volatility. It traces back to the fact that small values of a time series are followed by small values and vice versa large values are followed by large values. The second one is an incorporating dependence structure between the individual return sequences that is detected on the observed values from the past. The aim of this thesis is to provide a framework for generating multivariate returns that takes these two factors into consideration.

The traditional multivariate Time Series Approach (TSA) with standard normal distributed error terms is based on calculation of the future volatility and future mean by using the correlation structure of the past values. The usual linear correlation is not a satisfactory measure of the dependence between *financial* time series. Firstly, two return series might be uncorrelated but not independent. The uncorrelateness implies automatically independence only for normal distributed random vectors. Generally, financial time series do not follow normal distribution, they are mostly skewed and exhibit a high kurtosis and heavy tails. The problem of non-zero kurtosis and skewness can be handled by using an appropriate distribution for error terms, for example skewed standard normal or (skewed) standard Student distributions. Secondly, linear correlation measures only linear dependence but does not discover non-linear dependence. Another drawback of the TSA is that the correlation is assumed to be constant over the time period of interest. Moreover, this methodology implies in multivariate return series the same dependence for positive returns as well as for negative returns. In contrast, some empirical studies, for example Longin and Solnik (1995, 2001), have found significantly stronger positive correlations coefficients for negative returns than for positive returns. This phenomenon is called asymmetric tail dependence. Hence, the classical multivariate Time Series Approach seems be unable to capture all the properties and co-movements of the multivariate financial time series. An alternative way of modeling (asymmetric) dependence in stock market indices returns which can cover the dynamics of the univariate time series will be researched in this diploma thesis.

Copula theory provides a modern and more efficient approach to model the high-dimensional dependency. A copula is a function that combines univariate margins into their joint distribution. In this way, the nature of dependence can be modeled more generally compared to the techniques based on the linear correlation measure. Large numbers of parametric

bivariate copulas were developed for the modeling of various dependency forms. But only a small set of these copulas can be extended to the higher dimensions. Another disadvantage of the multivariate copulas is that they allow only one copula type, for example a Gaussian copula, for all marginal pairs. So called pair-copula constructions were developed to handle this problem. The advantage of this innovation lies in the ability to use miscellaneous dependency structures, i.e varying copula families are allowed for different marginal pairs. The idea of such a modeling scheme is based on a decomposition of a multivariate density into a cascade of the pair-copulas, applied to the original variables and to their conditional and unconditional distribution functions. The most appropriate copula type for the pair-copulas can be detected by special inference methods, which will be discussed in this thesis.

This thesis is based on an article *"Pair-copula constructions of multiple dependence"* of Aas, Czado, Frigessi, and Bakken (2009). It surveys a framework for analyzing and forecasting high-dimensional multivariate time series that we name TS-PAIR-COPULA approach throughout this work. This method combines pair-copula constructions with (ARMA-)GARCH methodology in an innovative way. At first, a set of (ARMA-)GARCH models will be fitted to marginal distributions to predict expected returns and volatility. Then the standardized residuals of these models will be used to create a joint distribution via semi parametric multivariate copula based on the pair-copula constructions. In this way, the vital factors mentioned at the beginning of this chapter can be captured: employment of the (ARMA-)GARCH models overcomes the problem of non-constant volatility; the appropriate innovations distribution captures the problem of high kurtosis and non-zero skewness; finally, utilization of the copula concept handles the tail dependence problem .

The TS-PAIR-COPULA approach will be applied to a FINANCIAL MARKET INDICES (FMI) data set consisting of daily time series log-returns of 10 European stock market indices. This data were provided by Deutsche Börse and cover the time period from January 3, 2006 to November 11, 2009.

This thesis is organized as follows:

- ✓ Chapter 1 provides a description of the data set used in this thesis including some of its descriptive statistics and time series properties.
- ✓ Chapter 2 introduces briefly the basic foundations of the theory of time series such as white noise, stationarity, autocorrelation and partial autocorrelation functions. The most important processes such as AR, MA, ARMA, ARCH, GARCH and ARMA-GARCH are also considered in this chapter. Further, useful tools for determining an appropriate model order and some methods of model checking, that will be demonstrated on an illustrative example, are given at the end of this chapter.
- ✓ The following Chapter 3 outlines the classical dependence measures like Pearson's and Kendall's correlation coefficients. Two graphical non-parametric methods for

detecting dependence structure, Chi- and K-plot, are also presented. The next section defines copulas and gives some of their basic properties. Various copula families, which are common in multivariate modeling of the financial markets, will be introduced in the following part. Further, pair-copula constructions are described at the end of this chapter.

- ✓ Chapter 4 provides an alternative methodology for goodness-of-fit testing for copulas that is based on Vuong and Clarke tests.
- ✓ Chapter 5 is dedicated to empirical results. The multivariate TS-PAIR-COPULA approach is applied to the financial data set introduced in Chapter 1. Goodness-of-prediction will be judged by contrasting the forecasted values with the true future observations.
- ✓ The thesis concludes with a summary of the application results.

In this thesis we provide many illustrative examples that makes the theory chapters more understandable. Furthermore, references to the R functions and packages are given to facilitate the introduced statistical inference and modeling tools. All numerical calculations presented in this diploma thesis have been computed using free statistical software R version 2.9.2, R Development Core Team (2009). Evaluation results obtained in R were automatically converted into LaTeX tables by use the R package `xtable`, Dahl (2009), in combination with a `Sweave()` function.

Chapter 1

Data description

The Financial Market Indices data set consists of daily time series log-returns of 10 European stock market indices and contains a total of 955 observations for each index. The sample covers the period of three years, from January 3, 2006 to November 11, 2009, due to constraints on data availability. The data were obtained from the Bloomberg database. The stock market indices of interest are AEX of Netherlands, ATX of Austria, BVLG of Portugal, FCHI of France, FTMIB of Italy, GDAXI of Germany, OMXC20 of Denmark, OMXSPI of Sweden, SMSI of Spain and STOXXER of Europe. Their short descriptions are presented below. These indices come from the Eurozone and were traded in euro. The problem of non-overlapping trading hours does not arise, because all exchanges, where these indices are traded, operate in the same time zone. In order to reduce numerical problems and to permit stable estimation results the returns are given in percentages.

The observed data were separated into a training and a testing set. The first one covers the time frame from 2006-01-02 to 2009-06-18, a total of 855 observations. It will be used for the estimating of model parameters. The test set contains 100 data points dating from 2009-06-19 to 2009-11-11. It is employed to compare predictions based on the estimated model with true observed values and, afterwards, to assess the appropriateness of the models studied and determine its correctness.

Some descriptive statistics of return sequences are given in Table 1.1. The sample means, rounded to 2 positions after decimal point, of all 10 series are close to zero. Sample standard deviations range between 1.31 and 2.13. According to the p-values of Shapiro and Jack-Berra statistics, the return series are not normally distributed. Moreover, they indicate a high positive kurtosis lying between 3.5 and 11.31 and a non-zero skewness. Both features are typical for financial time series. The estimated skewness is positive for FCHI, FTMIB, GDAXI and OMXSPI indices. For all other indices it is negative. Further, Ljung-Box statistics provides a useful tool for examining the null hypothesis of independence or, in other words, for testing a serial correlation in time series. Based on the p-values of the Ljung-Box statistics at lag 15, the observed data series are autocorrelated except ATX, GDAXI and OMXSPI. More details about Shapiro, Jack-Berra and Ljung-Box tests are

INVESTIGATED STOCK MARKET INDICES

AEX *Amsterdam Stock Index, Netherlands*; Is composed of a maximum of 25 of the most actively traded Netherlands securities, provides a fair representation of the Dutch economy

ATX *Austrian Traded Index, Austrian*; Is a most important stock market index of the Wiener Börse and is defined as a price index and currently consists of 20 stocks

BVLG *Lisbon BVL General Index, Portugal*; Is a market value-weighted index that tracks the daily total return performance of all stocks traded on the official market of the Lisbon Stock Exchange

FCHI *CAC 40 Index, France*; Is a benchmark French stock market index, represents a capitalization-weighted measure of the 40 most significant values among the 100 highest market caps on the Paris Bourse

FTMIB *FTSE MIB Milan Index, Italy*; Is a benchmark stock market index for the Borsa Italiana and measures the performance of the 40 most liquid and capitalized Italian shares

GDAXI *German Stock Index (DAX Index), Germany*; Is a stock market index consisting of the 30 major German companies trading on the Frankfurt Stock Exchange

OMXC20 *OMX Copenhagen 20 Index, Denmark*; Is a stock market index consisting of the 20 most actively traded shares on the Copenhagen Stock Exchange

OMXSPI *OMX Stockholm Index, Sweden*; Is a stock market index of all shares that trade on the Stockholm Stock Exchange

SMSI *Madrid Stock Exchange General Index, Spain*; Is a capitalization-weighted stock market index that measures the performance of a selected number of continuous market stocks

STOXXER *Dow Jones EURO STOXX 50, Europe*; Is a stock index of Eurozone stocks and provides a blue-chip representation of 50 Supersector leaders in the Eurozone

Table 1.1: Descriptive statistics of return series: sample mean, sample kurtosis, sample skewness, test statistics and p-values of Shapiro and Jarque-Bera tests for testing normality, test statistic and p-value of Ljung-Box test for testing autocorrelation at lag 15

INDEX	mean	sd	kurtosis	skewness	Shapiro		Jarque-Bera		Ljung-Box	
					stat.	p-val.	stat.	p-val.	stat.	p-val.
AEX	-0.03	1.77	7.14	-0.16	0.89	0	2031.87	0	41.72	0.00
ATX	-0.04	2.13	4.28	-0.23	0.94	0	737.35	0	11.46	0.72
BVLG	0.01	1.31	11.31	-0.41	0.87	0	5120.15	0	36.22	0.00
FCHI	-0.02	1.73	6.59	0.10	0.91	0	1727.08	0	51.59	0.00
FTMIB	-0.05	1.70	6.35	0.02	0.91	0	1606.96	0	60.52	0.00
GDAXI	0.00	1.68	6.98	0.19	0.91	0	1941.62	0	21.15	0.13
OMXC20	-0.02	1.71	5.55	-0.24	0.93	0	1234.15	0	27.57	0.02
OMXSPI	-0.01	2.02	3.50	0.10	0.95	0	488.48	0	16.69	0.34
SMSI	0.01	1.63	6.36	-0.08	0.92	0	1611.71	0	41.36	0.00
STOXXER	-0.01	1.62	5.71	-0.07	0.92	0	1300.27	0	30.90	0.01

Figure 1.1: Time series plots of stock market returns on the same scale $[-12, 12]$: the train data is marked with black color and the test data with gray color

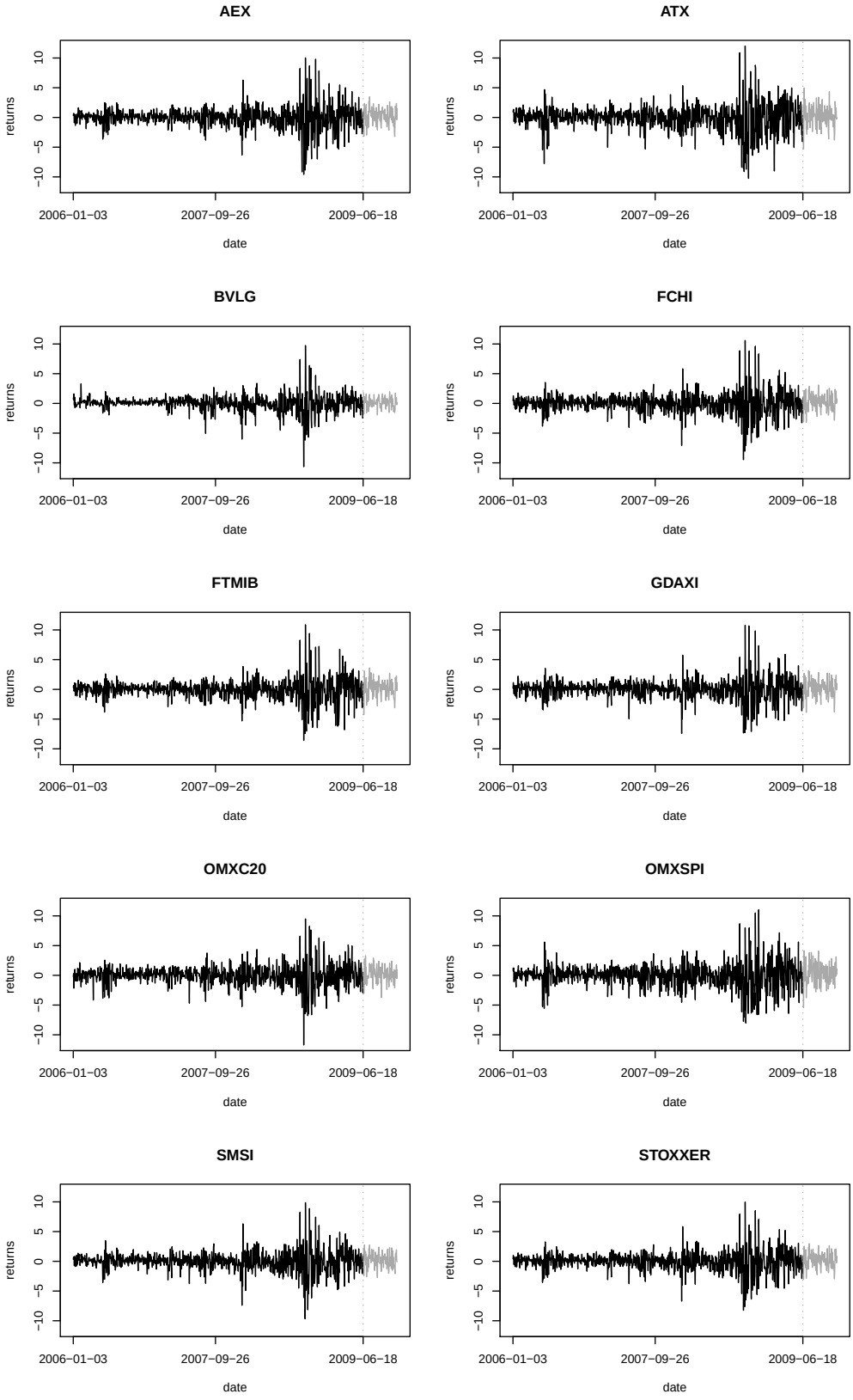
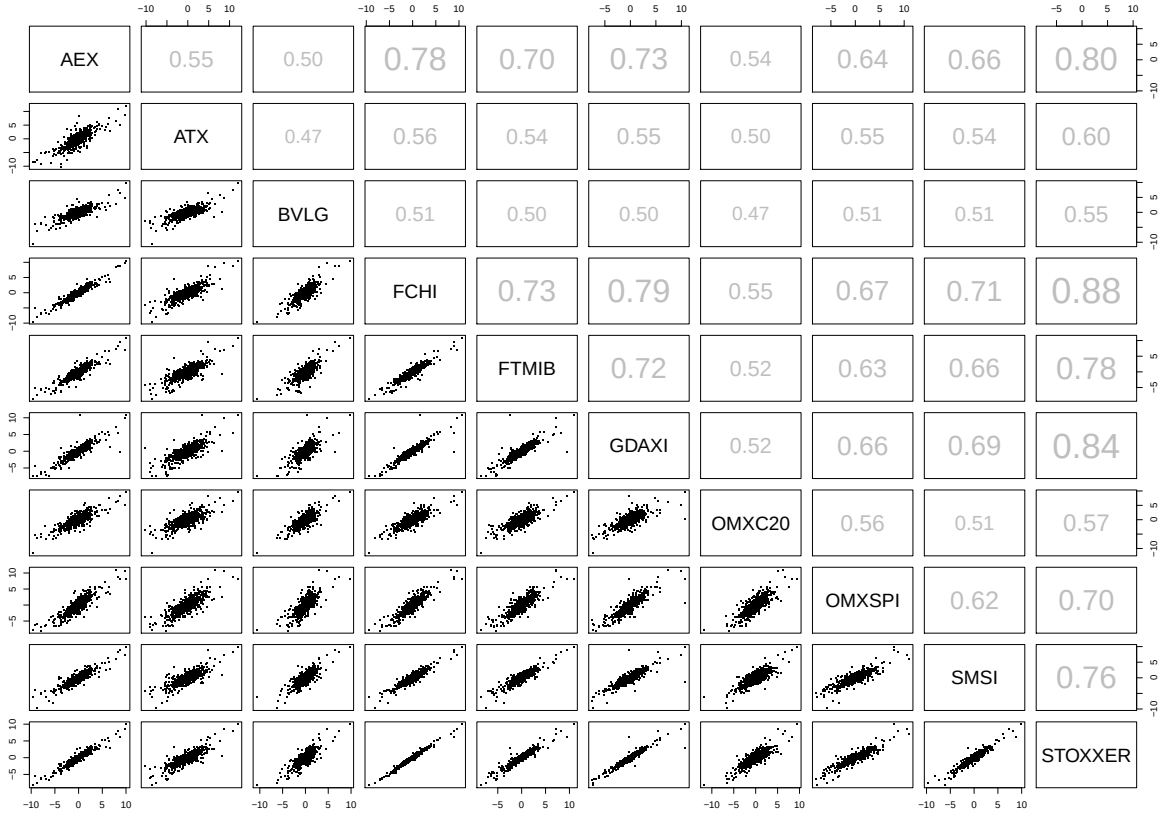


Figure 1.2: Visualization of dependences in the FINANCIAL MARKET INDICES data set: sample Kendall's correlation coefficients are on the upper triangle, bivariate scatter plots are on the lower triangle



given in Chapter 2, Section "*Model diagnostics for time series models*".

The time series plots of the 10 stock market returns are drawn in Figure 1.1, in black color for the train data and in gray color for the test data. All time series were plotted on the same scale. All data series exhibit a non-constant variance over the complete time period. The observed returns are more volatile over some time periods and less volatile over some other time periods. The most volatile period is observed at the end of 2008. Figure 1.2 exhibits pairwise scatter plots on the lower panel and sample Kendall's τ correlation coefficients on the upper panel. As we can see, there are positive dependencies of varying degree between the return series. The Kendall's assosiation coefficients are all positive and lie between 0.5 and 0.9.

Chapter 2

Common time series models

This chapter provides background on time series and introduces the notations that will be used throughout this thesis. The time series theory described here follows description given by Shumway and Stoffer (2006). The concept of the stationarity, white noise, mean, autocovariance and the autocorrelation functions are introduced in the first section. Next, the definitions of moving average (MA), autoregressive models (AR) and their hybrid ARMA models are given. In the next section, the ARCH and GARCH models for the time series models with the heteroscedastic variance are discussed. After that, the ARMA-GARCH models for simultaneous modeling of the conditional mean and conditional variance structure by combining an ARMA model with conditional heteroscedasticity are given. At the end of this chapter, we discuss how the autocorrelation functions and the partial autocorrelation functions (ACF and PACF) can be used for detecting an appropriate order of time series models and for model diagnostics. The tools for checking model assumptions are listed in the next section. Finally, we illustrate the techniques of residual analysis on an ARMA-GARCH model that were fitted to the real data.

2.1 Basic concepts

Te basic notion of this chapter is a white noise process. An uncorrelated sequence $\{\varepsilon_t\}$ of random variables is called *white noise* process, if it has a zero mean and a finite variance, i.e

$$\begin{aligned} &E[\varepsilon_t] = 0 \\ &\mathrm{Var}(\varepsilon_t) = \sigma^2 \\ &E[\varepsilon_{t_1}\varepsilon_{t_2}] = 0 \quad \forall t_1 \neq t_2 \ , \end{aligned}$$

An independent white noise process is defined by a stronger condition that ε_{t_1} and ε_{t_2} are independent $\forall t_1 \neq t_2$. In case of $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$, the random process $\{\varepsilon_t\}$ is called *Gaussian white noise process.*

A *time series* can be defined as a collection of random variables indexed according to the order they were obtained in time. For example, we may consider a time series as a sequence of random variables $x_1, x_2, x_3, \ldots$, wherein the random variable x_1 denotes the value taken by the series at the first time point t_1, the variable x_2 denotes the value for the second time period t_2, and so on. A time difference $h = t_2 - t_1$ between two time points t_2 and t_1, $t_2 > t_1$, is called *lag*.

The multidimensional distribution function (2.1) describes the data completely and is a natural tool for displaying and analyzing time series data.

$$F(x_1, x_2, \ldots, x_n) = P(X_{t_1} \leq x_1, X_{t_2} \leq x_2, \ldots, X_{t_n} \leq x_n) \tag{2.1}$$

Unfortunately, the multidimensional distribution function (2.1) can not usually be written easily, unless the random variables are independent and identical distributed. In case of independence, the joint distribution function can be expressed as a product of the marginal distributions

$$F(x_1, x_2, \ldots, x_n) = \prod_{t=1}^{n} F_t(x_t) = \prod_{t=1}^{n} P_t(X_t \leq x_t) \tag{2.2}$$

with one-dimensional distribution functions

$$F_t(x) = P_t(X_t \leq x)$$

and corresponding one-dimensional density functions

$$f_t(x) = \frac{\partial F_t(x)}{\partial x} .$$

As it is usual in statistics, the complete description involves the multivariate distribution function of the jointly sampled values $x_1, x_2, \ldots, x_n$, whereas more economical descriptions can be described in terms of the mean and autocorrelation functions. Because correlation is an essential feature of the time series analysis, the most useful descriptive measures can be expressed in terms of covariance and correlation functions.

Definition 2.1 (Mean function) *The* ***mean function*** *is defined as*

$$\mu_t = E[X_t] = \int_{-\infty}^{\infty} x f_t(x) dx , \tag{2.3}$$

provided it exists, where $E[\cdot]$ denotes the usual expected value operator.

The lack of *linear independence* between two points x_{t_1} and x_{t_2} of the same time series can be estimated numerically using the notions of covariance and correlation. Assuming the variance of x is finite, we have the following definition.

Definition 2.2 (Autocovariance function) *The* ***autocovariance function*** *is defined*

for all t_1 and t_2 as a second moment product

$$\gamma(t_1, t_2) = E\left[(X_{t_1} - \mu_{t_1})(X_{t_2} - \mu_{t_2})\right] . \tag{2.4}$$

Definition 2.3 (ACF) *The **autocorrelation function (ACF)** is defined as*

$$\rho(t_1, t_2) = \frac{\gamma(t_1, t_2)}{\sqrt{\gamma(t_1, t_1)\gamma(t_2, t_2)}} . \tag{2.5}$$

The autocovariance function measures the linear predictability of the series at time t_2 using only the values from time t_1. The ACF presents the scaled value of autocovariance function in the interval $[-1, 1]$. Equation $|\rho(t_1, t_2)| = 1$ means x_{t_2} can be *perfectly* predicted by x_{t_1} with a linear relationship $x_{t_2} = \beta_0 + \beta_1 x_{t_1}$.

To make statistical inference about the structure of a stochastic process on the basis of the observed sample $x_1, x_2, \ldots, x_n$ of that process, we must usually make some simplifying (and presumably reasonable) assumptions about that structure. Such most important assumption is that some sort of regularity exists over time in the behavior of a time series. We introduce the notion of regularity using a concept called *stationarity.*

Definition 2.4 (Strictly stationary) *A **strictly stationary** time series is one for which the probabilistic behavior of every sequence of values $x_{t_1}, \ldots, x_{t_k}$ is identical to that of time shifted set $x_{t_1+h}, \ldots, x_{t_k+h}$. That is,*

$$P\left(X_{t_1} \leq c_1, \ldots, X_{t_k} \leq c_k\right) = P\left(X_{t_1+h} \leq c_1, \ldots, X_{t_k+h} \leq c_k\right) \tag{2.6}$$

for all $k = 1, 2, \ldots$, all time points $t_1, t_2, \ldots$, all numbers $c_1, c_2, \ldots$, and all time shifts $h = 0, \pm 1, \pm 2, \ldots$.

The version of stationarity in (2.6) is too strong for most applications. Moreover, it is difficult to verify the stationarity from a single data set. Rather than assume conditions on all subset distributions of a time series, we will use a milder version that requires conditions only on the two first moments of the series. We now have the following definition.

Definition 2.5 (Weakly stationary) *A **weakly stationary** time series, $\{x_t\}$, is a finite variance process such that*

(i) the mean value function, μ_t, defined in (2.3) is constant and does not depend on time t, and
(ii) the covariance function, $\gamma(t_1, t_2)$, defined in (2.4) depends on t_1 and t_2 only through their difference $|t_1 - t_2|$.

From now on, we will use the term *stationary* to express the weak stationary. For the stationary time series, we have an equality for the mean function

$$E\left[X_t\right] = \mu_t = \mu \tag{2.7}$$

and the covariance function can be simplified with $t_2 = t_1 + h$ as

$$\begin{aligned}\gamma(t_1, t_2) = \gamma(t_1, t_+h) &= E\left[(X_{t_1+h} - \mu)(X_{t_1} - \mu)\right] \\ &= E\left[(X_h - \mu)(X_0 - \mu)\right] = \gamma(h, 0) = \gamma(h) \ ,\end{aligned} \tag{2.8}$$

since the covariance of a stationary time series, $\gamma(t_1, t_2)$, depends on t_1 and t_2 only through their difference $|t_1 - t_2|$.

Definition 2.6 *The **autocovariance function of a stationary time series** will be written as*

$$\gamma(h) = E\left[(X_{t+h} - \mu)(X_t - \mu)\right] \ . \tag{2.9}$$

Definition 2.7 *The **autocorrelation function (ACF) of a stationary time series** will be written as*

$$\rho(h) = \frac{\gamma(t+h, h)}{\sqrt{\gamma(t+h, t+h)\gamma(t, t)}} = \frac{\gamma(h)}{\gamma(0)} \ . \tag{2.10}$$

In practice, only the sample data are available for time series analyze. For such data, the assumption of stationarity can not be verified and the observations are not i.i.d. copies of the same random variable. Hence, only the average values can be used to estimate the population means and covariance functions. If a time series is stationary, the mean function is constant. So that we can estimate it by the sample mean in a following way

$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t \ . \tag{2.11}$$

And the theoretical autocovariance function in (2.9) is estimated by the sample autocovariance function.

Definition 2.8 *The **sample autocovariance function** of observed time series $\{x_t\}_{t=1\ldots n}$ is defined as*

$$\hat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-h} (x_{t+h} - \bar{x})\,(x_t - \bar{x}) \tag{2.12}$$

with $\hat{\gamma}(-h) = \hat{\gamma}(h)$ for all $h = 0, 1, \ldots, n-1$.

The sum in (2.12) is defined only till $n-h$ because x_{t+h} is not available for $t+h > n$. The estimator in (2.12) is generally preferred to the one that would be obtained by dividing by $n - h$ since (2.12) is a non-negative definite function (for more details see Shumway and Stoffer (2006, Section "*Estimation of Correlation*", pp. 29–34)). Note, an unbiased estimate of $\gamma(h)$ can not be obtained neither dividing by n nor $n - h$ in (2.12).

Definition 2.9 *The **sample autocorrelation function** of observed time series* $\{x_t\}_{t=1\ldots n}$ *is defined, analogously to* (2.10)

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} \,. \tag{2.13}$$

The sampling distribution of the sample autocorrelation function can be used to check whether correlations at some lags are significant. Precise details are given by Shumway and Stoffer (2006, Appendix A, pp. 513–521).

Property 2.10 (Large Sample Distribution of the ACF) *Under general regularity conditions, if a sequence* $\{x_t\}_{t=1\ldots n}$ *of random variables is a white noise, then for* n *large, the sample ACF,* $\hat{\rho}(h)$*, for* $h = 1, 2, \ldots, H$*, where* H *is fixed but arbitrary, is approximately normally distributed with zero mean and standard deviation given by*

$$\sigma_{\hat{\rho}(h)} = \frac{1}{n} \,. \tag{2.14}$$

Using this property, we can construct a confidence interval for empirical autocorrelation coefficients and check whether a correlation at some lag is significant. Under certain assumptions the interval $\pm\frac{2}{\sqrt{n}}$ can be viewed as a 95% confidence interval. Empirical autocorrelation coefficients which fall outside of this interval, are significantly different from zero.

The sample ACF can be computed and plotted in R with function `acf` from the standard package `stats`. The blue dashed lines on the pacf plot indicate the confidence interval (95% by default) and allow a simple judgment about statistical significance of sample correlation at some lag h. Note, the bars on the `acf` plot start at the lag 0, so that the first bar is also significant and ist equal to one.

2.2 ARMA models

Classical regression was developed for the static case where dependent variable is influenced by current values of independent variables. Such classical regression is often insufficient for modeling a time series and explaining all of their properties. Firstly, it is reasonable to allow the dependent variable to be influenced by the past values of the independent variables as well as by its own past values. Secondly, residuals of classical regression applied to time series are often correlated. It causes a pattern in the corresponding residuals plot. In this section we introduce alternative methods for the modeling and forecasting univariate time series.

Autoregressive models are based on the idea that the current value of a time series $\{x_t\}$, x_t, can be explained as a linear combination of p past values, $x_{t-1}, \ldots, x_{t-p}$. Hence, p

determines the number of steps into the past needed for sufficient prediction of the current value x_t. By contrast, the current value of the moving average model of order q, abbreviated $MA(q)$, will be explained by the linear combination of q white noise values $\varepsilon_{(t-1)}, \ldots, \varepsilon_{(t-q)}$ from the past.

Definition 2.11 *An **autoregressive model of order p**, abbreviated **AR(p)**, is of the form*

$$x_t = \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \ldots + \varphi_p x_{t-p} + \varepsilon_t \ , \tag{2.15}$$

where $\{x_t\}$ is a stationary random process, $\varphi_1, \ldots, \varphi_p$ are constants ($\varphi_p \neq 0$). Moreover, $\{\varepsilon_t\}$ is assumed to be a Gaussian white noise series with mean zero and variance σ_ε^2. Thus, the mean of x_t in (2.15) is also zero. If the mean, μ, of x_t is not zero, replace x_t by $x_t - \mu$ in (2.15), i.e.,

$$x_t - \mu = \varphi_1(x_{t-1} - \mu) + \varphi_2(x_{t-2} - \mu) + \ldots + \varphi_p(x_{t-p} - \mu) + \varepsilon_t \ ,$$

or write

$$x_t = \varphi_0 + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \ldots + \varphi_p x_{t-p} + \varepsilon_t \ ,$$

where $\varphi_0 = \mu(1 - \varphi_1 - \ldots - \varphi_p)$.

Definition 2.12 *The **moving average model of order q**, or **MA(q)** model, is defined as*

$$x_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \tag{2.16}$$

where there are q lags in the moving average and $\theta_1, \ldots, \theta_q$ ($\theta_q \neq 0$) are parameters. The noise ε_t is assumed to be a Gaussian white noise.

It is also possible to combine both AR and MA models to a general $ARMA$ model, that provides in some situations much more precise representation of the stochastic process. Now, we introduce the general definition of autoregressive (AR), moving average (MA), and mixed autoregressive moving average (ARMA) models for the stationary time series.

Definition 2.13 *A time series $\{x_t\}_{t=0,1,2,\ldots}$ is **ARMA(p,q)** if it is stationary and*

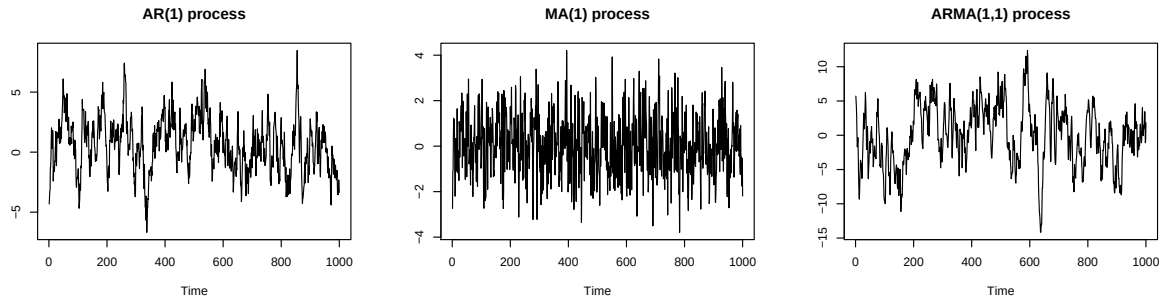
$$x_t = \varphi_1 x_{t-1} + \ldots + \varphi_p x_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \ , \tag{2.17}$$

with $\varphi \neq 0$, $\theta \neq 0$, and $\sigma_\varepsilon^2 > 0$. The parameters p and q are called the autoregressive and the moving average orders respectively. If x_t has a non-zero mean μ, we set $\varphi_0 = \mu(1 - \varphi_1 - \ldots - \varphi_p)$ and write a model as

$$x_t = \varphi_0 + \varphi_1 x_{t-1} + \ldots + \varphi_p x_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \ .$$

Unless stated otherwise, $\{\varepsilon_t\}$ with $t = 0, 1, 2, \ldots$, is a Gaussian white noise sequence.

Figure 2.1: Time series plot of an AR(1) ($\varphi = 0.9$), MA(1) ($\theta = 0.9$) and ARMA(1,1) ($\varphi = 0.9$, $\theta = 0.9$) processes sampled with a length of 1000



For $q = 0$, this model reduces to the autoregressive model of order p. And when $p = 0$, the model simplifies to the moving average model of order q. Note, stationarity of the time series poses a fundamental assumption for justification of the ARMA models. The ARMA models have a vital advantage against non-mixed AR or MA models that it is able to model an stationary time series process adequate involving fewer parameters as AR or MA models. Chatfield (2004, p. 47) calls this phenomenon a *Principe of Parsimony* that means, it is more effective to find a model with as few parameters as possible, but which gives an adequate representation of the data at hand.

Three sampled processes AR(1), MA(1) and ARMA(1,1) are displayed in Figure 2.1. As we can see, a time plot of time series alone gives us a little information about a type of the observed process. Some techniques, such as ACF and PACF plots, for detecting the kind of time series at hand are given in Section 2.6 "*Identifying ARMA-GARCH models*".

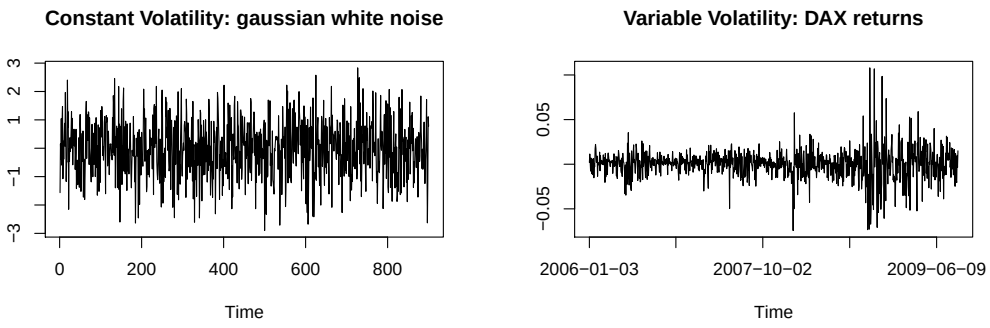
There are numerous methods for estimating and forecasting of ARMA models. More details for this subject may be found by Shumway and Stoffer (2006, chap. "*ARIMA Models*", pp. 110–140), Cryer and Chan (2008, chap. "*Parameter Estimation*", pp. 149–170 and chap. "*Forecasting*", pp. 191–213), Hamilton (1994) or Brockwell and Davis (1991, chap. "*Estimation for ARMA Models*", pp. 238–262 and chap. "*Prediction of Stationary Processes*", pp. 166–191).

The ARMA models are implemented in the R functions `arima` and `arma` from packages `stats` and `tseries` (Trapletti and Hornik (2009)) respectively.

2.3 ARCH and GARCH models

The ARMA models presented in the previous section were developed to model conditional mean structure and require a constant variance. Most financial time series are characterized by changing variance in time plot of the data. In such cases, it is conventional to talk

Figure 2.2: Time series with constant and variable volatility



about the *volatility* of the time series rather than about the variance. Volatility in a time series occurs in situations where the conditional variance of the time series varies over time. Under *volatility clustering* we understand the observation, that large changes tend to be followed by large changes and small changes tend to be followed by small changes.

Figure 2.2 displays plots of two time series with constant (left) and changing volatility (right) over the time period of length 900. Gaussian white noise is displayed on the left-hand side and DAX returns of the time period from 2006-01-03 to 2009-08-19 are displayed on the right-hand side of the figure.

Engle (1982) proposed to use *autoregressive conditional heteroscedasticity (ARCH)* model for modeling changing variance in the time series. The basic of his approach is that a current variance σ_t^2 of a time series $\{y_t\}$ is described by a linear function of squared lagged values of $\{y_t\}$. These models were later extended by Bollerslev (1986) to the generalized ARCH or the GARCH models. The advantage of GARCH models is that the current variance σ_t^2 is described not only by a function of squared lagged values y_t but also of itself lagged values.

Definition 2.14 *An* ***autoregressive conditionally heteroscedastic model of order m****, abbreviated* ***ARCH(m)****, is of the form*

$$\begin{aligned} y_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \omega + \alpha_1 y_{t-1}^2 + \ldots + \alpha_m y_{t-m}^2 \end{aligned} \tag{2.18}$$

where $\{\varepsilon_t\}$ *is a standard Gaussian white noise, i.e.* $\varepsilon_t \sim \mathcal{N}(0,1)$.

Definition 2.15 *An* ***generalized autoregressive conditionally heteroscedastic model of order (m,r)****, abbreviated* ***GARCH(m,r)****, is defined as*

$$\begin{aligned} y_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \omega + \sum_{j=1}^{m} \alpha_j y_{t-j}^2 + \sum_{j=1}^{r} \beta_j \sigma_{t-j}^2 \end{aligned} \tag{2.19}$$

where $\{\varepsilon_t\}$ *is a standard Gaussian white noise, i.e.* $\varepsilon_t \sim \mathcal{N}(0,1)$.

Error terms ε_t are also known as *innovation*. Generally, the innovation is a sequence of independently and identical distributed random variables with unit variance and zero mean. It may be shown, that the unconditional kurtosis of y_t is always larger that 3, see Tsay (2005, p.105) or Shumway and Stoffer (2006, p.282), and tail distribution of y_t is heavier than that of (standard) normal distribution.

Schoffer (2003) analyzed the of GARCH models for modeling of capital market returns. Such characteristics of financial time series as constant mean and volatility clustering can be well reproduced. In contrast, a typical high kurtosis of the financial time series can not be adequately simulated by this model type, although the kurtosis of conditional distribution allowed by GARCH model is somewhat larger than the kurtosis of the unconditional normal distribution. Another disadvantage of the GARCH process with the Gaussian white noise in modeling financial returns is that positive and negative returns $\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots$ have the same importance. This follows, since the volatility in (2.19) depends on the squared terms y_{t-j}, $j = 1, \ldots, m$. In practice, we observe so called *leverage effect*, that is bad news tends to have a larger influence on volatility than good news (Duffee (1995) and Black (1976)).

To handle the problems discussed above, non-Gaussian distributions for error terms ε_t in GARCH models have been proposed. Bollerslev (1987) suggested a GARCH model with Student's t-distribution for the standardized residuals. To capture both features of financial time series, skewness and kurtosis, Fernandez and Steel (1998) proposed the standardized skewed Student's t-distribution. (Quasi) Maximum Likelihood Estimation methods for GARCH models were precisely discussed by Straumann (2005).

The GARCH models are implemented in R in functions `garch` and `garchFit` of packages `tseries` and `fGarch` (Wuertz, with contribution from Michal Miklovic, Boudt, Chausse, et al. (2009)), respectively. The optional parameter `cond.dist` of function `fitGarch` allows several non-Gaussian conditional distributions for error terms ε_t such skewed normal (`dsnorm`) and (skewed) Student's t-distribution (`dstd` or `dsstd`). The default value is the standard normal distribution.

2.4 ARMA-GARCH models

The ARMA-GARCH models were developed to combine an ARMA model for describing the mean behavior and a GARCH model to reproduce the GARCH effect in the residuals from the ARMA model. We refer here to Lai and Xing (2008, pp. 155-156)

Definition 2.16 *The* ***ARMA(p,q)-GARCH(m,r)*** *model for* (y_t, σ_t) *where* $\{y_t\}$ *follows*

an ARMA model with GARCH innovations is defined as

$$\begin{aligned} y_t &= \mu + \sum_{i=1}^{p} \varphi_i y_{t-i} + \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} \\ \epsilon_t &= \sigma_t \varepsilon_t \ , \end{aligned} \tag{2.20}$$

where $\{\sigma_t\}$ *satisfies the recurrent equation*

$$\sigma_t^2 = \omega + \sum_{j=1}^{m} \alpha_j \epsilon_{t-j}^2 + \sum_{j=1}^{r} \beta_j \sigma_{t-j}^2 \ . \tag{2.21}$$

Error terms $\{\varepsilon_t\}$ *are i.i.d. with zero mean and unit variance.*

Equation (2.20) is called *conditional mean equation* and equation (2.21) is called *conditional variance equation.*

As in the case of simple GARCH model it is allowed for ε_t to follow standard normal, skewed standard normal, standard Student's t or skewed standard Student's t distribution. More details about these distributions are given in Appendix C. By modeling price time series, a GARCH model for volatility have to be combined with an ARMA model for mean, because expected price values are non-constant over the time period. By contrast, expected mean of return time series is (generally) constant and may be adequately estimated by the constant term μ of mean equation (2.20). So, the problem reduces to fitting a simple GARCH model. More details about MLE estimation of ARMA-GARCH models see follow references: Francq and Zakoïan (2004), Ling and Li (1997), Ling and Li Ling and Li (1998) or Ling and McAleer Ling and McAleer (2003).

The ARMA-GARCH models are implemented in R in function `garchFit` of the package `fGarch`.

2.4.1 Forecasting future mean and volatility

Given estimated parameters and using formulas (2.20) and (2.21), the one-step-ahead forecast for conditional mean and conditional volatility of y_t can be calculated by

$$\hat{y}_{t+1} = \mu + \sum_{i=1}^{p} \varphi_i y_{t+1-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t+1-i} \tag{2.22}$$

and

$$\hat{\sigma}_{t+1}^2 = \omega + \sum_{j=1}^{m} \alpha_j \epsilon_{t+1-j}^2 + \sum_{j=1}^{r} \beta_j \sigma_{t+1-j}^2 \ . \tag{2.23}$$

2.5 Partial autocorrelation function (PACF)

It may be shown, Shumway and Stoffer (2006, pp. 103–104), that autocorrelation function of MA(q) models will be zero after lag q and may not be zero till lag q inclusively. Hence, the ACF of some moving average process contains the whole information about its order. But there is another one situation by handling with ARMA or AR models. Namely, it can be evidenced, Shumway and Stoffer (2006, pp. 105–106), that ACF does not become zero after a certain number of lags for this model types. It dies off rather it cuts off. So, we need another function to help determine the order of an autoregressive model. Such a function is the *partial autocorrelation function (PACF)*. For more details, we refer to Shumway and Stoffer (2006, pp. 106–108). This function has similar behavior for AR models as the ACF for MA models. Namely, it will be zero after after lag p and is unequal to zero till lag p inclusively. The partial autocorrelation function at lag k measures a correlation between x_t and x_{t-k} after removing the effect of the variables $x_{t-1}, \ldots, x_{t-k+1}$ and will be denoted by ϕ_{kk}.

Definition 2.17 *The **partial autocorrelation function (PACF)** of a stationary process, x_t, denoted ϕ_{kk}, for $h = 1, 2, \ldots$, is*

$$\phi_{11} = cor(x_1, x_0) = \rho(1) \tag{2.24}$$

and

$$\phi_{hh} = cor(x_h - x_h^{h-1}, x_0 - x_0^{h-1}),\ h \geq 2\ . \tag{2.25}$$

Directly calculation of sample PACF at any lag is not possible, it can only be estimated. A detailed description of the estimation procedure based on the Yule-Walker equations is given by Cryer and Chan (2008). The estimator is called *sample partial autocorrelation function (sample PACF)* and will be denote it by $\hat{\phi}_{hh}$.

Quenoulle (1949) has shown that, under the hypothesis that the order of an AR(p) model is correctly defined, the sample partial autocorrelation at lags greater than p are approximately normally distributed with zero mean and variances $1/n$ for large n. To decide that the value of the PACF at lag h is zero, compare it with the standard deviation. If $|\hat{\phi}_{hh}| > 1.96\frac{1}{\sqrt{n}}$ for $h = p$ and $|\hat{\phi}_{hh}| \leq 1.96\frac{1}{\sqrt{n}}$ for $h > p$, then there is statistical evidence that the AR model is of order p.

The sample PACF can be computed and plotted in R with function `pacf` from the standard package `stats`. The blue dashed lines of PACF-plot indicates a confidence interval (95% by default) and allows a simple judgment about the statistical significance of $\hat{\phi}_{hh}$ at some lag h. In contrast to the `acf` function, the bars of the `pacf`-function start at lag 1.

2.6 Specification of ARMA-GARCH models

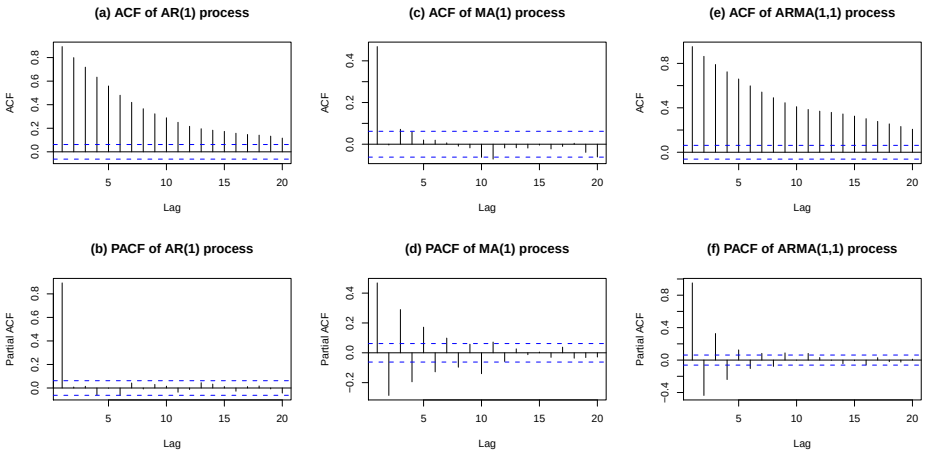
2.6.1 Order of ARMA models

The aim of this section is an investigation of some tools for choosing an appropriate order of AR, MA or ARMA models. In the previous section, we discussed the behavior of ACF and PACF functions on these time series processes. Following calculations given in Brockwell and Davis (1991), we get Table 2.1. This table can be very helpful for detecting a suitable order of pure AR or MA models. For an AR(p) process, the sample PACF cuts off after the lag p and the sample ACF tails off, see Figure 2.3, (a) and (b). For an MA(1) process, the sample ACF cuts off after the lag q and the sample PACF tails off, see panels (c) and (d) of Figure 2.3. The usage of ACF and PACF is not applicable for mixed ARMA models. It amiss in order identification order of ARMA model. Theoretical ACF and PACF have infinitely many non-zero values in this case, see panels (e) and (f) of Figure 2.3. Method based on *extended autocorrelation function (EACF)*, described by Cryer and Chan (2008, pp. 112–117), were developed for handling of this problem. Discussion of this method goes beyond the scope of this thesis.

Table 2.1: General behavior of autocorrelation and partial autocorrelation functions in AR, MA and ARMA processes

	AR(p)	**MA(q)**	**ARMA(p,q)**
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

Figure 2.3: ACF and PACF of simulated AR(1) ($\varphi = 0.9$), MA(1) ($\theta = 0.9$), ARMA(1,1) ($\varphi = 0.9$, $\theta = 0.9$) processes of length 1000



2.6.2 Order of GARCH models

Order specification of the GARCH models is generally not easy, since the volatility is not directly observable. In most applications, only the GARCH models of lower order such as GARCH(1,1), GARCH(2,1), and GARCH(1,2) are used. In many practical situation, the GARCH models of first order (1,1) provide adequate data explanation.

2.7 Model diagnostics for time series models

After the model fitting, it is necessary to check whether the model was correctly specified. Only when the model assumptions are fulfilled, it may be accepted and the results interpreted. Whether the assumptions of error terms such as distribution and independence are fulfilled should be verified by residuals analysis. Raw residuals

$$\hat{\varepsilon}_{t,raw} = y_t - \hat{y}_t \ , \tag{2.26}$$

where $\hat{y}_t$ are the fitted values and y_t are observed values, are not well suited to detect models misspecifications by the reason of conditional variance inequality. Hence, the standardized residuals

$$\hat{\varepsilon}_t = \hat{\varepsilon}_{t,stand} = \frac{\hat{\varepsilon}_{t,raw}}{\hat{\sigma}_{t|t-1}} \ , \tag{2.27}$$

where $\hat{\sigma}_{t|t-1}$ are estimated conditional volatilities from (2.21), have to be involved for checking the model specifications. If the ARMA(p,q)-GARCH(m,r) model is correctly specified, the standardized residuals $\{\hat{\varepsilon}_t\}$ should be approximately independently and identical distributed with assumed innovation distribution. These statements can be checked by following tools (that will be provided by functions `summary` and `plot` applied to an `garchFit` object):

(i) ***QQ-plot of the standardized residuals to examine the reliability of assumed distribution of error terms***: Under correct assumption, points of the QQ-plot should lie approximately on a straight line. If there is a significant deviation from the normal QQ-line, the normality assumption may not be appropriate. Symmetric anomalies in both tails could be corrected by choosing Student distribution in place of Gaussian. Discrepancy only in one tail could be liquidated with skewed normal or skewed t distributions.

(ii) ***Shapiro-Wilk or Jarque-Bera tests to validate normality of standardized residuals*** (Shapiro and Wilk (1965), Jarque and Bera (1987)): Both tests suggest normality of standardized residuals. Hence, if it holds, the p-values of test statistics χ^2 (Jarque-Bera test) and W (Shapiro-Wilk test) have to be larger as 0.05 for a conventional significance level $\alpha = 5\%$.

(iii) ***Time series plot of standardized residuals to verify the accuracy of estimated means and volatilities***: If estimated regression model holds, the standardized residuals have zero mean and unit variance. Furthermore, the majority of residuals lies between -2 and 2 and there is no pattern in the residual plot.

(iv) ***Ljung-Box test and ACF correlogramm of standardized residuals to validate uncorrelateness***: Firstly we define a lag-l sample autocorrelation of $\hat{\varepsilon}_t$ as

$$\hat{\rho}_l = \frac{\sum_{t=l+1}^{n} \hat{\varepsilon}_t \hat{\varepsilon}_{t-l}}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2} \, , 0 \leq l \leq n-1 \, .$$

As a test statistic for the null hypothesis

$$H_0 : \ \rho_l = \ldots = \rho_m$$

against the alternative hypothesis

$$H_1 : \ \rho_i \neq 0 \text{ for some } i \in \{1, \ldots, m\} \ ,$$

Box and Pierce (1970) suggested a Portmanteau statistic

$$Q^*(m) = n \sum_{l=1}^{m} \hat{\rho}_l^2 \ ,$$

that is asymptotically Chi-squared distributed with m degrees of freedom under assumption that $\hat{\varepsilon}_t$ is an i.i.d. sequence. Ljung and Box (1978) modified this $Q^*(m)$ statistic to $Q(m)$, defined as

$$Q(m) = n(n+2) \sum_{l=1}^{m} \frac{\hat{\rho}_l^2}{n-l} \ ;$$

with the motivation that $Q(m)$ is better approximated by χ^2_m than $Q^*(m)$. The decision rule is to reject H_0, if $Q(m) > \chi^2_{m,\alpha}$, where $\chi^2_{m,\alpha}$ denotes the $100(1-\alpha)$th percentile of a Chi-squared distribution with m degrees of freedom, or the p-value of the test statistic is less than accepted significance level $\alpha = 5\%$.

If the ARMA-GARCH model is successful at modeling the serial correlation structure in both terms, the conditional mean and the conditional variance, then there should be no autocorrelation left in the sequences of standardized residuals and squared standardized residuals.

Box-Pierce and Ljung-Box tests are implemented in R function `Box.test` from standard R package `stats`.

(v) ***Ljung-Box test and ACF correlogramm of squared standardized residuals to verify independence***: In the non-normality case, the uncorrelateness of residuals does not imply their independence. The idea of independence check is based on the fact that, in case of independence, not only $\hat{\varepsilon}_t$ and $\hat{\varepsilon}_{t+h}$ are uncorrelated, but also $f(\hat{\varepsilon}_t)$ and $f(\hat{\varepsilon}_{t+h})$ are uncorrelated for any function f (for example, $f(x) = x^2$). The Ljung-Box statistic for squared residuals will be denoted by $Q^2(m)$.

(vi) ***Engles's LM ARCH test for existence of ARCH effects***: The Lagrange multiplier test for ARCH effects was proposed by Engle (1982). It detects the remaining ARCH effects in the sequences of standardized residuals. The test statistic is given by TR^2, where R is the sample multiple correlation coefficient computed from the linear regression of $\hat{\varepsilon}_t$ on the squared $\hat{\varepsilon}_{t-1}, \ldots, \hat{\varepsilon}_{t-m}$, i.e.

$$\hat{\varepsilon}_t^2 = a_0 + \sum_{i=1}^{m} a_i \hat{\varepsilon}_{t-i}^2 \, , \tag{2.28}$$

and T is the sample size. Under the null hypothesis, which states that there are no ARCH effects of order m, i.e

$$H_0 : \; a_0 = a_{t-1} = \ldots = a_{t-m} \, ,$$

the test statistic is asymptotically Chi-square distributed with m degrees of freedom, i.e

$$TR^2 \sim \chi_m^2 \, .$$

The decision rule is to reject the null hypothesis, if $TR^2 > \chi_{m,\alpha}^2$, where $\chi_{m,\alpha}^2$ is the upper $100(1-\alpha)$th percentile of χ_m^2, or the p-value of the test statistic is less than accepted significance level $\alpha = 5\%$.

The LM ARCH test provides a good tool for checking the presence of ARCH effects in a time series of interest. On the other side, it can be used in residual analysis for checking the model specification. If there are no ARCH effects in the sequence of standardized residuals, the modeling of volatility, i.e. the order of ARCH model, can be assumed as adequate.

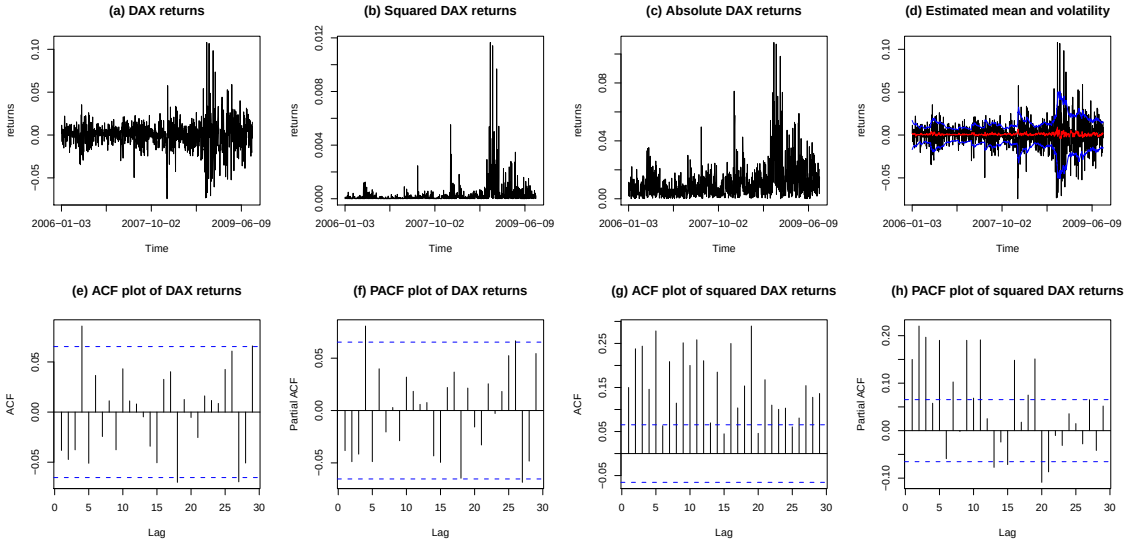
LM ARCH test is implemented in R in function `ArchTest` of package `FinTS` (Graves (2009)).

2.8 An illustrative example

We illustrate techniques mentioned in the previous section in an example. We analyze the data set of daily log-returns of the DAX German index, dating from January 3rd, 2006 to August 19th, 2009, in total 900 data points. Figure 2.4 (a) presents this return series. According to this graph, the volatility of the DAX returns is not constant over the time

period. The most volatile period is observed at the end of 2008. Volatility clustering may be more clearly seen by plotting the time sequence of squared returns or absolute returns, see Figure 2.4 (b) and (c). Graphs of sample ACFs and PACFs, panels (d) and (e) of Figure 2.4, do not exhibit significant peaks at any lags. In fact, they are close to zero. But it is well-known, that generally uncorelateness does not imply independence. If time series are truly independent, then non-linear, for example quadratic, transformation will also provide an uncorrelated sequence. Furthermore, looking at graphs of sample ACFs and PACFs of squared DAX returns, Figure 2.4 (g) and (h), we can detect some dependency in the squared returns despite of uncorrelateness. Both ACF and PACF functions tail off. The choice of an ARMA(1,1) model seems to be appropriate for modeling expected mean. A GARCH model of the order (1,1) will be used for modeling conditional variance.

Figure 2.4: Time series plot with estimated mean and volatility, sample ACF and sample PACF of DAX returns



Four ARMA(1,1)-GARCH(1,1) model of the form

$$y_t = \mu + \varphi y_{t-1} + \epsilon_t + \theta\epsilon_{t-1}, \qquad \epsilon_t = \sigma_t \varepsilon_t$$
$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

with different kinds of ε_t distribution (standard normal, skewed standard normal, standardized Student't and skewed standardized t) were fitted. Table 2.2 contains:

- Set of estimated time series model parameters (μ, φ, θ, ω, α, β)
- Estimated parameters (λ, ν) of error terms distribution
- Results of model checking such as the test statistics of Shapiro-, Wilk-, Ljung-Box and ML ARCH tests and the p-values of theses statistics

Graphical analysis of the standardized residuals as a QQ-plot, plot of standardized resid-

Table 2.2: Summaries of estimated ARMA(1,1)-GARCH(1,1) models fitted to the daily log-returns of German DAX index with four different kinds of error terms distribution: standard normal in model (a), skewed standard normal in model (b), standard Student and standard skewed Student in models (c) and (d), respectively. $\hat{\mu}$, $\hat{\varphi}$ and $\hat{\theta}$ are the estimated parameters of the ARMA part; $\hat{\omega}$, $\hat{\alpha}$, $\hat{\beta}$ are the estimated parameters of the GARCH part; $\hat{\lambda}$ and $\hat{\nu}$ are the estimated shape and skew distribution parameters. Residual analysis is based on: Shapiro-Wilk χ^2 and Jarque-Berra W test statistics for testing normality; Ljung-Box $Q(15)$ statistics for standardized residuals and Ljung-Box $Q^2(15)$ statistic for squared standardized residuals for testing uncorrelateness and independent; Engle's ARCH test statistic TR^2 for testing remaining ARCH effects

	MODEL	ARMA			GARCH			skew	shape	RESIDUALS ANALYSIS				
		$\hat{\mu}$	$\hat{\varphi}$	$\hat{\theta}$	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\nu}$	χ^2	W	Q(15)	$Q^2(15)$	TR^2
	stand. normal													
(a)	Est./Stat.	0.00	0.77	-0.82	0.00	0.13	0.86			250.86	0.97	18.61	11.78	12.81
	p-value	0.21	0.00	0.00	0.00	0.00	0.00			0.00	0.00	0.23	0.70	0.62
	skewed stand. normal													
(b)	Est./Stat.	0.00	0.72	-0.81	0.00	0.11	0.88	0.79		255.87	0.97	30.56	12.04	12.93
	p-value	0.06	0.00	0.00	0.01	0.00	0.00	0.00		0.00	0.00	0.01	0.68	0.61
	stand. Student't													
(c)	Est./Stat.	0.00	-0.97	0.95	0.00	0.10	0.90		6.33	388.95	0.97	9.88	12.22	13.27
	p-value	0.69	0.00	0.00	0.06	0.00	0.00		0.00	0.00	0.00	0.83	0.66	0.58
	skewed stand. Student													
(d)	Est./Stat.	0.00	0.60	-0.66	0.00	0.10	0.90	0.85	6.90	367.75	0.97	18.08	12.07	12.82
	p-value	0.17	0.02	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.26	0.67	0.62

uals and sample ACF of standardized residuals and squared standardized residuals are displayed on Figure 2.5. The DAX returns with estimated mean and estimated volatility are shown on the panel (d) of Figure 2.4. One can see, that ARMA(1,1)-GARCH(1,1) model reproduce expected values and conditional volatilities quite well.

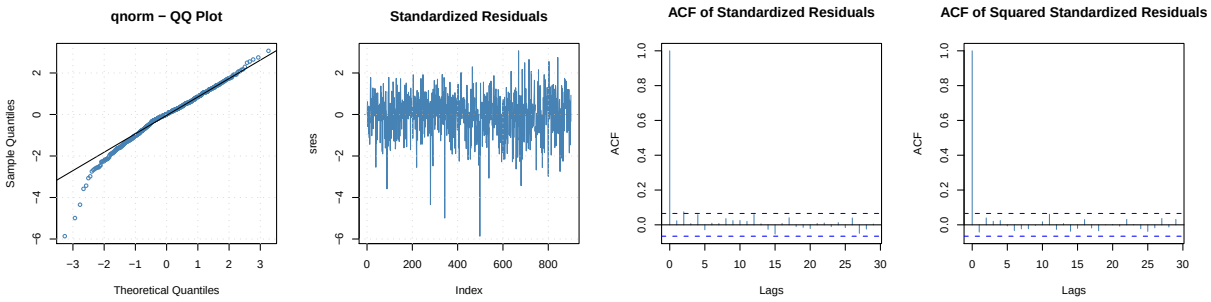
To sum up:

- Based on the p-values of Shapiro- and Wilk-test statistics obtained for the standardized residuals $\hat{\varepsilon}_t$, one can reject the null hypothesis of its normality for each significance level α. So, standardized residuals do not follow standard normal distribution. It holds for each of four models (a)-(d).

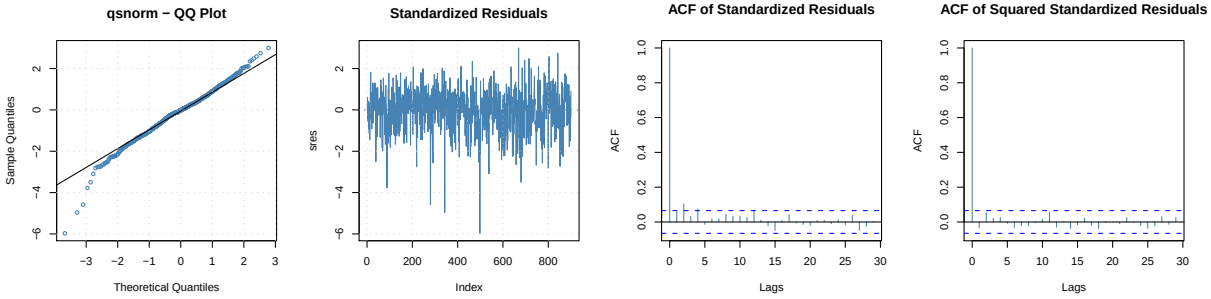
- Next, we identify an appropriate noise distribution. Only the QQ-plot of $\hat{\varepsilon}_t$ coming from the model with skewed Student's t innovations lie (approximately) on a straight line, except of the several points in the lower tail. The points of the QQ-plots for other models do not resemble a straight line. They indicate non-normal distribution with fat tails. The shape and skew parameters, given in Table 2.2, are significant at level $\alpha = 5\%$. Therefore, the skewed standard t distribution for error terms is the most adequate distribution.

- Next, we test whether $\{\hat{\varepsilon}_t\}$ of model (d) follow a white noise process, i.e. it is a i.i.d. sequence characterized by zero mean and unit variance. Based on the p-values of $Q(15)$ and $Q^2(15)$ Ljung-Box test statistics, that are larger then the accepted significance level $\alpha = 5\%$, the null hypothesis of independence in a sequence of standardized residuals can not be rejected. It justifies the sample ACF function

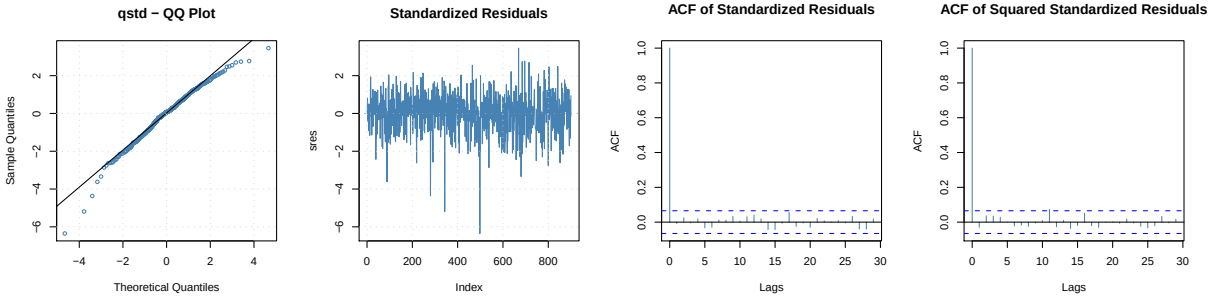
Figure 2.5: Residual analysis of the ARMA(1,1)-GARCH(1,1) models with four different kinds of innovation distributions fitted to the DAX daily log-returns: QQ-plots in the left column, time series plot of the standardized residuals in the second column, sample autocorrelation functions of standardized residuals in the third column and sample autocorrelation functions of squared standardized residuals in the last column



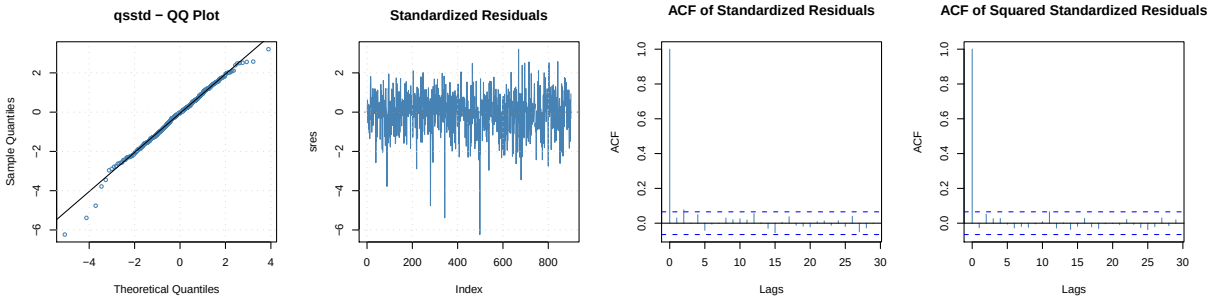
(a) Error terms follow standard normal distribution $\varepsilon_t \sim \mathcal{N}(0,1)$



(b) Error terms follow standard skewed normal distribution $\varepsilon_t \sim \mathcal{SN}(0,1,\lambda)$



(c) Error terms follow standard Student's t distribution $\varepsilon_t \sim t_\nu^{stand}$



(d) Error terms follow standard skewed Student's t distribution $\varepsilon_t \sim t_{\nu,\lambda}^{stand}$

of standardized residuals and squared standardized residuals, that lie not over the critical value (dashed line). Moreover, the time series plot of standardized residuals looks quite well: the majority of all observations lies between $+2$ and -2 and fluctuates about zero. According to these facts, we can not reject the hypothesis that standardized residuals $\{\hat{\varepsilon}_t\}$ form a strong white noise.

- The p-values of TR^2 statistics are much more higher then $\alpha = 5\%$ in all four models. So, there are no remaining ARCH effects in the sequences of standardized residuals and the GARCH models of order (1,1) are suffucient by modeling conditional volatility.

- The GARCH parameters α and β, except of overall parameter ω in models (c) and (d), are significant at level $\alpha = 5\%$ in all four models. The ARMA parameters φ and θ, except of overall parameter μ, are also significant.

Chapter 3

Modeling multivariate dependence: correlation coefficients, tail dependence, copulas and pair-copula constructions

At the beginning of this chapter, we introduce several dependence coefficients and discuss their benefits and disadvantages. Next, we present the notion of tail-dependence that is of vital importance in financial analysis. We show how the presence of tail-dependence in bivariate data can be detected using graphical tools or using non-parametric estimators. Further, we introduce most common parametric copula families and their properties. In addition, we study pair-copula constructions that are more suitable in modeling multivariate dependences in comparison to the multivariate copulas. Finally, we provide an algorithm for modeling multivariate returns that combines time series and pair-copula approaches.

3.1 Moments and their estimation

In this section we introduce moments, which are important characteristics of univariate random variables.

Definition 3.1 (Moments) *Let X be continuous random variable with distribution function F and density f. Its* ***moment of order k****, if it exists, is given by*

$$E[X^k] = \int_{-\infty}^{\infty} x^k \mathrm{d}F(x) = \int_{-\infty}^{\infty} x^k f(x) \mathrm{d}x \,. \tag{3.1}$$

The first moment

$$\mu = E[X] = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x$$

is called *mean*, *expected value* or *expectation* and will be denoted by μ. It provides the average value for the distribution of the random variable. If $x_1, \ldots, x_n$ are observations of X, then the kth moment of X can be estimated by

$$\hat{E}[X^k] = \frac{1}{n} \sum_{i=1}^{n} x_i^k \, .$$

Definition 3.2 (Central Moments) *The **k-th central moment** of a random variable X, provided it exists, is defined as*

$$\mu_k = E\left[(X-\mu)^k\right] = \int_{-\infty}^{\infty} (x-\mu)^k \mathrm{d}F(x) = \int_{-\infty}^{\infty} (x-\mu)^k f(x) \mathrm{d}x \, . \tag{3.2}$$

The second central moment of X

$$Var(X) = \sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) \mathrm{d}x$$

is called *variance* and will be denoted by $\sigma^2 = \mathrm{Var}(X)$. The variance gives information about the dispersion of the distribution around the expected value. Lower value of variance indicates that the distribution is concentrated close to the mean. In contrast, high value indicates that the distribution is spread out over a wider range of possible values. Thus, the variance is a scale parameter of the distribution. The square root of variance is called *standard deviation* and will be denoted by σ.

Another important distribution characteristics are based on the moments skewness and kurtosis. The *skewness* of a distribution indicates its asymmetry relative to mean. It is given by

$$\gamma = \frac{\mu_3}{\sigma^3} \, .$$

Symmetric distributions have skewness equal to zero, for example Normal and Student's t distributions. Negative value indicates that the distribution is shifted to the left. Positive value implies the opposite. The *kurtosis* of a distribution indicates its flatness with respect to the normal distribution and is a measure for probability of extreme outcomes. It is given by

$$\kappa = \frac{\mu_4}{\sigma^4} - 3 \, .$$

The normal distribution has a kurtosis of zero. Positive value of kurtosis indicates that the distribution is flatter and has fatter tails compared to the normal distribution. A negative kurtosis shows that extreme events are even less probable in comparison to the normal case and the distribution has a higher peak around the mean value.

3.2 Correlation coefficients

In this section, we discuss different types of correlation coefficients, their advantages and drawbacks. References for this part are Kurowicka and Cooke (2006, pp. 26–32) and Genest and Favre (2007).

Definition 3.3 (Independence) *Let $F_{X_1 \ldots X_n}$ be a joint cumulative distribution function (cdf) and $f_{X_1 \ldots X_n}$ a corresponding probability distribution function (pdf) of random variables $X_1, \ldots, X_n$ with marginal distributions $F_{X_1}, \ldots, F_{X_n}$ respectively. The random variables $X_1, \ldots, X_n$ are said to be* ***independent*** *if and only if*

$$F_{X_1 \ldots X_n}\left(X_1 \leq x_1, \ldots, X_n \leq x_n\right) = \prod_{i=1}^{n} F_{X_i}\left(X_i \leq x_i\right) \ , \tag{3.3}$$

or equivalently

$$f_{X_1 \ldots X_n}\left(x_1, \ldots, x_n\right) = \prod_{i=1}^{n} f_{X_i}\left(x_i\right) \ , \tag{3.4}$$

Definition 3.4 (Pearson's correlation) *The* ***Pearson's correlation coefficient*** *of two random variables X and Y with finite expectations $E[X]$, $E[Y]$ and finite variances $Var(X)$, $Var(Y)$ is defined as*

$$r\left(X, Y\right) = \frac{Cov(X, Y)}{\sqrt{Var(X)\ Var(Y)}} = \frac{E\left[XY\right] - E\left[X\right] E\left[Y\right]}{\sqrt{Var(X)\ Var(Y)}} \in [-1, 1] \ . \tag{3.5}$$

The Pearson's correlation is also called ***product moment correlation****.*

Given n realizations of X and Y, $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$, the Pearson correlation can be estimated by a sample correlation coefficient, i.e.

$$\hat{r}\left(X, Y\right) = \frac{\sum_{i=1}^{n} \left(x_i - \hat{\mu}_X\right)\left(y_i - \hat{\mu}_Y\right)}{\hat{\sigma}_X \hat{\sigma}_Y} \ ,$$

where $\hat{\mu}_X$ and $\hat{\mu}_Y$ are sample means, $\hat{\sigma}_X$ and $\hat{\sigma}_Y$ are sample standard deviations of X and Y, respectively.

We summarize the properties of the Pearson's correlation coefficient:

- X and Y independent $\Rightarrow r = 0$
- $r = 0 \quad \not\Rightarrow \quad X$ and Y independent
- Only if X and Y are bivariate normal $r = 0$, implies independence (see Appendix C)
- r requires the existence of $\mathrm{Var}(X)$ and $\mathrm{Var}(Y)$
- r measures only the *linear* dependence
- r is not invariant under monotone transformations
- If $Y = a + bX$, then $|r| = 1$.

The Pearson's correlation coefficient is a parametric statistic, it requires normality of the data. Another one drawback of r is that it can capture only a linear relationship. In practice, however, sampled data do not fulfill these criteria. An alternative in such situations is to use non-parametric techniques based on ranks. The notion *non-parametric* means that the dependence measure does not depend on the marginal probability distributions. There are two commonly used rank based methods, one due to Spearman, so called Spearman's ρ, and one due to Kendall, so called Kendall's τ. The most important feature of the rank based methods is their invariance under monotonic transformations. The Spearman's ρ, is exactly the same as the Pearson correlation r calculated on the ranks of the observations. For a definition of Kendall's τ, we need such notions as *concordant* and *discordant* pairs. Two pairs (X_1, Y_1) and (X_2, Y_2) are said to be concordant when $(X_1 - X_2)(Y_1 - Y_2) > 0$, and discordant when $(X_1 - X_2)(Y_1 - Y_2) < 0$.

Definition 3.5 (Kendall's τ) *Let (X_1, Y_1) and (X_2, Y_2) be two independent pairs of random variables with joint distribution F and marginal distributions F_X and F_Y. The* ***Kendall's*** *τ is given by*

$$\tau(X, Y) = P((X_1 - X_2)(Y_1 - Y_2) > 0) - P((X_1 - X_2)(Y_1 - Y_2) < 0) \in [-1, 1]\,. \quad (3.6)$$

For a bivariate sample (X, Y) of length n, τ may be estimated using an unbiased estimator $\hat{\tau}$, see Nelsen (2006, pg. 158)

$$\hat{\tau}(X, Y) = \frac{P_n - Q_n}{\binom{n}{2}}\,,$$

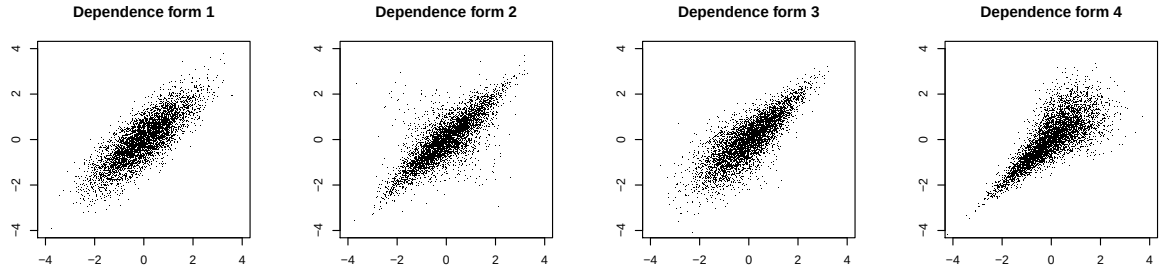
where P_n and Q_n are the number of concordant and discordant pairs, respectively.

There are some properties of Spearman's ρ and Kendall's τ:

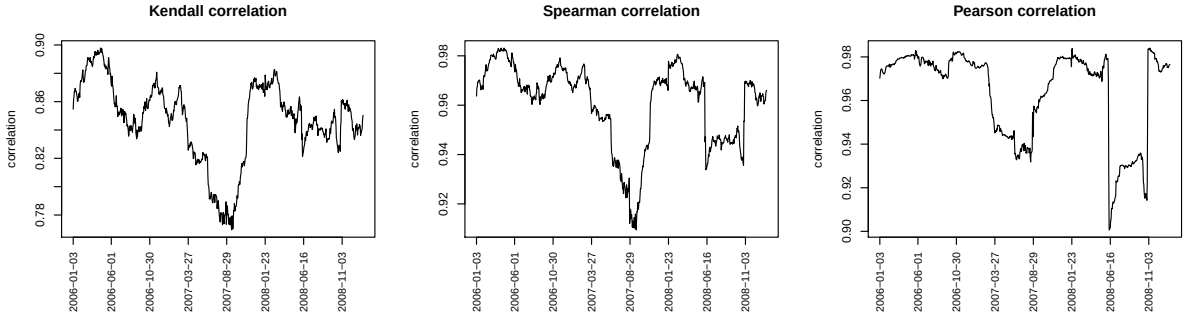
- X and Y independent $\Rightarrow \tau = 0$, $\rho = 0$
- $\tau(X, Y) = 1 \Leftrightarrow Y = T(X)$ with $T(\cdot)$ any increasing map (the same for ρ)
- $\tau(X, Y) = -1 \Leftrightarrow Y = T(X)$ with $T(\cdot)$ any decreasing map (the same for ρ)
- τ, ρ are invariant under monotone transformations
- τ, ρ measure the degree of monotonic dependence.

Figure 3.1 exhibits graphically some major disadvantages of correlation coefficients. Upper panel presents four random vector pairs that are characterized by significant different forms of the point clouds and same degree of association $\tau = 0.6$. On the basis of this

Figure 3.1: Drawbacks of describing dependence structure using correlation coefficients



(a) Different dependence forms with the same Kendall's correlation coefficient, $\tau = 0.6$



(b) Neither of correlation coefficients due to Kendall, Spearman and Pearson is constant over the time period

plots, it is clear that correlation measures only a degree of dependence but is unable to reflect it's structure. Lower panel shows a trend of correlation coefficients τ (left plot), ρ (middle plot) and r (right plot) of daily DAX and STOXXER returns over time period from 2006-01-02 to 2009-06-18 calculated for the time windows of 100 days. These plots illustrate a time-varying correlation phenomenon. In the next section we will examine other qualities of financial time series that can not be captured by correlation coefficients. Hence, another one measure is needed for complete description of dependency structure.

3.3 Tail dependence

Several empirical studies have shown that empirical joint distribution of return-pairs of stock indices is not symmetric. For example, Longin and Solnik (1995, 2001) have shown that correlation between stock return series tends to be higher in market downturns than in market upturns. Rather, it arises in high tail dependence in the lower-quadrant and lower tail dependence in the upper-quadrant of a bivariate distribution. This fact can not account standard symmetric models of multivariate stock returns. In many applications such as risk management, the presence of tail dependence is of vital importance. Below we will study a phenomenon of tail dependence quite precisely according to Nelsen (2006).

Definition 3.6 (Upper tail dependence) *Let X and Y be continuous random variables with distribution functions F_X and F_Y, respectively. The* ***upper tail dependence parameter*** *λ_U is the limit (if it exists) of the conditional probability that Y is greater than the $100t$-th percentile of F_Y, $F_Y^{-1}(t)$, given that X is grater than the $100t$-th percentile of F_X, $F_X^{-1}(t)$, as t approaches 1, i.e.*

$$\lambda_U = \lim_{t \to 1^-} P\left(Y > F_Y^{-1}(t) \mid X > F_X^{-1}(t)\right) . \tag{3.7}$$

Definition 3.7 (Lower tail dependence) *Similarly, the* ***lower tail dependence parameter*** *λ_L is the limit (if it exists) of the conditional probability that Y is less than or equal to the $100t$-th percentile of F_Y, $F_Y^{-1}(t)$, given that X s less than or equal to the $100t$-th percentile of F_X, $F_X^{-1}(t)$, as t approaches 1, i.e.*

$$\lambda_L = \lim_{t \to 0^+} P\left(Y \leq F_Y^{-1}(t) \mid X \leq F_X^{-1}(t)\right) . \tag{3.8}$$

Let $F_{X,Y}$ denote a joint distribution of random variables X and Y. Therefore the tail dependence can be interpreted as a measure for extreme co-movements in the lower and upper tails of $F_{X,Y}$, respectively. If $\lambda_U = 0$, we say that there is no upper tail dependence, and if $\lambda_L = 0$ we say that there is no lower tail dependence. In case of $\lambda_U = \lambda_L \neq 0$, we speak about tail dependence and denote it by λ.

The upper and lower tail dependence can be estimated empirically using formulas given by Čížek, Härdle, and Weron (2005, pp. 75-77)

$$\hat{\lambda}_{U,n} = \frac{1}{k} \sum_{i=1}^{n} \mathbb{1}\left(R(X_i) > n - k, R(Y_i) > n - k\right) \tag{3.9}$$

and

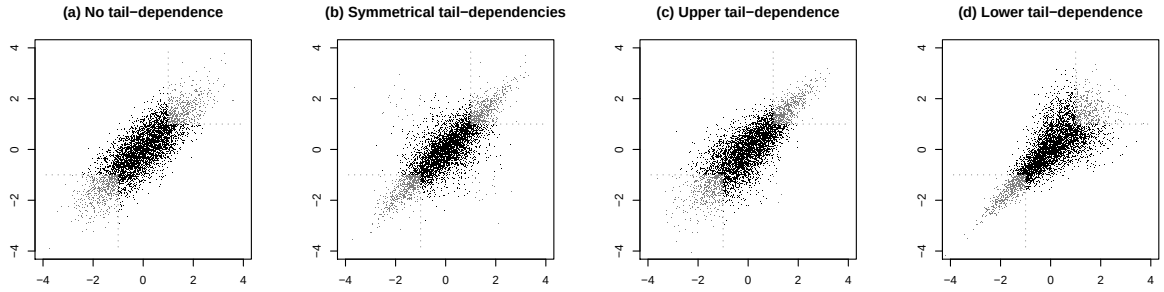
$$\hat{\lambda}_{L,n} = \frac{1}{k} \sum_{i=1}^{n} \mathbb{1}\left(R(X_i) \leq k, R(Y_i) \leq k\right) , \tag{3.10}$$

where (X_i, Y_i) is an observed bivariate data of length n, $R(x_i)$ is a rank of X_i in vector $X_1, \ldots, X_n$; k is chosen to satisfy following conditions

$$k = k(n), \ \frac{k}{n} \xrightarrow{n \to \infty} 0 . \tag{3.11}$$

The trivial graphical tool for detecting lower and/or upper tail dependence is a classical scatter plot. Figure 3.2 (a)-(d) demonstrates lower and upper tails of four bivariate data sets with the same positive correlation $\tau = 0.6$. As we can see, the points clouds are remarkable for significantly different patterns in the lower and upper quadrants. The images (a) and (b) are symmetrical in contrast to the images (c) and (d). There are distinguishable tail dependences on panel (b). The image (a) does not indicate any art

Figure 3.2: Lower and upper tails of different dependence forms with the same coefficient of association $\tau = 0.6$



of dependence in both tails. Upper and lower tails on the map (c) as well as on the map (d) are dissimilar. On the plot (c), clear upper tail but no lower tail dependence can be recognized. And vice versa for the panel (d): there is a strong lower tail but no upper tail dependence.

A scatter plot is not appropriate for small sample size or for weak pronounced lower and upper tail dependences. In the next section, we introduce another one graphical method for exploring lower or upper tail dependence.

3.4 Graphical tools for detecting (tail) dependence

Sometimes, dependence between two random variables is hard to detect using scatter plot of observation points. In this section, two additional graphical tool for assessing dependence will be described, Chi-plots and K-plots. Moreover, it will be shown how Chi-plot can be employed for detecting a kind of tail dependence, upper or lower. The Chi- and the K-plots will be later used to assess the (tail) dependence structure between pairs of transformed return series and for choosing a most suitable copula building block in a pair-copula construction.

3.4.1 Chi-plots

Chi-plots were first introduced by Fisher and Switzer (1985, 2001) and are based on the chi-square statistic for independence in a two way table. By presenting this section, we refer to the article *"Everything you always wanted to know about copula modeling but were afraid to ask"* of Genest and Favre (2007).

Let $(x_1, y_1), \ldots, (x_n, y_n)$ be a random sample for a pair of random variables (X, Y) with a joint distribution F_{XY} and marginal distributions F_X and F_Y, respectively. We estimate

F_{XY}, F_X and F_Y by empirical cumulative distribution functions

$$\hat{F}_{XY,i} = \hat{F}_{XY,i}(x_i, y_i) = P\left(j \neq i : X_j \leq x_i, Y_j \leq y_i\right) = \frac{\#\left\{j \neq i : x_j \leq x_i, y_j \leq y_i\right\}}{n-1} \quad (3.12)$$

$$\hat{F}_{X,i} = \hat{F}_{X,i}(x_i) = \frac{\#\left\{j \neq i : x_j \leq x_i\right\}}{n-1} \quad (3.13)$$

$$\hat{F}_{Y,i} = \hat{F}_{Y,i}(y_i) = \frac{\#\left\{j \neq i : y_j \leq y_i\right\}}{n-1} \quad (3.14)$$

Under independence, it holds $\hat{F}_{XY,i} \approx \hat{F}_{X,i} \cdot \hat{F}_{Y,i}$ or $\hat{F}_{XY,i} - \hat{F}_{X,i} \cdot \hat{F}_{Y,i} \approx 0$ for all $i = 1, \ldots, n$. The authors proposed to transform observed data $(x_1, y_1), \ldots, (x_n, y_n)$ into pairs $(\lambda_1, \chi_1), \ldots, (\lambda_n, \chi_n)$ with

$$\chi_i = \frac{\hat{F}_{XY,i} - \hat{F}_{X,i}\hat{F}_{Y,i}}{\sqrt{\hat{F}_{X,i}(1-\hat{F}_{X,i})\hat{F}_{Y,i}(1-\hat{F}_{Y,i})}} \quad (3.15)$$

$$\lambda_i = 4\text{sgn}\left(\tilde{F}_{X,i}, \tilde{F}_{Y,i}\right) \cdot \max\left(\tilde{F}^2_{X,i}, \tilde{F}^2_{Y,i}\right) , \quad (3.16)$$

where $\tilde{F}_{X,i} = \hat{F}_{X,i} - 0.5$ and $\tilde{F}_{Y,i} = \hat{F}_{Y,i} - 0.5$ for $i = 1, \ldots, n$.

Fisher and Switzer argued following statements:

- Chi-plot depends only on the ranks of the data
- λ_i and χ_i are scaled to the interval $[-1, 1]$ for all $i = 1, \ldots, n$
- λ_i measures a distance of a data point (x_i, y_i) to the center of bivariate data set
- χ_i accords to a correlation coefficient between dichotomized values of X and Y
- $\chi_i \sim \mathcal{N}(0, \frac{1}{n})$ and $\lambda_i \sim \mathcal{U}[-1, 1]$ under independence.

Further, values of χ_i close to zero indicate independence. For positively correlated margins, the pairs of (λ_i, χ_i) will tend to be located above the upper band, and vice versa for negatively correlated margins. To help identify whether the values χ_i lie close enough to zero, Fisher and Switzer suggested to draw so called control bounds. Pairs (λ_i, χ_i) lie beyond of the bound $\pm c_p/\sqrt{n}$ indicate a presence of dependence between X and Y. The factor c_p should be selected so that about $100p\%$ of the pairs (λ_i, χ_i) lie between lower and upper lines. For $p = 0.05$, Fisher and Switzer found the corresponding value $c_p = 1.78$ through simulations.

3.4.2 Kendall's process plots

The Kendall-process-plot (or K-plot) is another tool for detecting dependence in a bivariate data and based on the ranks. It was recently introduced by Genest and Boies (2003), we refer here to Genest and Favre (2007). Genest and Boies suggested to transform the pairs of data (x_i, y_i) to those $\left(W_{i:n}, \hat{F}_{XY,i:n}\right)$ for all $i = 1, \ldots, n$, where $\hat{F}_{XY,i:n}$ are the

order statistics of the terms $\hat{F}_{XY,i}$ introduced in (3.12). The $W_{i:n}$ are the expectations of the i^{th} order statistics from the random sample of $W_i = \hat{F}_{XY}(x_i, y_i)$ of size n under the null hypothesis of independence between X and Y. The $W_{i:n}$ can be calculated as follows

$$W_{i:n} = n\binom{n-1}{i-1} \int_0^1 \omega k_0(\omega) \left(K_0(\omega)\right)^{i-1} \left(1 - K_0(\omega)\right)^{n-i} d\omega \,, \tag{3.17}$$

where

$$K_0(\omega) = \omega - \omega \log(\omega)$$

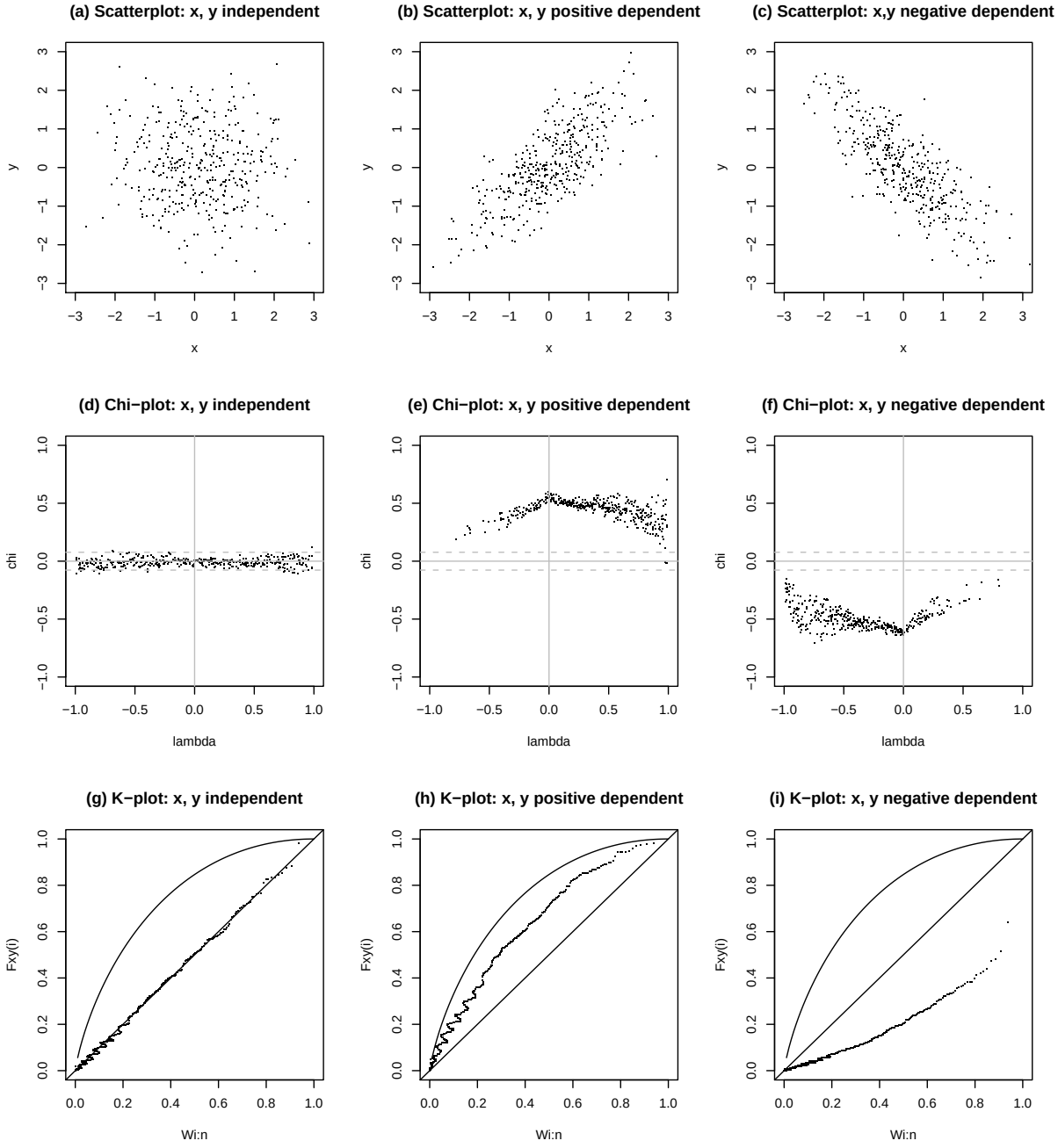
and $k_0(\cdot)$ is the corresponding density.

K-plot can be interpreted similar to the QQ-plots. If the points of a K-plot lie approximately on the diagonal $y = x$, then independence between X and Y can be assumed. Any deviation from the diagonal line points towards dependence. In case of the positive dependence, the points of the K-plot should be located above the diagonal line, and vice versa for negative association. The farther deviation from the diagonal, the greater is the degree of dependency. There is a perfect positive dependence, if points $\left(W_{i:n}, \hat{F}_{XY,i:n}\right)$ lie on the bent curve $K_0(\omega)$ that is located above the main diagonal. Positioning points $\left(W_{i:n}, \hat{F}_{XY,i:n}\right)$ on the x-axis indicates a perfect negative dependence between X and Y.

3.4.3 Examples of Chi- and K-plots

Figure 3.3 displays Chi- and K-plots created for independent, positive dependent and negative dependent data. Panel (a) is a scatter plot of an independent bivariate data set of length $n = 400$ with standard normal distributed margins. Panel (d) shows a Chi-plot and panel (g) a K-plot of this data. All points of the Chi-plot lie between control bounds that indicates independence. This fact is confirmed by the K-plot where points are located on a diagonal line. Next, panel (b) represents a scatter plot of a positive dependent data coming from bivariate standard normal distribution with dependence parameter $r = 0.8$. Panels (e) and (h) illustrate corresponding Chi- and K-plots. Dots of the Chi-plot lie above the upper control line and the points of the K-plot scatter above the diagonal line. Both facts indicate a presence of positive dependence. At last, we look at data coming from bivariate standard normal distribution with negative coefficient of linear dependence $r = -0.8$. The scatter plot(c) of this data displays positive association between X and Y that is confirmed by Chi-plot. Points of the K-plot follow a curve that deviates from the main diagonal and are located below it. The points of a Chi-plot lie as well below the lower control bound.

Figure 3.3: Examples of Chi- and K-plots: panels (a), (d) and (g) displays scatter, Chi- and K-plots for independent data respectively; panels (b), (e) and (h) displays scatter, Chi- and K-plots for positive dependent data respectively; panels (c), (f) and (i) displays scatter, Chi- and K-plots for negative dependent data respectively



3.4.4 Chi-plots for detecting tail dependences

Fisher and Switzer introduced a graphical method for assessing dependence in a bivariate data set that does not discuss a problem of tail dependence. This section is based on the work of Abberger (2004). Abberger has shown how presence of tail dependence in a bivariate data set and it's kind, upper or lower, can be detected with the help of Chi-plots.

As noted in the previous section, λ_i measures a distance of a data point (x_i, y_i) to the center of a bivariate data set. Data pairs (x_i, y_i) with λ_i close to -1 lie not far from a data center. Further, data points characterized by large values of λ_i close to $+1$ are far away from a data center. Hence, for assessing the presence of tail dependence we are interested in observations with positive λ_i's. According to Abberger, if there is no tail dependence, the χ_i values have to return to the zero line at the right edge of Chi-plot, i.e. for the λ values close to $+1$. And vice versa, a presence of tail dependence causes a deviation of χ values from the zero line for the λ values closed to $+1$.

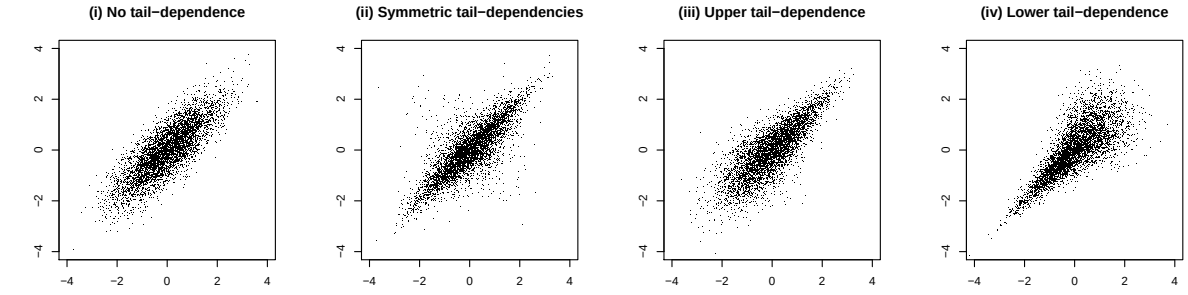
To check for tail dependences Abberger suggested to compute Chi-plots only for observations with positive λ_i values that lie separately in two quadrants, in the upper right corner and in the lower left corner of a bivariate scatter plot. Such separation enables us to recognize asymmetries in the tails dependences. If there is no upper or lower tail dependence the χ values rightmost of the Chi-plot have to return to the zero line.

Figure 3.4 illustrates the usefulness of Chi-plots for recognition of tail dependencies. Upper panel (a) shows scatter plots of bivariate data sets characterizing by different forms of dependency. Middle panel (b) contains corresponding Chi-plots of total data sets. Panel (c) exhibits Chi-plots of observations from upper right corner of scatter plot and panel (d) presents Chi-plots of observations from lower left corner of scatter plot. Chi-plots of the column (i) converge to the zero horizontal line on the right hand side which indicates no upper and no lower tail dependence. The column (ii) is the opposite to (i). Namely, the Chi-plot points for positive λ values that are closed to 1 do not incline to the zero line. That is an indication that of both tail dependences, upper and lower, are present. The Chi-plots (iii) and (iv) of panel (b) have rotated bell-shaped forms on the right hand side that points to a presence of tail dependence. The Chi-plots (iii) and (iv) make it clear what kinds of tail dependence are on hand. We can identify asymmetrical forms of dependence for upper and lower tails: there is an upper tail dependence for column (iii) and a lower tail dependence for column (iv).

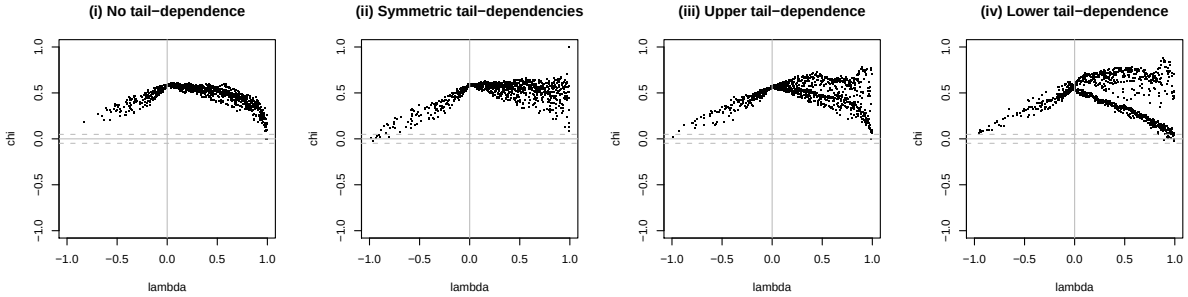
3.5 Copulas

In the previous section, several approaches for measuring dependence between two random vectors based on ranks of observed data were explored. In this section, we introduce a modern powerful technique, a copula concept, for modeling dependencies in the mul-

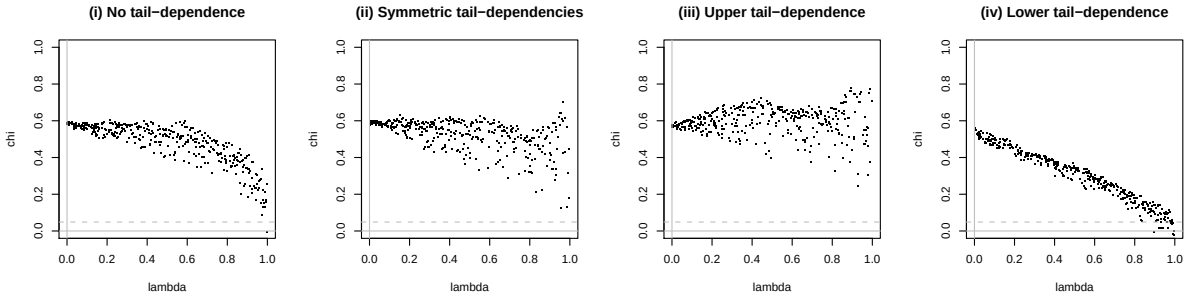
Figure 3.4: Illustration of detecting tail dependency and it kind, lower or upper, with the help of the Chi-plots: (a) bivariate scatter plots with different dependency forms; (b) Chi-plots for detecting tail dependency; (c) Chi-plots for detecting upper tail dependence; (d) Chi-plots for detecting lower tail dependence



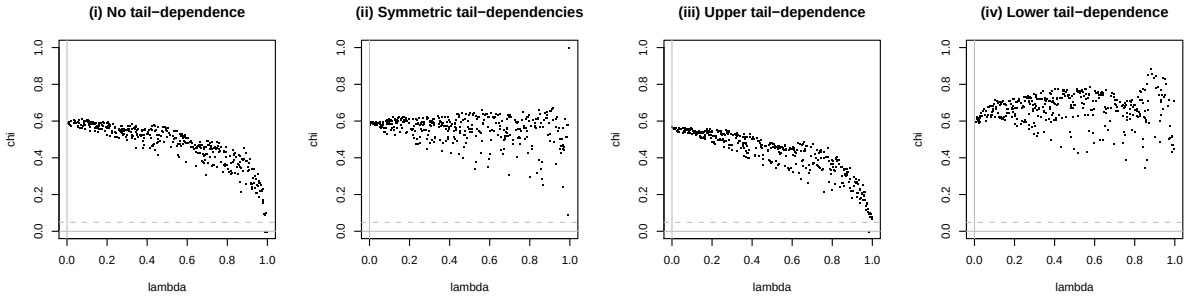
(a) Scatter plots of bivariate data sets characterized by different forms of dependency



(b) Chi-plots for different forms of dependency



(c) Detecting upper tail dependence: Chi-plots of observations from upper right corner of the scatter plot



(d) Detecting lower tail dependence: Chi-plots of observations from lower left corner of the scatter plot

tivariate data. An important advantage of this approach is that copula allows to model dependence structure independently of the marginal distributions.

3.5.1 Joint distributions

Definition 3.8 (Joint Distribution) *The **joint (cumulative) distribution function** of bivariate random variable (X, Y) is given by*

$$F_{XY}(x, y) = P\left(X \leq x, Y \leq y\right) \ . \tag{3.18}$$

The distributions of X and of Y are called *marginal distributions* and are denoted by F_X and F_Y, respectively. The following conditions are necessary and sufficient for a right-continuous function to be a bivariate cumulative distribution function:

(1) $\lim\limits_{x \to -\infty} F_{XY}(x, y) = 0$ and $\lim y \to -\infty \quad F_{XY}(x, y) = 0$

(2) $\lim\limits_{x,y \to \infty} F_{XY}(x, y) = 1$

(3) For all x_1, x_2, y_1, y_2 with $x_1 \leq x_2$, $y_1 \leq y_2$:

$$F_{XY}(x_2, y_2) - F_{XY}(x_2, y_1) - F_{XY}(x_1, y_2) + F_{XY}(x_1, y_1) \geq 0 \ .$$

Let a distribution function of (X, Y) be continuous with a sufficiently smooth second order mixed partial derivative. Then the density function can be evaluated as

$$f_{XY}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x, y) \ .$$

The purpose of this section is to study how the distribution of (X, Y) is related to the distributions of X and Y. The marginal distributions can be obtained from the joint distribution in a following way

$$F_X(x) = \lim_{y \to \infty} F_{XY}(x, y) = F_{XY}(x, \infty) \ ,$$

$$F_Y(y) = \lim_{x \to \infty} F_{XY}(x, y) = F_{XY}(\infty, y) \ .$$

And for the marginal density functions it holds

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy \quad \text{and} \quad f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx \ .$$

If the joint distribution is equal to the product of all the marginal distributions for every (x, y), so that

$$F_{XY}(x, y) = F_X(x) F_Y(y) \ ,$$

then the random variables X and Y are said to be *independent*. However, if X and Y are dependent, the joint distribution can not be determined from the marginal distributions. A copula function provides a connection between bivariate distribution function and their marginal distribution functions.

3.5.2 Definition and basic copula properties

The *copula* concept was developed by Sklar (1959), then extended and summarized to a book by Nelsen (2006) to that we refer in this section. A copula function can be defined generally for an arbitrary dimension d, but we focus our presentation only on a bivariate case. We assume also all distribution functions considered in this section to be continuous unless otherwise stated.

Definition 3.9 (Copula) *A (bivariate)* ***copula*** *is a function C from $[0,1] \times [0,1]$ to a unit interval $[0,1]$ with the following properties:*

(i) For every $u, v \in [0,1]$,

$$C(u,0) = 0 = C(0,v) \; ; \tag{3.19}$$

(ii) For every $u, v \in [0,1]$,

$$C(u,1) = u \quad \textit{and} \quad C(1,v) = v \; ; \tag{3.20}$$

(iii) For every $u_1, u_2, v_1, v_2 \in [0,1]$ such that $u_1 \leq u_2$ and $v_1 \leq v_2$,

$$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \geq 0 \; . \tag{3.21}$$

Looking at the copula properties (i)-(iii), one can recognize its identity with the properties (1)-(3) of the bivariate joint distribution mentioned in the previous section. Hence, a copula can be defined as a two-dimensional cdf with domain $[0,1]^2$ whose one-dimensional margins U and V are uniformly distributed on the unit interval, i.e $U, V \sim \mathcal{U}[0,1]$.

Next, we state a famous Sklar's theorem. This theorem shows how copula can be used to create a joint distribution when only marginal distributions are known.

Theorem 3.10 (Sklar's theorem) *Let F_{XY} be a joint distribution function of a bivariate random vector (X,Y) with marginal distributions $X \sim F_X$ and $Y \sim F_Y$. Then there exists a copula C such that for all $x, y \in \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$,*

$$F_{XY}(x,y) = C\left(F_X(x), F_Y(y)\right) \; . \tag{3.22}$$

If F_X and F_Y are continuous, the C is unique. And vice versa, if C is a copula and F_X and F_Y are distribution functions, then the function F_{XY} defined by (3.22) is a joint distribution function with margins F_X and F_Y.

Theorem 3.11 (Probability Integral Transformation) *Let X be a real valued random variable and $F(x)$ its continuous cumulative distribution function. Let U denote a random variable that is uniformly distributed on $[0,1]$. Then a random variable $F(X)$ is uniformly distributed on $[0,1]$. Furthermore, a random variable $F^{-1}(U)$ has cdf F where F^{-1} is the inverse distribution function.*

With help of the Probability Integral Transformation Theorem we can now connect the copula definition above and equation (3.22). In fact, for $X \sim F_X$, $Y \sim F_Y$ and $U, V \sim \mathcal{U}[0,1]$ from

$$F_X(x) = u \quad \text{and} \quad F_Y(y) = v$$

follows

$$F_X^{-1}(u) = x \quad \text{and} \quad F_Y^{-1}(v) = y \ ,$$

where F_X^{-1} and F_Y^{-1} are the inverse distribution functions of the margins. Thus, the copula from (3.22) has a form

$$C(u,v) = F_{XY}\left(F_X^{-1}(u), F_Y^{-1}(v)\right) \quad u, v \in [0,1] \ .$$

Note, a density function f_{XY} of a bivariate distribution can be computed under smoothness conditions by using a distribution function F_{XY} as $f_{XY}(x,y) = \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$. Similarly, we define a copula density.

Definition 3.12 (Copula density) *Let C be a twice partially differentiable copula. Then the function $c : [0,1] \times [0,1] \to [0,1]$ with*

$$c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v}$$

is called the ***density of the copula****.*

Let $f_X(x)$ and $f_Y(y)$ be marginal densities of bivariate distribution with joint density $f_{XY}(x,y)$, respectively, then we have

$$\begin{aligned} f_{XY}(x,y) &= \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y} = \frac{\partial^2 C\left(F_X(x), F_Y(y)\right)}{\partial x \partial y} \\ &= \frac{\partial^2 C\left(F_X(x), F_Y(y)\right)}{\partial F_X(x) \partial F_Y(y)} \frac{\partial F_X(x)}{\partial x} \frac{\partial F_Y(y)}{\partial y} \\ &= \frac{\partial^2 C(u,v)}{\partial u \partial v} = c(u,v) f_X(x) f_Y(y) \ . \end{aligned}$$

The last equation means that the joint density $f_{XY}(x,y)$ can be expressed in terms of the copula density and marginal densities and vice versa.

$$f_{XY}(x,y) = c(F_X(x), F_Y(y)) f_X(x) f_Y(y) \tag{3.23}$$

$$c(u,v) = \frac{f_{XY}\left(F_X^{-1}(u), F_Y^{-1}(v)\right)}{f_X\left(F_X^{-1}(u)\right) f_Y\left(F_Y^{-1}(v)\right)} \tag{3.24}$$

Very important is a fact that copula is invariant under increasing continuous transformations of margins. Proof of the following property can be found in Embrechts, Lindskog, and McNeil (2001).

Proposition 3.13 *Let C be a copula of (x, y) and T_1, T_2 are increasing continuous functions, then $T_1(x)$ and $T_2(y)$ also have copula C.*

In the next section, we study copulas C that are parameterized by a *copula parameter* $\boldsymbol{\theta}$. If $\boldsymbol{\theta}$ has a dimension greater than one, we call it multivariate copula parameter and note it by a bold symbol $\boldsymbol{\theta}$, in the one-dimensional case by a non-bold symbol θ.

3.5.3 Copula and dependence measures

Following the calculations given in Kurowicka and Cooke (2006, pp. 71–72), the rank based dependence measures introduced in the previous section can be expressed in terms of the copula.

Proposition 3.14 *Let F_{XY} be a joint distribution function of a bivariate continuous random variable (X, Y) with margins F_X and F_Y. Further, let C be a copula distribution associated with F_{XY} that is $F_{XY}(x, y) = C(F_X(x), F_Y(y))$. Then the rank based dependence measures, Kendall's τ and Spearman's ρ, can be linked to the copula C in the following way*

$$\tau(X,Y) = 4 \int_0^1 \int_0^1 C(u,v)\mathrm{d}C(u,v) - 1 \,, \tag{3.25}$$

and

$$\rho(X,Y) = 12 \int_0^1 \int_0^1 C(u,v)\mathrm{d}y\,\mathrm{d}s - 3 \,. \tag{3.26}$$

Table 3.1 gives an overview about formulas for computing Kendall's τ for several copula families that will be introduced in the next section.

In Section 3.3 we introduced a concept of tail dependence that means extreme market movements are likely to occur together. Surprisingly, the lower and upper tail dependence coefficients can be embraced within the copula theory (for proof see Nelsen (2006)).

Proposition 3.15 *Let (X, Y) be a continuous bivariate random variable and let C be a copula associated with their joint distribution function F_{XY}. Then the coefficients of lower*

Table 3.1: The different copula families, their parameters, Kendall's τ and coefficients of upper and lower tail dependences

Class	Copula	Parameter	Kendall's τ	$\tau \in$	λ_U	λ_L
Elliptical	Gauss	$r \in [-1,1]$	$\tau = \frac{2}{\pi}\arcsin(r)$	[-1,1]	0	0
	Student	$r \in [-1,1]$, $\nu > 0$	$\tau = \frac{2}{\pi}\arcsin(r)$	[-1,1]	$2t_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-r}{1+r}}\right)$	λ_U
Archimedean	Independent			0	0	0
	Gumbel	$\theta \geq 1$	$\tau = 1 - \frac{1}{\theta}$	[0,1]	$2 - 2^{\frac{1}{\theta}}$	0
	Clayton	$\theta > 0$	$\tau = \frac{\theta}{\theta+2}$	[0,1]	0	$2^{\frac{1}{\theta}}$
	Frank [1]	$\theta \in \mathbb{R} \setminus \{0\}$	$\tau = 1 + \frac{4}{\theta}(D_1(\theta) - 1)$	[-1,1]	0	0
	BB1	$\theta > 0$, $\delta \geq 1$	$\tau = 1 - \frac{2}{\delta(\theta+2)}$	[0,1]	$2 - 2^{\frac{1}{\delta}}$	$2^{-\frac{1}{\delta\theta}}$
	BB7 [2]	$\theta \geq 1$, $\delta > 0$	$\tau = 1 - \frac{2}{\delta(2-\theta)} + \frac{4}{\theta^2\delta}B(a,b)$	[0,1]	$2 - 2^{\frac{1}{\theta}}$	$2^{-\frac{1}{\delta}}$
Other	Plackett	$\theta > 0$	no closed form	[-1,1]	0	0

[1] Here is $D_k(x)$ with $k = 1$ a so called Debye function, defined as

$$D_k(x) = \begin{cases} \frac{k}{x^k}\int_0^x \frac{t^k}{e^t-1}\mathrm{d}t, & \text{if } x \geq 0\,, \\ \frac{k|x|}{1+k} + \frac{k}{|x|^k}\int_x^0 \frac{t^k}{e^t-1}\mathrm{d}t, & \text{if } x < 0\,. \end{cases}$$

[2] Here are $a = \frac{2-2\theta}{\theta} + 1$, $b = \delta + 2$ and $B(\cdot,\cdot)$ is a Beta-function given by $B(x,y) = \int_0^1 t^{x+1}(1-t)^{y-1}\mathrm{d}t$.

and upper tail dependence defined in (3.8) *and* (3.7) *can be calculated by using a copula* C *in a following way*

$$\lambda_L(X,Y) = \lim_{u \downarrow 0} \frac{C(u,u)}{u}\,, \tag{3.27}$$

$$\lambda_U(X,Y) = \lim_{u \uparrow 1} \frac{1 - 2u + C(u,u)}{1-u}\,. \tag{3.28}$$

3.5.4 Most common parametric copula families

In this section we introduce the most popular copula families. We restrict our presentation to the bivariate case. We begin with a class of *elliptical copulas*, Normal and Student's t. They both do not have a simple closed form and are expressed through their distribution functions. Hence, they belong to the class of *implicit* copulas. Another one importance class of copulas are *Archimedean copulas*, Clayton, Gumbel and Frank. Archimedean copulas are also called *explicit* copulas because of their simple closed form. Some definitions of this sections are based on the multivariate Gauss and Student's t-distributions, for more informations about these topics we refer to Appendix C. The presentation of the next three subsections follows Jondeau, Poon, and Rockinger (2007), Joe (1997) and Čížek, Härdle, and Weron (2005).

Elliptical copulas

Definition 3.16 (Gaussian (normal) copula) *Let $\Phi_r(\cdot,\cdot)$ denote the cumulative distribution function of the bivariate standard normal distribution with correlation $r \in [-1,1]$. Further, let Φ^{-1} be the inverse of the univariate standard normal cumulative distribution function Φ. Based on the Sklar's theorem, the* ***Gauss copula*** *with dependence parameter $r \in [-1,1]$ is defined as*

$$C_r(u,v) = \Phi_r\left(\Phi^{-1}(u), \Phi^{-1}(v)\right) \tag{3.29}$$

for all $(u,v) \in [0,1]^2$.

The copula parameter in this case is univariate and $\theta = r$. The expression (3.29) can be written in form

$$C_r(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi-1(v)} \frac{1}{2\pi\sqrt{1-r^2}} \left(-\frac{s^2 - 2rst + t^2}{2(1-r^2)}\right) dtds \tag{3.30}$$

and the bivariate normal copula has a density

$$c_r(u,v) = \frac{1}{\sqrt{1-r^2}} \exp\left(-\frac{r^2(u^2+v^2) - 2ruv}{2(1-r^2)}\right) . \tag{3.31}$$

Definition 3.17 (Bivariate Student's t Copula) *Let $t_{\nu,r}(\cdot,\cdot)$ denote the cumulative distribution function of the bivariate standard t-distribution with degrees of freedom $\nu > 0$ and dependence parameter $r \in [-1,1]$. Further, let t_ν^{-1} be the inverse of the cdf of the standard univariate t-distribution with ν degrees of freedom. By using Sklar's theorem, the* ***Student's t copula*** *with parameter ν and r is defined as*

$$C_{r,\nu}(u,v) = t_{\nu,r}\left(t_\nu^{-1}(u), t_\nu^{-1}(v)\right) \tag{3.32}$$

for all $(u,v) \in [0,1]^2$.

The copula parameter is in this case bivariate and $\boldsymbol{\theta} = (\theta_1, \theta_2) = (\nu, r)$. The expression (3.32) can be also written as

$$C_{r,\nu}(u,v) = \int_{-\infty}^{t_\nu^{-1}(u)} \int_{-\infty}^{t_\nu^{-1}(v)} \frac{1}{2\pi\sqrt{1-r^2}} \left(1 + \frac{s^2 - 2rst + t^2}{\nu(1-r^2)}\right)^{-\frac{\nu+2}{2}} dtds ,$$

and the density of this copula is given by (3.32) can be written as

$$c_{r,\nu}(u,v) = \frac{1}{\sqrt{1-r^2}} \frac{\Gamma\left(\frac{\nu+2}{2}\right)\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu+2}{2}\right)^2} \frac{\left(1 + \frac{s^2-2rst+t^2}{\nu(1-r^2)}\right)^{-\frac{\nu+2}{2}}}{\left(1+\frac{s^2}{\nu}\right)^{-\frac{\nu+1}{2}} \left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}} , \tag{3.33}$$

where $s = t_\nu^{-1}(u)$ and $t = t_\nu^{-1}(v)$.

Archimedean copulas

In contrast to the previous two copula families, the Archimedean ones are not derived from any multivariate distributions. The Archimedean copulas have a simple closed form, but they are difficult to establish to higher dimensions. The general definition of Archimedean family of copulas is based on the following theorem provided by Nelsen (2006, pg. 110).

Theorem 3.18 *Let $\phi : [0,1] \to [0,\infty)$ be a continuous, strictly decreasing function such that $\varphi(1) = 0$. Further, let φ^{-1} be the inverse function of φ. Then, the function $C : [0,1]^2 \to [0,1]$ given by*

$$C(u,v) = \varphi^{-1}\left(\varphi(u) + \varphi(v)\right) \tag{3.34}$$

is a copula, if and only if φ is convex.

The copula created from the expression (3.34) is called *Archimedean copula*. The function φ is called a *generator* of the copula. The copula density can be written as

$$c(u,v) = \frac{\varphi^{-1''}\left(\varphi(u) + \varphi(v)\right)}{\varphi^{-1'}\left(\varphi'(u)\right) \cdot \varphi^{-1'}\left(\varphi'(v)\right)} \tag{3.35}$$

if φ^{-1} is twice continuously differentiable. The following expression for computing Kendall's τ using the generator φ can be proved (see Kurowicka and Cooke (2006, pg. 75))

$$\tau = 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} \mathrm{d}x \ . \tag{3.36}$$

Definition 3.19 *The **product or independent copula** is given by*

$$C(u,v) = uv \ . \tag{3.37}$$

This copula has no dependence between variates. Its density function is unity everywhere and its generator is given as

$$\varphi(t) = -\log t \ .$$

Definition 3.20 *The **Clayton copula** with dependence parameter $\theta \in (0,\infty)$ is given by*

$$C_\theta(u,v) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-\frac{1}{\theta}} \ . \tag{3.38}$$

The perfect dependence is observed at $\theta \to \infty$, while $\theta \to 0$ implies independence. Note that the Clayton copula can not be used for modeling negative dependence. This copula has a generator

$$\varphi(t) = \frac{t^{-\theta} - 1}{\theta}$$

and density

$$c_\theta(u, v) = (1 + \theta)(uv)^{-\theta-1}(u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}} .$$

Definition 3.21 *The **Gumbel copula** with dependence parameter $\theta \in [1, \infty)$ is given by*

$$C_\theta(u, v) = exp\left(-\left((-log\ u)^\theta + (-log\ v)^\theta\right)^{\frac{1}{\theta}}\right) . \tag{3.39}$$

The perfect dependence is observed at $\theta \to \infty$, while $\theta = 1$ implies independent. Similar to the Clayton copula, the Gumbel copula captures only positive dependence. A generator of this copula is

$$\varphi(t) = (-\log t)^\theta$$

and a density function

$$c_\theta(u, v) = \frac{C_\theta(u, v)(\log u \cdot \log v)}{uv\left((-\log v)^\theta + (-\log u)^\theta\right)^{2-\frac{1}{\theta}}} \left\{\left((-\log u)^\theta + (-\log v)^\theta\right)^{\frac{1}{\theta}} + \theta - 1\right\} .$$

Definition 3.22 *The **Frank copula** with real dependence parameter $\theta \in (-\infty, \infty) \setminus \{0\}$ has a form*

$$C_\theta(u, v) = -\frac{1}{\theta} log\left(1 + \frac{\left(e^{-\theta u} - 1\right)\left(e^{-\theta v} - 1\right)}{e^{-\theta} - 1}\right) . \tag{3.40}$$

The values of δ closed to $\pm\infty$ relate to independence. In opposition to Gumbel and Clayton copulas, Frank copula allows both positive and negative dependence structures. The Frank copula is generated by

$$\varphi(t) = \log(e^{-\theta} - 1) - \log(e^{-\theta t} - 1)$$

and has a density function

$$c_\theta(u, v) = \frac{\theta(1 - e^{-\theta})e^{-\theta(u+v)}}{\left(1 - e^{-\theta} - (1 - e^{-\theta u})(1 - e^{-\theta v})\right)^2} .$$

Next, we introduce 2-parameter Archimedean copulas that can be used to capture more than one type of dependence.

Definition 3.23 *The **BB1 copula** with parameter $\boldsymbol{\theta} = (\theta, \delta)^t$, $\theta > 0$ and $\delta \geq 1$, is given by*

$$C_{\theta,\delta}(u,v) = \left\{1 + \left((u^{-\theta} - 1)^{\delta} + (v^{-\theta} - 1)^{\delta}\right)^{\frac{1}{\delta}}\right\}^{-\frac{1}{\theta}} \tag{3.41}$$

and its generator is

$$\varphi(t) = \left(t^{-\theta} - 1\right)^{\delta} .$$

Definition 3.24 *The **BB7 copula** with parameter $\boldsymbol{\theta} = (\theta, \delta)^t$, $\delta > 0$ and $\theta \geq 1$, is given by*

$$C_{\theta,\delta}(u,v) = 1 - \left\{1 - \left[(1-(1-u)^{\theta})^{-\delta} + (1-(1-v)^{\theta})^{-\delta} - 1\right]^{-\frac{1}{\delta}}\right\}^{\frac{1}{\theta}} \tag{3.42}$$

and has a generator

$$\varphi(t) = \left(1 - (1-t)^{\theta}\right)^{-\delta} - 1 .$$

The ability of dependence modeling of BB1 and BB7 copula is restricted to positive case. Modeling of negative dependence is not possible. The Gumbel copula is the limiting case of BB1 with $\theta \to 0$. The Clayton arises from BB1 when $\delta = 1$. The expressions for BB1 and BB7 copula densities are not trivial and can be found in Schepsmeier (2010).

Other copulas

Finally, we introduce a Plackett copula, see Joe (1997, pg. 91), that belongs neither to Archimedean nor to elliptical families.

Definition 3.25 *The **Plackett copula** with association parameter $\theta > 0$ is defined as*

$$C_{\theta}(u,v) = \frac{1}{2(\theta-1)}\left(1 + (\theta-1)(u+v) - \sqrt{\left[1 + (\theta-1)(u+v)\right]^2 - 4uv\theta(\theta-1)}\right) \tag{3.43}$$

Its density function is given by

$$c_{\theta}(u,v) = \frac{\theta\left(1 + (u - 2uv + v)(\theta-1)\right)}{\left(\left[1 + (\theta-1)(u+v)\right]^2 - 4uv\theta(\theta-1)\right)^{\frac{3}{2}}} .$$

Like the Normal and Student copulas, the Plackett copula allows to model positive ($\theta > 1$) as well as negative ($\theta < 1$) dependency. The case $\theta = 1$ implies independence.

There are several R packages where different copula families are implemented in. The well-known is the R package `copula` that contains estimation function and a lot of other useful tools for handling with Normal, Student, Gumbel, Clayton, Plackett, Frank and Independent copulas. For more details above R packages for copula fitting see Appendix B.

Tail dependence in specific copula families

Figure 3.5 displays a visualization of data generated from bivariate Gaussian, Frank, Plackett copulas and Student's t copula with 2 degrees of freedom with on the unit interval uniformly distributed margins: scatter plot on the left panel (a), theoretical contour plot of copula density on the middle panel and 3D perspective copula density plot on the right panel. All bivariate samples have the same Kendall's τ, namely $\tau = 0.6$. As we can see from this figure, the copula density explodes for extremely small and extremely large values of margins. Besides, a lot of points on the scatter plots are concentrated in the left lower and right upper corner which makes the exploration of tail dependences difficult.

In practice, it is usual to create visualization plots of copulas whose margins are transformed to standard normal. The plots obtained in a such way are more suitable for detecting the dependence form, absence or presence of the tail dependences. A confirmation of this procedure is based on the copula property 3.13 that states that the copula is invariant under increasing transformations of margins. Figures 3.6–3.8 show scatter plots, theoretical and empirical contour plots of copula density for the most common copula families with standard normal margins: Gauss, Frank, Plackett, independent copula, Student's t, Gumbel, Clayton, BB1 and BB7. In contrast to the case with uniform margins, dependence contours and tails on these plots are more clearly depicted.

In practice, a true copula of observed data is unknown. So, it is impossible to create a theoretical contour plot for the bivariate case. In such situations, we are forced to use empirical contour plots. As we can see from Figures 3.6–3.8, the empirical contour plots are not able to reflect the dependence form just like the theoretical contour plots do it. But they are adequate enough to detect the most appropriate copula class.

Panel (a) of Figure 3.6 shows visualization plots of a data set generated from a bivariate Gaussian copula with the dependence parameter $r = 0.6$. Panels (b) and (c) display visualization plots of data sets generated from the bivariate Frank and Plackett copulas respectively. Tails of the Frank copula illustrated on the scatter plot tend to be relatively weak compared to the Gaussian ones. Another one remarkable feature of the Frank copula is that a strongest dependence is centered in the middle of the data. There are no observable tail dependences in the right upper as well as in the left lower corner of the scatter plots. According to Malevergne and Sornette (2006, pp. 170–172), it can be shown for these three copulas

$$\lambda(X, Y) = \lambda_L(X, Y) = \lambda_U(X, Y) = 0 \, .$$

The panel (d) of Figure 3.6 represents data from the independent copula. Neither tail dependences nor any other art of association can be observed in this case.

Both upper panels (a) and (b) of Figure 3.7 display data from the bivariate Student's t copula with different number of degrees of freedom $df = 2$ and $df = 4$. Symmetric tails are observed in this case. In contrast to the Gaussian copula, the coefficients of lower and

upper tail dependence for Student copula are different from zero and are equal to

$$\lambda(X,Y) = \lambda_L(X,Y) = \lambda_U(X,Y) = 2t_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-r}{1+r}}\right) ,$$

where $t_{\nu+1}$ denotes the distribution function of the univariate t distribution with $(\nu+1)$ degrees of freedom. The stronger the dependence parameter r and the lower the degree of freedom, the stronger is λ.

The tails of the Gumbel copula, see panel (c) of Figure 3.7, are asymmetric. Lower tail but no upper tail dependence is oberved in this case. According to Malevergne and Sornette (2006, pp. 170–172), it holds

$$\lambda_L(X,Y) = 0 \quad \text{and} \quad \lambda_U(X,Y) = 2 - 2^{\frac{1}{\delta}} .$$

Similar to the Gumbel case, the Clayton copula is also asymmetric, see panels (d) of Figure 3.7. Observations are strongly correlated in the left lower and do not correlated in the right upper corner of the scatter plot. It points to the presence of the lower and absence of upper tail dependency. Following calculations in Malevergne and Sornette (2006, pp. 170–172), it can be shown

$$\lambda_U(X,Y) = 0 \quad \text{and} \quad \lambda_L(X,Y) = 2^{\frac{1}{\delta}} .$$

The last Figure 3.8 demonstrates visualisation plots of data from BB1 and BB7 copulas. These copulas are more flexible and allow asymmetric non-zero tail dependences. Panels (a) and (c) exhibit data from BB1 and BB7 copulas with the strong upper and weak lower tail dependence. In contrast, panels (b) and (d) illustrate data with strong lower and weak upper tail dependence. Joe (1997, pg. 150 and pg. 153) gives following expressions for BB1 copula

$$\lambda_U = 2 - 2^{\frac{1}{\delta}} \quad \text{and} \quad \lambda_L = 2^{-\frac{1}{\delta\theta}}$$

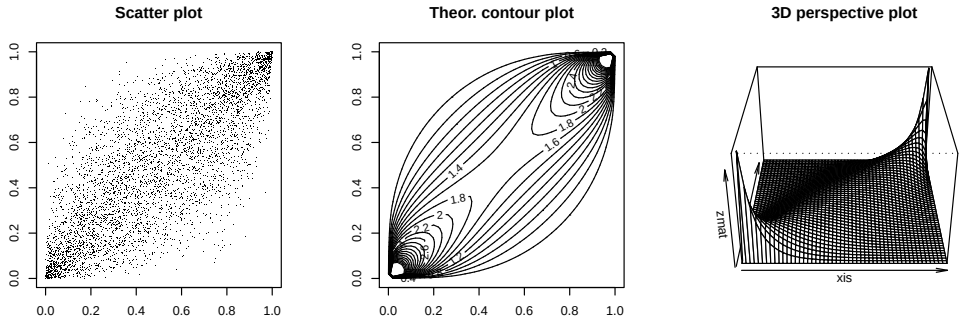
and for BB7 copula

$$\lambda_U = 2 - 2^{\frac{1}{\theta}} \quad \text{and} \quad \lambda_L = 2^{-\frac{1}{\delta}} .$$

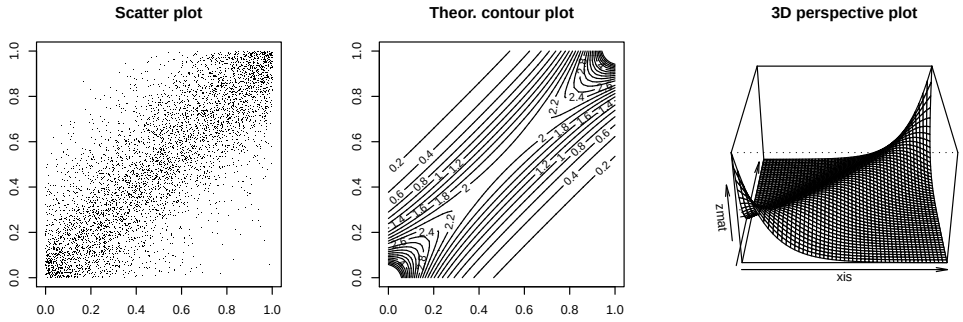
3.6 Pair-copula constructions and vines

One of the current techniques for dependence modeling is the joint normal transformation. It comprises following steps: standard normal transformation of the margins, measuring a dependence structure and transforming it back. This method was implemented by Ghost and Henderson (2002) and by Iman and Helton (1985). The primary drawback of this approach is that the joint normal distribution allows only for symmetric dependence structures and no tail dependence.

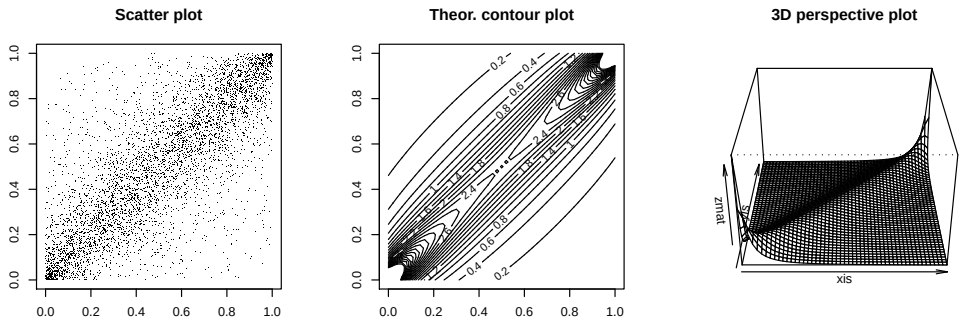
Figure 3.5: Graphical representation of Gaussian, Frank, Plackett and Student's t copula families with univariate margins and coefficient of association $\tau = 0.6$: scatter plot on the left panel, theoretical contour plot in the middle and empirical contour plot on the right panel. Copula parameters and coefficients of tail dependence were calculated (numerically for Frank and Plackett copulas) using τ according to the expressions given in Table 3.1.



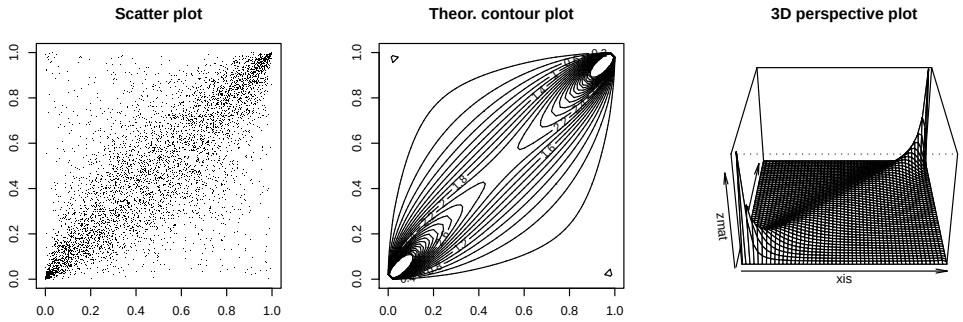
(a) Gaussian copula, $\tau = 0.6$, $r = 0.809017$, $\lambda_U = \lambda_L = 0$



(b) Frank copula, $\tau = 0.6$, $\tilde{\theta} = 7.929643$, $\lambda_U = \lambda_L = 0$

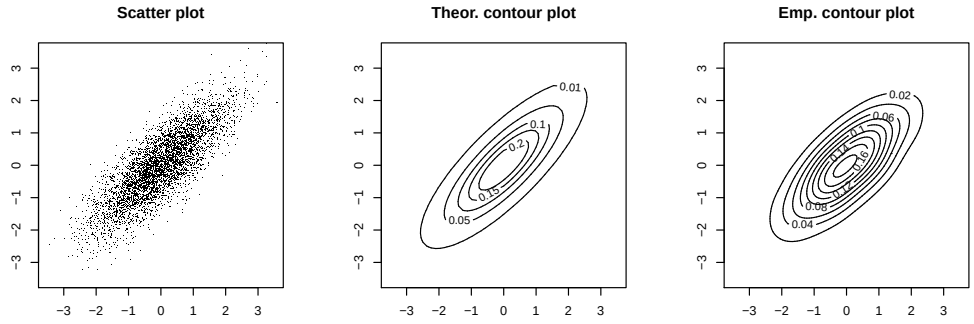


(c) Plackett copula, $\tau = 0.6$, $\tilde{\theta} = 21.1318$, $\lambda_U = \lambda_L = 0$

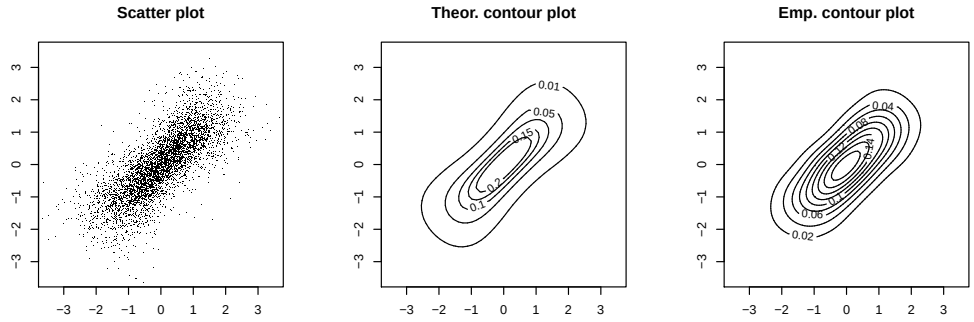


(d) Student's t copula, $\tau = 0.6$, $df = 2$ $\lambda_U = \lambda_L = 0.4501849$

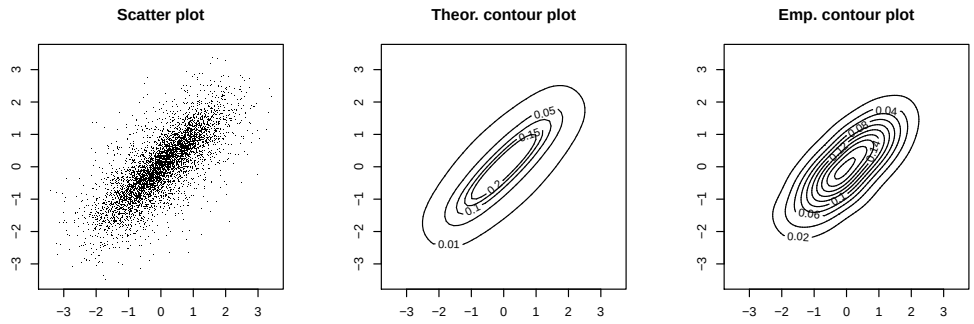
Figure 3.6: Graphical representation of Gaussian, Frank, Plackett and Independent copula families with standard normal margins and coefficient of association $\tau = 0.6$ (except of Independent copula): scatter plot on the left panel, theoretical contour plot in the middle and empirical contour plot on the right panel. Copula parameters and coefficients of tail dependence were calculated (numerically for Frank and Plackett copulas) using τ according to the expressions given in Table 3.1.



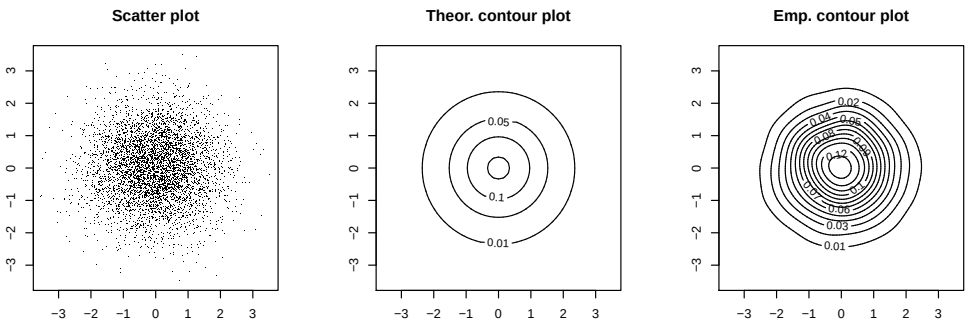
(a) Gaussian copula, $\tau = 0.6$, $r = 0.809017$, $\lambda_U = \lambda_L = 0$



(b) Frank copula, $\tau = 0.6$, $\tilde{\theta} = 7.929643$, $\lambda_U = \lambda_L = 0$

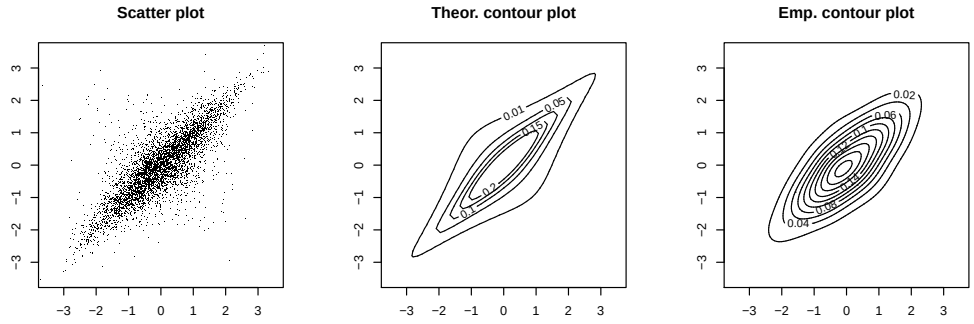


(c) Plackett copula, $\tau = 0.6$, $\tilde{\theta} = 21.1318$, $\lambda_U = \lambda_L = 0$

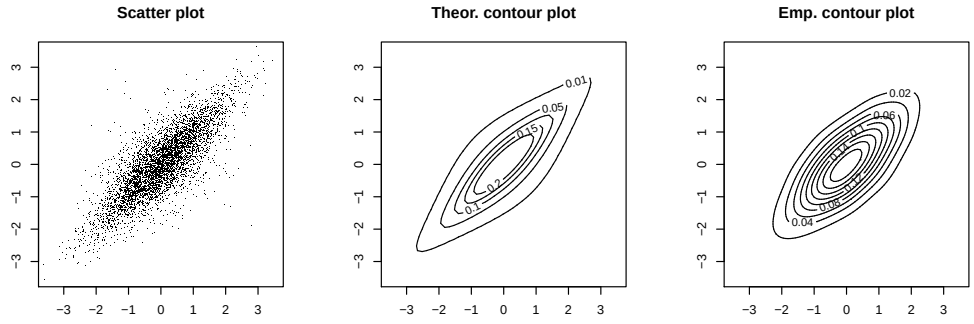


(d) Independent copula, $\tau = 0$, $\lambda_U = \lambda_L = 0$

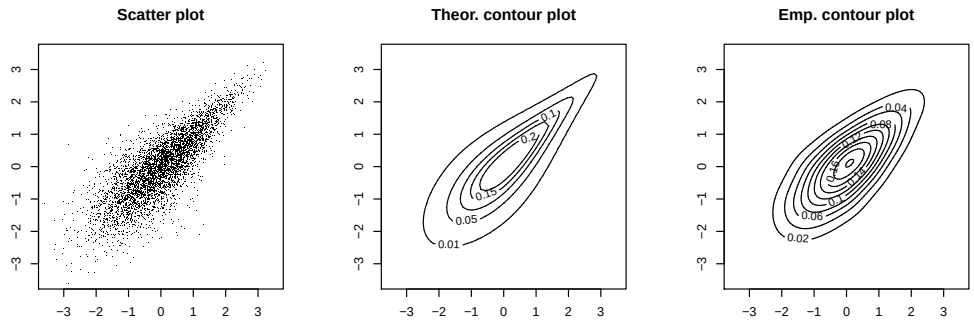
Figure 3.7: Graphical representation of Student's t, Gumbel and Clayton copula families with standard normal margins and coefficient of association $\tau = 0.6$: scatter plot on the left panel, theoretical contour plot in the middle and empirical contour plot on the right panel. Copula parameters and coefficients of tail dependence were calculated using τ according to the expressions given in Table 3.1.



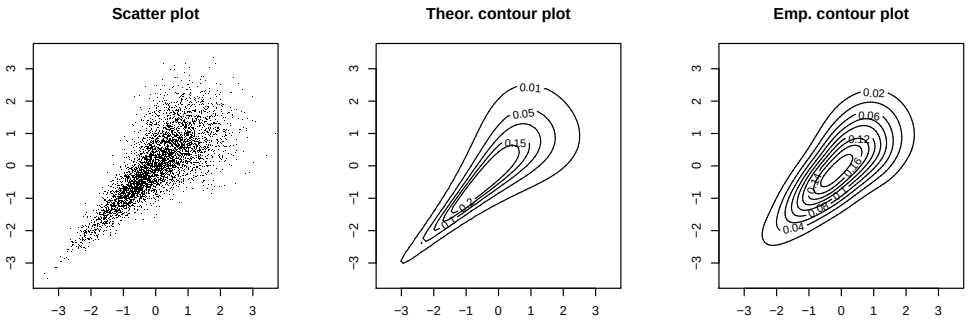
(a) Student copula, $\tau = 0.6$, $df = 2$, $r = 0.809017$, $\lambda_U = \lambda_L = 0.6129021$



(b) Student copula, $\tau = 0.6$, $df = 4$, $r = 0.809017$, $\lambda_U = \lambda_L = 0.5000811$

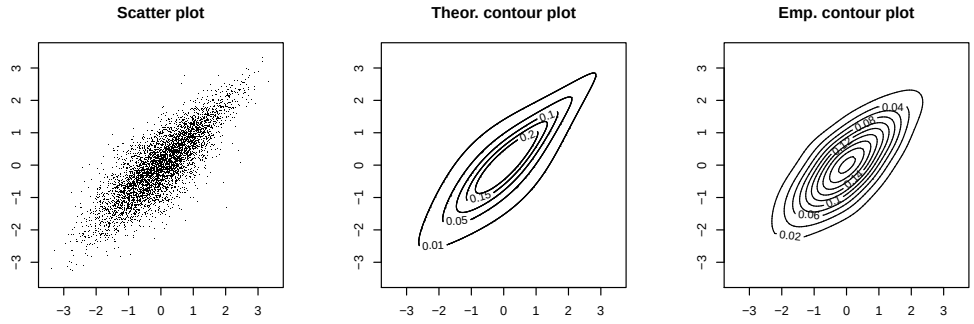


(c) Gumbel copula, $\tau = 0.6$, $\theta = 2.5$, $\lambda_U = 0.6804921$, $\lambda_L = 0$

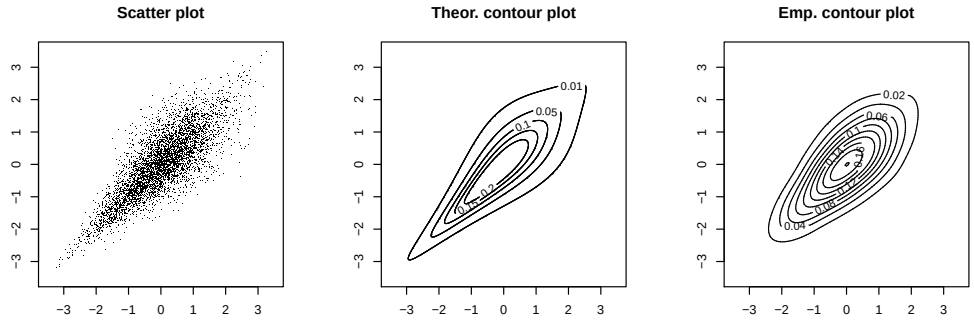


(d) Clayton copula, $\tau = 0.6$, $\theta = 3$, $\lambda_U = 0$, $\lambda_L = 0.7937005$

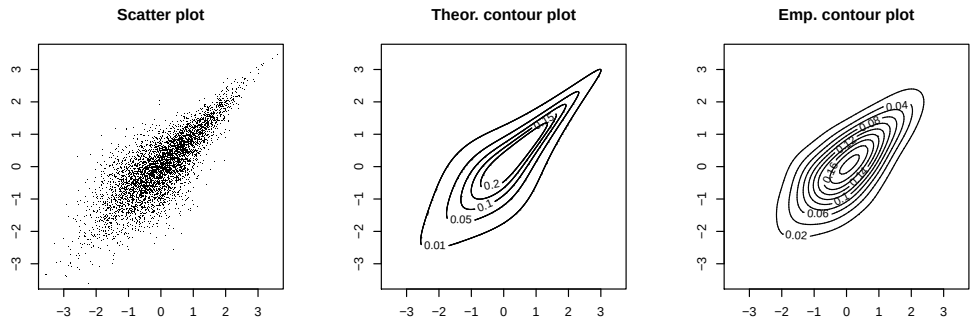
Figure 3.8: Graphical representation of BB1 and BB7 copula families with standard normal margins: scatter plot on the left panel, theoretical contour plot in the middle and empirical contour plot on the right panel. We choose parameters for BB1 and BB7 copula to make Kendall's τ approximately equal to 0.6. Coefficients of tail dependence and Kendall's τ were calculated according to the expressions given in Table 3.1.



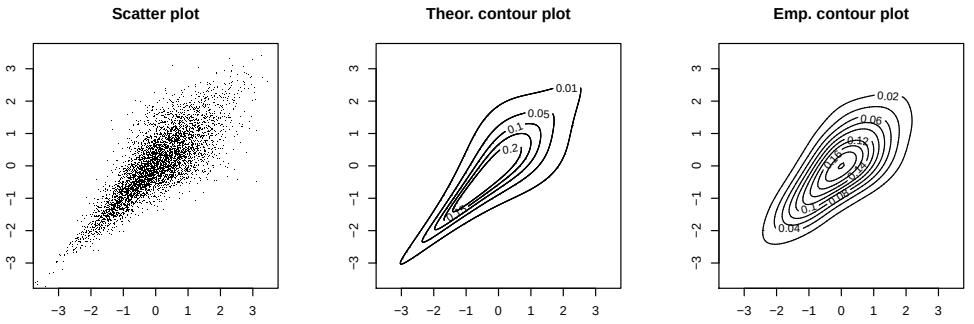
(a) BB1 copula, $\tau = 0.5929996$, $(\theta, \delta) = (0.33, 2.3)$, $\lambda_U = 0.6089344$, $\lambda_L = 0.378782$



(b) BB1 copula, $\tau = 0.6112731$, $(\theta, \delta) = (1.5, 1.47)$, $\lambda_U = 0.397558$, $\lambda_L = 0.7302616$



(c) BB7 copula, $\tau = 0.6035885$, $(\theta, \delta) = (3.2, 0.8)$, $\lambda_U = 0.7581422$, $\lambda_L = 0.4204482$



(d) BB7 copula, $\tau = 0.6036062$, $(\theta, \delta) = (1.5, 2.8)$, $\lambda_U = 0.4125989$, $\lambda_L = 0.7807092$

An alternative popular method for modeling multivariate dependencies is provided by the copula approach. There are a huge number of bivariate copula families that allow modeling of different dependence structures. An extension of the bivariate copulas to higher dimensions is in general not trivial. It requires additional parameter restrictions that causes flexibility reduction. Besides, the extension up to the higher dimension of some copula families is even not possible, for example in case of the Plackett copula.

A so called pair-copula approach provides very flexible tool for constructing multivariate dependency. The idea of this concept is to decompose the joint density function into a cascade of *building blocks* of the bivariate copulas and their conditional and unconditional distribution functions. The building blocks are called *pair-copulas*. The most important advantage of this methodology lies in the ability to comprise miscellaneous dependency structures, i.e varying copula families are allowed for different marginal pairs. In this chapter we present a most important topics of the pair-copula theory based on the article of Aas, Czado, Frigessi, and Bakken (2009). Further, we will show how the graphical concept of vines, introduced by Bedford and Cooke (2001, 2002), can be helpful when determinating an appropriate decomposition of a joint density.

3.6.1 Pair-copula constructions

Let $\boldsymbol{X} = (X_1, \ldots, X_d)^t$ be a vector of random variables with a joint density $f(x_1, \ldots, x_d)$ and marginal densities $f_1(x_1), \ldots, f_d(x_d)$ and marginal distributions $F_1(x_1), \ldots, F_d(x_d)$. For the joint density, we have the functions of subvectors $(X_r, \ldots, X_d)^t$ with $r \in \{2, \ldots, d-1\}$

$$\begin{aligned} &f(x_2, \ldots, x_d) = f(x_3, \ldots, x_d) f(x_2 | x_3, \ldots, x_d) \\ &\ldots \\ &f(x_r, \ldots, x_d) = f(x_{r+1}, \ldots, x_d) f(x_r | x_{r+1}, \ldots, x_d) \\ &\ldots \\ &f(x_{d-1}, x_d) = f_d(x_d) f(x_{d-1} | x_d) \end{aligned} \tag{3.44}$$

Using (3.44) we can factorize the joint density $f(x_1, \ldots, x_d)$ as

$$\begin{aligned} f(x_1, \ldots, x_d) &= f(x_2, \ldots, x_d) f(x_1 | x_2, \ldots, x_d) \\ &= f(x_3, \ldots, x_d) f(x_2 | x_3, \ldots, x_d) f(x_1 | x_2, \ldots, x_d) \\ &= \ldots \\ &= f_d(x_d) f(x_{d-1} | x_d) f(x_{d-2} | x_{d-1}, x_d) \cdots f(x_1 | x_2, \ldots, x_d) \end{aligned} \tag{3.45}$$

The factorization (3.45) is unique up to re-labeling of the variables. In subsection 3.5.2 we have shown for the bivariate case how the joint density of two random variables can be expressed in terms of a copula and marginal densities. Thus, we get using the notations of this section

$$f(x_1, x_2) = c_{12}(F_1(x_1), F_2(x_2)) f_1(x_1) f_2(x_2) \, . \tag{3.46}$$

The generalization of (3.46) to the d-dimensional case is also possible

$$f(x_1, \ldots, x_d) = c_{1\ldots d}(F_1(x_1), \ldots, F_d(x_d)) f_1(x_1) \cdots f_d(x_d) \; . \tag{3.47}$$

For the conditional density $f(x_1|x_2)$ we get the unique expression

$$\begin{aligned} f(x_1|x_2) &= \frac{f(x_1, x_2)}{f_2(x_2)} = \frac{c_{12}(F_1(x_1), F_2(x_2)) f_1(x_1) f_2(x_2)}{f_2(x_2)} \\ &= c_{12}(F_1(x_1), F_2(x_2)) f_1(x_1) \end{aligned} \tag{3.48}$$

In contrast to the previous case, the conditional density $f(x_1|x_2, x_3)$ allows for two different presentations

$$\begin{aligned} f(x_1|x_2, x_3) &= \frac{f(x_1, x_2, x_3)}{f(x_2, x_3)} = \frac{f(x_1, x_2|x_3) f(x_3)}{f(x_2|x_3) f(x_3)} = \frac{f(x_1, x_2|x_3)}{f(x_2|x_3)} \\ &\overset{(3.46)}{=} \frac{c_{12|3}\left(F(x_1|x_3), F(x_2|x_3)\right) f(x_1|x_3) f(x_2|x_3)}{f(x_2|x_3)} \\ &= c_{12|3}\left(F(x_1|x_3), F(x_2|x_3)\right) f(x_1|x_3) \\ &\overset{(3.48)}{=} c_{12|3}\left(F(x_1|x_3), F(x_2|x_3)\right) c_{13}(F_1(x_1), F_3(x_3)) f_1(x_1) \end{aligned} \tag{3.49}$$

and similarly

$$f(x_1|x_2, x_3) = c_{13|2}\left(F(x_1|x_2), F(x_3|x_2)\right) c_{12}(F_1(x_1), F_2(x_2)) f_1(x_1) \; . \tag{3.50}$$

For example, for three variables X_1, X_2 and X_3 we can decompose the joint density as

$$\begin{aligned} f(x_1, x_2, x_3) &\overset{(3.45)}{=} f_3(x_3) f(x_2|x_3) f(x_1|x_2, x_3) \\ &= f_3(x_3) \cdot \underbrace{f_2(x_2) c_{23}\left(F_2(x_2), F_3(x_3)\right)}_{\overset{(3.48)}{=} f(x_2|x_3)} \\ &\quad \cdot \underbrace{c_{12|3}\left(F(x_1|x_3), F(x_2|x_3)\right) c_{13}(F_1(x_1), F_3(x_3)) f_1(x_1)}_{\overset{(3.49)}{=} f(x_1|x_2, x_3)} \; . \end{aligned}$$

So, one of the possible pair-copula decompositions of $f(x_1, x_2, x_3)$ is

$$\begin{aligned} f(x_1, x_2, x_3) = f_1(x_1) f_2(x_2) f_3(x_3) &\cdot c_{23}\left(F_2(x_2), F_3(x_3)\right) c_{13}(F_1(x_1), F_3(x_3)) \\ &\cdot c_{12|3}\left(F(x_1|x_3), F(x_2|x_3)\right) \end{aligned} \tag{3.51}$$

The factorization (3.45) for an arbitrary dimension d of random vector $\boldsymbol{X}$ can be obtained by applying the general formula for the conditional densities

$$f(x|\boldsymbol{\nu}) = c_{x\nu_j|\boldsymbol{\nu}_{-j}}\left(F(x|\boldsymbol{\nu}_{-j}), F(\nu_j|\boldsymbol{\nu}_{-j})\right) \cdot f(x|\boldsymbol{\nu}_{-j}) \; , \tag{3.52}$$

where $\boldsymbol{\nu}$ is a vector of dimension $m < d$ and $j \in \{1, \ldots, k\}$. In the following, we denote with $\boldsymbol{\nu}$ always the conditioning variables. By the symbol $\boldsymbol{\nu}_{-j}$ we denote the vector $\boldsymbol{\nu}$ without its jth component. Note, that there are many different re-parameterizations of this decomposition. Next, we introduce rules that can be used for building such pair-copula decompositions.

For evaluating the marginal conditional distribution functions $F(x|\boldsymbol{\nu})$, one uses the formula

$$F(x|\boldsymbol{\nu}) = \frac{\partial C_{x\nu_j|\boldsymbol{\nu}_{-j}}\left(F(x|\boldsymbol{\nu_{-j}}), F(\nu_j|\boldsymbol{\nu_{-j}})\right)}{\partial F(\nu_j|\boldsymbol{\nu_{-j}})}\,, \tag{3.53}$$

where $C_{x\nu_j|\boldsymbol{\nu}_{-j}}$ is a bivariate copula. To make the interpretation of this formula easier, we consider first a univariate ν. For this case we have with $x = x_1$ and $\nu = x_2$

$$F(x_1|x_2) = \frac{\partial C_{12}\left(F_1(x_1), F_2(x_2)\right)}{\partial F_2(x_2)}\,. \tag{3.54}$$

And for bivariate $\boldsymbol{\nu} = (x_2, x_3)^t$, the expression (3.53) acquires a form

$$F(x_1|x_2, x_3) = \frac{\partial C_{12|3}\left(F(x_1|x_3), F(x_2|x_3)\right)}{\partial F(x_2|x_3)}\,.$$

We use the notion $h_{\boldsymbol{\theta}}(x, \nu)$ to represent the relation (3.54) when x and ν are uniform, where $\boldsymbol{\theta} \in \Theta$ states a parameter set of specified parametric copula family of the joint distribution of x and ν. I.e for $x, \nu \sim \mathcal{U}[0, 1]$

$$h_{\boldsymbol{\theta}}(x, \nu) = \frac{\partial C_{\boldsymbol{\theta}}\left(F_x(x), F_\nu(\nu)\right)}{\partial F_\nu(\nu)} = \frac{\partial C_{\boldsymbol{\theta}}(x, \nu)}{\partial v}\,, \tag{3.55}$$

where $\boldsymbol{\theta} \in \Theta$ is known or estimated copula parameter. Using h-functions one can calculate the conditional distribution of an arbitrary order. For instance, if $x_1, x_2, x_3, x_4 \sim \mathcal{U}[0, 1]$, we get

$$F(x_1|x_2) = h_{\theta_{1,2}}(x_1, x_2)$$

and

$$F(x_1|x_2, x_3) = h_{\boldsymbol{\theta}_{1,2|3}}(F(x_1|x_3), F(x_2|x_3)) = h_{\boldsymbol{\theta}_{1,2|3}}(h_{\boldsymbol{\theta}_{1,3}}(x_1, x_3), h_{\boldsymbol{\theta}_{2,3}}(x_2, x_3))$$

or

$$F(x_1|x_2, x_3) = h_{\boldsymbol{\theta}_{1,3|2}}(F(x_1|x_2), F(x_3|x_2)) = h_{\boldsymbol{\theta}_{1,3|2}}(h_{\boldsymbol{\theta}_{1,2}}(x_1, x_2), h_{\boldsymbol{\theta}_{3,2}}(x_3, x_2))$$

further

$$\begin{aligned} &F(x_1|x_2, x_3, x_4) = h_{\boldsymbol{\theta}_{1,4|2,3}}(F(x_1|x_2, x_3), F(x_4|x_2, x_3)) \\ &= h_{\boldsymbol{\theta}_{1,4|2,3}}(h_{\boldsymbol{\theta}_{1,3|2}}(h_{\boldsymbol{\theta}_{1,2}}(x_1, x_2), h_{\boldsymbol{\theta}_{3,2}}(x_3, x_2)), h_{\boldsymbol{\theta}_{4,3|2}}(h_{\boldsymbol{\theta}_{4,2}}(x_4, x_2), h_{\boldsymbol{\theta}_{3,2}}(x_3, x_2))) \end{aligned}$$

or

$$F(x_1|x_2, x_3, x_4) = h_{\boldsymbol{\theta}_{1,2|3,4}}(F(x_1|x_3, x_4), F(x_2|x_3, x_4)) = \cdots$$

or

$$F(x_1|x_2,x_3,x_4) = h_{\boldsymbol{\theta}_{1,3|2,4}}(F(x_1|x_2,x_4),F(x_3|x_2,x_4)) = \cdots \ .$$

In the following, we simplify a notation of h-functions by omitting a term $\boldsymbol{\theta}$, i.e. $h_{\boldsymbol{\theta}_{1,2}}$ gets a form $h_{1,2}$. Below, we give expressions for h-functions of some copula families. For its presentation, we refer to the article of Aas, Czado, Frigessi, and Bakken (Gauss and Student) and to the diploma thesis of Ulf Schepsmeier Schepsmeier (2010) (BB1 and BB7).

Proposition 3.26 *Let $C_r(u,v)$ be a BB1 copula from Definition 3.29. The h-function of this copula is given by*

$$h_r(u,v) = \Phi_r\left(\frac{\Phi^{-1}(u) - r\Phi^{-1}(v)}{\sqrt{1-r^2}}\right) \ . \tag{3.56}$$

Proposition 3.27 *Let $C_{r,\nu}(u,v)$ be a BB1 copula from Definition 3.32. The h-function of this copula is given by*

$$h_{r,\nu}(u,v) = t_{\nu+1}\left(\frac{t_\nu^{-1}(u) - rt_\nu^{-1}(v)}{\sqrt{\frac{(\nu+(t_\nu^{-1}(v))^2)(1-r^2)}{\nu+1}}}\right) \tag{3.57}$$

Proposition 3.28 *Let $C_{\theta,\delta}(u,v)$ be a BB1 copula from Definition 3.41. The h-function of this copula is given by*

$$\begin{aligned} h_{\theta,\delta}(u,v) = &\left[1+\left((u^{-\theta}-1)^\delta + (v^{-\theta}-1)^\delta\right)^{\frac{1}{\delta}}\right]^{-\frac{1}{\theta}-1} \\ &\cdot \left[(u^{-\theta}-1)^\delta + (v^{-\theta}-1)^\delta\right]^{\frac{1}{\delta}-1} \cdot (v^{-\theta}-1)^{\delta-1} \cdot v^{-\theta-1} \ . \end{aligned} \tag{3.58}$$

Proposition 3.29 *Let $C_{\theta,\delta}(u,v)$ be a BB7 copula from Definition 3.42. The h-function of this copula is given by*

$$\begin{aligned} h_{\theta,\delta}(u,v) = &\left[1-\left\{\left(1-(1-u)^\theta\right)^{-\delta} + \left(1-(1-v)^\theta\right)^{-\delta} - 1\right\}^{-\frac{1}{\delta}}\right]^{\frac{1}{\theta}-1} \\ &\cdot \left[\left(1-(1-u)^\theta\right)^{-\delta} + \left(1-(1-v)^\theta\right)^{-\delta} - 1\right]^{-\frac{1}{\delta}-1} \\ &\cdot (1-(1-v)^\theta)^{-\delta-1} \cdot (1-v)^{\theta-1} \ . \end{aligned} \tag{3.59}$$

3.6.2 D-vines

As we saw in the previous section, the representation of conditional distribution is not unique. Hence, the factorization of joint density (3.44) is also not well-defined. Bedford and Cooke (2001,2002) develops a graphical structure, a so called *regular vine*, that is helpful by determining the decomposition of joint density. In this section we present only a special case of regular vines, a *D-vine*, which was introduced by Kurowicka and Cooke (2006).

The D-vine is one special method of decomposing the joint density. The factorization provided in a such manner is not unique, but it restricts the number of all possible decomposition enormously. The d-dimensional D-vine is specified by $d-1$ nested trees T_j, $j = 1, \ldots, d-1$, each of them consists of $d-j+1$ nodes and $d-j$ edges. Each node of the D-vine is connected to at most 2 other nodes of the same tree. Each edge of the D-vine corresponds to a bivariate copula density in a pair copula construction. When determining a D-vine for given data, one has to solve two problems: the evaluation of (pseudo-) data for each tree and copula identification for each edge. One builds the first tree with original data. For computing data for tree T_j, $j = 2, \ldots, d-1$, one needs (conditional) data from the previous tree. This problem can be solved by nesting h-functions. The second problem is not trivial and its solving requires some techniques of goodness-of-fit testing for copulas (see Chapter 4). Generally, density decomposition corresponding to the D-vine can be written in the form of

$$\begin{aligned} f(x_1, \ldots, x_d) &= \prod_{k=1}^{d} f(x_k) \\ &\cdot \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{i,i+j|i+1,\ldots,i+j-1} \left(F(x_i|x_{i+1}, \ldots, x_{i+j-1}), F(x_{i+j}|x_{i+1}, \ldots, x_{i+j-1}) \right) . \end{aligned} \tag{3.60}$$

Here, the index j corresponds to the D-vine trees and index i corresponds to the edges. Figure 3.9 illustrates a four-dimensional D-vine for data set $\boldsymbol{u} = (u_1, u_2, u_3, u_4)$, where $u_k \sim \mathcal{U}[0,1]$ for each $k = 1, 2, 3, 4$. Corresponding notations for the pair-copulas together with its parameters are placed below the edges, and whereas variables required for specific building block are placed above the edges. The first tree T_1 consists of the data $\boldsymbol{u}$ itself. For this step, we need only to determine pair-copulas $c_{1,2}$, $c_{2,3}$, $c_{3,4}$ and their parameter vectors $\boldsymbol{\theta}_{1,2}$, $\boldsymbol{\theta}_{2,3}$, $\boldsymbol{\theta}_{3,4}$. For constructing the second tree T_2 we need conditional variables $u_{1|2}$ and $u_{3|2}$ for edge $c_{1,3|2}$ and conditional variables $u_{2|3}$ and $u_{4|3}$ for edge $c_{2,4|3}$. Using results from the previous section, we get

$$u_{1|2} = h_{1,2}(u_1, u_2)$$

$$u_{3|2} = h_{3,2}(u_3, u_2) = h_{2,3}(u_3, u_2)$$

Figure 3.9: Example of a D-vine with 4 variables, 3 trees and 6 edges. Each edge will be associated with a pair-copula building block.

$$\mathbf{T_1}: \quad \mathbf{1} \xrightarrow[C_{\theta_{1,2}}]{u_2 \quad u_1} \mathbf{2} \xrightarrow[C_{\theta_{2,3}}]{u_2 \quad u_2} \mathbf{3} \xrightarrow[C_{\theta_{3,4}}]{u_2 \quad u_3} \mathbf{4}$$

$$\mathbf{T_2}: \quad \mathbf{1,2} \xrightarrow[C_{\theta_{1,3|2}}]{u_3|u_2 \quad u_1|u_2} \mathbf{2,3} \xrightarrow[C_{\theta_{2,4|3}}]{u_4|u_3 \quad u_2|u_3} \mathbf{3,4}$$

$$\mathbf{T_3}: \quad \mathbf{1,3|2} \xrightarrow[C_{\theta_{1,4|2,3}}]{u_4|u_2,u_3 \quad u_1|u_2,u_3} \mathbf{2,4|3}$$

and

$$u_{2|3} = h_{2,3}(u_2, u_3)$$
$$u_{4|3} = h_{4,3}(u_4, u_3) = h_{3,4}(u_4, u_3) \ .$$

At last, for evaluation of the third tree, we need conditional data

$$u_{1|2,3} = h_{1,3|2}\left(u_{1|2}, u_{3|2}\right)$$

and

$$u_{4|2,3} = h_{4,2|3}\left(u_{4|3}, u_{2|4}\right)$$

for the edge $1, 4|2, 3$.

Generally, the algorithm for sequential estimating pair-copula parameters consists in following steps:

(1) Estimate parameters of the pair-copulas using original data that is uniformly i.i.d.
(2) Compute conditional observations for the second tree using h-functions and estimated pair-copula parameters from step (1).
(3) Estimate pair-copula parameters for the second tree
(4) Repeat steps (2)-(3) for the next trees sequentially.

Pair-copula parameter achieved in a such way can be used as starting values in numerical maximization of the log-likelihood.

3.6.3 Inference for D-vine

Algorithm 1 can be used for calculating the copula part of the D-Vine log-likelihood defined in (3.60). The whole expression can be obtained by adding the logarithms of the marginal densities.

3.6.4 Simulation from D-vine

In this subsection, we present an algorithm for sampling from the D-vine according to Aas, Czado, Frigessi, and Bakken (2009). This algorithm is based on the Probability Integral Transformation Theorem 3.11 and requires the iterative application of the inverse h-functions.

Problem: Sample a set of uniform variables $x_1, \ldots, x_n \sim \mathcal{U}[0,1]$ with by the D-vine pre-defined dependence structure.

Idea: Simulate n independent variables $w_1, \ldots, w_n \sim \mathcal{U}[0,1]$. Variables $x_1, \ldots, x_n$ with

$$\begin{aligned} x_1 &= w_1 \\ x_2 &= F_{2|1}^{-1}(w_2|x_1) \\ x_3 &= F_{3|1,2}^{-1}(w_3|x_1, x_2) \\ \ldots &= \ldots \\ x_n &= F_{n|1,2,\ldots,n-1}^{-1}(w_n|x_1, \ldots, x_{n-1}) \end{aligned} \tag{3.61}$$

fulfill the pre-defined conditions. To calculate the condition distribution function

$$F_{j|1,\ldots,j-1}(w_n|x_1, \ldots, x_{j-1})$$

for each $j = 2, \ldots, n$, we use the definition of the h-function

$$h_{\boldsymbol{\theta}}(x, \nu) = \frac{\partial C_{\boldsymbol{\theta}}\left(F_x(x), F_\nu(\nu)\right)}{\partial F_\nu(\nu)} ,$$

and relationship

$$F(x|\boldsymbol{\nu}) = \frac{\partial C_{x\nu_j|\boldsymbol{\nu}_{-j}}\left(F(x|\boldsymbol{\nu_{-j}}), F(\nu_j|\boldsymbol{\nu_{-j}})\right)}{\partial F(\nu_j|\boldsymbol{\nu_{-j}})} ,$$

recursively. From the previous section we know, that the expression for $F(x|\boldsymbol{\nu})$ is not unique as long as $\boldsymbol{\nu}$ is not univariate. Though, $F(x|\boldsymbol{\nu})$ has an unique representation in case of D-vine, i.e.

$$F(x_j|x_1, \ldots, x_{j-1}) = \frac{\partial C_{j,1|2,\ldots,j-1}\left(F(x_j|x_2, \ldots, x_{j-1}, F(x_1|x_2, \ldots, x_{j-1})\right)}{\partial F(x_1|x_2, \ldots, x_{j-1})} .$$

Algorithm 1 Likelihood evaluation for the D-vine with uniformly distributed margins

input dependent variables $x_1, \ldots, x_n \sim \mathcal{U}[0,1]$
input parameters $\boldsymbol{\theta}_{i,j}$ of pair-copula densities $c_{i,i+j|i+1,\ldots,i+j-1}(\cdot,\cdot)$
log-likelihood $= 0$
for $i = 1, 2, \ldots, n$ **do**
  $\boldsymbol{v}_{0,i} = \boldsymbol{x}_i$
end for
for $i = 1, 2, \ldots, n-1$ **do**
  log-likelihood = log-likelihood + $L_{\boldsymbol{\theta}_{1,i}}(\boldsymbol{v}_{0,1}, \boldsymbol{v}_{0,i+1})$
end for
$\boldsymbol{v}_{1,1} = h_{\boldsymbol{\theta}_{1,1}}(\boldsymbol{v}_{0,1}, \boldsymbol{v}_{0,2})$
for $k = 1, 2, \ldots, n-3$ **do**
  $\boldsymbol{v}_{1,2k} = h_{\boldsymbol{\theta}_{1,k+1}}(\boldsymbol{v}_{0,k+2}, \boldsymbol{v}_{0,k+1})$
  $\boldsymbol{v}_{1,2k+1} = h_{\boldsymbol{\theta}_{1,k+1}}(\boldsymbol{v}_{0,k+1}, \boldsymbol{v}_{0,k+2})$
end for
$\boldsymbol{v}_{1,2n-4} = h_{\boldsymbol{\theta}_{1,n-1}}(\boldsymbol{v}_{0,n}, \boldsymbol{v}_{0,n-1})$
for $j = 2, \ldots, n-1$ **do**
  for $i = 1, 2, \ldots, n-j$ **do**
    log-likelihood = log-likelihood + $L_{\boldsymbol{\theta}_{i,j}}(\boldsymbol{v}_{j-1,2i-1}, \boldsymbol{v}_{j-1,2i})$
  end for
  if $j == (n-1)$ **then**
    Stop
  end if
  $\boldsymbol{v}_{j,1} = h_{\boldsymbol{\theta}_{j,1}}(\boldsymbol{v}_{j-1,1}, \boldsymbol{v}_{j-1,2})$
  if $n > 4$ **then**
    for $i = 1, 2, \ldots, n-j-2$ **do**
      $\boldsymbol{v}_{j,2i} = h_{\boldsymbol{\theta}_{j,i+1}}(\boldsymbol{v}_{j-1,2i+2}, \boldsymbol{v}_{j-1,2i+1})$
      $\boldsymbol{v}_{j,2i+1} = h_{\boldsymbol{\theta}_{j,i+1}}(\boldsymbol{v}_{j-1,2i+1}, \boldsymbol{v}_{j-1,2i+2})$
    end for
  end if
  $\boldsymbol{v}_{j,2n-2j-2} = h_{\boldsymbol{\theta}_{j,n-j}}(\boldsymbol{v}_{j-1,2n-2j}, \boldsymbol{v}_{j-1,2n-2j-1})$
end for
output log-likelihood

Algorithm 2 Simulation algorithm for D-vine: generates a set of uniform variables $x_1, \ldots, x_n \sim \mathcal{U}[0,1]$ with by the D-vine pre-defined dependence structure.

input independent variables $w_1, \ldots, w_n \sim \mathcal{U}[0,1]$
input parameters $\boldsymbol{\theta}_{i,j}$ of h-functions of corresponding copula density $c_{i,i+j|i+1,\ldots,i+j-1}(\cdot,\cdot)$

$x_1 = v_{1,1} = w_1$
$x_2 = v_{2,1} = h^{-1}_{\theta_{1,1}}(w_2, v_{1,1})$
$v_{2,2} = h_{\theta_{1,1}}(v_{1,1}, v_{2,1})$
for $i \leftarrow 3, 4, \ldots, n$ **do**
  $v_{i,1} = w_i$
  for $k \leftarrow i-1, i-2, \ldots, 2$ **do**
    $v_{i,1} = h^{-1}_{\theta_{k,i-k}}(v_{i,1}, v_{i-1,2k-2})$
  end for
  $v_{i,1} = h^{-1}_{\theta_{i,i-1}}(v_{i,1}, v_{i-1,1})$
  $x_i = v_{i,1}$
  if i==n **then**
    Stop
  end if
  $v_{i,2} = h_{\theta_{1,i-1}}(v_{i-1,1}, v_{i,1})$
  $v_{i,3} = h_{\theta_{1,i-1}}(v_{i,1}, v_{i-1,1})$
  if i $>$ 3 **then**
    for $j \leftarrow 2, 3, \ldots, i-2$ **do**
      $v_{i,2j} = h_{\theta_{j,i-j}}(v_{i-1,2j-2}, v_{i,2j-1})$
      $v_{i,2j+1} = h_{\theta_{j,i-j}}(v_{i,2j-1}, v_{i-1,2j-2})$
    end for
  end if
  $v_{i,2i-2} = h_{\theta_{i-1,1}}(v_{i-1,2i-4}, v_{i,2i-3})$
end for
output dependent variables $x_1, \ldots, x_n \sim \mathcal{U}[0,1]$

Given the h-function, we can calculate the inverse of conditional distribution through the inverse h^{-1}, i.e.

$$F^{-1}(x_j|x_1,\ldots,x_{j-1}) = h^{-1}_{j,1|2,\ldots,j-1}(x_j,(x_1,\ldots,x_{j-1}))\ .$$

Algorithm 2 provides a procedure for sampling from the D-vine. The index $\theta_{i,j}$ of the h-function denotes the parameter vector of corresponding copula density $c_{i,j|i+1,\ldots,i+j-1}$.

3.7 Time series pair-copula approach

In the previous sections, we introduced time series models for univariate time series and the features of copulas for modeling dependency in multivariate data. At last, we studied how bivariate copulas can be extended to the higher dimensions via pair-copula constructions. In this section we present a methodology for modeling multiple dependence, a so called TS-PAIR-COPULA approach (TSPCA). This method is performed in two steps. The first step, we apply (ARMA-) GARCH filter to each of the univariate return series to model conditional mean and conditional volatility. Resulting standardized residuals of each model should be identical and independent distributed under true model specification. The second step, we analyze multiple dependence between standardized residuals transformed to the interval $[0,1]$ via a pair-copula approach. In such way, univariate margins and dependencies are modeled separately. For modeling future returns, we do it backwards. At the first step, we sample random vectors on $[0,1]^d$ from the estimated pair-copula construction and transform them to standardized residuals. Further, these quasi standardized residuals should be rescaled to the future returns by using forecasted means and variances that have been estimated with the marginal (ARMA-) GARCH models.

Let $\boldsymbol{r}_{1\ldots t} = (\boldsymbol{r}_1,\ldots,\boldsymbol{r}_d)$ be a d-dimensional data set of interest with components $\boldsymbol{r}_i = (r_{i1},\ldots,r_{it})^t$ for all $i = 1\ldots d$. Each $\boldsymbol{r}_i$ consists of log-returns of any financial index observed from the time period $[1,t]$. One is interested in simulating future multivariate returns

$$\tilde{\boldsymbol{r}}_{t+1\ldots t+m} = \begin{pmatrix} \tilde{r}_{1,t+1} & \cdots & \tilde{r}_{d,t+1} \\ & \cdots & \\ \tilde{r}_{1,t+m} & \cdots & \tilde{r}_{d,t+m} \end{pmatrix}$$

for the time period $[t+1,t+m]$. We denote by $\hat{\cdot}$ estimated values and by $\tilde{\cdot}$ forecasted values. Following algorithm summarize the proposed procedures.

I. Modeling multiple dependence between time series

(1) Fit an appropriate (ARMA-) GARCH model with assumed innovation distributions F_i to vector $\boldsymbol{x}_i$ with fitted conditional mean $\hat{\boldsymbol{\mu}}_i = (\hat{\mu}_{i1},\ldots,\hat{\mu}_{it})^t$ and conditional volatility $\hat{\boldsymbol{\sigma}}_i = (\hat{\sigma}_{i1},\ldots,\hat{\sigma}_{it})^t$ for each $i = 1\ldots d$. A residual analysis is performed to check the model assumptions and fit the (ARMA-) GARCH model.

(2) Compute standardized residuals

$$\hat{\varepsilon}_{ij} := \frac{x_{ij} - \hat{\mu}_{ij}}{\hat{\sigma}_{ij}} \quad \text{for all} \quad i = 1 \dots d,\ j = 1 \dots t\ .$$

(3) Apply the estimated innovation distribution $\hat{F}_i$ and the Probability Integral Transformation to transform the standardized residuals to the unit interval

$$u_{ij} := \hat{F}_i(\hat{\varepsilon}_{ij}) \sim \mathcal{U}(0,1) \quad \text{for all} \quad i = 1 \dots d,\ j = 1 \dots t\ .$$

(4) Fit a pair-copula construction to the set $\boldsymbol{u} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_d)$ with $\boldsymbol{u}_i = (u_{i1}, \dots, u_{it})^t$ for all $i = 1 \dots d$. Appropriate pair-copulas will be chosen using the techniques discussed in Chapter 4.

(5) Save estimated parameters of the decomposed model from step (4).

II. Simulation of multivariate returns

(1) Generate a random sample $\tilde{\boldsymbol{u}} = (\tilde{\boldsymbol{u}}_1, \dots, \tilde{\boldsymbol{u}}_d)$ from the estimated pair-copula model with $\tilde{\boldsymbol{u}}_i = (\tilde{u}_{i,t+1}, \dots, \tilde{u}_{i,t+m})^t$ and $\tilde{u}_{ij} \sim \mathcal{U}(0,1)$ for all $i = 1 \dots d$, $j = t+1, \dots, t+m$.

(2) Transform each $\tilde{\boldsymbol{u}}_i$ to the vector of standardized residuals by applying the inverse of $\hat{F}_i$

$$\tilde{\varepsilon}_{ij} := \hat{F}_i^{-1}(\tilde{u}_{ij}) \quad \text{for all} \quad i = 1 \dots d,\ j = t+1, \dots, t+m\ .$$

Due to the Probability Integral Transformation Theorem 3.11, it holds for $\tilde{\boldsymbol{\varepsilon}}_i = (\tilde{\varepsilon}_{i,t+1}, \dots, \tilde{\varepsilon}_{i,t+m})^t$

$$\tilde{\varepsilon}_{ij} \sim \hat{F}_i \quad \text{for all} \quad i = 1 \dots d, \quad j = t+1 \dots t+m\ .$$

(3) Forecast future return and future conditional volatility at th time point $t+1$ using equations (2.22) and (2.23), respectively, and estimated (ARMA-) GARCH model parameters. For example, for an ARMA(1,1)-GARCH(1,1) model we get expressions

$$\tilde{r}_{i,t+1} := \hat{\mu}_i + \hat{\varphi}_i \cdot r_{i,t} + \hat{\theta}_i \cdot \hat{\epsilon}_{i,t} = \hat{\mu}_i + \hat{\varphi}_i \cdot r_{i,t} + \hat{\theta}_i \cdot \hat{\sigma}_{i,t} \cdot \hat{\varepsilon}_{i,t}$$

$$\tilde{\sigma}^2_{i,t+1} := \hat{\omega}_i + \hat{\alpha}_i \cdot \hat{\epsilon}^2_{i,t} + \hat{\beta}_i \cdot \hat{\sigma}^2_{i,t} = \hat{\omega}_i + \hat{\alpha}_i \cdot (\hat{\sigma}_{i,t} \cdot \hat{\varepsilon}_{i,t})^2 + \hat{\beta}_i \cdot \hat{\sigma}^2_{i,t}\ ,$$

and for an GARCH(1,1) model

$$\tilde{\sigma}^2_{i,t+1} := \hat{\omega}_i + \hat{\alpha}_i \cdot r^2_{i,t} + \hat{\beta}_i \cdot \hat{\sigma}^2_{i,t}\ ,$$

$$\tilde{r}_{i,t+1} := \tilde{\sigma}_{i,t+1} \cdot \tilde{\varepsilon}_{i,t+1}\ .$$

(4) Using forecasted values $\tilde{r}_{i,t+1}$, $\tilde{\sigma}_{i,t+1}$, and $\tilde{\varepsilon}_{i,t+1}$ from the previous steps, predict future return and future conditional volatility at time point $t+2$.
For an ARMA(1,1)-GARCH(1,1) model we get

$$\tilde{r}_{i,t+2} := \hat{\mu}_i + \hat{\varphi}_i \cdot \tilde{r}_{i,t+1} + \hat{\theta}_i \cdot \tilde{\sigma}_{i,t+1} \cdot \tilde{\varepsilon}_{i,t+1} \ ,$$

$$\tilde{\sigma}^2_{i,t+2} := \hat{\omega}_i + \hat{\alpha}_i \cdot (\tilde{\sigma}_{i,t+1} \cdot \tilde{\varepsilon}_{i,t+1})^2 + \hat{\beta}_i \cdot \tilde{\sigma}^2_{i,t+1} \ ,$$

and for an GARCH(1,1) model

$$\tilde{\sigma}^2_{i,t+2} := \hat{\omega}_i + \hat{\alpha}_i \cdot \tilde{r}^2_{i,t+1} + \hat{\beta}_i \cdot \tilde{\sigma}^2_{i,t+1} \ ,$$

$$\tilde{r}_{i,t+2} := \tilde{\sigma}_{i,t+2} \cdot \tilde{\varepsilon}_{i,t+2} \ .$$

Note, if one is interested only on the *one step ahead* forecasting, the true values $r_{i,t+2}$ should be used instead off predicted values $\tilde{r}_{i,t+2}$ for calculating conditional volatility at time point $t+2$.

(5) Similar to the step (4), forecast future returns and volatilities for the next time points $t+3, \ldots, t+m$.

Chapter 4

Goodness-of-fit tests for copulas

If one wants to investigate the dependence structure in multivariate data, one needs to solve two different problems. The first one is choosing an appropriate parametric copula family. In this case, the validity of the null hypotheses $H_0 : C \in \mathcal{C}$ for some class $\mathcal{C}$ of copulas should be tested. For example, $\mathcal{C} = \{\text{Gauss}, \text{Student}, \text{Clayton}, \ldots\}$. The second one is the problem of estimating the dependence parameter $\boldsymbol{\theta}$ on the chosen copula class $C = \{C_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\}$, where Θ is the range of all possible values of $\boldsymbol{\theta}$. There are many proceedings for goodness-of-fit testing of copula models, an overview of which is given by Genest, Rémillard, and Beaudoin (2009). The tests presented in this paper require pseudo-copula-observations as an input and test whether the data really comes from estimated copula. In some situations, tests proposed by Genest, Rémillard, and Beaudoin (2009) do not reject the null hypothesis for more than one copula family or the null hypothesis is rejected for all copula families of interest. In the first case, one should make a subjective judgment to decide whether the chosen copula is most suitable. In the second case, one can not make any decision. In this section we provide an alternative methodology for goodness-of-fit testing that is based on so called Vuong or Clarke test. The Vuong and Clarke tests are methods that were developed for the comparison of non-nested models. Our approach favors either only one copula or states the equivalence of several copulas. Additionally, we study the power of the pairwise Vuong and Clarke tests and investigate the accuracy of the proposed score method. This will depend on sample size, degree of association and underlying copula family.

4.1 Vuong Test

For the comparison of two copulas fitted to the same data, we utilize the Vuong test proposed by Vuong (1989). The Vuong test compares two regression models which need not to be nested and is based on the Kullback-Leibler information criterion (KLIC), see Kullback and Leibler (1951). KLIC measures the "distance" between two statistical

models. Consider a data set of the length n consisting of a dependent variable Y_i and a possible set of explanatory variables $\boldsymbol{x}_i$. We have

$$KLIC := E_0\left[\log h_0(Y_i|\boldsymbol{x}_i)\right] - E_0\left[\log f(Y_i|\boldsymbol{x}_i, \hat{\boldsymbol{\beta}})\right] , \tag{4.1}$$

with $h_0(\cdot)$ the true (but unknown) conditional density of Y_i given $\boldsymbol{x}_i$, $i = 1 \dots n$, and E_0 expectation given the true model. Further, $f(\cdot)$ is conditional density of the fitted model with estimated parameters $\hat{\boldsymbol{\beta}}$. Generally, the one of two models with the smaller KLIC is the better one, for it is closer to the true, but unknown, specification. For comparison of two models with conditional densities $f_1(Y_i|\boldsymbol{x}_i, \hat{\boldsymbol{\beta}}^1)$ and $f_2(Y_i|\boldsymbol{x}_i, \hat{\boldsymbol{\beta}}^2)$ one has to look at their KLIC values. If model 1 is better than model 2, its KLIC statistic is smaller and the following inequality holds

$$KLIC_1 < KLIC_2$$

or in terms of (4.1)

$$E_0\left[\log h_0(Y_i|\boldsymbol{x}_i)\right] - E_0\left[\log f_1(Y_i|\boldsymbol{x}_i, \hat{\boldsymbol{\beta}}^1)\right] < E_0\left[\log h_0(Y_i|\boldsymbol{x}_i)\right] - E_0\left[\log f_2(Y_i|\boldsymbol{x}_i, \hat{\boldsymbol{\beta}}^2)\right]$$

This expression reduces to

$$E_0\left[\log\left(\frac{f_1(Y_i|\boldsymbol{x}_i, \hat{\boldsymbol{\beta}}^1)}{f_2(Y_i|\boldsymbol{x}_i, \hat{\boldsymbol{\beta}}^2)}\right)\right] > 0 . \tag{4.2}$$

In other words, the model 1 is favored over the model 2, if its log-likelihood values are significantly larger. Vuong proposed the following statistic

$$m_i := \log\left(\frac{f_1(Y_i|\boldsymbol{x}_i, \hat{\boldsymbol{\beta}}^1)}{f_2(Y_i|\boldsymbol{x}_i, \hat{\boldsymbol{\beta}}^2)}\right) \quad \text{with} \quad i = 1 \dots n .$$

Then $\boldsymbol{m} := (m_1, \dots, m_n)^t$ is a random vector with expectation

$$E_0[\boldsymbol{m}] := \boldsymbol{\mu}_0^m = (\mu_1^m, \dots, \mu_n^m)^t ,$$

if $h_0(\cdot)$ is the true probability mass function. If both models are equally close to the true specification, it holds $\boldsymbol{\mu}_0^m = 0$. Hence, we formulate our test problem as

$$H_0 : \ \boldsymbol{\mu}_0^m = 0 \quad \text{against} \quad H_1 : \ \boldsymbol{\mu}_0^m \neq 0 . \tag{4.3}$$

The quantity $\boldsymbol{\mu}_0^m$ is unknown. Based on $\boldsymbol{m}$, Vuong defined a test statistic ν

$$\nu := \frac{\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n m_i\right)}{\sqrt{\frac{1}{n}\sum_{i=1}^n (m_i - \bar{m})^2}} \quad \text{with} \quad \bar{m} := \frac{1}{n}\sum_{i=1}^n m_i \tag{4.4}$$

and has shown that under H_0

$$\nu \xrightarrow{D} \mathcal{N}(0,1) \ .$$

So, we formulate a decision rule for (4.3) in a following way

VUONG TEST: Reject the null hypothesis of equivalence of the two models

$$H_0 : \boldsymbol{\mu}_0^m = 0 \ \textbf{\textit{against}} \ H_1 : \ \boldsymbol{\mu}_0^m \neq 0$$

at significance level α***, if*** $|\nu| \geq z_{1-\frac{\alpha}{2}}$***, where*** $z_{1-\frac{\alpha}{2}}$ ***is a*** $\left(1-\frac{\alpha}{2}\right)$***-quantile of the standard normal distribution.***

The test favors model 1 over model 2, if $\nu \geq z_{1-\frac{\alpha}{2}}$. This is reasonable since significantly high values of ν indicate a higher KLIC of model 1 compared to model 2 according to the formula (4.1). And vice versa, test chooses the model 2, if $\nu \leq z_{1-\frac{\alpha}{2}}$. Both models are equivalent if $-z_{1-\frac{\alpha}{2}} < \nu < z_{1-\frac{\alpha}{2}}$.

The Vuong test is not an exact test, since the test statistic is normally distributed only asymptotically. Simulations demonstrate that it is not very powerful in small samples, Clarke (2003).

The Vuong statistic is sensitive to the number of estimated coefficients in both models. It is based only on the individual log-likelihood values and does not incorporate a number of model parameters in the computation. Vuong suggested to improve his test by adding a so called Schwarz's correction term

$$\frac{p_1}{2}\log n - \frac{p_2}{2}\log n \ , \tag{4.5}$$

where p_1 and p_2 are the numbers of parameters in model 1 and model 2 respectively, n is a number of observations. Thus, the adjusted Vuong statistic with Schwarz's correction is defined as

$$\tilde{\nu} := \frac{\sqrt{n}\left[\left(\frac{1}{n}\sum_{i=1}^n m_i\right) - \left(\frac{p_1}{2}\log n - \frac{p_2}{2}\log n\right)\right]}{\sqrt{\frac{1}{n}\sum_{i=1}^n (m_i - \bar{m})^2}} \tag{4.6}$$

Note, the correction of individual log-likelihoods of model 1 by the term $\frac{p_1}{2n}\log n$ and the model 2 by the term $\frac{p_2}{2n}\log n$ leads to the same adjusted test statistic $\tilde{\nu}$, i.e

$$\tilde{\nu} := \frac{\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n \tilde{m}_i\right)}{\sqrt{\frac{1}{n}\sum_{i=1}^n (\tilde{m}_i - \bar{\tilde{m}})^2}} \tag{4.7}$$

with

$$\begin{aligned} \tilde{m}_i &= m_i - \left(\frac{p_1}{2n}\log n - \frac{p_2}{2n}\log n\right) \\ &= \left[\log f_1(Y_i|\boldsymbol{x}_i, \hat{\boldsymbol{\beta}}^1) - \frac{p_1}{2}\log n\right] - \left[\log f_2(Y_i|\boldsymbol{x}_i, \hat{\boldsymbol{\beta}}^2) - \frac{p_2}{2}\log n\right] \ . \end{aligned} \tag{4.8}$$

An alternative correction term due to Akaike can be also used. For individual log-likelihoods

it is defined as

$$(p_1 - p_2)/n \ . \tag{4.9}$$

The choice of a significance level α can be deferred to the user by considering a p-value. The p-value is the smallest α value at which the null hypothesis can be rejected for the data at hand. Before one starts with the testing of any hypothesis, the so called *significance level* α should be chosen. The significance level is a fixed probability of rejecting the null hypothesis wrongly, if it is in fact true. If the calculated p-value is less than the chosen significance level then the null hypothesis should be rejected. We can calculate the p-value of the Vuong test as follows: the smallest α at which the test rejects is given by

$$|\nu| = z_{1-\frac{\alpha}{2}} \ . \tag{4.10}$$

Let $\Phi(\cdot)$ denote the distribution function of the standard normal distribution. Applying $\Phi(\cdot)$ to both sides of (4.10) and using symmetry of $\Phi(\cdot)$, we become

$$\begin{aligned} & \Phi(|\nu|) = 1 - \frac{\alpha}{2} \\ \Leftrightarrow \ & \frac{\alpha}{2} = 1 - \Phi(|\nu|) \\ \Leftrightarrow \ & \alpha = 2(1 - \Phi(|\nu|)) \\ \Leftrightarrow \ & \alpha = 2\Phi(-|\nu|) \end{aligned} \tag{4.11}$$

The last expression of (4.11) provides a formula for calculating the p-value for the Vuong statistic

$$\mathrm{p} = 2\Phi(-|\nu|) \ . \tag{4.12}$$

The R code of the Vuong test is given in Appendix D.

4.2 Clarke Test

An alternative for comparing of non-nested models is distribution-free test proposed by Clarke (2007). Likewise the Vuong test, it is based on the Kullback-Leibler Information Criterion (4.1) and compares whether the log-likelihood of one model is significantly larger than the log-likelihood of the other model. Similar to the Vuong test, the null hypothesis of this test postulates the equivalence of both models. With the notions from the previous section we formulate the null hypothesis of the Clarke test as

$$H_0 : P\left(\log\left(\frac{f_1(Y_i|\boldsymbol{x}_i^1, \hat{\boldsymbol{\beta}}^1)}{f_2(Y_i|\boldsymbol{x}_i^2, \hat{\boldsymbol{\beta}}^2)} > 0\right)\right) = 0.5 \ . \tag{4.13}$$

This equation means that the individual log-likelihoods are equally distributed around zero under the null hypothesis, i.e. one half of the individual log-likelihood ratios should

be greater than zero and other half should be less than zero. We introduce d_i, $i = 1 \dots n$

$$m_i = \log f_1(Y_i|\boldsymbol{x}_i^1, \hat{\boldsymbol{\beta}}^1) - \log f_2(Y_i|\boldsymbol{x}_i^2, \hat{\boldsymbol{\beta}}^2) \tag{4.14}$$

and define a test statistic as

$$B = \sum_{i=1}^{n} \mathbb{1}_{(0,\infty)}(m_i)$$

where $\mathbb{1}(\cdot)$ denotes an indicator function. B corresponds to the number of positive differences d_i and can be interpreted as binomial distributed random variable with parameters n and $p = 0.5$ under the null hypothesis, i.e.

$$B \sim \text{Binom}(n, p) \ .$$

Model 1 characterized by the parameter vector $\hat{\boldsymbol{\beta}}^1$ is equivalent to the model 2 characterized by the parameter vector $\hat{\boldsymbol{\beta}}^2$, if B is equal to the expectation $np = n/2$ under the null hypothesis. One can assume the equivalence of both models at significance level α, if

$$B \in \left(\frac{n}{2} - \varepsilon_\alpha, \frac{n}{2} + \varepsilon_\alpha\right) \tag{4.15}$$

for ε_α ample small. Using $c_{\alpha+} := \frac{n}{2} + \varepsilon_\alpha$ and $c_{\alpha-} := \frac{n}{2} - \varepsilon_\alpha$, expression (4.15) has the form

$$B \in (c_{\alpha-}, c_{\alpha+}) \tag{4.16}$$

If model 1 is better then model 2, B will be significantly larger then its expected value $\frac{n}{2}$ under the null hypothesis. Generally, it is difficult to construct a two-sided test for the testing problem (4.13). Usage of one-sided tests makes sense in such situations. We split the testing problem

$$H_0 : \ B = \frac{n}{2} \text{ against } H_1 : \ B \neq \frac{n}{2}$$

into two cases: the upper tail test

$$H_0 : \ B = \frac{n}{2} \text{ versus } H_1 : \ B > \frac{n}{2} \tag{4.17}$$

and the lower tail test

$$H_0 : \ B = \frac{n}{2} \text{ versus } H_1 : \ B < \frac{n}{2} \ . \tag{4.18}$$

Next, we need to determine a rejection region for (4.17) and (4.18), respectively. The upper tail testing problem has a significance level α, if the error probability of type 1 is not larger then α, i.e.

$$P\left(B \geq c_{\alpha+}\right) \leq \alpha \ . \tag{4.19}$$

Using the formula

$$P\left(B \leq a\right) = \sum_{i=1}^{a} \binom{n}{i} p^i (1-p)^{n-i} \tag{4.20}$$

for $B \sim Binom(n, p)$, the expression (4.19) can be written as

$$\sum_{c=c_{\alpha+}}^{n} \binom{n}{c} 0.5^n \leq \alpha \ . \tag{4.21}$$

UPPER TAIL CLARKE TEST: ***Reject the null hypothesis of equivalence***

$$H_0 : \ B = \frac{n}{2} \ \textbf{\textit{versus}} \ H_1 : \ B > \frac{n}{2}$$

at significance level α***, if*** $B \geq c_{\alpha+}$***, where*** $c_{\alpha+}$ ***is chosen to be the smallest integer such that*** (4.20) ***holds.***

For determining $c_{\alpha+}$ we use the equals sign in (4.19)

$$\begin{aligned}
& P\left(B \geq c_{\alpha+}\right) = \alpha \\
\Leftrightarrow \ & 1 - P\left(B < c_{\alpha+}\right) = \alpha \\
\Leftrightarrow \ & P\left(B < c_{\alpha+}\right) = 1 - \alpha \\
\Leftrightarrow \ & P\left(B \leq c_{\alpha+} - 1\right) = 1 - \alpha \\
\Leftrightarrow \ & \mathcal{B}\left(c_{\alpha+} - 1\right) = 1 - \alpha \\
\Leftrightarrow \ & c_{\alpha+} - 1 = z_{bin}(1 - \alpha) \\
\Leftrightarrow \ & c_{\alpha+} = 1 + z_{bin}(1 - \alpha) \ ,
\end{aligned} \tag{4.22}$$

where $\mathcal{B}$ denotes the distribution function of $B \sim Binom(n, 0.5)$ and z_{bin} is its quantile function. The p-value is the smallest significance level at which the test rejects H_0. Hence, we compute the p-value of the upper tail test by setting

$$\begin{aligned}
& B = c_{\alpha+} \\
\Leftrightarrow \ & B = 1 + z_{bin}(1 - \alpha) \\
\Leftrightarrow \ & B - 1 = z_{bin}(1 - \alpha) \\
\Leftrightarrow \ & \mathcal{B}(B - 1) = 1 - \alpha \\
\Leftrightarrow \ & \alpha = 1 - \mathcal{B}(B - 1) \\
\Leftrightarrow \ & p = 1 - \mathcal{B}(B - 1)
\end{aligned} \tag{4.23}$$

The lower tail testing problem has a significance level α, if the error probability of type 1 is not larger then α, i.e.

$$P\left(B \leq c_{\alpha-}\right) \leq \alpha \tag{4.24}$$

or

$$\sum_{c=0}^{c_{\alpha-}} \binom{n}{c} 0.5^n \leq \alpha \ . \tag{4.25}$$

So, we formulate the lower tail Clarke test as

LOWER TAIL CLARKE TEST: ***Reject the null hypothesis of equivalence***

$$H_0 : \ B = \frac{n}{2} \ \textbf{\textit{versus}} \ H_1 : \ B < \frac{n}{2}$$

at significance level α***, if*** $B \leq c_{\alpha-}$***, where*** $c_{\alpha-}$ ***is chosen to be the smallest integer such that*** (4.24) ***holds.***

For determining $c_{\alpha-}$ we use the equals sign in (4.23)

$$\begin{aligned} & P\left(B \leq c_{\alpha-}\right) = \alpha \\ \Leftrightarrow \ & \mathcal{B}\left(c_{\alpha-}\right) = \alpha \\ \Leftrightarrow \ & c_{\alpha-} = z_{bin}(\alpha) \ , \end{aligned} \tag{4.26}$$

We evaluate the p-value of the lower tail test by setting

$$\begin{aligned} & B = c_{\alpha-} \\ \Leftrightarrow \ & B = z_{bin}(\alpha) \\ \Leftrightarrow \ & \mathcal{B}(B) = \alpha \\ \Leftrightarrow \ & p = \mathcal{B}(B) \ . \end{aligned} \tag{4.27}$$

The Clarke test is also sensible to the number of estimated parameters in each model. Since the distribution-free test works with the individual log-likelihood ratios, we can not apply the Schwarz's correction term (4.5) as in the Vuong test with the *summed* log-likelihood ratio. Clarke suggested to use an *average* correction. Individual log-likelihoods of the model 1 and model 2 should to be corrected by the terms $\frac{p_1}{2n} \log n$ and $\frac{p_2}{2n} \log n$, respectively. The adjusted test statistic becomes a form

$$\tilde{B} = \sum_{i=1}^{n} \mathbb{1}_{(0,\infty)}(\tilde{m}_i)$$

where $\tilde{m}_i$ is defined as in (4.8).

4.3 Simulation study and comparison of the Vuong and Clarke tests

In this section, we present results of the simulation study where we used the Vuong and Clarke tests for the bivariate goodness-of-fit testing for copulas:

- We simulated data from each copula $C_i \in \mathcal{C}$ with true parameter vector $\boldsymbol{\theta}_i$, where $i = 1, \ldots 8$ and $\mathcal{C} = \{Normal, Student, Clayton, Gumbel, Frank, Plackett, BB1, BB7\}$. For the used fixed degrees of freedom for the Student copula $df = 4$ and determined BB1 and BB7 copula so that $\lambda_U = 0.3$

- We fitted every copula C_j to the simulated data and denoted the estimated parameters by $\hat{\boldsymbol{\theta}}_j$, $j = 1, \ldots, 8$.
- We calculated the Vuong and Clarke statistics to test the null hypothesis of equivalence of two models, C_i with true parameter vector $\boldsymbol{\theta}_i$ and C_j with estimated parameter vector $\hat{\boldsymbol{\theta}}_j$, $i, j = 1, \ldots, 8$.
- We repeated sampling procedure $N = 1000$ times for each copula $C_i \in \mathcal{C}$ and evaluated percentages of non-rejecting of the null hypothesis, percentages of rejecting the null hypothesis in favor of true model and percentages of rejecting the null hypothesis in favor of estimated model, which is not a true model.
- For the Vuong test as well as for the Clarke test we used Schwarz's correction term.
- We repeated our simulation study for two different choices of sample length $n \in \{150, 1000\}$ and three different choices of Kendall's $\tau \in \{0.25, 0.5, 0.75\}$.

Summarizing results of this simulation study are given in Tables 4.2-4.4. Looking at these tables, we state:

- Decision results of the bivariate Vuong and Clarke tests depend on the sample size and on the Kendall's τ.
- Error rate of the Clarke test is much more higher than the error rate of the Vuong test. It holds for all possible combinations of n and τ. Under *error rate* we understand percentages of rejection of the null hypothesis in favor of false model.
- Accuracy of both tests increases with the rising sample length. Under *accuracy* we understand percentages of rejection of the null hypothesis in favor of true model.
- For large sample size and strong association, the chance of choosing the true copula is close to 100% for both tests in exeption of several cases.

4.4 Distances between different copula families

As mentioned above, the results of the pairwise Vuong and Clarke tests depend substantially on the the combination of τ and sample size n as well as on the two copulas to be compared. To investigate the influence of copula type, we measure the distance between every two copula classes from the set $\{Normal, Student, Clayton, Gumbel, Frank, Plackett, BB1, BB7\}$ for different values of Kendall's $\tau \in \{0.25, 0.5, 0.75\}$. We do it by applying the Kullback-Leibler information criterion (KLIC) that we have introduced in Section 4.1. For given bivariate data set, we are interested whether the copula family $\mathcal{C}_1$ or the copula family $\mathcal{C}_2$ is the most appropriate for describing dependence structure within the data. Let $\boldsymbol{\theta}_1$

and $\boldsymbol{\theta}_2$ be (estimated) parameters of copula families $\mathcal{C}_1$ and $\mathcal{C}_2$ respectively. Hence, the distance between two copulas is measured by KLIC in a following way

$$KLIC(C_1, C_2) = \int_0^1 \int_0^1 \log\left(\frac{c_{\boldsymbol{\theta}_1}(x_1, x_2)}{c_{\boldsymbol{\theta}_2}(x_1, x_2)}\right) c_{\boldsymbol{\theta}_1}(x_1, x_2)\mathrm{d}x_1 \, \mathrm{d}x_2 \ , \tag{4.28}$$

where $c_{\boldsymbol{\theta}_1}$ and $c_{\boldsymbol{\theta}_2}$ are the densities of copulas 1 and 2. Given i.i.d. observations (x_{1i}, x_{2i}), $i = 1, \ldots, n$, from copula with the true or estimated density $c_{\boldsymbol{\theta}_1}$, the KLIC between this and some other copula with density $c_{\boldsymbol{\theta}_2}$ can be estimated as

$$\widehat{KLIC}(C_1, C_2) = \frac{1}{n}\sum_{i=1}^{n} \log\left(\frac{c_{\boldsymbol{\theta}_1}(x_{1i}, x_{2i})}{c_{\boldsymbol{\theta}_2}(x_{1i}, x_{2i})}\right) \ , \tag{4.29}$$

the parameter vectors $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are either known or will be estimated from the sample data. Note, it holds following inequality

$$KLIC(C_1, C_2) \neq KLIC(C_2, C_1) \ .$$

Figure 4.1 (a) illustrates KLIC distances that were estimated empirically using formula 4.29. KLICs were calculated for data sets of length $n = 10000$ sampled from different copula families: Normal, Student ($df = 4$), Clayton, Gumbel, Frank, Plackett, BB1 ($\lambda_U = 0.3$) and BB7 ($\lambda_U = 0.3$). Besides, Figure 4.1 (b) displays KLIC distances that were evaluated by numerical integration of expression 4.28. These calculations were executed by Brechmann and can be found in his diploma thesis Brechmann (2010). The results for both figures were obtained for different Kendall's $\tau \in \{0.25, 0.5, 0.75\}$. Points with notation $C_1 - C_2$ on the x-axes correspond to the KLIC distances between true copula C_1 and another copula C_2 which was fitted to the data from C_1. We abbreviate names of copula families by its first letters except of BB1 and BB7 copulas. Hence, N corresponds to Normal copula, S to Student, P to Plackett, F to Frank, C to Clayton and G to Gumbel. According to this notation, points with label $N - P$ display the KLIC distances between two copulas, Normal and Plackett, where Plackett copula was fitted to the data from Normal copula. Black circular points represent the KLICs for data with $\tau = 0.25$. Blue square points accord to the KLICs evaluated for the data with $\tau = 0.5$. At last, red triangle points correspond to the bivariate data characterized by $\tau = 0.75$.

Looking at Figure 4.1, we state:

- There are some differences between KLICs obtained empirically or numerically. We restrict our further interpretation only on the KLIC values estimated by numerical integration.
- KLIC measures are generally not symmetrical: the distance between copula C_1 and copula C_2 is not equal to the distance between copula C_2 and copula C_1.

- Dissimilarities between copula families increase with rising degree of dependence τ, except of some special cases. Firstly, distances $N-S$, $BB1-BB7$ and $BB7-BB1$ do not depend on the Kendall's τ. Secondly, distances $C-BB1$, $BB1-C$, $C-BB7$ and $BB7-C$ decrease with rising degree of dependence.
- Distances $S-N$, $N-G$, $G-N$, $S-P$, $P-S$, $S-G$, $G-S$, $F-P$, $P-F$, $G-P$, $C-BB1$, $BB1-C$, $C-BB7$ and $BB7-C$ do not change essentially with increasing Kendall's τ.

4.5 Relationship between accuracy of the Vuong/Clarke test and KLIC values

Now, we interpret Tables 4.2-4.4 taking into account the distances between specific copula families. Examination of all possible copula pairs and all combinations of τ and n is not possible because of a huge number. Therefore, we do it in following examples:

- **Clayton - BB1**: KLIC distances for this copula pair decrease with increasing Kendall's τ. Hence, we expect that accuracy of the bivariate Vuong and Clarke tests falls when Kendall's τ goes up. For the small sample size $n = 150$ the percentages of choosing true copula are 51.17, 41.29, 35.34 for the Vuong test and 55.06, 55.29, 62.86 for the Clarke test, when Kendall's τ takes values 0.25, 0.5, 0.75 respectively. For the large sample size $n = 1000$ the percentages of choosing true copula are 58.96, 52.12, 51.61 for the Vuong test and 53.26, 53.53, 62.05 for the Clarke test, when Kendall's τ takes values 0.25, 0.5, 0.75 respectively. So, our presumption will be fulfilled only for the Vuong test and it depends on the sample length: the accuracy of the Vuong test declines more rapidly for small samples as for large samples. Surprisingly, the accuracy of the Clarke test goes up when Kendall's τ increases for small as well as for large samples.
- **Student - Normal**: KLIC values barely change, when Kendall's τ goes up. Hence, we expect the same behavior for accuracy of the Vuong and Clarke tests. For the small sample size $n = 150$ the percentages of choosing true copula are 2, 2.8, 3.2 for the Vuong test and 30.7, 34.9, 35.8 for the Clarke test when Kendall's τ takes values 0.25, 0.5, 0.75 respectively. For the large sample size $n = 1000$ the percentages of choosing true copula are 80, 82.5, 78.6 for the Vuong test and 99.9, 98.3, 96.3 for the Clarke test, when Kendall's τ takes values 0.25, 0.5, 0.75 respectively. As we can see, there is a strong dependence between the accuracy and the sample length for both the Vuong and the Clarke test. The results do not contradict our presumption: there is no depict association between Kendall's τ and accuracy of both tests.

Based on these proposals, we conclude our results:
The accurateness of the bivariate Vuong and Clarke tests depends on the KLIC distances

between copulas to be compared. This relationship is more stronger for small sample sizes than for large sample sizes. Moreover, the Vuong test seems to have harder connection to the KLIC distances in comparison to the Clarke test.

4.6 Comparison to Genest study

Next, we are interested whether the bivariate Vuong and Clarke tests are more efficient than goodness-of-fit tests proposed by Genest, Rémillard, and Beaudoin (2009). The authors give a critical overview of common bivariate goodness-of-fit techniques for copulas and suggest the new ones, a total of seven test statistics. Additionally, the authors study the finite-sample properties of the presented goodness-of-fit methodologies in depending on degree of association Kendall's τ and specific copula classes to be compared. For every $\tau \in \{0.25, 0.5, 0.75\}$, they generated $N = 10000$ random samples of length $n = 150$ from each copula $C_i \in \mathcal{C} = \{Gauss, Student\ (df=4), Clayton, Gumbel, Frank, Plackett\}$, $i = 1, \ldots, 6$, and tested the null hypothesis H_0 : *data comes from* C_j with $C_j \in \mathcal{C}$, $j = 1, \ldots, 6$, according to each of the seven test statistics. Two issues were of interest:

- A chance of rejecting a copula under H_0, when it is a true copula. It should be consistent with the pre-defined significance level α.
- A chance of rejecting a copula under H_0 when it is a false copula.

We extend the results of Genest study presented in Tables 4.5-4.7 by the bivariate Vuong and Clarke tests taken from Tables 4.2-4.4, where both tests use Schwarz's correction. Genest, Rémillard, and Beaudoin executed their calculations only for Normal, Student, Clayton, Gumbel, Plackett and Frank copula families. Hence, we restrict our extension on these copula classes and leave BB1 and BB7 copulas out. Note, the directly comparison of Genest's to our results is not possible because of non-identity of null hypothesis. Namely, in the tests provided by Genest, Rémillard, and Beaudoin (2009), one tests whether the data of interest comes really from the estimated copula family $C_k \in \mathcal{C}$, $k = 1, \ldots, 6$. In case of the Vuong and Clarke tests, one decides whether any copula $C_k \in \mathcal{C}$ fitted to the data of interest is equivalent to any other copula $C_m \in \mathcal{C}$ fitted to the same data, $k, m = 1, \ldots, 6$. Additionally, it is not possible to calculate entries in columns (I)>(II) and (I)=(II) for the Vuong and Clarke tests for the rows where copula families in columns (I) and (II) are identical. Entries in columns (I)<(II) can be calculated from Tables 4.2-4.4 by building an average value of entries in rows denoted by (I)<(II) according to specific true copula family. We demonstrate it on example of Normal true copula, $\tau = 0.25$ and $n = 150$ (see first three rows of Table 4.2):

- For this case, it holds (I)=Normal and (II) is one of the following copulas: Student (df=4), Clayton, Gumbel, Frank or Plackett.
- The average error, i.e. percentages of favoring the false copula (II), for the Vuong

test is

$$\frac{0+0.3+0+0.3+0.2}{5} = 0.16$$

and for the Clarke test

$$\frac{17.45+54.1+73.6+50+21.1}{5} = 43.3\ .$$

Analyzing Tables 4.5-4.7, we state:
The error rate of the Vuong test with Schwarz's correction is much more smaller than the error rates of the Genest's goodness-of-fit tests. It however does not hold for the Clarke test with Schwarz's correction, error rate of which is much more higher for some copulas than the error rates of the Genest's goodness-of-fit tests. In some situations, the percentages of rejecting the false copula by the Vuong or by the Clarke test is higher than by the Genest's test and in the other situations it is lower. Generally, the percentages of false decision tend to decrease with increasing Kendall's τ for the Vuong as well as for the Clarke test.

4.7 Goodness-of-fit test for copulas based on distribution free tests

The Vuong as well as the Clarke test compares two copulas fitted to the same bivariate data. In this section, we show how the Vuong and Clarke tests can be used for comparison of more than two copulas. Let $\mathcal{C} = \{C_1, \ldots, C_m\}$ denote a collection of all possible copula families to be compared. For example, $\mathcal{C} = \{Gauss, Student, Clayton, \ldots\}$. We illustrate our idea on an bivariate data set generated from Gaussian copula with dependence parameter $r = 0.75$. Method presented in this section consists of four steps:

(i) Estimate parameter vector $\boldsymbol{\theta}_i$ of every copula C_i, $i = 1, \ldots, m$, from $\mathcal{C}$, for the bivariate data of interest via maximum-likelihood estimation or using empirical Kendall's τ. Note, that for Gauss, Gumbel, Clayton, Frank and Plackett copula classes, there is unique relationship between copula parameter and Kendall's τ, see Table 3.1 in Chapter 3. So, for this copula families, we need only to determine Kendall's τ empirically. For Student, BB1 and BB7 copulas we use MLE's. We denote estimated parameter vector of copula C_i by $\hat{\boldsymbol{\theta}}_i$, $i = 1, \ldots, m$.

(ii) For each $i = 1, \ldots, m$, compare copula $C_{\hat{\boldsymbol{\theta}}_i}$ with the remaining copulas $\mathcal{C} \setminus C_{\hat{\boldsymbol{\theta}}_i}$ using the bivariate Vuong or Clarke test. Resulting output of this step for the example data is presented in Table 4.8 for the Vuong test and in Table 4.9 for the Clarke test.

(iii) Calculate score of each row in Tables 4.8 and 4.9 in a following way: weight every test outcome (I)>(II) with coefficient 1, every test outcome (I)<(II) with coefficient -1 and every test outcome (I)=(II) with factor 0. The score value results from the

Table 4.1: An example illustrating calculation of *better*, *equivalent* and *worse* entries with data from Gaussian copula

data	t	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel	better	equiv.	worse
data 1	5	3	-2	*7*	-2	-2	-7	-2			
data 2	4	2	-2	*5*	*5*	-2	-6	-1	25 %	50 %	25 %
data 3	1	-2	-3	*4*	*4*	*4*	-4	3			
data 4	1	-2	-3	4	*5*	4	-4	3			

sum of the evaluated weights. For example, the number of (I)>(II) outcomes of the first row (Gaussian) in Table 4.8 is 7 and the number of (I)<(II) outcomes is 0. The score value of this row is equal to $7 + 0 = 7$. For the Clarke test, the score of Gaussian copula is equal to 4.

(iv) Make decision according to the rule: *copula with the highest score is the most suitable one among all copula types considered.* For instance, the Gaussian copula in Table 4.8 fitted to the example data is characterized by a highest score, so it is the most appropriate copula for our data. It is different for the Clarke test. In this case, a highest score corresponds to a Frank copula that is not a true one.

Next, we are interested whether the score approach based on the Vuong or on the Clarke test provides accurate results subject to specific copula family chosen. We simulate bivariate data from a fixed copula class $N = 1000$ times and evaluate scores N times according to the steps (i)-(iv) described above. Then we calculate percentages of cases where the score test favored the true copula and denote such entries by *better*. Further, we evaluate percentages of cases where the score test outputted highest score value for the true copula, but there were some other copulas with the same score. We denote such entries by *equivalent*. Finally, we compute the error rate, i.e. percentages of false decision. Table 4.1 demonstrates this procedure for data from the bivariate Normal copula with number of repetitions $N = 4$. As we can see for the first case, the test decision is unique and it is a (true) Normal copula. For the second and third cases, the score test can not distinguish between Normal, Frank and Plackett copulas. And in the last case, the test favors (false) Frank copula.

Table 4.10 reports results of the score test study for two sample sizes $n = 150$ and $n = 1000$. This study contains computations for different values of $\tau \in \{0.25, 0.5, 0.75\}$ for the Vuong and Clarke tests, both with Schwarz's correction. Note, we use the same data in this study as for Tables 4.2–4.4. For example, Table 4.10 shows that when testing the null hypothesis

$$H_0 : \textit{true copula is Gauss}$$

for data from Normal copula with $\tau = 0.25$ and sample size $n = 150$, there were approximately a of chance 45 % that the score test based on the Vuong test will assume the null hypothesis. And in 31 % of all cases, the score test could not favor the Gaussian copula over some others. These results are plausible because there is hardly any difference

between Gauss and other copulas, particularly Frank and Plackett. It is different for the Clarke test. For the same combination of n and τ, a chance of choosing the true copula for the Normal case as well as the chance of equivalence with some copula(s) is equal to zero.

One can easy see from Table 4.10, that the score procedure based on the Vuong as well as on the Clarke tests depends vigorously on the copula family and the combinations of n and τ. In case of Plackett copula, the test increases its accuracy, i.e. percentages of "better" decisions, with increasing Kendall's τ and as sample size goes from $n = 150$ to $n = 1000$. The accuracy of the test for Gumbel data with small sample size does not change depending on τ. The accuracy of the test based on the Vuong statistic for the Clayton data is approximately of the same degree for small as well for large sample size. Moreover, there is an increasing effect of accuracy depending on increasing τ for the Clarke case. Another remarkable fact is that the accuracy for BB1 and BB7 copulas is very small. It is an explainable fact, because BB1 and BB7 with weak upper and vigorously lower tail dependencies are very similar to dependence structure of the Clayton model.

Similar to Table 4.10, Table 4.11 contains results of the simulations study for score approach. The distinction is, that we used here the Akaike's correction term in the Vuong and Clarke tests. The differences between these two tables are minimal.

As mentioned in the previous section, the pairwise Vuong test has significantly lower error rate in comparison to the Clarke test. Hence, we will use the score method based on the Vuong test for goodness-of-fit testing and choosing an appropriate copula for D-vine building blocks in the following chapter.

Table 4.2: Comparison efficiency of the Vuong and Clarke tests for pairwise goodness-of-fit testing for copulas: percentages of non-rejection of H_0 : (I)=(II) and its rejection in favor of true model, (I)>(II), and in favor of false model, (I)<(II) with number of repetitions $N = 1000$. Results were obtained for bivariate data sets of length $n \in \{150, 1000\}$ arising from different copula families with $\tau = 0.25$. Both tests use Schwarz's correction.

(I)↓, (II)→	n	test decision	Normal		Student [4]		Clayton		Gumbel		Frank		Plackett		BB1 [4]		BB7 [4]	
			Vuong	Clarke	Vuong	Clarke	Vuong	Clarke	Vuong	Clarke	Vuong	Clarke	Vuong	Clarke	Vuong	Clarke	Vuong	Clarke
Normal	150	(I)>(II)	0	49.1	53.82	47.31	18.1	13.4	6.5	8.9	7.1	18.6	9.7	21	27.01	41.12	24.17	35.55
		(I)=(II)	99.7	0.6	46.18	35.23	81.6	32.5	93.5	17.5	92.6	31.4	90.1	57.9	72.99	40.91	75.83	42.2
		(I)<(II)	0.3	50.3	0	17.46	0.3	54.1	0	73.6	0.3	50	0.2	21.1	0	17.97	0	22.26
Student [1]	150	(I)>(II)	2	30.7	0	12.94	7.2	52.7	0.9	28.4	5.2	49.9	2.6	40.2	1.21	40.28	0.7	38.97
		(I)=(II)	95.1	59.6	99	38.92	91.3	41.5	94.3	59.4	93.6	40.6	94.2	31.8	96.96	44.23	96.98	41.19
		(I)<(II)	2.9	9.7	1	48.14	1.5	5.8	4.8	12.2	1.2	9.5	3.2	28	1.82	15.49	2.32	19.84
Clayton	150	(I)>(II)	26.6	67.8	48.99	39.17	0	47.6	50.8	29.9	31.1	41.3	30.1	23.4	51.17	55.06	51.02	56.81
		(I)=(II)	73.3	20.2	51.01	42.91	98.6	0.6	49.2	47.6	68.9	40.6	69.9	57.2	48.83	22.68	48.98	21.24
		(I)<(II)	0.1	12	0	17.91	1.4	51.8	0	22.5	0	18.1	0	19.4	0	22.27	0	21.95
Gumbel	150	(I)>(II)	12.7	75.3	36.93	46.42	46.2	55.3	0	43.2	15	48.7	14.3	32.5	48.27	55.61	26.67	57.37
		(I)=(II)	87.1	9.7	63.07	30.88	53.8	26.4	99.7	4	84.7	27.9	85.4	42.4	51.73	20.41	73.33	18.28
		(I)<(II)	0.2	15	0	22.7	0	18.3	0.3	52.8	0.3	23.4	0.3	25.1	0	23.98	0	24.34
Frank	150	(I)>(II)	2.6	60.2	44.08	67.86	20.1	32.6	11.7	21.8	0	48.1	1.5	27.3	28.94	66.56	29.25	72.27
		(I)=(II)	97.1	25.3	55.92	18.88	79.8	37.4	88.1	39.7	100	2.9	98.2	61.5	71.06	24.64	70.75	23.08
		(I)<(II)	0.3	14.5	0	13.27	0.1	30	0.2	38.5	0	49	0.3	11.2	0	8.79	0	4.66
Plackett	150	(I)>(II)	2	31.9	43.96	74.21	20.5	52.7	8.5	38.2	0.7	9.3	0	31	31.41	82.88	32.23	86.4
		(I)=(II)	96.5	54	56.04	12.59	79.5	35.8	91.4	36.9	98	63.4	100	24.9	68.59	10.94	67.77	10.67
		(I)<(II)	1.5	14.1	0	13.2	0	11.5	0.1	24.9	1.3	27.3	0	44.1	0	6.18	0	2.92
BB1 [2]	150	(I)>(II)	0.3	10.7	2.43	27.91	18.4	29.5	0	1.4	0.9	17.2	0.6	14.4	0	31.91	0	27.39
		(I)=(II)	87.1	60.1	96.87	16.78	81.3	34.8	9.6	23.8	94.2	38.7	93	25.6	99.8	9.48	98.79	22.96
		(I)<(II)	12.6	29.2	0.71	55.31	0.3	35.7	90.4	74.8	4.9	44.1	6.4	60	0.2	58.61	1.21	49.65
BB7 [3]	150	(I)>(II)	0	9.1	1.52	21.28	7.6	28.6	0	8	1.4	12.6	0.6	11	0.4	27.59	0	28.31
		(I)=(II)	84.4	62.4	97.77	35.46	90.9	33	73.3	25.2	95.3	49	94.8	33.5	98.49	17.02	99.2	13.05
		(I)<(II)	15.6	28.5	0.71	43.26	1.5	38.4	26.7	66.8	3.3	38.4	4.6	55.5	1.11	55.39	0.8	58.63
Normal	1000	(I)>(II)	0	49.2	64.94	34.76	93.3	2.2	67.9	0	39.8	2.8	46.3	8.5	52.94	13.08	60.34	21.59
		(I)=(II)	99.8	0	35.06	21.24	6.7	15.7	32.1	0.6	60.2	7.8	53.7	40.7	47.06	21.6	39.66	40.16
		(I)<(II)	0.2	50.8	0	44	0	82.1	0	99.4	0	89.4	0	50.8	0	65.31	0	38.25
Student [1]	1000	(I)>(II)	80	99.9	0	37.74	94.9	99.9	62.5	99.9	91.4	99.1	77.9	89.6	35.9	81.44	31.49	86.46
		(I)=(II)	20	0.1	99.8	6.41	5.1	0.1	37.5	0	8.6	0.8	22.1	4.8	64	13.69	68.41	8.93
		(I)<(II)	0	0	0.2	55.86	0	0	0	0.1	0	0.1	0	5.6	0.1	4.87	0.1	4.61
Clayton	1000	(I)>(II)	95.7	96.6	96.6	40.44	0.2	50.4	99.9	31.6	97.8	54.8	97	12.8	58.96	53.26	59.07	54.54
		(I)=(II)	4.3	1.9	3.4	22.92	98.2	0	0.1	38.1	2.2	30.2	3	50.4	41.04	4.99	40.93	4.54
		(I)<(II)	0	1.5	0	36.64	1.6	49.6	0	30.3	0	15	0	36.8	0	41.75	0	40.93
Gumbel	1000	(I)>(II)	71.2	99.5	75.7	41.9	100	87.5	0	51.6	84.2	73.6	79.9	22.5	63.14	56.42	37.96	65.79
		(I)=(II)	28.8	0.1	24.3	6.9	0	8.8	100	0	15.8	13.3	20.1	31.7	36.86	4.48	62.04	12.96
		(I)<(II)	0	0.4	0	51.2	0	3.7	0	48.4	0	13.1	0	45.8	0	39.1	0	21.26
Frank	1000	(I)>(II)	34.6	97.6	69.13	95.45	97.6	53.2	84.7	18.6	0	49.3	4.3	19.6	79.51	93.24	88	99.4
		(I)=(II)	65.4	1.5	30.87	4.05	2.4	24.9	15.3	30.8	100	0	95.1	57.2	20.49	6.45	12	0.6
		(I)<(II)	0	0.9	0	0.51	0	21.9	0	50.6	0	50.7	0.6	23.2	0	0.31	0	0
Plackett	1000	(I)>(II)	34.1	86.4	63.35	98.8	97.4	96.1	78.8	84.4	3.3	41.4	0	49.9	75.74	99.69	84.04	100
		(I)=(II)	65.8	10.6	36.65	0.8	2.6	2.7	21.2	8.2	96.2	43.3	99.9	0	24.26	0.2	15.96	0
		(I)<(II)	0.1	3	0	0.4	0	1.2	0	7.4	0.5	15.3	0.1	50.1	0	0.1	0	0
BB1 [2]	1000	(I)>(II)	32.6	97.5	34.2	36.3	99.4	79.1	0	34.5	65	61	55	12.2	0	44.7	1.91	53.68
		(I)=(II)	67.2	1.3	65.8	5.1	0.6	12.6	17.2	3.2	35	18.7	45	19.3	99.29	2.12	97.08	11.48
		(I)<(II)	0.2	1.2	0	58.6	0	8.3	82.8	62.3	0	20.3	0	68.5	0.71	53.18	1.01	34.84
BB7 [3]	1000	(I)>(II)	27.3	93.6	20	22.8	93.3	74.5	9.7	28.1	76.3	53.9	67.4	11	2.42	32.53	0	42.08
		(I)=(II)	72.4	4.2	79.7	10	6.7	16.3	89.7	20.7	23.7	30.1	32.6	18	96.97	11.01	99.8	4.84
		(I)<(II)	0.3	2.2	0.3	67.2	0	9.2	0.6	51.2	0	16	0	71	0.61	56.46	0.2	53.08

[1] Student's t copula with 4 degrees of freedom
[2] BB1 copula with upper tail dependence $\lambda_U = 0.3$
[3] BB7 copula with upper tail dependence $\lambda_U = 0.3$
[4] Fitted copula with parameters which were both estimated

Table 4.3: Comparison efficiency of the Vuong and Clarke tests for pairwise goodness-of-fit testing for copulas: percentages of non-rejection of H_0 : (I)=(II) and its rejection in favor of true model, (I)>(II), and in favor of false model, (I)<(II) with number of repetitions $N = 1000$. Results were obtained for bivariate data sets of length $n \in \{150, 1000\}$ arising from different copula families with $\tau = 0.5$. Both tests use Schwarz's correction.

(I)↓, (II)→	n	test decision	Normal		Student [4]		Clayton		Gumbel		Frank		Plackett		BB1 [4]		BB7 [4]	
			Vuong	Clarke	Vuong	Clarke	Vuong	Clarke	Vuong	Clarke	Vuong	Clarke	Vuong	Clarke	Vuong	Clarke	Vuong	Clarke
Normal	150	(I)>(II)	0	49.8	46.44	43.76	69.8	43.2	19	29.4	30.9	17.2	35.5	30.6	21.54	51.7	23.9	62
		(I)=(II)	99.2	0	53.56	30.65	30.1	46.2	81	49.1	69.1	39.3	64.5	66.7	78.36	36.97	75.8	19.9
		(I)<(II)	0.8	50.2	0	25.59	0.1	10.6	0	21.5	0	43.5	0	2.7	0.1	11.32	0.3	18.1
Student[1]	150	(I)>(II)	2.8	34.9	0	20.62	52.8	56.3	5.5	28.7	23.4	50.5	8.3	19.6	0.3	27.43	1.3	47.7
		(I)=(II)	93.2	50.8	98.8	32.03	47.2	39.1	91.7	43	76.3	18.8	90.7	54.6	94.29	37.24	93.59	35.07
		(I)<(II)	4	14.3	1.2	47.35	0	4.6	2.8	28.3	0.3	30.7	1	25.8	5.41	35.34	5.11	17.23
Clayton	150	(I)>(II)	69.2	51.6	74.25	47.08	0.1	47.5	92	51.6	76.3	51.4	75.6	60.6	41.29	55.29	41.35	58.95
		(I)=(II)	30.8	43.1	25.75	51.11	96.9	0.7	8	47.7	23.7	45.9	24.4	39.3	58.51	19.23	58.45	16.2
		(I)<(II)	0	5.3	0	1.81	3	51.8	0	0.7	0	2.7	0	0.1	0.2	25.48	0.2	24.85
Gumbel	150	(I)>(II)	28.7	46	39.66	49	93.8	73.4	0	45.8	47.6	39.4	43.8	33.4	40.24	53.22	18.12	68.07
		(I)=(II)	71.1	31.5	60.34	36.14	6.2	26.3	99.1	1.2	52.4	41.3	56.2	63.3	59.76	17.2	81.88	13.51
		(I)<(II)	0.2	22.5	0	14.86	0	0.3	0.9	53	0	19.3	0	3.3	0	29.58	0	18.42
Frank	150	(I)>(II)	10.5	69.6	40.27	89.45	76.5	40	33.5	37.9	0	49	12.9	14.4	42.91	91.76	55.21	99.3
		(I)=(II)	89.4	16.7	59.73	5.53	23.5	56.2	66.4	48.5	99.9	0.4	86.9	80.5	57.09	6.93	44.79	0.7
		(I)<(II)	0.1	13.7	0	5.02	0	3.8	0.1	13.6	0.1	50.6	0.2	5.1	0	1.31	0	0
Plackett	150	(I)>(II)	12.9	30.5	31.96	64.03	73	41.6	26.5	32.6	7.2	28.8	0	27.5	34.53	76.18	45.4	94.3
		(I)=(II)	86.8	64.3	68.04	31.46	27	57.6	73.4	62	92.4	61.3	99.9	38.1	65.47	22.82	54.6	5.7
		(I)<(II)	0.3	5.2	0	4.51	0	0.8	0.1	5.4	0.4	9.9	0.1	34.4	0	1	0	0
BB1[2]	150	(I)>(II)	13.6	30.6	19.24	22.55	9.5	21.2	45.5	24.9	37.8	23.8	30.2	14.7	0	31.88	0.5	38.2
		(I)=(II)	84.5	29	80.26	43.49	88	41.4	54.4	58.1	62	47.9	69.5	79.2	98.58	14.07	95.8	19
		(I)<(II)	1.9	40.4	0.5	33.97	2.5	37.4	0.1	17	0.2	28.3	0.3	6.1	1.42	54.05	3.7	42.8
BB7[3]	150	(I)>(II)	32.9	42.5	34.53	19.22	1.8	12	63.9	26.2	55.4	25.1	47	19.1	0.7	26.98	0	31.83
		(I)=(II)	66.6	26.7	65.07	60.06	90.6	39.9	36	66.1	44.6	59.6	52.9	78.2	97.69	17.45	98.2	13.31
		(I)<(II)	0.5	30.8	0.4	20.72	7.6	48.1	0.1	7.7	0	15.3	0.1	2.7	1.6	55.57	1.8	54.85
Normal	1000	(I)>(II)	0	48.5	57.31	37.51	100	70.3	97.1	36.1	95.9	1.6	99.4	68.7	69.17	51.45	88.68	88.98
		(I)=(II)	99.4	0	42.69	17.62	0	25.2	2.9	41	4.1	13.8	0.6	31.3	30.83	29.73	11.32	6.81
		(I)<(II)	0.6	51.5	0	44.87	0	4.5	0	22.9	0	84.6	0	0	0	18.82	0	4.21
Student[1]	1000	(I)>(II)	82.5	98.3	0	41.2	100	99.9	88.4	91.1	99.4	83	89.4	34.6	11.82	69.14	37.1	98.4
		(I)=(II)	17.5	1.5	99.6	6.4	0	0.1	11.6	6.4	0.6	9.1	10.6	46.7	87.78	14.33	62.9	1.2
		(I)<(II)	0	0.2	0.4	52.4	0	0	0	2.5	0	7.9	0	18.7	0.4	16.53	0	0.4
Clayton	1000	(I)>(II)	100	86.5	100	71.2	0.8	50.5	100	98	100	94.6	100	99.8	52.12	53.53	51.96	56.08
		(I)=(II)	0	13	0	28.4	97.1	0	0	2	0	5.3	0	0.2	47.78	4.54	48.04	3.22
		(I)<(II)	0	0.5	0	0.4	2.1	49.5	0	0	0	0.1	0	0	0.1	41.94	0	40.7
Gumbel	1000	(I)>(II)	96.4	72.5	94.3	45.6	100	99.9	0	51.1	99.8	63.4	99	74.3	56.11	54.81	44.42	94.87
		(I)=(II)	3.6	19.6	5.7	32.6	0	0.1	99.2	0	0.2	25.8	1	25.5	43.89	4.11	55.58	2.11
		(I)<(II)	0	7.9	0	21.8	0	0	0.8	48.9	0	10.8	0	0.2	0	41.08	0	3.02
Frank	1000	(I)>(II)	93.4	98.5	97.18	100	100	86.5	99.9	68.2	0	48.9	66.9	7.1	99.5	100	100	100
		(I)=(II)	6.6	0.5	2.82	0	0	13.5	0.1	29	100	0	33.1	86.5	0.5	0	0	0
		(I)<(II)	0	1	0	0	0	0	0	2.8	0	51.1	0	6.4	0	0	0	0
Plackett	1000	(I)>(II)	98.7	86.6	94.9	97.6	100	96.7	99.9	85.8	64.4	95.4	0	49.4	99.5	100	99.8	100
		(I)=(II)	1.3	13.2	5.1	2.4	0	3.3	0.1	14	35.6	4.4	99.9	0	0.5	0	0.2	0
		(I)<(II)	0	0.2	0	0	0	0	0	0.2	0	0.2	0.1	50.6	0	0	0	0
BB1[2]	1000	(I)>(II)	94.9	79	88.1	12.7	86.8	72.5	100	67.1	100	46.8	99.7	66.3	0	43.65	2.2	71.6
		(I)=(II)	5.1	8.7	11.9	28.3	13.2	20	0	30.4	0	38	0.3	33.4	98.79	4.54	95	7.6
		(I)<(II)	0	12.3	0	59	0	7.5	0	2.5	0	15.2	0	0.3	1.21	51.81	2.8	20.8
BB7[3]	1000	(I)>(II)	99.7	78.6	98.9	11.5	54.4	62.1	100	77.8	100	62.3	100	87.4	3.2	25.13	0	43.3
		(I)=(II)	0.3	13.8	1.1	56.5	45.5	22.1	0	21.6	0	34.6	0	12.6	95.5	6.61	98.9	4.2
		(I)<(II)	0	7.6	0	32	0.1	15.8	0	0.6	0	3.1	0	0	1.3	68.27	1.1	52.5

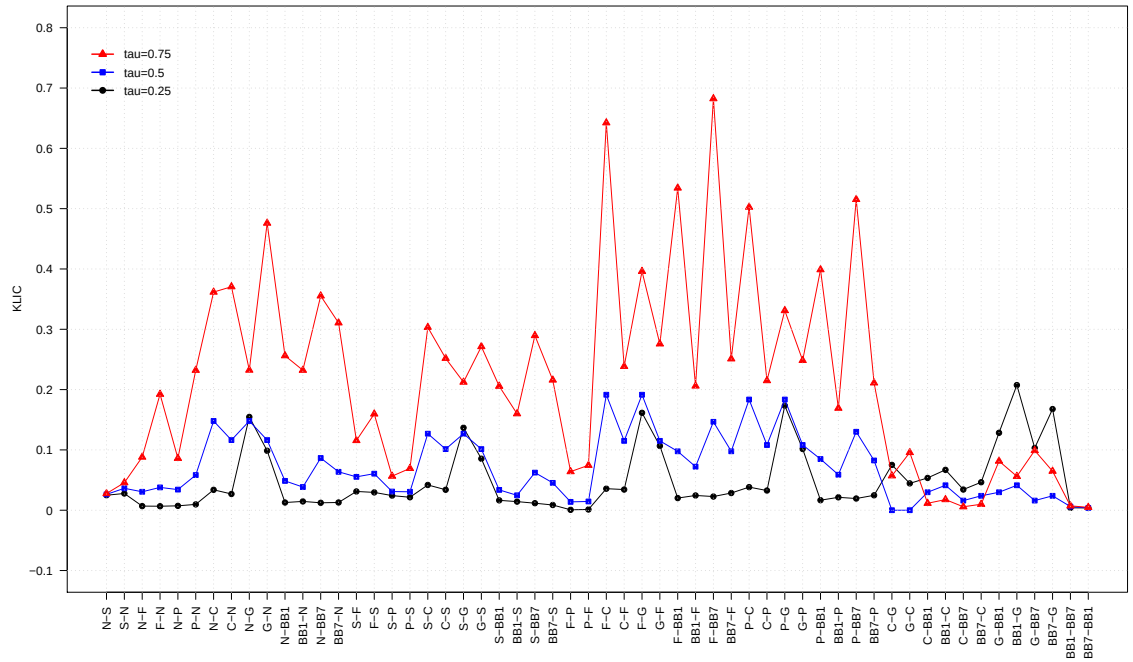
[1] Student's t copula with 4 degrees of freedom
[2] BB1 copula with upper tail dependence $\lambda_U = 0.3$
[3] BB7 copula with upper tail dependence $\lambda_U = 0.3$
[4] Fitted copula with parameters which were both estimated

Table 4.4: Comparison efficiency of the Vuong and Clarke tests for pairwise goodness-of-fit testing for copulas: percentages of non-rejection of H_0 : (I)=(II) and its rejection in favor of true model, (I)>(II), and in favor of false model, (I)<(II) with number of repetitions $N = 1000$. Results were obtained for bivariate data sets of length $n \in \{150, 1000\}$ arising from different copula families with $\tau = 0.75$. Both tests use Schwarz's correction.

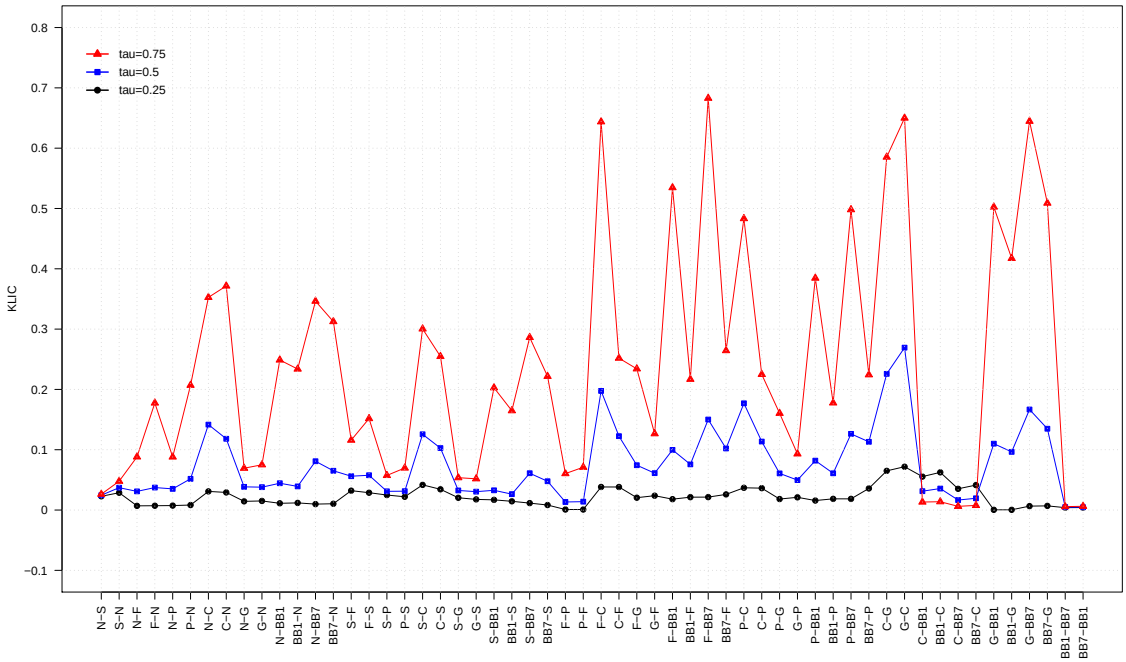
(I)↓, (II)→	n	test decision	Normal		Student [4]		Clayton		Gumbel		Frank		Plackett		BB1 [4]		BB7 [4]	
			Vuong	Clarke	Vuong	Clarke	Vuong	Clarke	Vuong	Clarke	Vuong	Clarke	Vuong	Clarke	Vuong	Clarke	Vuong	Clarke
Normal	150	(I)>(II)	0	49	41.4	48.59	96.3	91.2	36.7	65.2	65.9	70.2	70.5	56.7	20.32	66.27	44.41	97.46
		(I)=(II)	99.4	0	58.5	27.11	3.7	8.8	63.3	33.7	34.1	29.5	29.5	43.2	79.28	26.83	55.49	2.54
		(I)<(II)	0.6	51	0.1	24.3	0	0	0	1.1	0	0.3	0	0.1	0.4	6.91	0.1	0
Student [1]	150	(I)>(II)	3.2	35.8	0	28.74	91.5	53.3	18	5.3	61.4	25.4	30.3	44.4	0.3	18.04	13.75	68.56
		(I)=(II)	92.7	45	98.19	31.26	8.5	46.3	80.8	67.6	38.6	68.7	69.6	55.4	92.08	46.39	84.07	29.68
		(I)<(II)	4.1	19.2	1.81	40	0	0.4	1.2	27.1	0	5.9	0.1	0.2	7.62	35.57	2.17	1.76
Clayton	150	(I)>(II)	89.4	53.5	89.75	94.47	0.6	45	99.2	91.3	89.1	56.7	87.6	94.9	35.34	62.86	37.6	67.5
		(I)=(II)	10.6	46.2	10.25	5.53	95	1.6	0.8	8.7	10.9	43.1	12.4	5.1	64.46	16.42	62.2	11.4
		(I)<(II)	0	0.3	0	0	4.4	53.4	0	0	0	0.2	0	0	0.2	20.72	0.2	21.1
Gumbel	150	(I)>(II)	37.4	7.8	43.49	61.72	99.3	93.6	0	44.6	74.4	39.1	68.3	65.9	32.66	60.73	25.84	82.81
		(I)=(II)	62.2	66.7	56.41	34.97	0.7	6.4	98.7	1.5	25.6	57.4	31.7	34.1	67.34	16.65	73.37	15.28
		(I)<(II)	0.4	25.5	0.1	3.31	0	0	1.3	53.9	0	3.5	0	0	0	22.63	0.79	1.91
Frank	150	(I)>(II)	49.6	9.9	67.04	86.25	98.6	84.8	72	49.9	0	47.1	50.6	20.7	74.82	98.12	92.7	99.8
		(I)=(II)	50.4	82.3	32.96	13.65	1.4	15.2	28	48.4	99.8	0.8	49.4	78.3	25.18	1.88	7.3	0.2
		(I)<(II)	0	7.8	0	0.1	0	0	0	1.7	0.2	52.1	0	1	0	0	0	0
Plackett	150	(I)>(II)	35.6	36.7	42.08	35.02	97.3	41.7	53.8	27.1	35.6	52	0.5	23.1	52.82	49.9	79.36	82.28
		(I)=(II)	64.3	61.9	57.82	63.98	2.7	58	46.2	71	64.2	42	98.7	54	47.08	49.9	20.64	17.72
		(I)<(II)	0.1	1.4	0.1	1.01	0	0.3	0	1.9	0.2	6	0.8	22.9	0.1	0.2	0	0
BB1 [2]	150	(I)>(II)	72.2	17.5	71.63	74.45	3.1	6.7	95.5	69.6	82.1	20.3	75.5	82.1	0	33.17	1.1	51.7
		(I)=(II)	27.8	78.6	28.37	25.45	85.2	33.4	4.5	30.4	17.9	77.5	24.4	17.9	97.39	29.76	95.2	34.4
		(I)<(II)	0	3.9	0	0.1	11.7	59.9	0	0	0	2.2	0.1	0	2.61	37.07	3.7	13.9
BB7 [3]	150	(I)>(II)	83.5	33	81.64	83.35	0	17.8	97.7	81.1	87.2	31.4	81.6	89	1.1	27.88	0	40.8
		(I)=(II)	16.5	65.7	18.36	16.65	70	20.6	2.3	18.9	12.8	67.7	18.4	11	95.79	34	96.7	21.1
		(I)<(II)	0	1.3	0	0	30	61.6	0	0	0	0.9	0	0	3.11	38.11	3.3	38.1
Normal	1000	(I)>(II)	0	48.5	51.24	43.6	100	100	99.7	99.7	100	99.8	100	100	80.58	84.21	99.79	100
		(I)=(II)	99.1	0	48.76	13.64	0	0	0.3	0.3	0	0.2	0	0	19.42	13.18	0.21	0
		(I)<(II)	0.9	51.5	0	42.77	0	0	0	0	0	0	0	0	0	2.62	0	0
Student [1]	1000	(I)>(II)	78.6	96.3	0	44.84	100	100	97.3	36	100	88.8	99.6	99.1	8.95	52.92	91.13	100
		(I)=(II)	21.4	3	99.5	8.81	0	0	2.7	57.7	0	11.1	0.4	0.9	90.34	24.04	8.87	0
		(I)<(II)	0	0.7	0.5	46.35	0	0	0	6.3	0	0.1	0	0	0.7	23.04	0	0
Clayton	1000	(I)>(II)	100	100	100	100	0.7	52.2	100	100	100	99.9	100	100	51.61	62.05	49.95	63.06
		(I)=(II)	0	0	0	0	96.5	0	0	0	0	0.1	0	0	48.29	2.81	49.95	2.1
		(I)<(II)	0	0	0	0	2.8	47.8	0	0	0	0	0	0	0.1	35.14	0.1	34.83
Gumbel	1000	(I)>(II)	99.4	14.6	98.2	86.5	100	100	0	50.2	100	83.8	100	99.9	49.05	62.66	82.86	100
		(I)=(II)	0.6	63.8	1.8	13.1	0	0	99	0	0	16.2	0	0.1	50.95	4.8	17.14	0
		(I)<(II)	0	21.6	0	0.4	0	0	1	49.8	0	0	0	0	0	32.53	0	0
Frank	1000	(I)>(II)	100	25	100	100	100	100	100	98.9	0	49.8	100	58.4	100	100	100	100
		(I)=(II)	0	72.5	0	0	0	0	0	1.1	100	0	0	41.6	0	0	0	0
		(I)<(II)	0	2.5	0	0	0	0	0	0	0	50.2	0	0	0	0	0	0
Plackett	1000	(I)>(II)	100	99.3	99.8	86.26	100	98.8	100	91.3	100	99.8	0.2	48.6	100	100	100	100
		(I)=(II)	0	0.7	0.2	13.64	0	1.2	0	8.7	0	0.2	99.3	0	0	0	0	0
		(I)<(II)	0	0	0	0.1	0	0	0	0	0	0	0.5	51.4	0	0	0	0
BB1 [2]	1000	(I)>(II)	100	90.4	100	100	41.6	22.1	100	100	100	96.2	100	100	0	43.44	14.7	87.5
		(I)=(II)	0	9.6	0	0	58.3	50.6	0	0	0	3.8	0	0	98.5	11.81	83.2	10.1
		(I)<(II)	0	0	0	0	0.1	27.3	0	0	0	0	0	0	1.5	44.74	2.1	2.4
BB7 [3]	1000	(I)>(II)	100	99.2	100	100	7.5	42.6	100	100	100	99.4	100	100	7.72	19.46	0	52
		(I)=(II)	0	0.8	0	0	91	14.9	0	0	0	0.6	0	0	90.07	23.27	98.2	5.4
		(I)<(II)	0	0	0	0	1.5	42.5	0	0	0	0	0	0	2.21	57.27	1.8	42.6

[1] Student's t copula with 4 degrees of freedom
[2] BB1 copula with upper tail dependence $\lambda_U = 0.3$
[3] BB7 copula with upper tail dependence $\lambda_U = 0.3$
[4] Fitted copula with parameters which were both estimated

Figure 4.1: Estimated KLIC distances between different copulas calculated for data sets of length $n = 10000$ samples from copula families Normal, Student ($df = 4$), Clayton, Gumbel, Frank, Plackett, BB1 ($\lambda_U = 0.3$) and BB7 ($\lambda_U = 0.3$). Points with notation $C_1 - C_2$ on the x-axes correspond to the KLIC distance between true copula C_1 and another copula C_2 which was estimated to the data from C_1 and is not a true copula.



(a) KLIC distances estimated empirically



(b) KLIC distances estimated by numerical integration

Table 4.5: Extended Genest study for data sets of size $n = 150$, coefficient of association $\tau = 0.25$ and true model denoted by (I). Genest study: Percentages of rejection H_0 : *data comes from model (II) against* H_1 : *data comes from model (I)* by various tests for data from different copula models and number of repetitions $N = 10000$. Extension: Percentages of non-rejection H_0 : $(I) = (II)$ and its rejection in favor of model (I), denoted by (I)>(II), or in favor of model (II), denoted by (I)<(II), by the Vuong and Clarke tests for data from different copula models and number of repetitions $N = 1000$. Both tests use Schwarz's correction.

Copula		**Test based on statistic**							**Vuong**			**Clarke**		
(II)	**(I)**	S_n	T_n	$S_n^{(K)}$	$T_n^{(K)}$	$S_n^{(B)}$	$S_n^{(C)}$	A_n	**(I)$\wedge$(II)**	**(I)$\|\|$(II)**	**(I)$\vee$(II)**	**(I)$\wedge$(II)**	**(I)$\|\|$(II)**	**(I)$\vee$(II)**
Clayton	**Clayton**	**4.6**	**4.8**	**4.1**	**4.5**	**4.7**	**4.8**	**4.9**			**0**			**18**
	Gumbel	86.1	62.4	57.9	42.7	80.9	76.7	22.4	46.2	53.8	0	55.3	26.4	18.3
	Frank	56.3	32.7	37.4	26.4	42.8	36.2	6.2	20.1	79.8	0.1	32.6	37.4	30
	Plackett	56.0	31.2	33.7	23.4	43.9	39.0	8.1	20.5	79.5	0	52.7	35.8	11.5
	Normal	50.2	27.5	24.5	16.8	41.8	34.6	6.4	18.1	81.6	0.3	13.4	32.5	54.1
	Student [1]	56.5	32.3	23.2	15.5	51.0	52.7	32.9	7.2	91.3	1.5	52.7	41.5	5.8
Gumbel	Clayton	72.1	62.6	92.3	82.1	65.1	60.5	8.3	50.8	49.2	0	29.9	47.6	22.5
	Gumbel	**5.0**	**5.0**	**4.7**	**5.1**	**5.1**	**5.0**	**5.0**			**0.16**			**20.9**
	Frank	15.4	15.4	19.9	15.1	12.9	10.0	5.9	11.7	88.1	0.2	21.8	39.7	38.5
	Plackett	14.3	14.7	18.9	14.7	12.5	10.6	4.9	85	91.4	0.1	38.2	36.9	24.9
	Normal	10.1	11.7	24.4	18.9	10.2	7.5	5.9	65	93.5	0	8.9	17.5	73.6
	Student [1]	14.1	12.9	29.8	26.2	14.3	18.2	17.4	0.9	94.3	4.8	28.4	59.4	12.2
Frank	Clayton	40.0	36.8	77.3	70.6	36.2	36.1	9.6	31.1	68.9	0	41.3	40.6	18.1
	Gumbel	33.4	18.5	9.1	6.1	27.8	29.5	12.4	15	84.7	0.3	48.7	27.9	23.4
	Frank	**5.3**	**5.1**	**5.1**	**5.0**	**4.9**	**4.9**	**5.1**			**0.18**			**21.5**
	Plackett	5.7	5.2	5.4	5.1	5.2	6.1	6.6	0.7	98	1.3	9.3	63.4	27.3
	Normal	7.8	7.3	10.5	9.9	6.2	6.3	5.3	7.1	92.6	0.3	18.6	31.4	50
	Student [1]	18.5	11.4	22.0	19.7	14.6	23.0	40.7	5.2	93.6	1.2	49.9	40.6	9.5
Plackett	Clayton	37.6	34.2	69.8	60.5	33.3	31.9	6.2	30.1	69.9	0	23.4	57.2	19.4
	Gumbel	30.4	16.6	7.2	5.4	24.6	24.8	6.8	14.3	85.4	0.3	32.5	42.4	25.1
	Frank	5.0	5.2	4.8	5.1	5.0	4.2	6.5	1.5	98.2	0.3	27.3	61.5	11.2
	Plackett	**5.2**	**5.0**	**4.8**	**4.8**	**4.5**	**4.7**	**5.0**			**0.38**			**18.2**
	Normal	6.8	6.8	8.2	7.6	6.1	5.4	5.7	9.7	90.1	0.2	21	57.9	21.1
	Student [1]	14.1	9.8	15.6	14.4	10.1	15.6	26.2	2.6	94.2	3.2	40.2	31.8	28
Normal	Clayton	31.6	26.6	56.9	45.8	33.3	33.0	7.2	26.6	73.3	0.1	67.8	20.2	12
	Gumbel	23.8	11.9	7.1	5.5	24.7	27.0	8.9	12.7	87.1	0.2	75.3	9.7	15
	Frank	7.9	7.2	5.6	5.3	7.2	7.0	5.5	2.6	97.1	0.3	60.2	25.3	14.5
	Plackett	7.9	6.8	4.4	4.4	8.2	9.4	6.0	2	96.5	1.5	31.9	54	14.1
	Normal	**5.1**	**5.0**	**4.7**	**5.2**	**4.7**	**5.0**	**4.8**			**0.16**			**43.3**
	Student [1]	10.5	6.8	7.4	7.4	16.6	27.8	29.9	2	95.1	2.9	30.7	59.6	9.7
Student [2]	Clayton	27.7	26.2	52.1	39.0	25.1	17.4	11.2	49	51	0	39	43	18
	Gumbel	19.1	11.4	7.4	6.0	17.3	11.5	9.5	36.9	63	0	46	31	23
	Frank	9.1	8.2	9.5	7.6	8.9	4.5	23.3	44	56	0	68	19	13
	Plackett	7.7	7.7	7.3	6.2	6.6	3.6	13.9	44	56	0	74.2	12.6	13.2
	Normal	4.9	5.9	5.4	5.0	7.9	3.1	23.0	53.8	46.2	0	47.3	35.2	17.5
	Student [1]	**4.8**	**5.3**	**4.6**	**4.7**	**4.5**	**4.8**	**5.4**			**2.7**			**13**

[1] Student's t copula with 4 degrees of freedom
[2] Fitted Student's t copula: in Genest study the degrees of freedom are fixed, $df = 4$; in Vuong and Clarke studies the degrees of freedom are estimated

Table 4.6: Extended Genest study for data sets of size $n = 150$, coefficient of association $\tau = 0.5$ and true model denoted by (I). Genest study: Percentages of rejection H_0 : *data comes from model (II) against* H_1 : *data comes from model (I)* by various tests for data from different copula models and number of repetitions $N = 10000$. Extension: Percentages of non-rejection H_0 : $(I) = (II)$ and its rejection in favor of model (I), denoted by (I)>(II), or in favor of model (II), denoted by (I)<(II), by the Vuong and Clarke tests for data from different copula models and number of repetitions $N = 1000$. Both tests use Schwarz's correction.

Copula		Test based on statistic							Vuong			Clarke		
(II)	(I)	S_n	T_n	$S_n^{(K)}$	$T_n^{(K)}$	$S_n^{(B)}$	$S_n^{(C)}$	A_n	(I) $\wedge$ (II)	(I) $\|$ (II)	(I) $\vee$ (II)	(I) $\wedge$ (II)	(I) $\|$ (II)	(I) $\vee$ (II)
Clayton	**Clayton**	**5.3**	**5.0**	**4.5**	**4.5**	**5.1**	**5.0**	**5.0**			**0**			**2.12**
	Gumbel	99.9	98.3	98.5	91.4	99.7	99.5	78.3	93.8	6.2	0	73.4	26.3	0.3
	Frank	95.7	81.2	89.5	74.9	94.4	90.3	37.2	76.5	23.5	0	40	56.2	3.8
	Plackett	95.8	77.7	83.5	63.5	92.9	90.4	62.0	73	27	0	41.6	57.6	0.8
	Normal	93.7	74.1	75.1	53.7	89.0	85.5	35.2	69.8	30.1	0.1	43.2	46.2	10.6
	Student [1]	94.8	78.0	75.0	54.4	87.9	87.6	50.4	52.8	47.2	0	56.3	39.1	4.6
Gumbel	Clayton	99.6	98.4	99.9	99.0	99.7	99.5	33.4	92	8	0	51.6	47.7	0.7
	Gumbel	**4.6**	**5.0**	**4.6**	**4.9**	**4.5**	**4.9**	**5.0**			**0.04**			**12**
	Frank	39.8	37.5	42.4	28.4	52.1	37.0	9.3	33.5	66.4	0.1	37.9	48.5	13.6
	Plackett	29.8	27.2	32.0	23.1	43.2	37.0	21.6	26.5	73.4	0.1	32.6	62	5.4
	Normal	18.3	21.1	37.7	27.4	33.7	25.2	4.9	19	81	0	29.4	49.1	21.5
	Student [1]	21.8	21.2	40.6	31.7	29.7	31.9	10.0	5.5	91.7	2.8	28.7	43	28.3
Frank	Clayton	89.1	84.9	98.6	96.3	86.9	90.4	13.3	76.3	23.7	0	51.4	45.9	2.7
	Gumbel	63.0	39.6	28.3	15.8	44.1	57.6	9.2	47.6	52.4	0	39.4	41.3	19.3
	Frank	**4.8**	**5.1**	**4.8**	**5.2**	**4.8**	**4.8**	**5.1**			**0.08**			**8.24**
	Plackett	8.4	6.3	7.5	6.8	10.5	19.9	12.5	7.2	92.4	0.4	28.8	61.3	9.9
	Normal	19.9	15.0	22.6	17.3	8.9	14.4	4.8	30.9	69.1	0	17.2	39.3	43.5
	Student [1]	35.1	19.6	37.2	27.2	22.9	44.3	19.1	23.4	76.3	0.3	50.5	18.8	30.7
Plackett	Clayton	83.9	78.4	95.5	86.4	79.6	78.0	12.5	75.6	24.4	0	60.6	39.3	0.1
	Gumbel	48.8	28.1	16.4	10.1	29.1	30.4	8.1	43.8	56.2	0	33.4	63.3	3.3
	Frank	6.8	7.8	8.2	8.0	10.2	3.9	10.5	12.9	86.9	0.2	14.4	80.5	5.1
	Plackett	**5.0**	**5.3**	**5.0**	**5.1**	**4.9**	**5.2**	**4.7**			**0.4**			**5.16**
	Normal	9.8	11.2	9.4	7.9	6.9	5.1	12.3	35.5	64.5	0	30.6	66.7	2.7
	Student [1]	15.1	11.4	15.1	10.6	7.4	11.7	7.4	8.3	90.7	1	19.6	54.6	25.8
Normal	Clayton	80.0	68.8	90.3	75.2	90.8	88.2	7.8	69.2	30.8	0	51.6	43.1	5.3
	Gumbel	38.3	17.8	16.1	10.8	42.0	44.4	5.7	28.7	71.1	0.2	46	31.5	22.5
	Frank	20.2	14.3	17.4	14.1	13.4	8.5	8.7	10.5	89.4	0.1	69.6	16.7	13.7
	Plackett	13.2	9.7	6.8	6.6	18.0	22.7	18.1	12.9	86.8	0.3	30.5	64.3	5.2
	Normal	**4.9**	**5.0**	**4.9**	**5.2**	**5.0**	**5.3**	**4.8**			**0.02**			**20.78**
	Student [1]	8.2	5.3	5.9	5.2	20.4	32.1	8.8	2.8	93.2	4	34.9	50.8	14.3
Student [2]	Clayton	77.3	70.5	90.6	73.2	84.9	74.9	6.0	74.3	25.7	0	47.1	51.1	1.8
	Gumbel	33.9	18.2	17.3	11.8	30.3	20.9	4.9	39.7	60.3	0	49	36.1	14.9
	Frank	26.9	18.9	29.3	20.7	24.2	8.1	6.0	40.3	59.7	0	89.5	5.5	5
	Plackett	13.8	11.0	11.6	9.5	10.2	6.9	10.4	32	68	0	64	31.5	4.5
	Normal	5.2	6.4	5.9	6.1	9.9	2.9	6.7	46.4	53.6	0	43.8	30.7	25.6
	Student [1]	**5.0**	**4.9**	**4.9**	**5.0**	**5.1**	**5.2**	**4.9**			**1.62**			**20.74**

[1] Student's t copula with 4 degrees of freedom
[2] Fitted Student's t copula: in Genest study the degrees of freedom are fixed, $df = 4$; in Vuong and Clarke studies the degrees of freedom are estimated

Table 4.7: Extended Genest study for data sets of size $n = 150$, coefficient of association $\tau = 0.75$ and true model denoted by (I). Genest study: Percentages of rejection H_0 : *data comes from model (II) against* H_1 : *data comes from model (I)* by various tests for data from different copula models and number of repetitions $N = 10000$. Extension: Percentages of non-rejection H_0 : $(I) = (II)$ and its rejection in favor of model (I), denoted by (I)>(II), or in favor of model (II), denoted by (I)<(II), by the Vuong and Clarke tests for data from different copula models and number of repetitions $N = 1000$. Both tests use Schwarz's correction.

Copula		**Test based on statistic**							**Vuong**			**Clarke**		
(II)	**(I)**	S_n	T_n	$S_n^{(K)}$	$T_n^{(K)}$	$S_n^{(B)}$	$S_n^{(C)}$	A_n	**(I)$\wedge$(II)**	**(I)$\|\|$(II)**	**(I)$\vee$(II)**	**(I)$\wedge$(II)**	**(I)$\|\|$(II)**	**(I)$\vee$(II)**
Clayton	**Clayton**	**5.4**	**5.0**	**4.9**	**5.1**	**5.1**	**5.2**	**5.0**			**0**			**0.1**
	Gumbel	99.9	99.9	99.9	98.7	99.9	99.9	49.1	99.3	0.7	0	93.6	6.4	0
	Frank	99.1	86.2	97.0	81.2	99.9	99.7	76.7	98.6	1.4	0	84.8	15.2	0
	Plackett	99.5	89.1	93.6	73.6	99.6	99.5	64.1	97.3	2.7	0	41.7	58	0.3
	Normal	99.8	91.7	94.9	77.7	99.5	99.6	23.8	96.3	3.7	0	91.2	8.8	0
	Student [1]	99.8	95.1	94.3	79.4	99.0	99.1	18.2	91.5	8.5	0	53.3	46.3	0.4
Gumbel	Clayton	99.9	99.5	99.9	99.2	99.9	99.9	29.0	99.2	0.8	0	91.3	8.7	0
	Gumbel	**4.5**	**4.7**	**4.4**	**4.6**	**5.2**	**4.8**	**4.9**			**0.1**			**6.46**
	Frank	51.7	45.4	61.6	38.0	83.8	72.4	75.0	72	28	0	49.9	48.4	1.7
	Plackett	25.8	20.3	29.8	17.9	67.8	62.8	39.6	53.8	46.2	0	27.1	71	1.9
	Normal	12.3	17.0	29.4	18.6	60.7	53.6	5.9	36.7	63.3	0	65.2	33.7	1.1
	Student [1]	16.1	17.4	32.9	19.8	54.8	52.0	3.9	18	80.8	1.2	5.3	67.6	27.1
Frank	Clayton	96.6	91.7	99.6	95.5	99.7	99.7	26.8	89.1	10.9	0	56.7	43.1	0.2
	Gumbel	81.9	43.6	53.2	27.1	59.9	74.2	40.0	74.4	25.6	0	39.1	57.4	3.5
	Frank	**4.7**	**4.7**	**4.5**	**4.7**	**5.0**	**5.1**	**5.2**			**0**			**2.12**
	Plackett	20.6	8.0	15.4	8.8	18.6	36.0	7.9	35.6	64.2	0.2	52	42	6
	Normal	40.9	21.2	40.2	20.5	18.4	30.1	49.8	65.9	34.1	0	70.2	29.5	0.3
	Student [1]	59.4	26.0	56.0	27.9	34.4	58.2	42.3	61.4	38.6	0	25.4	68.7	5.9
Plackett	Clayton	89.8	86.8	97.7	78.6	99.5	99.1	18.8	87.6	12.4	0	94.9	5.1	0
	Gumbel	45.8	23.4	19.1	11.4	35.5	29.4	37.4	68.3	31.7	0	65.9	34.1	0
	Frank	14.9	15.4	18.5	15.3	9.7	3.6	10.9	50.6	49.4	0	20.7	78.3	1
	Plackett	**4.7**	**5.0**	**4.9**	**5.1**	**4.9**	**5.2**	**5.2**			**0.08**			**2.12**
	Normal	7.7	12.9	7.7	6.0	2.5	1.2	44.3	70.5	29.5	0	56.7	43.2	0.1
	Student[1]	11.0	12.3	11.4	6.7	4.3	3.6	45.2	30.0	69.6	0.1	44.4	55.4	0.2
Normal	Clayton	91.8	82.4	97.3	75.4	99.9	99.9	8.2	89.4	10.6	0	53.5	46.2	0.3
	Gumbel	38.5	13.2	17.9	10.6	55.5	54.0	4.7	37.4	62.2	0.4	7.8	66.7	25.5
	Frank	42.2	22.9	41.4	24.6	32.8	20.1	70.2	49.6	50.4	0	9.9	82.9	7.8
	Plackett	16.5	7.6	7.0	7.0	23.0	30.6	30.0	35.6	64.3	0.1	36.7	61.9	1.4
	Normal	**4.9**	**4.4**	**4.4**	**4.8**	**4.9**	**4.6**	**5.1**			**0.02**			**5.16**
	Student [1]	6.6	4.3	4.9	4.5	12.3	18.3	4.9	3.2	92.7	4.1	35.8	45	19.2
Student[2]	Clayton	90.6	86.6	97.7	78.6	99.9	99.7	10.9	89.8	10.2	0	94.5	5.5	0
	Gumbel	33.9	15.1	19.2	11.5	48.4	39.3	4.6	43.5	56.4	0.1	61.7	35	3.3
	Frank	48.2	30.5	53.9	32.4	39.3	20.3	81.8	67	33	0	86.3	13.6	0.1
	Plackett	15.7	8.9	11.0	9.7	16.4	17.2	43.5	42.1	57.8	0.1	35	64	1
	Normal	4.1	5.7	5.1	6.0	5.0	2.1	5.9	41.4	58.5	0.1	48.6	27.1	24.3
	Student [1]	**4.9**	**4.7**	**4.8**	**4.9**	**5.6**	**5.3**	**4.5**			**1.08**			**10.56**

[1] Student's t copula with 4 degrees of freedom

[2] Fitted Student's t copula: in Genest study the degrees of freedom are fixed, $df = 4$; in Vuong and Clarke studies the degrees of freedom are estimated

Table 4.8: Example of the score method for goodness-of-fit testing for copulas based on the pairwise Vuong test with Schwarz's correction. Sampled data of length $n = 1000$ comes from the bivariate Gaussian copula with dependence parameter $r = 0.75$. Test checks the equivalence of both models $H_0 : \ (I) = (II)$. The outcome (I)>(II) means that model (I) is better than model (II) and (I)<(II) vice versa.

(I)↓, (II)→	result	Gauss	Frank	Plackett	Clayton	Gumbel	Student	BB1	BB7	$\sum_{(I)>(II)}$	$\sum_{(I)<(II)}$	score
Gauss	decision		(I)>(II)	(I)>(II)	(I)>(II)	(I)>(II)	(I)>(II)	(I)>(II)	(I)>(II)	7	0	**7**
	ν		3.32	3.49	7.62	3.96	3.40	2.78	3.87			
	p-value		0.001	0	0	0	0.001	0.005	0			
Frank	decision	(I)<(II)		(I)=(II)	(I)>(II)	(I)=(II)	(I)<(II)	(I)=(II)	(I)=(II)	1	-2	-1
	ν	-3.32		0.48	5.89	1.01	-2.92	-1.07	0.26			
	p-value	0.001		0.62	0	0.31	0.003	0.28	0.79			
Plackett	decision	(I)<(II)	(I)=(II)		(I)>(II)	(I)=(II)	(I)<(II)	(I)=(II)	(I)=(II)	1	-2	-1
	ν	-3.49	-0.48		6.01	0.90	-3.28	-1.57	0.06			
	p-value	0	0.62		0	0.36	0.001	0.11	0.94			
Clayton	decision	(I)<(II)	(I)<(II)	(I)<(II)		(I)<(II)	(I)<(II)	(I)<(II)	(I)<(II)	0	-7	-7
	ν	-7.62	-5.89	-6.01		-4.63	-7.54	-7.04	-6.32			
	p-value	0	0	0		0	0	0	0			
Gumbel	decision	(I)<(II)	(I)=(II)	(I)=(II)	(I)>(II)		(I)<(II)	(I)<(II)	(I)=(II)	1	-3	-2
	ν	-3.96	-1.01	-0.90	4.63		-3.78	-3.13	-1.05			
	p-value	0	0.31	0.36	0		0	0.002	0.29			
Student	decision	(I)<(II)	(I)>(II)	(I)>(II)	(I)>(II)	(I)>(II)		(I)>(II)	(I)>(II)	6	-1	5
	ν	-3.40	2.92	3.28	7.54	3.78		2.57	3.80			
	p-value	0.001	0.003	0.001	0	0		0.01	0			
BB1	decision	(I)<(II)	(I)=(II)	(I)=(II)	(I)>(II)	(I)>(II)	(I)<(II)		(I)>(II)	3		1
	ν	-2.78	1.07	1.57	7.04	3.13	-2.57		5.12			
	p-value	0.005	0.28	0.11	0	0.002	0.01		0			
BB7	decision	(I)<(II)	(I)=(II)	(I)=(II)	(I)>(II)	(I)=(II)	(I)<(II)	(I)<(II)		1	-3	-2
	ν	-3.87	-0.26	-0.06	6.32	1.05	-3.80	-5.12				
	p-value	0	0.79	0.94	0	0.29	0	0				

Table 4.9: Example of the score method for goodness-of-fit testing for copulas based on the pairwise Clarke test with Schwarz's correction. Sampled data of length $n = 1000$ comes from the bivariate Gaussian copula with dependence parameter $r = 0.75$. Test checks the equivalence of both models $H_0 : \ (I) = (II)$. The outcome (I)>(II) means that model (I) is better than model (II) and (I)<(II) vice versa.

(I)↓, (II)→	result	Gauss	Frank	Plackett	Clayton	Gumbel	Student	BB1	BB7	$\sum_{(I)>(II)}$	$\sum_{(I)<(II)}$	score
Gauss	decision		(I)<(II)	(I)>(II)	(I)>(II)	(I)=(II)	(I)>(II)	(I)>(II)	(I)>(II)			
	B		416	541	593	524	587	543	579	5	-1	4
	p-value		0	0.005	0	0.06	0	0.004				
Frank	decision	(I)>(II)		(I)=(II)	(I)>(II)	(I)=(II)	(I)>(II)	(I)>(II)	(I)>(II)			
	B	584		508	570	503	595	618	654	5	0	**5**
	p-value	0		0.31	0	0.43	0	0	0			
Plackett	decision	(I)<(II)	(I)=(II)		(I)>(II)	(I)=(II)	(I)<(II)	(I)>(II)	(I)>(II)			
	B	459	492		546	503	467	543	584	3	-2	1
	p-value	0.005	0.31		0.002	0.43	0.02	0.004	0			
Clayton	decision	(I)<(II)	(I)<(II)	(I)<(II)		(I)<(II)	(I)<(II)	(I)<(II)	(I)>(II)			
	B	407	430	454		421	414	440	529	1	-6	-5
	p-value	0	0	0.002		0	0	0	0.03			
Gumbel	decision	(I)=(II)	(I)=(II)	(I)=(II)	(I)>(II)		(I)=(II)	(I)>(II)	(I)>(II)			
	B	476	497	497	579		524	583	625	3	0	3
	p-value	0.06	0.43	0.43	0		0.06	0	0			
Student	decision	(I)<(II)	(I)<(II)	(I)>(II)	(I)>(II)	(I)=(II)		(I)>(II)	(I)>(II)			
	B	413	405	533	586	476		548	728	4	-2	2
	p-value	0	0	0.02	0	0.06		0	0			
BB1	decision	(I)<(II)	(I)<(II)	(I)<(II)	(I)>(II)	(I)<(II)	(I)<(II)		(I)>(II)			
	B	457	382	457	560	417	452		709	2	-5	-2
	p-value	0	0	0	0	0.06	0		0			
BB7	decision	(I)<(II)	(I)<(II)	(I)<(II)	(I)<(II)	(I)<(II)	(I)<(II)	(I)<(II)				
	B	261	346	416	471	375	272	291	0	-7	-7	
	p-value	0	0	0	0	0	0	0				

Table 4.10: Simulation study with number of repetitions $N = 1000$ for goodness-of-fit testing via score approach based on the Vuong and on the Clarke tests. Both tests use Schwarz's correction term. The results were evaluated for different combinations of sample size $n \in \{150, 1000\}$ and Kendall's coefficient of association $\tau \in \{0.25, 0.5, 0.75\}$.

n	test	τ	decision	Gauss	Plackett	Frank	Gumbel	Clayton	Student [1]	BB1 [2]	BB7 [3]
150	Vuong	0.25	better	47.8	14.0	11.4	67.9	66.9	1.6	0.0	1.5
			equivalent	25.9	40.3	42.7	11.8	15.0	24.3	6.9	16.6
			worse	26.3	45.7	45.9	20.3	18.1	74.1	93.1	81.9
		0.5	better	75.0	41.0	38.6	70.1	60.7	4.8	7.0	11.5
			equivalent	15.9	26.0	21.8	14.4	20.7	45.2	46.5	47.3
			worse	9.1	33.0	39.6	15.5	18.6	50.0	46.5	41.2
		0.75	better	68.2	60.7	75.7	60.8	55.9	0.0	25.5	2.5
			equivalent	19.6	23.2	14.7	25.9	31.4	0.0	56.4	54.7
			worse	12.2	16.1	9.6	13.3	12.7	100.0	18.1	42.8
	Clarke	0.25	better	0.0	39.1	1.8	51.1	29.2	39.1	0.0	10.5
			equivalent	0.0	24.4	25.8	19.5	19.2	15.1	2.1	5.8
			worse	100.0	36.5	72.4	29.4	51.6	45.8	97.9	83.7
		0.5	better	12.7	33.5	62.6	52.4	49.2	29.3	11.5	15.5
			equivalent	11.9	11.5	24.3	19.7	16.1	23.0	20.0	11.2
			worse	75.4	55.0	13.1	27.9	34.7	47.7	68.5	73.3
		0.75	better	67.5	39.3	39.3	55.5	78.0	0.0	11.9	11.7
			equivalent	17.7	11.9	36.1	17.1	8.3	0.0	22.1	8.0
			worse	14.8	48.8	24.6	27.4	13.7	100.0	66.0	80.3
1000	Vuong	0.25	better	85.3	27.4	23.0	88.1	80.5	74.1	3.9	18.0
			equivalent	10.7	56.2	55.3	9.0	14.2	18.6	21.1	47.9
			worse	4.0	6.4	21.6	2.9	5.3	7.3	75.0	34.1
		0.5	better	77.4	79.5	80.3	84.5	75.6	42.7	19.4	17.6
			equivalent	14.8	20.2	19.5	12.0	21.1	48.4	78.2	79.9
			worse	7.8	0.3	0.2	3.5	3.3	8.9	2.4	2.5
		0.75	better	71.9	99.6	100.0	75.8	74.2	0.0	43.8	15.9
			equivalent	26.3	0.4	0.0	22.5	23.6	0.0	55.4	79.6
			worse	1.8	0.0	0.0	1.7	2.2	100.0	0.8	4.5
	Clarke	0.25	better	0.0	99.1	0.3	17.4	13.5	96.1	1.1	1.2
			equivalent	0.0	0.7	0.3	8.4	7.8	2.7	1.5	2.2
			worse	100.0	0.2	99.4	74.2	78.7	1.2	97.4	96.6
		0.5	better	0.6	98.2	68.5	55.8	53.8	88.9	22.2	0.7
			equivalent	6.2	1.1	25.7	10.0	2.8	8.3	12.6	2.0
			worse	93.2	0.7	5.8	34.2	43.4	2.8	65.2	97.3
		0.75	better	37.7	99.8	67.0	65	69.9	0.0	14.1	6.9
			equivalent	31.5	0.2	31.4	10.0	1.9	0.0	70.7	14.2
			worse	30.8	0.0	16.0	25.0	28.2	100.0	15.2	78.9

[1] Student's t copula with 4 degrees of freedom
[2] BB1 copula with upper tail dependence $\lambda_U = 0.3$
[3] BB7 copula with upper tail dependence $\lambda_U = 0.3$

Table 4.11: Simulation study with number of repetitions $N = 1000$ for goodness-of-fit testing via score approach based on the Vuong and on the Clarke tests. Both tests use Akaike's correction term. The results were evaluated for different combinations of sample size $n = 150$ and Kendall's coefficient of association $\tau \in \{0.25, 0.5, 0.75\}$.

n	test	τ	decision	Gauss	Plackett	Frank	Gumbel	Clayton	Student [1]	BB1 [2]	BB7 [3]
150	Vuong	0.25	better	49.4	14.0	11.5	67.0	66.8	1.6	0.0	1.5
			equivalent	25.5	38.7	41.5	12.1	15.1	24.3	6.9	16.6
			worse	25.1	47.3	47.0	20.9	18.1	74.1	93.1	81.9
		0.5	better	75.4	41.0	37.7	70.1	60.7	4.8	7.0	11.5
			equivalent	15.8	25.8	21.8	14.2	20.8	45.1	46.4	46.9
			worse	8.80	33.2	40.5	15.7	18.5	50.1	46.6	41.6
		0.75	better	68.9	60.7	75.6	60.7	55.9	0.0	25.5	2.5
			equivalent	18.9	23.2	15.0	25.9	31.9	0.0	56.4	54.7
			worse	12.2	16.1	9.4	13.4	12.7	0.0	18.1	42.8
	Clarke	0.25	better	0.0	39.0	1.9	51.2	29.7	39.1	0.0	10.5
			equivalent	0.0	23.7	25.8	19.3	19.0	15.1	2.1	5.8
			worse	100.0	37.3	72.3	29.5	51.3	45.8	97.9	83.7
		0.5	better	14.4	33.5	63.1	52.2	49.4	29.3	11.7	15.5
			equivalent	11.6	11.4	23.9	20.1	16.0	22.8	20.0	11.0
			worse	74.0	55.1	13.0	27.7	34.6	47.9	68.3	73.5
		0.75	better	69.1	39.3	39.3	55.2	78.0	0.0	11.9	11.7
			equivalent	16.3	11.9	35.5	17.3	8.3	0.0	22.0	8.0
			worse	14.6	48.8	25.2	27.5	13.7	100	66.1	80.3

[1] Student's t copula with 4 degrees of freedom
[2] BB1 copula with upper tail dependence $\lambda_U = 0.3$
[3] BB7 copula with upper tail dependence $\lambda_U = 0.3$

Chapter 5

Application: Multivariate modeling of international European stock market index returns

In this chapter we simulate future returns for FINANCIAL MARKET INDICES data set by applying TS-PAIR-COPULA approach. At first, we model conditional mean and volatility via Time Series models. We see, that the GARCH(1,1) models with skewed standard t distribution for error terms are most successful at modeling the serial correlation in the conditional mean and the conditional variance. We perform the further analysis on the transformed standardized residuals and use them as an input data for pair-copula construction with copulas from different families. The structure of the pair-copula decomposition of joint density is specified by applying a D-vine rule. We estimate the starting values of the D-vine parameters sequentially using the hierarchical algorithm in Section 3.6.2 that is based on the h-functions. Additionally, after fitting the D-vine, we validate the bivariate distribution of the variable pairs, which were not explicitly modeled in the PCC. Next, we simulate scenarios for future returns and investigate its time series properties. Finally, we validate our results by comparison the dependence structure within the simulated and the true data.

We introduce some notions that will be used through this chapter. We denote by n the number of observations in each return time series and $i = 1, \ldots, n$. Further, d means the dimension of the FINANCIAL MARKET INDICES data set, i.e. $d = 10$ and $j = 1, \ldots, d$.

5.1 Univariate modeling of margins using GARCH approach

In this section, we apply the TIME SERIES APPROACH discussed in Chapter 2 to model marginal distributions. According to Table 2.1 in Section 2.6.1 "*Order of ARMA Models*", ACF and PACF provide graphical tools that can be useful when detecting an appropriate order of an ARMA model. Sample ACF and PACF of 10 return series are given in Figure 5.1. As we can see, neither ACF nor PACF cuts off after any lag. On the basis of these plots, it is difficult to guess the correct order of the observed processes. We have fitted four models for each return series from the data set FINANCIAL MARKET INDICES, namely GARCH(1,1), AR(1)-GARCH(1,1), MA(1)-GARCH(1,1) and ARMA(1,1)-GARCH(1,1). In Chapter 1, it was discovered that the return series itself are non-normally distributed because of high kurtosis and non-zero skewness. So, it is reasonable to assume that the error terms in GARCH models follow a standard skewed t distribution. The term *standard* means that the distribution has zero mean and unit variance. The full summaries of univariate TS models are provided in Table 5.2. This table contains estimated model parameters, model checking results and goodness-of-fit measures. Graphical tools of residual analysis such a QQ-plot and time series plot of standardized residuals plus ACF of standardized and squared standardized residuals are presented in Figures 5.2 and 5.3. Model checking techniques used here were explicitly described in Chapter 2 in Section 2.8 "*An illustrative example*".

The next step after the model fitting is its diagnostics. It implies the residual analysis as well as model comparison. In practice, we are interested in a model with smallest number of parameters. If our model assumptions are correct, the series of standardized residuals have to be independently identical distributed. According to the Akaike Information Criterion given in Table 5.2, none of the extended models AR-GARCH, MA-GARCH or ARMA-GARCH fit the data better than the simplest GARCH(1,1) model: the AICs in extended models are nearly equal to the AICs in the GARCH models for all indices. The other goodness-of-fit criteria such as BIC and SIC yield the same conclusion. Thus, we restrict our following analysis only to the GARCH models without any ARMA extension.

Next, we are interested in a distribution of standardized residuals of the univariate TS models, which we assumed to follow standard skewed t distribution. According to the p-values of Shapiro- and Jarque-Bera statistics presented in Table 5.2, the null hypothesis of normality can be rejected for every significance level. The QQ-plots make it possible to judge correctness of innovation distributions. For AEX, BVLG, FTMIB and OMXSPI residuals the QQ-points lie approximately on a QQ-line that validates the model assumption of $\varepsilon_i \sim t^{stand}_{\nu,\lambda}$. For residual series of another indices, there are small deviations on the upper and lower tails. However, these discrepancies are not of a high degree. Additionally, the shape and skew parameters, ν and λ, are significant in all 10 GARCH models. Hence, we accept a standard skewed t distribution $t^{stand}_{\nu,\lambda}$ for error terms as correct.

Another point of interest is whether any correlations in the residual series are present. The sample autocorrelation functions for standardized residuals as well as the sample autocorrelation functions for squared standardized residuals show few significant peaks till lag 30. If we study these peaks outside the confidence band more precisely, we see that the deviations from the upper bound are low. Moreover, the standardized residuals are not only uncorrelated but also independent until lag 15 according to the p-values of $Q(15)$ and $Q^2(15)$ statistics of Ljung-Box test.

Further, points on the time series plots of the standardized residuals fluctuate randomly about zero line and do not show any pattern. Estimated means and standard deviations of all residual series are close to 0 and 1 respectively (see Table 5.2). So, residuals seem to be independent identical distributed. It is also remarkable by virtue of the great p-values of the TR^2 statistics that there are no ARCH effects in the series of standardized residuals (see Table 5.2).

Summarizing the conclusions presented above, we assume the GARCH(1,1) models with skewed Student's t innovations to be suitable for modeling the time-changing volatility of the univariate return series.

5.2 Analysing dependence structure within standardized residuals pairs of the fitted GARCH models

As mentioned in Chapter 3, the rank based dependence measures are invariant under monotone increasing transformations. Such dependence measures are Kendall's τ, Chi- and K-plots. Pair-copula constructions requires as input data, which is independently identically distributed on the unit interval. So, the standardized residuals $\hat{\varepsilon}_i$ should be transformed into the interval $[0, 1]$ before we start the estimation procedure. We do it by applying estimated cumulative distribution functions $\hat{F}_{\boldsymbol{\theta}_j}$, $j = 1, \ldots, d$ to each residual sequence. $\hat{F}_{\boldsymbol{\theta}_j}$ is assumed distribution of error terms in time series models. Its parameters will be estimated in a GARCH estimation procedure. Thereby, we become sequences $u_{ij} \sim \mathcal{U}[0, 1]$, where

$$u_{ij} = \hat{F}_{\boldsymbol{\theta}_j}(\hat{\varepsilon}_{ij}) \quad \forall i, j \; . \tag{5.1}$$

The (empirical) cumulative distribution function is a monotone increasing mapping. Hence, associations described by the rank based dependence measures are as aforesaid the same for the original data vectors $\hat{\varepsilon}_j$ as well as for the transformed residuals u_j, $j = 1, \ldots, d$.

Figure 5.4 represents scatter and empirical contour plots of all possible pairs $u_{j_1} \sim u_{j_2}$ with $j_1, j_2 \in \{1, \ldots, d\}$ and $j_1 \neq j_2$, a total of 45 cases. Chi- and K-plots are displayed on Figure 5.5 on its upper and lower panel respectively. Figure 5.7 illustrates estimated Kendall's τ correlation matrix in the form of an image matrix. Each cell of the image plot stands for Kendall's correlation coefficient and is colored. A dark color marks a comparatively strong dependence. And vice versa, a bright color stands for the dependence close

to 0. As we can see from Chi- and Kendall plots, there are exclusive positive dependences of different degrees between the pairs u_j, $j = 1, \ldots, d$. The empirical Kendall's τs confirm this fact.

It is also remarkable, that STOXXER transformed residuals show a stronger dependence with another data vectors as any other index. It can be explained by its calculation method. The STOXXER stock index is based on the most stable European stocks, that are also included in the calculation formula of national stock indices. The highest dependency is observed between the residuals of FCHI and STOXXER whose Kendall's τ is equal to 0.88. The two pairs BVLG vs. ATX and BVLG vs. OMXC20 are characterized by the lowest dependency with corresponding Kendall's τ equal to 0.43.

By modeling financial time series, we are also interested in co-movement of univariate returns. As mentioned in Chapter 3, tail dependence and particularly lower tail dependence is an important feature of financial return series. To decide whether there are tail dependences between the pairs of transformed standardized residuals, we look at the Chi-plots. This methodology is extensively described in Section 3.4.4. If points of Chi-plot do not return to the horizontal zero line for lambda values close to 1, it indicates a presence of tail dependence. We see the presence of tail dependence for each pair from the Chi-plots presented in Figure 5.5. The kind of such dependence, lower or upper one, is not clear. So, we have to look at the Chi-plot that were computed separately for the data from right upper and left lower corners of bivariate scatter plots. Figure 5.6 makes it clear that lower tail dependence is presented for all data pairs. In contrast, upper tail dependence is also observed, but not for all data pairs. For example, there is no upper tail dependence between OMXC20 and any other data vector including STOXXER.

5.3 D-vine identification

For a D-vine construction, we need to choose only an order of 10 variables for the first D-vine tree. It determines the other trees on the D-vine. There are 12! possible orders of the portfolio index set. We constrain our analysis only to one D-vine that we determine using the rule *put strongest correlation on top* for the first tree. We do it by looking at the image plot of the estimated Kendall's τ correlation matrix of the transformed data u_j, $j = 1 \ldots d$. We build a first tree of our D-vine by sequentially adding an edge that corresponds to the strongest possible correlation of all correlations that are left without violating the D-vine rules. As first we find an index pair with the strongest correlations, it is

$$4 - 10$$

with Kendall's correlation $\hat{\tau} = 0.88$. On the next step, we determine further index that is characterized by strongest correlation either with index 10 or with index 4. Such one is 6

whose correlation with 10 is $\hat{\tau} = 0.85$. So, we extend our sequence to

$$4 - 10 - 6 \ .$$

The next strongest correlated pair one part of which is either 4 or 6 is $5 - 4$ with corresponding empirical Kendall's $\hat{\tau} = 0.79$. Hence, the tree will be upgraded by the new root

$$5 - 4 - 10 - 6 \ .$$

Continuing this procedure, we get a resulting sequence that is a first tree of our D-vine

$$2 - 9 - 5 - 4 - 10 - 6 - 1 - 8 - 7 - 3 \ . \tag{5.2}$$

Finally, the first tree (5.2) determines completely the whole structure of the D-vine that is illustrated on Figure 5.8.

5.4 D-vine estimation

5.4.1 Sequential estimation

In this section, we estimate pair-copula parameters of the D-vine sequentially by applying the algorithm described in Section 3.6.2 on the page 59. The formulas for determining conditional data on each D-vine tree are given in Appendix E. For selection of the appropriate copula for each building block, we combine following techniques which were described in the theoretical chapters:

- Scatter plot of the data
- Scatter plot of the data with standard normal margins
- Empirical contour plot of the data with standard normal margins
- Chi-plots for the upper and lower tails
- Empirical coefficients of the lower and upper tail dependence
- Score test based on the Vuong statistic with the Schwarz's correction.

Selecting procedures containing these steps for each building block of all D-vine trees are given in Appendix F. We illustrate this approach on the two examples given below. Data in the both examples comes from the first D-vine tree.

GDAXI ~ AEX: transformed standardized residuals of univariate time series models for indices GDAXI and AEX, Figure 5.9. As we can see, the score method based on the Vuong comparison prefers a Student copula. The tails on both scatter plots are symmetrical. Pursuant to the Chi-plots for the lower and upper tails, we can conclude a presence of the symmetrical middle dependences in both tails. Tail dependences estimated empirically according to the formulas (3.9)-(3.10) are nearly

equal. Note, these estimators for λ_U and λ_L are not stable, they depend on the sample size and have a tendency to overestimation. Here, we are not interested in absolute values of estimators but on the difference between them. Additionally, we calculate tail dependences from estimated Student and BB1 models and compare them with the empirical ones. The Student as well as the BB1 copula seems to reproduce the difference between λ_L and λ_U similarly adequately. Finally, we choose the Student copula. Our decision does not contradict conclusions above.

OMXC20 ~ BVLG: transformed standardized residuals of univariate time series models for indices OMXC20 and BVLG, Figure 5.10. We choose an appropriate copula similar to the previous example. At first, the score test favor the BB1 copula over all others. Looking at the scatter and empirical contour plots, one can detect asymmetrical tail dependences. This fact will be also confirmed by the Chi-plots. There is no upper tail dependence but clearly depicted lower tail dependence. The estimated lower tail dependence is nearly twice as big as the upper one. This fact can not be explained by the symmetrical Student copula, but by the BB1 copula. So, we select the BB1 model as the most appropriate one.

5.4.2 Joint MLE estimation

Finally, we use the parameter values estimated sequentially as input for the optimization function. In this step, the log-likelihood function of the pair-copula decomposition determined by the 10-dimensional D-vine will be maximized by joint optimization of all parameters. Table 5.3 represents starting values (column SEQ) and resulting parameter values after optimization procedure (column MLE). As we can see, there are inconsiderable differences between SEQ and MLE estimations. Additionally, we fit D-vine with all Gaussian blocks to our data.

We apply the Vuong test to examine whether the MLE model is better than the SEQ model. Note, we do not need any correction in this case because of the same number of parameters in both models. According to the value of the test statistic -1.18 and it's p-value 0.23 when testing the null hypothesis $MLE = SEQ$, we can not reject the null hypothesis of equivalence. So, if one does not possess any procedure for joint estimation of D-vine, the sequential estimation provides right well results. Finally, we state that both mixed D-vine models, MLE and SEQ, fit the data better than the D-vine with all Gaussian blocks.

5.5 Validation results

Next, it is interesting whether the dependence structure of the new data, sampled from the estimated pair-copula construction is similar to the dependence structure within the

Table 5.1: Testing the null hypothesis about equivalence of models (I) and (II) via Vuong test without any correction term in cases MLE vs. SEQ or SEQ vs. MLE and with Akaike's and with Schwarz's correction in other cases

	(II) SEQ			**(II) D-vine with all GAUSS blocks**					
	none correction			**Schwarz's correction**			**Akaike's correction**		
	stat.	**p-val.**	**decision**	**stat.**	**p-val.**	**decision**	**stat.**	**p-val.**	**decision**
(I) MLE	−1.18	0.23	(I)=(II)	2.99	0.003	(I)>(II)	3.72	0	(I)>(II)
(I) SEQ	1.18	0.23	(I)=(II)	3.19	0.001	(I)>(II)	3.92	0	(I)>(II)

transformed standardized residuals. Note, not all variable pairs were explicitly modeled in the pair-copula decomposition. To investigate the bivariate distribution of the simulated data, we

(i) generate a set of N samples of length n from the pair-copula decomposed model using estimated parameters
(ii) compare simulated variables with the observed data according to average values of following statistics: the sample Kendall's $\hat{\tau}$, empirical coefficients of the lower and upper tail dependences $\hat{\lambda}_L$ and $\hat{\lambda}_U$.

Table 5.4 presents the results of this comparison with the number of random samples $N = 2000$ each of the length $n = 855$. The diagonal entries of this table correspond to the data pairs which were modeled directly within the pair-copula construction. All other data pairs were not explicitly modeled. As we can see from this table, all directly modeled dependences are captured quite well. The data pairs with indirectly modeled dependences shows features similar to the observed data either. Based on this study, we state that the Kendall's τ are reproduced excellently for all index pairs. The coefficients of tail dependence are reflected for the data pairs from the main diagonal of Table 5.4 quite well and for the remaining data pairs a little worse.

5.6 Simulation of future returns and validation results

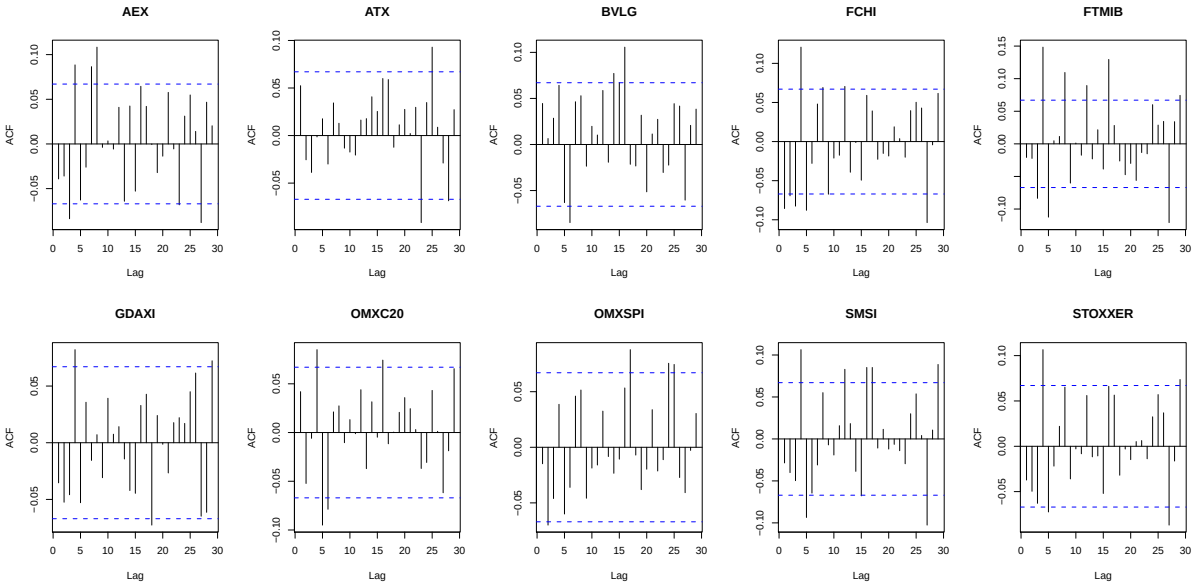
For the simulation of the future returns, we use the second part of the algorithm described in Section 3.7. Based on the one-step-ahead forecasting, we simulate $N = 2000$ scenarios of length 50 for each index. Multivariate data sets created on this way feature the same dependence structure as the data from the past and should have the same time series properties as the observed return series. Sets of time series plots of simulated future returns are displayed by Figure 5.11 separately for each index.

We validate the time series properties of simulated future returns and summarize results in Table 5.5. Here, we compute such descriptive statistics as sample mean, sample standard

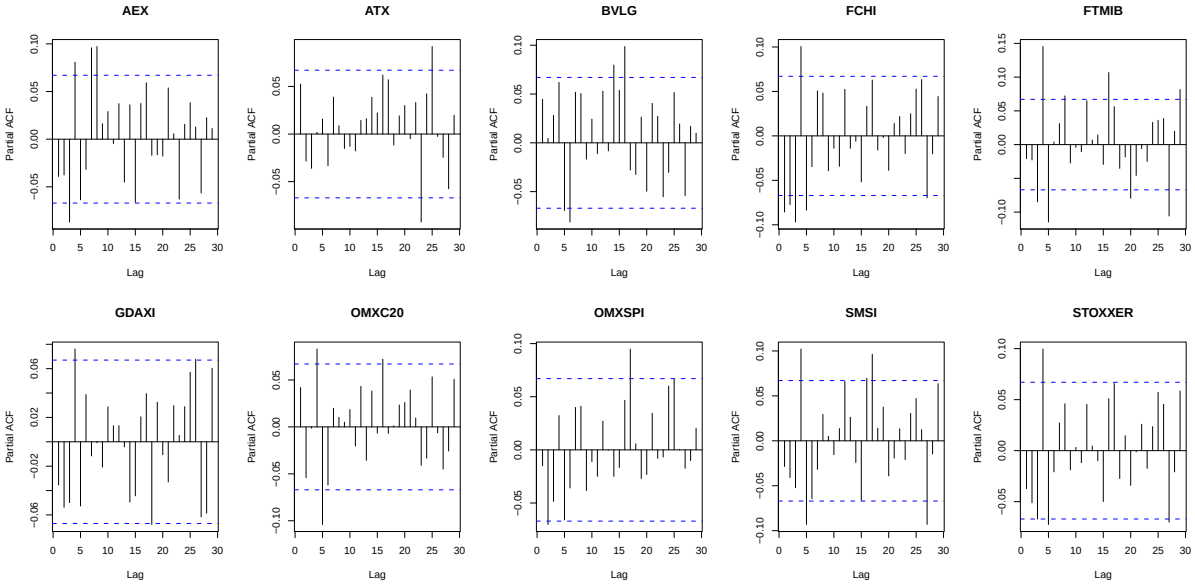
deviation, sample kurtosis and sample skewness for the true data and evaluate its average values over the all of the scenarios for each stock market index separately. As we can see, the standard deviation, kurtosis ans skewness is reproduced by the univariate series quite well. There are little differences in the reproduction of the expected mean: the expectation of the simulated returns is a little bit higher as the expectation of test data.

Finally, we are interested whether the dependence structure within the simulated returns is similar to the dependence structure within the observed future returns. For validation we use the test data and simulated scenarios and sum up the results in Table 5.6. We calculate Kendall's τ and tail dependence coefficients empirically for all possible pairs of the test data for the validation. Additionally, we evaluate average values of these statistics over all of the scenarios for the same pairs of the simulated data. Note, the calculation of empirical estimators for τ, λ_U and λ_L is based on the small samples of the length 50. Hence, the estimators may be deviant from the true values. As we can see, the dependence structure within the simulated data is reproduced quite well.

Figure 5.1: Sample autocorrelation and sample partial autocorrelation functions of stock market indices from data set FINANCIAL MARKET INDICES



(a) Estimated ACFs of return series till lag 30



(b) Estimated PACFs of return series till lag 30

Table 5.2: Summaries, residual analysis and goodness-of-fit of estimated ARMA(1,1)-GARCH(1,1) models with standard skewed Student's t distribution for error terms applied to the data set FINANCIAL MARKET INDICES

MODEL SPECIFICATION			AR	MA	GARCH			$t_{\nu,\lambda}$		RESIDUALS ANALYSIS							GOODNESS-OF-FIT		
(1) AEX		μ	ϕ	θ	ω	α	β	λ	ν	χ^2	W	Q(15)	$Q^2(15)$	TR^2	mean	sd	AIC	BIC	SIC
	Estimate/Statistic	0.03			0.02	0.12	0.88	0.88	8.6	64.51	0.99	15.47	6.91	6.82	-0.06	1	3.37	3.4	3.37
GARCH(1,1)	stand. deviation	0.03			0.01	0.02	0.02	0.04	2.49										
	p-value	0.41			0.06	0	0	0	0	0	0	0.42	0.96	0.96					
	Estimate/Statistic	0.03	-0.04		0.02	0.12	0.88	0.87	8.74	65.83	0.99	17.72	6.57	6.37	-0.06	1	3.37	3.41	3.37
AR(1)-GARCH(1,1)	stand. deviation	0.03	0.04		0.01	0.02	0.02	0.04	2.57										
	p-value	0.39	0.29		0.06	0	0	0	0	0	0	0.28	0.97	0.97					
	Estimate/Statistic	0.03		-0.04	0.02	0.12	0.88	0.87	8.74	65.86	0.99	17.82	6.56	6.36	-0.06	1	3.37	3.41	3.37
MA(1)-GARCH(1,1)	stand. deviation	0.03		0.04	0.01	0.02	0.02	0.04	2.57										
	p-value	0.39		0.29	0.06	0	0	0	0	0	0	0.27	0.97	0.97					
	Estimate/Statistic	0.01	0.66	-0.7	0.02	0.12	0.88	0.86	9.07	67.94	0.99	21.17	6.6	6.34	-0.07	1	3.37	3.42	3.37
ARMA(1,1)-GARCH(1,1)	stand. deviation	0.01	0.2	0.19	0.01	0.02	0.02	0.04	2.77										
	p-value	0.37	0	0	0.06	0	0	0	0	0	0	0.13	0.97	0.97					
(2) ATX		μ	ϕ	θ	ω	α	β	λ	ν	χ^2	W	Q(15)	$Q^2(15)$	TR^2	mean	sd	AIC	BIC	SIC
	Estimate/Statistic	0.05			0.05	0.14	0.85	0.81	10	121.97	0.98	6.06	11.98	12.57	-0.05	1	3.84	3.87	3.84
GARCH(1,1)	stand. deviation	0.05			0.02	0.03	0.03	0.04	2.65										
	p-value	0.31			0.01	0	0	0	0	0	0	0.98	0.68	0.64					
	Estimate/Statistic	0.05	-0.02		0.05	0.14	0.85	0.81	10	123.5	0.98	6.83	11.73	12.29	-0.05	1	3.84	3.88	3.84
AR(1)-GARCH(1,1)	stand. deviation	0.05	0.04		0.02	0.03	0.03	0.05	2.64										
	p-value	0.31	0.61		0.01	0	0	0	0	0	0	0.96	0.7	0.66					
	Estimate/Statistic	0.05		-0.02	0.05	0.14	0.85	0.81	10	123.61	0.98	6.88	11.71	12.27	-0.05	1	3.84	3.88	3.84
MA(1)-GARCH(1,1)	stand. deviation	0.04		0.04	0.02	0.03	0.03	0.05	2.64										
	p-value	0.31		0.6	0.01	0	0	0	0	0	0	0.96	0.7	0.66					
	Estimate/Statistic	0.01	0.75	-0.79	0.05	0.14	0.85	0.81	10	127.07	0.98	9.68	11.2	11.75	-0.06	1	3.84	3.89	3.84
ARMA(1,1)-GARCH(1,1)	stand. deviation	0.01	0.17	0.16	0.02	0.03	0.03	0.05	2.63										
	p-value	0.37	0	0	0.01	0	0	0	0	0	0	0.84	0.74	0.7					
(3) BVLG		μ	ϕ	θ	ω	α	β	λ	ν	χ^2	W	Q(15)	$Q^2(15)$	TR^2	mean	sd	AIC	BIC	SIC
	Estimate/Statistic	0.09			0.01	0.17	0.84	0.94	4.94	227.86	0.97	34.06	13.05	13.08	-0.06	0.99	2.71	2.74	2.71
GARCH(1,1)	stand. deviation	0.02			0	0.04	0.03	0.05	0.86										
	p-value	0			0.05	0	0	0	0	0	0	0	0.6	0.6					
	Estimate/Statistic	0.09	0.03		0.01	0.17	0.85	0.95	4.95	220.06	0.97	29.21	13.93	14	-0.06	0.99	2.71	2.75	2.71
AR(1)-GARCH(1,1)	stand. deviation	0.02	0.04		0	0.04	0.03	0.05	0.87										
	p-value	0	0.43		0.05	0	0	0	0	0	0	0.02	0.53	0.53					
	Estimate/Statistic	0.09		0.02	0.01	0.17	0.85	0.95	4.93	222.5	0.97	30.32	13.79	13.86	-0.06	0.99	2.71	2.75	2.71
MA(1)-GARCH(1,1)	stand. deviation	0.02		0.03	0	0.04	0.03	0.05	0.86										
	p-value	0		0.49	0.05	0	0	0	0	0	0	0.01	0.54	0.54					
	Estimate/Statistic	0.02	0.75	-0.7	0.01	0.17	0.85	0.97	4.96	217.22	0.98	14.12	14.87	14.88	-0.06	0.99	2.71	2.75	2.71
ARMA(1,1)-GARCH(1,1)	stand. deviation	0.02	0.14	0.14	0	0.04	0.03	0.05	0.87										
	p-value	0.13	0	0	0.06	0	0	0	0	0	0	0.52	0.46	0.46					
(4) FCHI		μ	ϕ	θ	ω	α	β	λ	ν	χ^2	W	Q(15)	$Q^2(15)$	TR^2	mean	sd	AIC	BIC	SIC
	Estimate/Statistic	0.04			0.02	0.1	0.89	0.9	7.68	217.9	0.98	15.51	16.31	17.23	-0.06	1.01	3.47	3.5	3.47
GARCH(1,1)	stand. deviation	0.04			0.01	0.02	0.02	0.04	1.83										
	p-value	0.32			0.04	0	0	0	0	0	0	0.42	0.36	0.31					
	Estimate/Statistic	0.04	-0.09		0.02	0.1	0.89	0.89	7.7	233.99	0.98	10.77	16.75	17.3	-0.06	1.01	3.46	3.5	3.46
AR(1)-GARCH(1,1)	stand. deviation	0.04	0.03		0.01	0.02	0.02	0.04	1.84										
	p-value	0.3	0.01		0.05	0	0	0	0	0	0	0.77	0.33	0.3					
	Estimate/Statistic	0.04		-0.1	0.02	0.1	0.89	0.89	7.73	234.36	0.98	11.02	16.89	17.37	-0.06	1.01	3.46	3.5	3.46
MA(1)-GARCH(1,1)	stand. deviation	0.03		0.04	0.01	0.02	0.02	0.04	1.85										
	p-value	0.29		0.01	0.05	0	0	0	0	0	0	0.75	0.33	0.3					
	Estimate/Statistic	0.01	0.63	-0.73	0.02	0.1	0.89	0.85	8.42	221.15	0.97	13.87	18.36	18.51	-0.07	1.01	3.46	3.5	3.46
ARMA(1,1)-GARCH(1,1)	stand. deviation	0.01	0.12	0.11	0.01	0.02	0.02	0.05	2.15										
	p-value	0.21	0	0	0.04	0	0	0	0	0	0	0.54	0.24	0.24					

Continuing on the next page

Summaries, residual analysis and goodness-of-fit of estimated ARMA(1,1)-GARCH(1,1) models with standard skewed Student's t distribution for error terms applied to the data set FINANCIAL MARKET INDICES – Continuing

MODEL SPECIFICATION			AR	MA	GARCH			$t_{\nu,\lambda}$		RESIDUALS ANALYSIS							GOODNESS OF FIT		
(5) FTMIB		μ	ϕ	θ	ω	α	β	λ	ν	χ^2	W	Q(15)	$Q^2(15)$	TR^2	mean	sd	AIC	BIC	SIC
	Estimate/Statistic	0.01			0.01	0.1	0.89	0.85	8.32	91.07	0.98	11.54	23.81	27.14	-0.06	1	3.31	3.34	3.31
GARCH(1,1)	stand. deviation	0.03			0.01	0.02	0.02	0.04	2.31										
	p-value	0.73			0.08	0	0	0	0	0	0	0.71	0.07	0.03					
	Estimate/Statistic	0.01	-0.09		0.01	0.1	0.9	0.82	9.08	91.51	0.98	8.52	24.43	27	-0.06	1.01	3.31	3.34	3.31
AR(1)-GARCH(1,1)	stand. deviation	0.03	0.04		0.01	0.02	0.02	0.04	2.71										
	p-value	0.75	0.01		0.08	0	0	0	0	0	0	0.9	0.06	0.03					
	Estimate/Statistic	0.01		-0.09	0.01	0.1	0.9	0.82	9.11	91.48	0.98	9.29	24.52	27.04	-0.06	1.01	3.31	3.34	3.31
MA(1)-GARCH(1,1)	stand. deviation	0.03		0.04	0.01	0.02	0.02	0.04	2.73										
	p-value	0.74		0.01	0.08	0	0	0	0	0	0	0.86	0.06	0.03					
	Estimate/Statistic	0.01	-0.24	0.15	0.01	0.1	0.9	0.82	9.03	91.37	0.98	7.22	24.31	26.98	-0.06	1.01	3.31	3.35	3.31
ARMA(1,1)-GARCH(1,1)	stand. deviation	0.04	0.57	0.58	0.01	0.02	0.02	0.04	2.68										
	p-value	0.76	0.67	0.8	0.08	0	0	0	0	0	0	0.95	0.06	0.03					
(6) GDAXI		μ	ϕ	θ	ω	α	β	λ	ν	χ^2	W	Q(15)	$Q^2(15)$	TR^2	mean	sd	AIC	BIC	SIC
	Estimate/Statistic	0.07			0.02	0.1	0.9	0.89	6.02	424.42	0.97	14.9	11.48	12.77	-0.06	1.01	3.4	3.43	3.4
GARCH(1,1)	stand. deviation	0.04			0.01	0.02	0.02	0.04	1.25										
	p-value	0.06			0.07	0	0	0	0	0	0	0.46	0.72	0.62					
	Estimate/Statistic	0.07	-0.07		0.02	0.1	0.9	0.87	6.07	412.92	0.97	15.73	11.3	12.4	-0.06	1.01	3.39	3.43	3.39
AR(1)-GARCH(1,1)	stand. deviation	0.04	0.03		0.01	0.02	0.02	0.04	1.27										
	p-value	0.05	0.05		0.08	0	0	0	0	0	0	0.4	0.73	0.65					
	Estimate/Statistic	0.07		-0.07	0.02	0.1	0.9	0.87	6.08	410.11	0.97	16.06	11.35	12.44	-0.06	1.01	3.39	3.43	3.39
MA(1)-GARCH(1,1)	stand. deviation	0.04		0.03	0.01	0.02	0.02	0.04	1.28										
	p-value	0.05		0.05	0.08	0	0	0	0	0	0	0.38	0.73	0.65					
	Estimate/Statistic	0.03	0.62	-0.68	0.02	0.1	0.9	0.86	6.48	364.27	0.97	18.28	11.89	12.98	-0.07	1.01	3.4	3.44	3.4
ARMA(1,1)-GARCH(1,1)	stand. deviation	0.02	0.21	0.19	0.01	0.02	0.02	0.04	1.46										
	p-value	0.14	0	0	0.07	0	0	0	0	0	0	0.25	0.69	0.6					
(7) OMXC20		μ	ϕ	θ	ω	α	β	λ	ν	χ^2	W	Q(15)	$Q^2(15)$	TR^2	mean	sd	AIC	BIC	SIC
	Estimate/Statistic	0.07			0.03	0.13	0.87	0.91	6.97	328.87	0.97	8.01	8.38	8.48	-0.06	1.01	3.48	3.51	3.48
GARCH(1,1)	stand. deviation	0.04			0.02	0.03	0.03	0.04	1.51										
	p-value	0.07			0.04	0	0	0	0	0	0	0.92	0.91	0.9					
	Estimate/Statistic	0.07	-0.01		0.03	0.13	0.87	0.91	6.96	334.67	0.97	8.33	8.26	8.36	-0.06	1.01	3.48	3.52	3.48
AR(1)-GARCH(1,1)	stand. deviation	0.04	0.03		0.02	0.03	0.03	0.04	1.5										
	p-value	0.07	0.77		0.04	0	0	0	0	0	0	0.91	0.91	0.91					
	Estimate/Statistic	0.07		-0.01	0.03	0.13	0.87	0.91	6.96	335.3	0.97	8.38	8.25	8.34	-0.06	1.01	3.48	3.52	3.48
MA(1)-GARCH(1,1)	stand. deviation	0.04		0.04	0.02	0.03	0.03	0.04	1.5										
	p-value	0.07		0.76	0.04	0	0	0	0	0	0	0.91	0.91	0.91					
	Estimate/Statistic	0.13	-0.86	0.87	0.03	0.13	0.86	0.91	6.94	335.31	0.97	8.24	8.2	8.29	-0.06	1.01	3.48	3.53	3.48
ARMA(1,1)-GARCH(1,1)	stand. deviation	0.08	0.37	0.34	0.02	0.03	0.03	0.04	1.5										
	p-value	0.09	0.02	0.01	0.04	0	0	0	0	0	0	0.91	0.92	0.91					

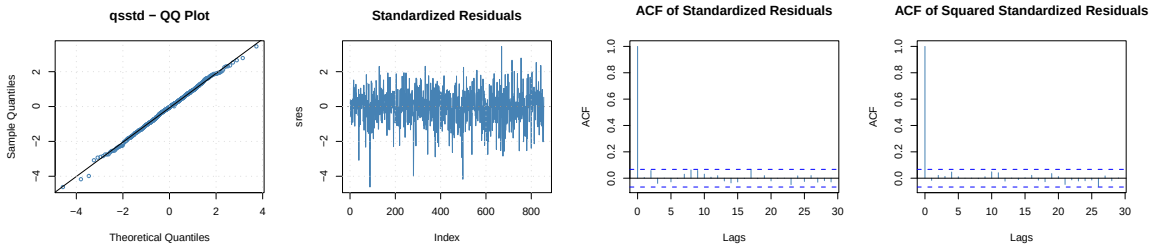
Continuing on the next page

Summaries, residual analysis and goodness-of-fit of estimated ARMA(1,1)-GARCH(1,1) models with standard skewed Student's t distribution for error terms applied to the data set FINANCIAL MARKET INDICES – Continuing

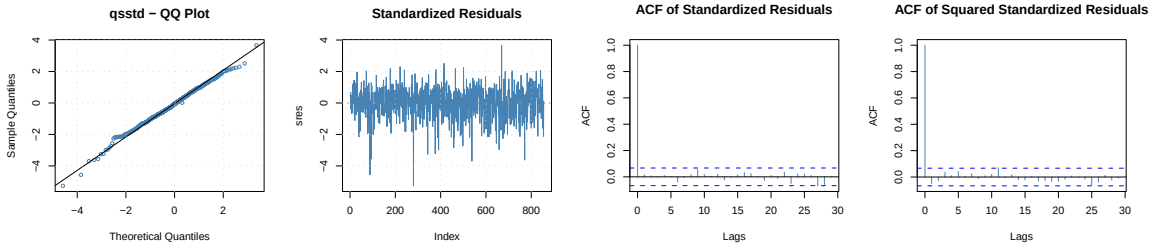
MODEL SPECIFICATION			AR	MA	GARCH			$t_{\nu,\lambda}$		RESIDUALS ANALYSIS							GOODNESS OF FIT		
(8) OMXSPI		μ	ϕ	θ	ω	α	β	λ	ν	χ^2	W	Q(15)	$Q^2(15)$	TR^2	mean	sd	AIC	BIC	SIC
	Estimate/Statistic	0.05			0.02	0.11	0.89	0.87	7.54	444.22	0.97	5.54	11.66	11.47	-0.06	1.01	3.81	3.85	3.81
GARCH(1,1)	stand. deviation	0.05			0.01	0.02	0.02	0.04	1.79										
	p-value	0.24			0.1	0	0	0	0	0	0	0.99	0.7	0.72					
	Estimate/Statistic	0.05	-0.04		0.02	0.11	0.89	0.86	7.54	450.14	0.97	5.71	11.41	11.18	-0.06	1.01	3.81	3.85	3.81
AR(1)-GARCH(1,1)	stand. deviation	0.05	0.03		0.01	0.02	0.02	0.04	1.8										
	p-value	0.25	0.21		0.1	0	0	0	0	0	0	0.98	0.72	0.74					
	Estimate/Statistic	0.05		-0.05	0.02	0.11	0.89	0.86	7.54	451.05	0.97	5.85	11.41	11.18	-0.06	1.01	3.81	3.85	3.81
MA(1)-GARCH(1,1)	stand. deviation	0.04		0.04	0.01	0.02	0.02	0.04	1.8										
	p-value	0.24		0.19	0.1	0	0	0	0	0	0	0.98	0.72	0.74					
	Estimate/Statistic	0.02	0.58	-0.64	0.02	0.1	0.9	0.85	7.65	473.13	0.97	7.52	11.73	11.54	-0.06	1.01	3.81	3.86	3.81
ARMA(1,1)-GARCH(1,1)	stand. deviation	0.02	0.19	0.18	0.01	0.02	0.02	0.04	1.86										
	p-value	0.27	0	0	0.09	0	0	0	0	0	0	0.94	0.7	0.71					
(9) SMSI		μ	ϕ	θ	ω	α	β	λ	ν	χ^2	W	Q(15)	$Q^2(15)$	TR^2	mean	sd	AIC	BIC	SIC
	Estimate/Statistic	0.08			0.02	0.12	0.88	0.9	5.85	503.51	0.96	12.66	13.55	13.4	-0.07	1.01	3.31	3.34	3.31
GARCH(1,1)	stand. deviation	0.04			0.01	0.03	0.02	0.04	1.16										
	p-value	0.03			0.07	0	0	0	0	0	0	0.63	0.56	0.57					
	Estimate/Statistic	0.08	-0.05		0.02	0.12	0.88	0.89	5.9	496.96	0.96	13.1	13.7	13.44	-0.07	1.01	3.31	3.35	3.31
AR(1)-GARCH(1,1)	stand. deviation	0.04	0.03		0.01	0.03	0.02	0.04	1.17										
	p-value	0.02	0.15		0.07	0	0	0	0	0	0	0.59	0.55	0.57					
	Estimate/Statistic	0.08		-0.05	0.02	0.12	0.88	0.89	5.91	495.64	0.96	13.29	13.71	13.45	-0.07	1.01	3.31	3.35	3.31
MA(1)-GARCH(1,1)	stand. deviation	0.03		0.04	0.01	0.03	0.02	0.04	1.18										
	p-value	0.02		0.14	0.07	0	0	0	0	0	0	0.58	0.55	0.57					
	Estimate/Statistic	0.04	0.5	-0.56	0.02	0.12	0.88	0.88	6.11	484.66	0.96	15.12	13.73	13.52	-0.07	1.01	3.31	3.35	3.31
ARMA(1,1)-GARCH(1,1)	stand. deviation	0.03	0.33	0.32	0.01	0.03	0.02	0.05	1.27										
	p-value	0.2	0.13	0.08	0.07	0	0	0	0	0	0	0.44	0.55	0.56					
(10) STOXXER		μ	ϕ	θ	ω	α	β	λ	ν	χ^2	W	Q(15)	$Q^2(15)$	TR^2	mean	sd	AIC	BIC	SIC
	Estimate/Statistic	0.06			0.02	0.11	0.89	0.87	7.28	336.97	0.97	12.05	12.88	13.73	-0.07	1.01	3.31	3.35	3.31
GARCH(1,1)	stand. deviation	0.03			0.01	0.02	0.02	0.04	1.65										
	p-value	0.08			0.06	0	0	0	0	0	0	0.68	0.61	0.55					
	Estimate/Statistic	0.06	-0.06		0.02	0.11	0.89	0.86	7.44	347.55	0.97	12.01	12.84	13.46	-0.07	1.01	3.31	3.35	3.31
AR(1)-GARCH(1,1)	stand. deviation	0.03	0.03		0.01	0.02	0.02	0.04	1.71										
	p-value	0.07	0.06		0.07	0	0	0	0	0	0	0.68	0.61	0.57					
	Estimate/Statistic	0.06		-0.07	0.02	0.11	0.89	0.86	7.46	347.69	0.97	12.4	12.85	13.45	-0.07	1.01	3.31	3.35	3.31
MA(1)-GARCH(1,1)	stand. deviation	0.03		0.04	0.01	0.02	0.02	0.04	1.71										
	p-value	0.07		0.05	0.07	0	0	0	0	0	0	0.65	0.61	0.57					
	Estimate/Statistic	0.02	0.65	-0.73	0.02	0.11	0.89	0.83	8.02	332.76	0.97	16.53	13.24	13.78	-0.08	1.02	3.31	3.36	3.31
ARMA(1,1)-GARCH(1,1)	stand. deviation	0.01	0.13	0.12	0.01	0.02	0.02	0.05	1.96										
	p-value	0.09	0	0	0.06	0	0	0	0	0	0	0.35	0.58	0.54					

End of table

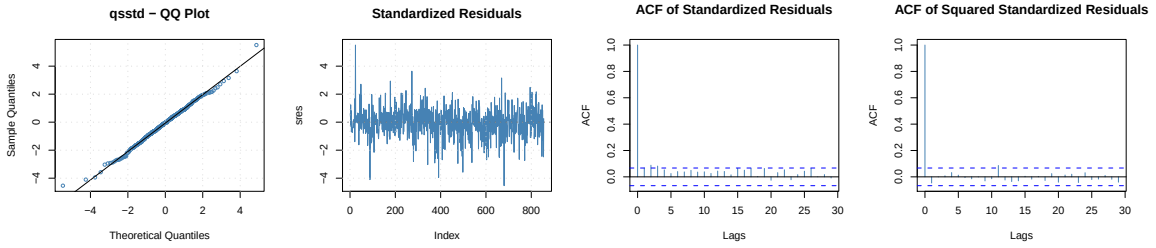
Figure 5.2: Residual analysis of the GARCH(1,1) models for daily returns of stock market indexes with assumption of standardized skewed t innovations $\varepsilon \sim t_{\nu,\lambda}^{stand}$



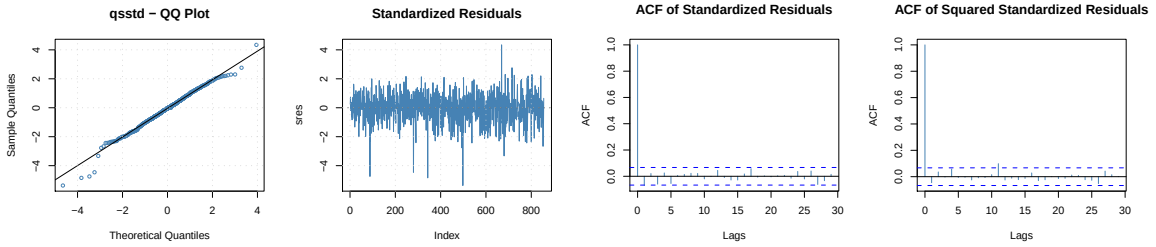
(a) Residual analysis for AEX time series model



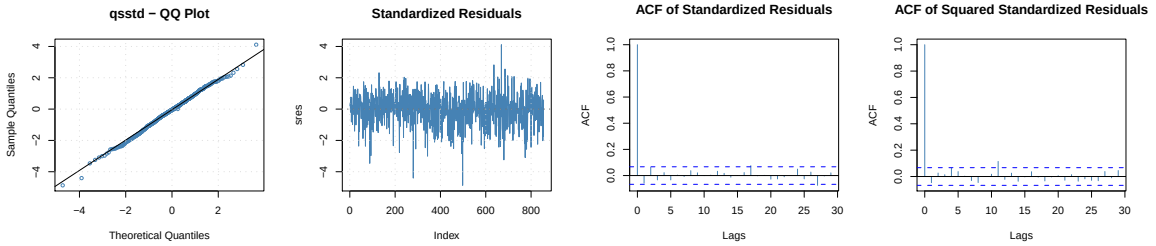
(b) Residual analysis for ATX time series model



(c) Residual analysis for BVLG time series model

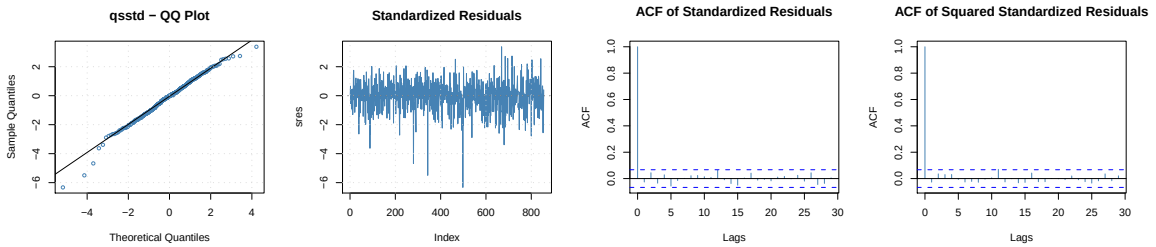


(d) Residual analysis for FCHI time series model

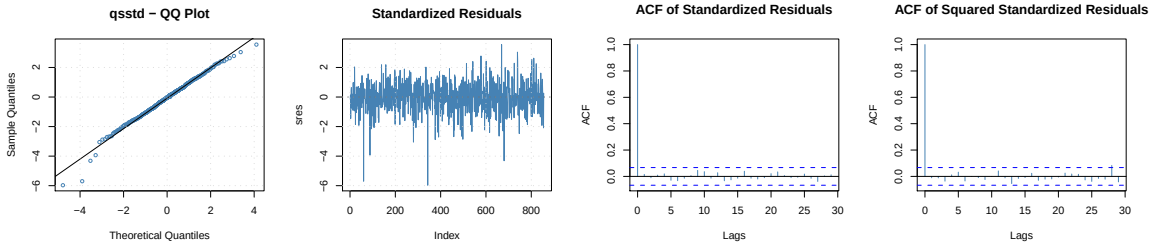


(e) Residual analysis for FTMIB time series model

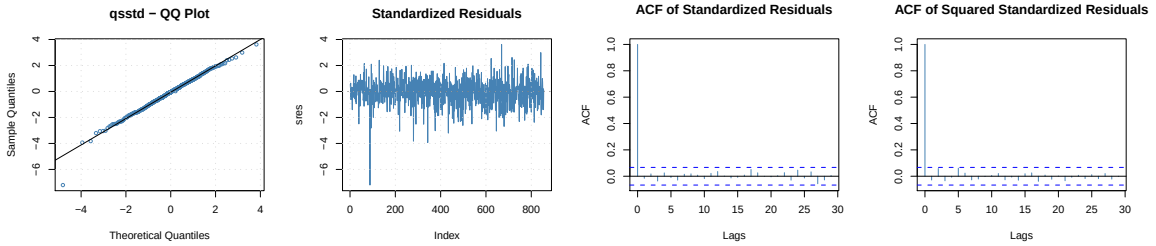
Figure 5.3: Residual analysis of the GARCH(1,1) models for daily returns of stock market indexes with assumption of standardized skewed t innovations $\varepsilon \sim t^{stand}_{\nu,\lambda}$



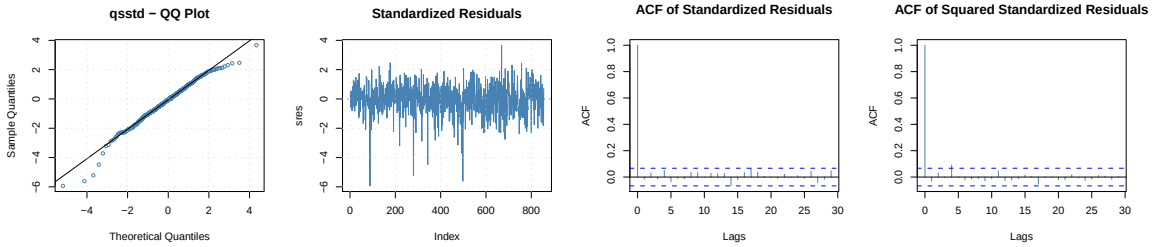
(a) Residual analysis for GDAXI time series model



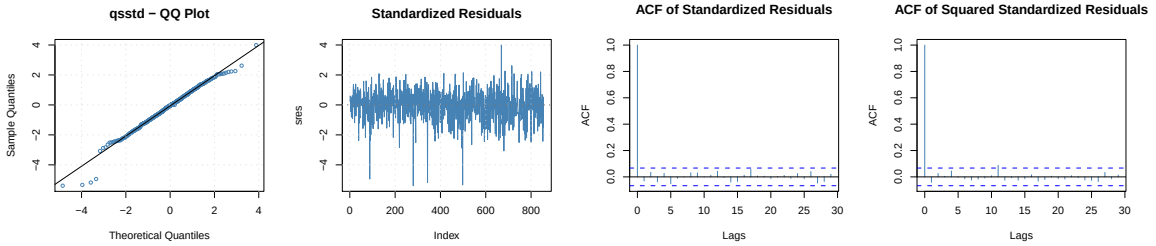
(b) Residual analysis for OMXC20 time series model



(c) Residual analysis for OMXSPI time series model



(d) Residual analysis for SMSI time series model



(e) Residual analysis for STOXXER time series model

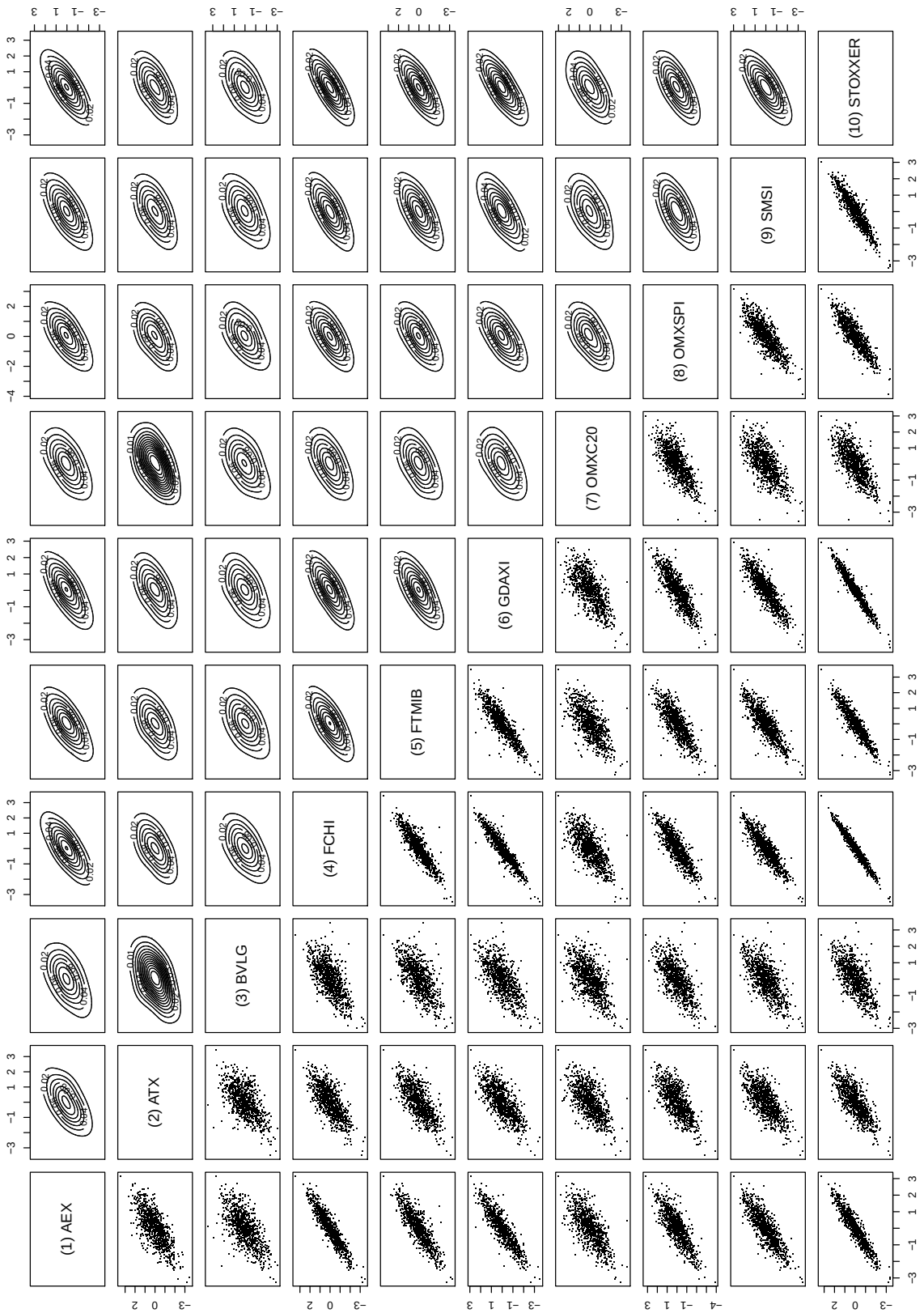


Figure 5.4: Scatter and contour plots for all possible pairs of transformed GARCH(1,1) standardized residuals. Margins were transformed to standard normal sequences.

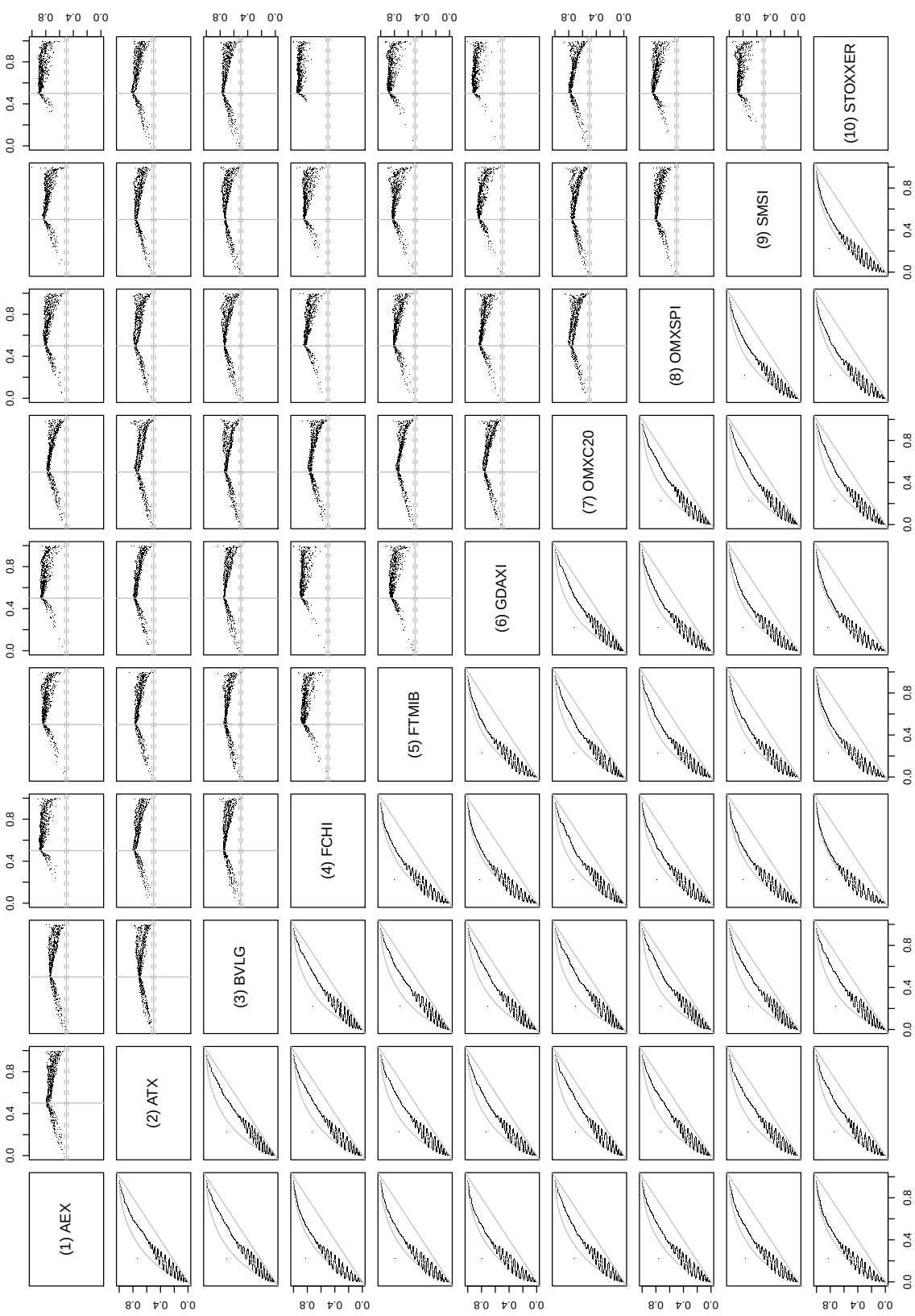


Figure 5.5: Chi- and K-plots of all possible transformed GARCH(1,1) standardized residual pairs: Chi-plots are provided for the interval $[-1, 1]$ on the upper panel and K-plots are presented on the interval $[0, 1]$on the lower panel.

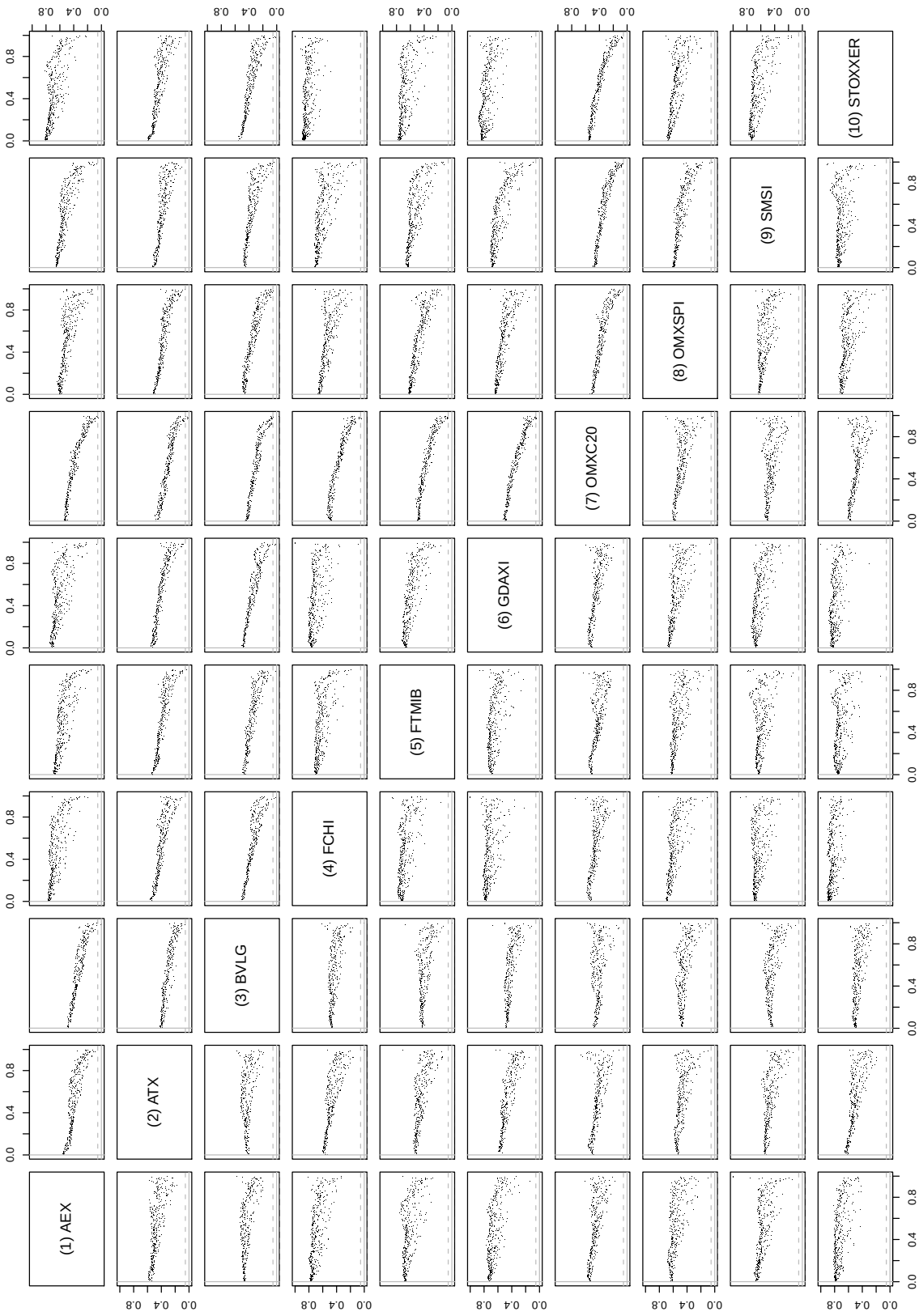


Figure 5.6: Chi-plots for right upper and left lower quadrants of the pairs of the transformed standardized GARCH(1,1) residuals: upper panel represent the data of right upper quadrant and lower panel represent the data for the left lower quadrant.

Figure 5.7: Visualization of the empirical Kendall's τ correlation matrix between all pairs.

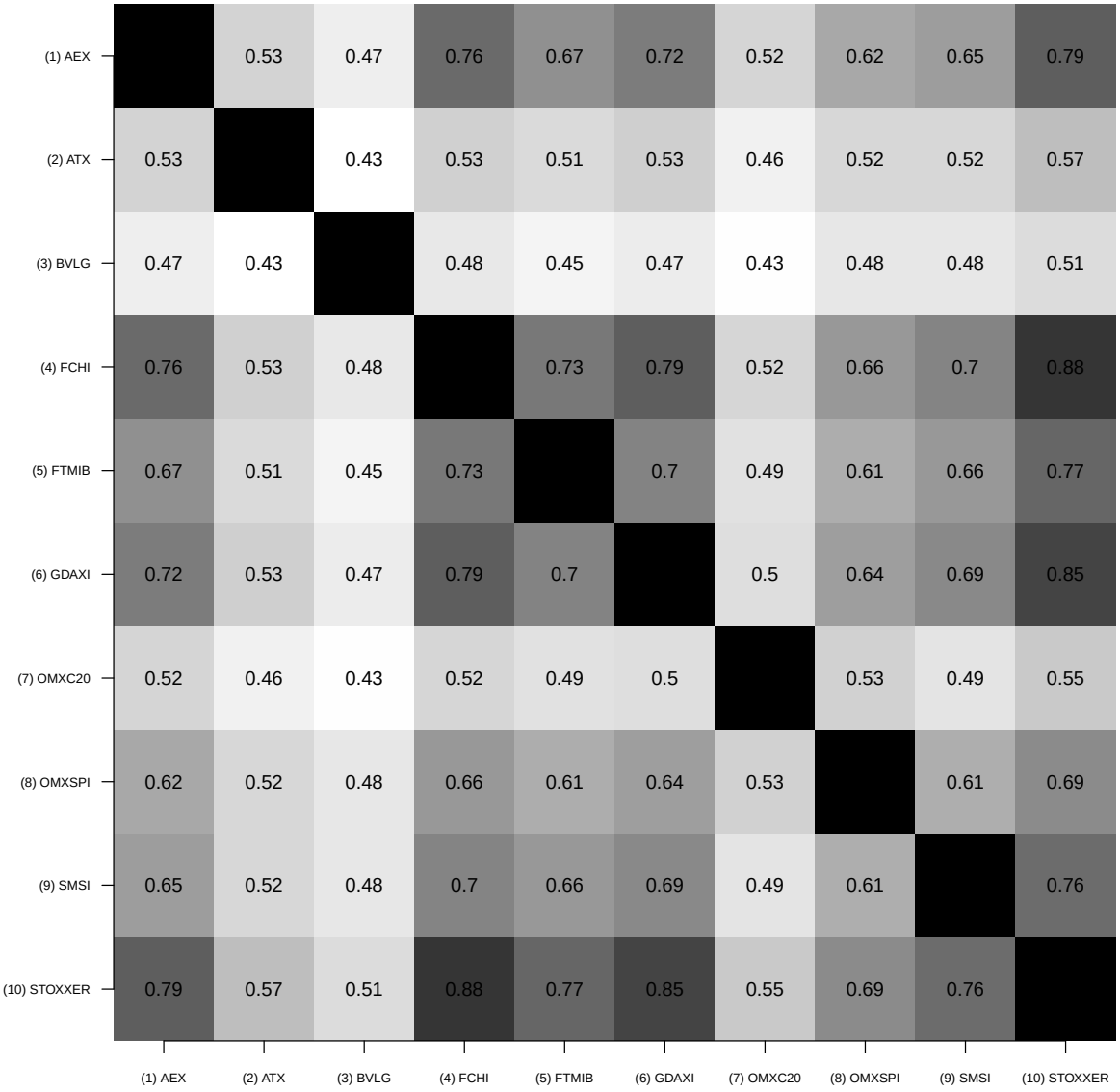


Figure 5.8: A D-vine for transformed standardized residuals with 10 variables and 9 trees; each edge may be associated with a pair-copula building block.

$$\mathbf{T_1}: \quad 2 \xrightarrow[C_{2,9}]{BB1} 9 \xrightarrow[C_{9,5}]{BB1} 5 \xrightarrow[C_{5,4}]{Student} 4 \xrightarrow[C_{4,10}]{Student} 10 \xrightarrow[C_{10,6}]{Student} 6 \xrightarrow[C_{6,1}]{Student} 1 \xrightarrow[C_{1,8}]{BB1} 8 \xrightarrow[C_{8,7}]{BB1} 7 \xrightarrow[C_{7,3}]{BB1} 3$$

$$\mathbf{T_2}: \quad 2,9 \xrightarrow[C_{2,5|9}]{Gauss} 9,5 \xrightarrow[C_{9,4|5}]{BB1} 5,4 \xrightarrow[C_{5,10|4}]{Student} 4,10 \xrightarrow[C_{4,6|10}]{Gauss} 10,6 \xrightarrow[C_{10,1|6}]{BB1} 6,1 \xrightarrow[C_{6,8|1}]{Gauss} 1,8 \xrightarrow[C_{1,7|8}]{Gauss} 8,7 \xrightarrow[C_{8,3|7}]{Gauss} 7,3$$

$$\mathbf{T_3}: \quad 2,5|9 \xrightarrow[C_{2,4|9,5}]{Gauss} 9,4|5 \xrightarrow[C_{9,10|5,4}]{BB1} 5,10|4 \xrightarrow[C_{5,6|4,10}]{Gauss} 4,6|10 \xrightarrow[C_{4,1|10,6}]{Indep.} 10,1|6 \xrightarrow[C_{10,8|6,1}]{Gauss} 6,8|1 \xrightarrow[C_{6,7|1,8}]{Indep.} 1,7|8 \xrightarrow[C_{1,3|8,7}]{Gauss} 8,3|7$$

$$\mathbf{T_4}: \quad 2,4|9,5 \xrightarrow[C_{2,10|9,5,4}]{Gauss} 9,10|5,4 \xrightarrow[C_{9,6|5,4,10}]{Gauss} 5,6|4,10 \xrightarrow[C_{5,1|4,10,6}]{Gauss} 4,1|10,6 \xrightarrow[C_{4,8|10,6,1}]{Gauss} 10,8|6,1 \xrightarrow[C_{10,7|6,1,8}]{Gauss} 6,7|1,8 \xrightarrow[C_{6,3|1,8,7}]{Gauss} 1,3|8,7$$

$$\mathbf{T_5}: \quad 2,10|9,5,4 \xrightarrow[C_{2,6|9,5,4,10}]{Gauss} 9,6|5,4,10 \xrightarrow[C_{9,1|5,4,10,6}]{Gauss} 5,1|4,10,6 \xrightarrow[C_{5,8|4,10,6,1}]{Gauss} 4,8|10,6,1 \xrightarrow[C_{4,7|10,6,1,8}]{Gauss} 10,7|6,1,8 \xrightarrow[C_{10,3|6,1,8,7}]{Gauss} 6,3|1,8,7$$

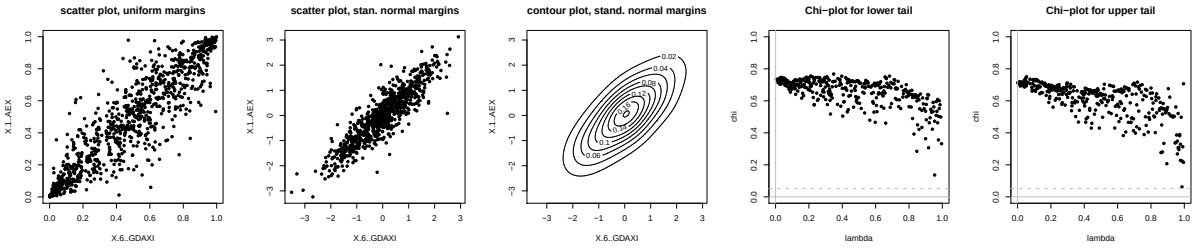
$$\mathbf{T_6}: \quad 2,6|9,5,4,10 \xrightarrow[C_{2,1|9,5,4,10,6}]{Indep.} 9,1|5,4,10,6 \xrightarrow[C_{9,8|5,4,10,6,1}]{Indep.} 5,8|4,10,6,1 \xrightarrow[C_{5,7|4,10,6,1,8}]{Indep.} 4,7|10,6,1,8 \xrightarrow[C_{4,3|10,6,1,8,7}]{Gauss} 10,3|6,1,8,7$$

$$\mathbf{T_7}: \quad 2,1|9,5,4,10,6 \xrightarrow[C_{2,8|9,5,4,10,6,1}]{Gauss} 9,8|5,4,10,6,1 \xrightarrow[C_{9,7|5,4,10,6,1,8}]{Gauss} 5,7|4,10,6,1,8 \xrightarrow[C_{5,3|4,10,6,1,8,7}]{Gauss} 4,3|10,6,1,8,7$$

$$\mathbf{T_8}: \quad 2,8|9,5,4,10,6,1 \xrightarrow[C_{2,7|9,5,4,10,6,1,8}]{Gauss} 9,7|5,4,10,6,1,8 \xrightarrow[C_{9,3|5,4,10,6,1,8,7}]{Indep.} 5,3|4,10,6,1,8,7$$

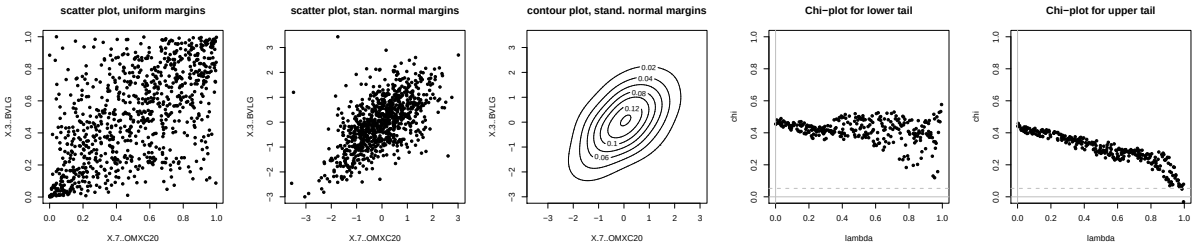
$$\mathbf{T_9}: \quad 2,7|9,5,4,10,6,1,8 \xrightarrow[C_{2,3|9,5,4,10,6,1,8,7}]{Gauss} 9,3|5,4,10,6,1,8,7$$

Figure 5.9: 1ST D-VINE TREE: choosing copula for building block $C_{6,1}$
GDAXI $\sim$ AEX



	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	***6***	4	-1	2	-3	1	-7	-2
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.62$		$\hat{\lambda}_U^{t} = 0.60$		$\hat{\lambda}_U^{BB1} = 0.69$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.65$		$\hat{\lambda}_L^{t} = 0.60$		$\hat{\lambda}_L^{BB1} = 0.68$			
Chi-plots	symmetrical strong tail dependences							
Kendall's $\hat{\tau}$	0.72***							
Chosen copula	*Student* with parameters $\hat{\rho} = 0.90$, $\hat{\nu} = 5.13$							

Figure 5.10: 1ST D-VINE TREE: choosing copula for building block $C_{7,3}$
OMXC20 $\sim$ BVLG



	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Vuong test	4	***5***	3	-2	-3	1	-1	-7
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.21$		$\hat{\lambda}_U^{t} = 0.30$		$\hat{\lambda}_U^{BB1} = 0.27$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.48$		$\hat{\lambda}_L^{t} = 0.30$		$\hat{\lambda}_L^{BB1} = 0.48$			
Chi-plots	not symmetrical tail dependences							
Kendall's $\hat{\tau}$	0.43***							
Choosen copula	*BB1* with parameters $\hat{\theta} = 0.75$, $\hat{\delta} = 1.26$							

Table 5.3: Estimated parameters for 10-dimensional D-vine: column SEQ contains estimations obtained sequentially and column MLE contains estimations obtained by Maximum-likelihood optimization.

block	family	par.	SEQ	MLE
$C_{2,9}$	BB1	$\hat{\theta}$	0.66	0.69
		$\hat{\delta}$	1.53	1.52
$C_{9,5}$	BB1	$\hat{\theta}$	0.73	0.76
		$\hat{\delta}$	2.10	2.08
$C_{5,4}$	Student	$\hat{\nu}$	4.69	5.07
		$\hat{r}$	0.91	0.91
$C_{4,10}$	Student	$\hat{\nu}$	6.91	7.10
		$\hat{r}$	0.98	0.98
$C_{10,6}$	Student	$\hat{\nu}$	5.63	6.74
		$\hat{r}$	0.97	0.97
$C_{6,1}$	Student	$\hat{\nu}$	5.13	7.15
		$\hat{r}$	0.90	0.91
$C_{1,8}$	BB1	$\hat{\theta}$	0.77	0.76
		$\hat{\delta}$	1.87	1.89
$C_{8,7}$	BB1	$\hat{\theta}$	0.81	0.76
		$\hat{\delta}$	1.46	1.49
$C_{7,3}$	BB1	$\hat{\theta}$	0.75	0.73
		$\hat{\delta}$	1.26	1.26
$C_{2,5\|9}$	Gauss	$\hat{r}$	0.26	0.26
$C_{9,4\|5}$	BB1	$\hat{\theta}$	0.37	0.40
		$\hat{\delta}$	1.30	1.29
$C_{5,10\|4}$	Student	$\hat{\nu}$	8.40	6.94
		$\hat{r}$	0.53	0.53
$C_{4,6\|10}$	Gauss	$\hat{r}$	−0.14	−0.17
$C_{10,1\|6}$	BB1	$\hat{\theta}$	0.49	0.50
		$\hat{\delta}$	1.37	1.37
$C_{6,8\|1}$	Gauss	$\hat{r}$	0.39	0.38
$C_{1,7\|8}$	Gauss	$\hat{r}$	0.28	0.28
$C_{8,3\|7}$	Gauss	$\hat{r}$	0.43	0.41

block	family	par.	SEQ	MLE
$C_{2,4\|9,5}$	Gauss	$\hat{r}$	0.15	0.16
$C_{9,10\|5,4}$	BB1	$\hat{\theta}$	0.31	0.31
		$\hat{\delta}$	1.36	1.35
$C_{5,6\|4,10}$	Gauss	$\hat{r}$	−0.15	p0.18
$C_{4,1\|10,6}$	Indep.			
$C_{10,8\|6,1}$	Gauss	$\hat{r}$	0.39	0.40
$C_{6,7\|1,8}$	Indep.			
$C_{1,3\|8,7}$	Gauss	$\hat{r}$	0.19	0.19
$C_{2,10\|9,5,4}$	Gauss	$\hat{r}$	0.38	0.37
$C_{9,6\|5,4,10}$	Gauss	$\hat{r}$	−0.28	−0.28
$C_{5,1\|4,10,6}$	Gauss	$\hat{r}$	−0.10	−0.11
$C_{4,8\|10,6,1}$	Gauss	$\hat{r}$	−0.10	−0.10
$C_{10,7\|6,1,8}$	Gauss	$\hat{r}$	0.19	0.18
$C_{6,3\|1,8,7}$	Gauss	$\hat{r}$	0.12	0.11
$C_{2,6\|9,5,4,10}$	Gauss	$\hat{r}$	−0.24	−0.20
$C_{9,1\|5,4,10,6}$	Gauss	$\hat{r}$	−0.27	−0.28
$C_{5,8\|4,10,6,1}$	Gauss	$\hat{r}$	−0.09	−0.09
$C_{4,7\|10,6,1,8}$	Gauss	$\hat{r}$	−0.11	−0.11
$C_{10,3\|6,1,8,7}$	Gauss	$\hat{r}$	0.19	0.19
$C_{2,1\|9,5,4,10,6}$	Indep.			
$C_{9,8\|5,4,10,6,1}$	Indep.			
$C_{5,7\|4,10,6,1,8}$	Indep.			
$C_{4,3\|10,6,1,8,7}$	Gauss	$\hat{r}$	−0.10	−0.10
$C_{2,8\|9,5,4,10,6,1}$	Gauss	$\hat{r}$	0.07	0.09
$C_{9,7\|5,4,10,6,1,8}$	Gauss	$\hat{r}$	−0.08	−0.06
$C_{5,3\|4,10,6,1,8,7}$	Gauss	$\hat{r}$	−0.08	−0.07
$C_{2,7\|9,5,4,10,6,1,8}$	Gauss	$\hat{r}$	0.09	0.08
$C_{9,3\|5,4,10,6,1,8,7}$	Indep.	$\hat{r}$		
$C_{2,3\|9,5,4,10,6,1,8,7}$	Gauss	$\hat{r}$	0.09	0.09

Table 5.4: Validation results by comparison some descriptive statistics of observed transformed standardized residuals with descriptive statistics of the data sets sampled from the MLE D-vine model

index	statistic	(9) SMSI		(5) FTMIB		(4) FCHI		(10) STOXXER		(6) GDAXI		(1) AEX		(8) OMXSPI		(7) OMXC20		(3) BVLG	
		obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.
	$\hat{\tau}$	*0.52*	*0.51*	0.51	0.50	0.53	0.52	0.57	0.56	0.53	0.52	0.53	0.53	0.52	0.50	0.46	0.43	0.43	0.42
(2) ATX	$\hat{\lambda}_U$	*0.41*	*0.46*	0.41	0.43	0.41	0.45	0.44	0.50	0.34	0.45	0.34	0.45	0.31	0.43	0.21	0.34	0.27	0.30
	$\hat{\lambda}_L$	*0.38*	*0.54*	0.41	0.50	0.44	0.52	0.51	0.57	0.38	0.51	0.44	0.51	0.58	0.50	0.55	0.43	0.41	0.42
	$\hat{\tau}$			*0.66*	*0.65*	0.70	0.70	0.76	0.75	0.69	0.68	0.65	0.64	0.61	0.61	0.49	0.47	0.48	0.46
(9) SMSI	$\hat{\lambda}_U$			*0.51*	*0.63*	0.58	0.68	0.62	0.74	0.41	0.65	0.41	0.60	0.38	0.56	0.21	0.39	0.34	0.34
	$\hat{\lambda}_L$			*0.62*	*0.65*	0.68	0.69	0.75	0.75	0.51	0.65	0.65	0.60	0.51	0.59	0.48	0.48	0.38	0.46
	$\hat{\tau}$					*0.73*	*0.73*	0.77	0.77	0.70	0.70	0.67	0.67	0.61	0.63	0.49	0.49	0.45	0.45
(5) FTMIB	$\hat{\lambda}_U$					*0.68*	*0.70*	0.72	0.75	0.62	0.66	0.65	0.62	0.55	0.58	0.27	0.40	0.31	0.33
	$\hat{\lambda}_L$					*0.62*	*0.67*	0.68	0.72	0.55	0.64	0.58	0.6	0.51	0.59	0.44	0.47	0.34	0.43
	$\hat{\tau}$							*0.88*	*0.88*	0.79	0.79	0.76	0.76	0.66	0.65	0.52	0.49	0.48	0.47
(4) FCHI	$\hat{\lambda}_U$							*0.92*	*0.87*	0.75	0.75	0.58	0.71	0.58	0.61	0.27	0.41	0.34	0.35
	$\hat{\lambda}_L$							*0.82*	*0.84*	0.75	0.73	0.79	0.69	0.65	0.62	0.48	0.49	0.41	0.46
	$\hat{\tau}$									*0.85*	*0.84*	0.79	0.78	0.69	0.68	0.55	0.52	0.51	0.49
(10) STOXXER	$\hat{\lambda}_U$									*0.79*	*0.83*	0.58	0.75	0.58	0.64	0.27	0.43	0.34	0.37
	$\hat{\lambda}_L$									*0.68*	*0.80*	0.82	0.73	0.65	0.66	0.58	0.53	0.41	0.49
	$\hat{\tau}$											*0.72*	*0.73*	0.64	0.63	0.50	0.47	0.47	0.46
(6) GDAXI	$\hat{\lambda}_U$											*0.62*	*0.68*	0.55	0.58	0.17	0.39	0.34	0.34
	$\hat{\lambda}_L$											*0.65*	*0.66*	0.58	0.60	0.44	0.47	0.34	0.45
	$\hat{\tau}$													*0.62*	*0.62*	0.52	0.50	0.47	0.45
(1) AEX	$\hat{\lambda}_U$													*0.58*	*0.59*	0.27	0.41	0.31	0.33
	$\hat{\lambda}_L$													*0.62*	*0.62*	0.48	0.51	0.38	0.46
	$\hat{\tau}$															*0.53*	*0.51*	0.48	0.46
(8) OMXSPI	$\hat{\lambda}_U$															*0.31*	*0.44*	0.27	0.34
	$\hat{\lambda}_L$															*0.58*	*0.55*	0.41	0.48
	$\hat{\tau}$																	*0.43*	*0.42*
(7) OMXC20	$\hat{\lambda}_U$																	*0.21*	*0.31*
	$\hat{\lambda}_L$																	*0.48*	*0.48*

Figure 5.11: Sets of simulated future returns calculated using one step ahead forecasting method: simulated data are plotted in gray color and true data in black color

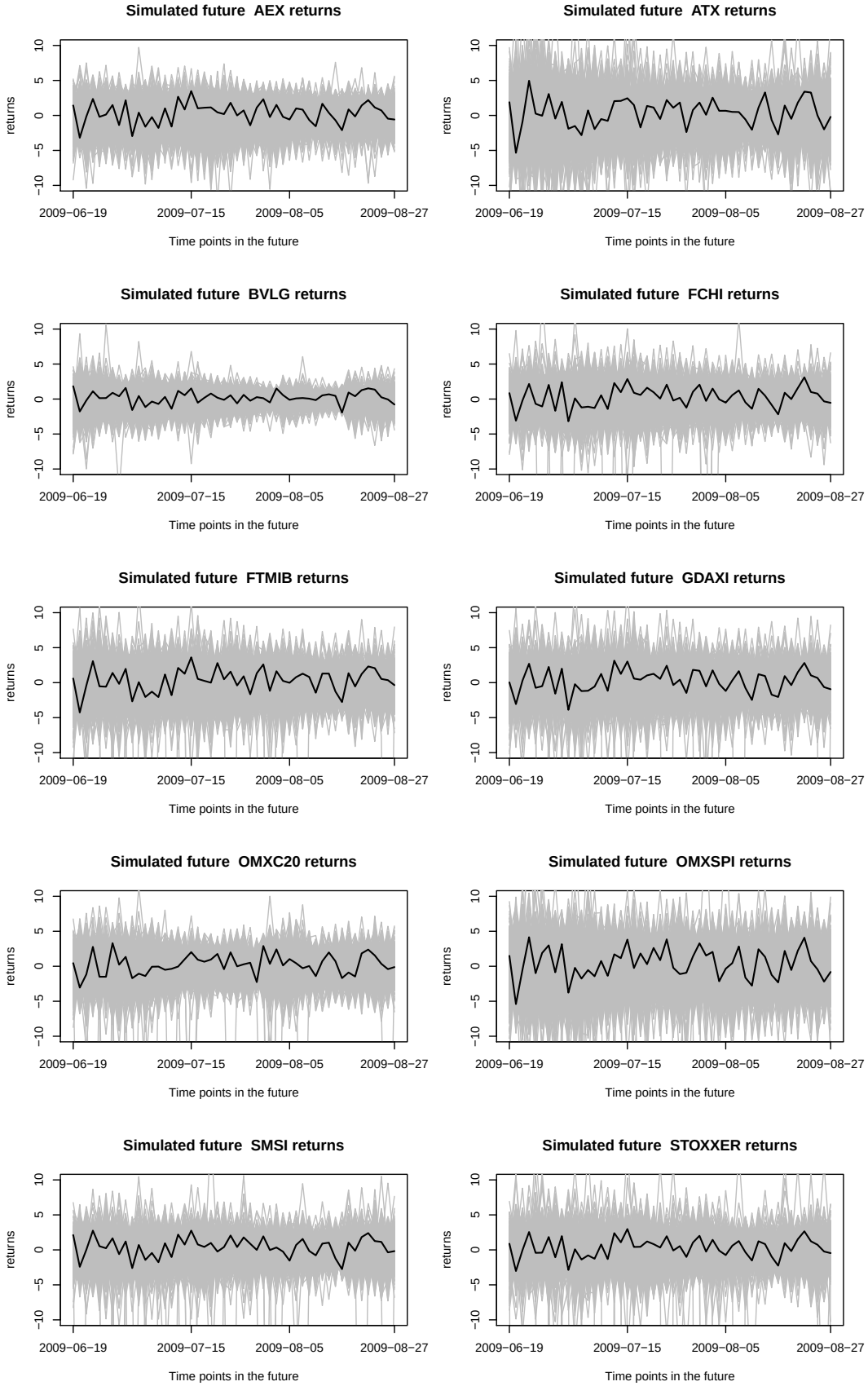


Table 5.5: Validation time series properties of simulated future returns. We compare the descriptive statistics (sample mean, sample standard deviation, sample kurtosis and sample skewness) of the observed test data to its average values over all of the scenarios

index	data	mean	st.deviation	kurtosis	skewness
ATX	observed	0.42	1.96	0.17	-0.32
	simulated	0.00	1.99	0.78	-0.21
SMSI	observed	0.40	1.33	-0.27	-0.43
	simulated	-0.01	1.41	0.38	-0.17
FTMIB	observed	0.31	1.63	-0.04	-0.43
	simulated	0.00	1.70	0.39	-0.25
FCHI	observed	0.27	1.44	-0.40	-0.24
	simulated	0.00	1.48	0.34	-0.19
STOXXER	observed	0.30	1.37	-0.30	-0.28
	simulated	0.00	1.41	0.82	-0.14
GDAXI	observed	0.25	1.59	-0.38	-0.27
	simulated	0.00	1.63	0.57	-0.22
AEX	observed	0.32	1.44	-0.25	-0.32
	simulated	0.01	1.53	0.23	-0.34
OMXSPI	observed	0.43	2.13	-0.36	-0.23
	simulated	-0.01	2.19	0.37	-0.26
OMXC20	observed	0.27	1.42	-0.52	0.04
	simulated	-0.01	1.50	0.34	-0.24
BVLG	observed	0.21	0.87	-0.06	-0.43
	simulated	0.00	0.99	0.61	-0.30

Table 5.6: Validation results by comparison the dependence structure within true and simulated future returns

index	statistic	ATX		BVLG		FCHI		FTMIB		GDAXI		OMXC20		OMXSPI		SMSI		STOXXER	
		obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.
AEX	$\hat{\tau}$	0.63	0.51	0.60	0.49	0.84	0.52	0.72	0.56	0.78	0.52	0.47	0.52	0.70	0.5	0.72	0.43	0.85	0.41
	$\hat{\lambda}_U$	0.57	0.60	0.57	0.58	0.99	0.61	0.85	0.65	0.85	0.61	0.71	0.61	0.85	0.59	0.85	0.51	0.99	0.49
	$\hat{\lambda}_L$	0.71	0.57	0.85	0.55	0.85	0.58	0.85	0.61	0.57	0.57	0.42	0.56	0.71	0.55	0.71	0.5	0.99	0.49
ATX	$\hat{\tau}$			0.55	0.65	0.60	0.70	0.62	0.75	0.59	0.68	0.56	0.64	0.56	0.60	0.56	0.46	0.63	0.46
	$\hat{\lambda}_U$			0.57	0.75	0.71	0.80	0.57	0.84	0.71	0.77	0.99	0.74	0.71	0.69	0.57	0.55	0.71	0.54
	$\hat{\lambda}_L$			0.57	0.67	0.57	0.71	0.71	0.76	0.57	0.69	0.28	0.65	0.71	0.63	0.57	0.53	0.71	0.52
BVLG	$\hat{\tau}$					0.58	0.72	0.57	0.76	0.54	0.69	0.42	0.66	0.46	0.62	0.55	0.48	0.61	0.44
	$\hat{\lambda}_U$					0.57	0.81	0.71	0.85	0.57	0.78	0.42	0.75	0.42	0.71	0.71	0.57	0.57	0.52
	$\hat{\lambda}_L$					0.71	0.70	0.85	0.73	0.42	0.67	0.42	0.65	0.57	0.63	0.71	0.53	0.85	0.5
FCHI	$\hat{\tau}$							0.76	0.88	0.85	0.79	0.51	0.75	0.69	0.65	0.68	0.49	0.91	0.47
	$\hat{\lambda}_U$							0.85	0.99	0.99	0.89	0.85	0.84	0.99	0.74	0.85	0.58	0.78	0.54
	$\hat{\lambda}_L$							0.71	0.86	0.71	0.77	0.42	0.73	0.57	0.67	0.57	0.55	0.85	0.53
FTMIB	$\hat{\tau}$									0.74	0.84	0.49	0.78	0.64	0.68	0.61	0.52	0.79	0.49
	$\hat{\lambda}_U$									0.71	0.95	0.57	0.87	0.71	0.77	0.71	0.6	0.85	0.56
	$\hat{\lambda}_L$									0.57	0.82	0.42	0.76	0.57	0.70	0.71	0.57	0.85	0.55
GDAXI	$\hat{\tau}$											0.47	0.72	0.69	0.63	0.65	0.47	0.87	0.46
	$\hat{\lambda}_U$											0.71	0.82	0.85	0.73	0.71	0.56	0.99	0.54
	$\hat{\lambda}_L$											0.42	0.70	0.57	0.65	0.57	0.53	0.57	0.52
OMXC20	$\hat{\tau}$													0.48	0.62	0.42	0.49	0.51	0.45
	$\hat{\lambda}_U$													0.85	0.71	0.71	0.58	0.85	0.52
	$\hat{\lambda}_L$													0.42	0.65	0.42	0.55	0.42	0.52
OMXSPI	$\hat{\tau}$															0.58	0.51	0.72	0.46
	$\hat{\lambda}_U$															0.71	0.59	0.99	0.54
	$\hat{\lambda}_L$															0.57	0.58	0.71	0.53
SMSI	$\hat{\tau}$																	0.73	0.42
	$\hat{\lambda}_U$																	0.85	0.48
	$\hat{\lambda}_L$																	0.71	0.52

Conclusion

In this thesis, we modeled and forecasted the multivariate return series of 10 European stock market indices using Time Series Pair-Copula approach. At first, we applied the (ARMA-) GARCH filter with the standard skewed Students's t innovations to get the i.i.d. sequences of the univariate returns. We saw that the simple GARCH(1,1) models fit data as good as the extended AR-GARCH, MA-GARCH or ARMA-GARCH models. Next, we transformed the standardized residuals of the GARCH models to unit intervals by applying the cumulative distribution function of the standard skewed Student's t distribution. The Probability Integral Transformation Theorem confirms this methodical approach. At last, we used the transformed residuals to estimate parameters of the pair-copula construction, the structure of which was determined by the 10-dimensional D-vine.

In this thesis, we also studied the distribution free Vuong and Clarke tests as tool of bivariate goodness-of-fit testing for copulas. We satisfied the suitability of these tests for copula selection in the simulation study which were executed for three different choices of Kendall's tau, for large and small sample sizes separately. We saw that the error rate of the Clarke test is much more higher than the error rate of the Vuong test. Moreover, the accuracy of the both tests depends on the sample length and the association degree measured in Kendal's τ. Generally, the correctness of the Vuong as well as of the Clarke test increases when Kendall's τ and sample length rise. Additionally, we assessed that results of the both tests depend on the KLIC distance between two copulas to be compared. This relationship is for the Vuong test stronger than for the Clarke test. It can be explained by the fact that the Vuong statistic is based on the KLIC distance, whereas the Clarke statistic is based on the transformed KLIC distance. Finally, we compared our tests with the common goodness-of-fit tests presented by Genest, Rémillard, and Beaudoin (2009) and found out that the Vuong test is more efficient.

Further, we generalized bivariate Vuong and Clarke tests to enable the comparison of more than two models. We call this extension score test or scoring. We validated our approach by simulation study which we performed for three different choices of Kendall's τ, for large and small sample sizes separately. The scoring accuracy depends on the sample size as well as on the Kendall's τ. The error rate of the score test based on the Vuong statistic is much more lower than the error rate of the score test based on the Clarke statistic. Moreover, the sets of copulas to be compared impact the results of the score test strongly. Hence, we recommend to leave out similar copulas classes before starting the selecting

procedure. For example, if there are no significant distinctions between modeling the data with Gaussian or with Frank copula, one of them should be leaved out.

For choosing an appropriate copula for each building block of the D-vine, we combine following techniques: scatter plots and empirical contour plots of the transformed standardized residuals with the uniform and standard normal margins; score test based on the Vuong statistic with Schwarz's correction; the empirical Kendall's τ and empirical coefficients of lower and upper tail dependences as well as coefficients of lower and upper tail dependences of estimated Student and BB1 copulas. As it is well-known, the empirical estimators of the tail dependence coefficients are not stable, so we were interested on the ratio between λ_U and λ_L but not on its absolute values. The combination of these methodologies turns out to be a sufficient tool for copula selection.

Finally, we simulated future returns of each stock market index by applying the TIME SERIES PAIR-COPULA APPROACH backwards. Firstly, we sampled data from the estimated D-vine and transformed this data to the quasi error sequences by applying the inverse of the cumulative distribution function. We estimated the D-vine parameters in two ways: sequentially (SEQ) and by using the MLE optimization procedure (MLE). Based on the Voung test, we ascertained that the both models were equivalent to each other and the both fitted the data better than a simple D-vine with all Gaussian blocks. The sequences of error terms obtained in a such way incorporate the dependence structure within the future data. Secondly, we forecasted future returns using parameters of the GARCH models and simulated error terms. In conclusion, we validated our results and found out that scenarios of the future returns reproduce the times series properties as well as the dependences within the true data quite well.

References

Aas, K., C. Czado, A. Frigessi, and H. Bakken (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics 44* (2), 182–198.

Abberger, K. (2004). A simple graphical method to explore tail-dependence in stock-return pairs. Working paper, University of Konstanz, Germany.

Almeida, D. C. and U. Schepsmeier (2010). *CDVineMLE: MLE Estimator and a few other useful functions for a PCC specification (D-Vine or canonical vine)*. R package version 1.0.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics 12*, 171–178.

Azzalini, A. (1986). Further results on a class of distributions which includes the normal one. *Statistica 46*, 199–208.

Bedford, T. and R. M. Cooke (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence 32*, 245–268.

Bedford, T. and R. M. Cooke (2002). Vines – a new graphical model for dependent random variables. *The Annals of Statistics 30(4)*, 1031–1068.

Black, F. (1976). Studies of stock price volatility changes. *Proceeding of the American Statistical Association*, 177–181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics 31*, 307–327.

Bollerslev, T. (1987). A conditionally heteroscedastic time series model for speculative prices and rates of returns. *Review of Econometrics and Statistics 69*, 542–547.

Box, G. E. P. and D. Pierce (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of American Statistical Associations 65*, 1509–1526.

Brechmann, E. (2010). *Truncated Canonical Vine Copulas*. Diploma thesis, Zentrum Mathematik Technische Universität München.

Brockwell, P. J. and R. A. Davis (1991). *Time Series: Theory and Methods* (2nd ed.). Springer Series in Statistics. New York: Springer.

Chatfield, C. (2004). *The Analysis of Time Series: An Introduction* (6th ed.). Texts in Statistical Science. Chapman & Hall/CRC.

Clarke, K. A. (2003). Nonparametric model discrimination in international relations. *Journal of Conflict Resolution 47(1)*, 72–93.

Clarke, K. A. (2007). A simple distribution-free test for nonnested model selection. *Political Analysis 15*, 347–363.

Cryer, J. D. and K.-S. Chan (2008). *Time Series Analysis with Applications in R* (2nd ed.). Springer Texts in Statistics. Springer.

Dahl, D. B. (2009). *xtable: Export tables to LaTeX or HTML*. R package version 1.5-5.

Duffee, G. R. (1995). Stock returns and volatility: A firm-level analysis. *Journal of Financial Economics 52*, 399–420.

Embrechts, P., F. Lindskog, and A. McNeil (2001). Modelling dependence with copulas and applications to risk management. Working paper, Department of Mathematics, ETH Zürich, Switzerland.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of u.k. inflation. *Econometrica 50*, 987–1007.

Fernandez, C. and M. Steel (1998). On bayesian modeling of fat tails and skewness. *Journal of American Statistical Association 93*, 359–371.

Fisher, N. I. and P. Switzer (1985). Chi-plots for assessing dependence. *Biometrika 72*, 253–265.

Fisher, N. I. and P. Switzer (2001). Graphical assessment of dependence: is a picture worth 100 test? *The American Statistician 55*, 233–239.

for S-Plus original; R port by Scott Ulman, A. M. (2008). *QRMlib: Provides R-language code to examine Quantitative Risk Management concepts*. R package version 1.4.4.

Francq, C. and J. M. Zakoïan (2004). Maximum likelihood estimation of pure garch and arma-garch processes. *Bernuolli 10*, 605–637.

Garcia, J. and V. Gonzalez-Lopez (2006). *mlCopulaSelection: Copula selection and fitting using maximum likelihood*. R package version 1.3.

Gelman, A., J. Hill, M. Yajima, Y.-S. Su, and M. G. Pittau (2009). *mi: Missing Data Imputation and Model Checking*. R package version 0.08-04.02.

Genest, C. and J.-C. Boies (2003). Detecting dependence with kendall plots. *The American Statistician 57*, 275–284.

Genest, C. and A.-C. Favre (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering 12*, 347–368.

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

Ghost, S. and S. Henderson (2002). Properties of the notra method in higher dimensions. Proceeding of the 2002 winter simulation conference, San Diego, USA.

Gonzalez-Lopez, V. A. (2009). *fgac: Generalized Archimedean Copula*. R package version 0.6-1.

Graves, S. (2009). *FinTS: Companion to Tsay (2005) Analysis of Financial Time Series*. R package version 0.4-3.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton, New Jersey: Princeton University Press.

Hansen, B. E. (1994). Autoregressive conditional density estimation. *International Economic Review 35*, 705–730.

Iman, R. and J. Helton (1985). A comparison of uncertainty and sensitivity analysis techniques for computer models. Technical report, nureg/cr-3904 sand84-1461 rg, Sandia National Laboratories, Albuquerque, New Mexico, USA.

Jarque, C. M. and A. K. Bera (1987). A test for normality of observations and regression residuals. *International Statistical Review 55*(2), 163–172.

Joe, H. (1997). *Multivariate Models and Dependence Concept*. Chapman & Hall/CRC.

Johnson, R. A. and D. W. Wichern (2007). *Applied Multivariate Statistical Analysis* (6th ed.). Pearson Education International.

Jondeau, E., S.-H. Poon, and M. Rockinger (2007). *Financial Modeling Under Non-Gaussian Distributions*. Springer Finance Textbook. Springer.

Kotz, S. and S. Nadarajah (2004). *Multivariate t-Distributions and Their Applications*. Cambridge University Press.

Kullback, S. and R. Leibler (1951). On information and sufficiencys. *Annals of Mathematical Statistics 2*, 79–86.

Kurowicka, D. and R. M. Cooke (2006). *Uncertainty Analysis with High Dimensional Dependence Modelling*. Wiley Series in Probability and Statistics. Wiley & Sons.

Lai, T. L. and H. Xing (2008). *Statistical Models and Methods for Financial Markets*. Springer Texts in Statistics. Springer.

Ling, S. and W. K. Li (1997). On fractionally integrated autoregressive moving-average time series model with conditional heteroscedasticity. *Journal of the American Statistical Association 92*, 1184–1194.

Ling, S. and W. K. Li (1998). Limiting distributions of maximum likelihood estimations for unstable arma models with garch errors. *Annals of Statistics 26*, 84–125.

Ling, S. and M. McAleer (2003). Asymptotic theory for a vector arma-garch model. *Econometric Theory 19*, 280–310.

Ljung, G. and G. E. P. Box (1978). On a measure of lack of fit in time series models. *Biometrika 66*, 67–72.

Longin, F. and B. Solnik (1995). Is the correlation in international equity returns constant: 1960-1990? *Journal of International Money and Finance 14*, 3–26.

Longin, F. and B. Solnik (2001). Extreme correlation of international equity markets. *Journal of Finance 56*, 649–676.

Malevergne, Y. and D. Sornette (2006). *Extrem Financial Risks: From Dependence to Risk Management*. Springer.

Nelsen, R. B. (2006). *An Introduction to Copulas* (2nd ed.). Springer Series in Statistics. Springer.

Quenoulle, M. H. (1949). Approximate tests of corelation in time series. *Journal of the Royal Statistical Society B 11*, 68–84.

R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Schepsmeier, U. (2010). *Maximum likelihood estimation of C-vine pair-copula constructions based on bivariate copulas from different families*. Diploma thesis, Zentrum Mathematik Technische Universität München.

Schoffer, O. (2003). *Modellierung von Kapitalmarktrenditen mittels asymmetrischer GARCH-Modelle*. Dissertation, Fachbereich Statistik, Universität Dortmund.

Shapiro, S. and M. Wilk (1965). An analysis of variance test for normality (complete samples). *Biometrika 52*(3–4), 591–611.

Shumway, R. H. and D. S. Stoffer (2006). *Time Series Analysis and Its Applications with R Examples* (2nd ed.). Springer Texts in Statistics. Springer.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Punlications de l'Institut de Statistique de l'Université de Paris 8*, 229–231.

Straumann, D. (2005). *Estimation in Conditionally Heteroscedastic Time Series Models*, Volume 181 of *Lecture Notes in Statistics*. Springer.

Trapletti, A. and K. Hornik (2009). *tseries: Time Series Analysis and Computational Finance*. R package version 0.10-21.

Tsay, R. S. (2005). *Analysis of Financial Time Series* (2nd ed.). Wiley Series in Probability and Statistics. Wiley-Interscience.

Čížek, P., W. Härdle, and R. Weron (2005). *Statistical Tools for Finance and Insurance*. Springer.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica 57(2)*, 307–333.

Wuertz, D. and Y. Chalabi (2009). *timeSeries: Rmetrics - Financial Time Series Objects*. R package version 2100.84.

Wuertz, D., Y. C. with contribution from Michal Miklovic, C. Boudt, P. Chausse, et al. (2009). *fGarch: Rmetrics - Autoregressive Conditional Heteroskedastic Modelling*. R package version 2110.80.

Wuertz, D., Y. C. with contributions from Martin Maechler, J. W. Byers, , et al. (2009). *timeDate: Rmetrics - Chronological and Calendarical Objects*. R package version 2100.86.

Yan, J. (2007). Enjoy the joy of copulas: with a package copula. *Journal of Statistical Software 21*(4), 1–21.

Appendix A

Notations

notation	definition/meaning
X	random variable
x	realization of random variable X
$\boldsymbol{X}$	$\boldsymbol{X} = (X_1, \ldots, X_n)$ random vector
$\boldsymbol{x}$	$\boldsymbol{x} = (x_1, \ldots, x_n)$ realization of random variable X
$\boldsymbol{\mu}$	mean vector of $\boldsymbol{X}$
$\boldsymbol{\Sigma}$	variance-covariance matrix of $\boldsymbol{X}$
n	length of sample size, length of random vector
d	dimension
t, t_i	time point index
$f(\cdot)$	density function
$F(\cdot)$	distribution function
$f_X(\cdot)$	marginal density of X
$f_Y(\cdot)$	marginal density of Y
$f_{XY}(\cdot)$	joint density function of random variables X and Y
$F_X(\cdot)$	marginal distribution function of X
$F_Y(\cdot)$	marginal distribution function of Y
$F_{XY}(\cdot)$	joint distribution function of random variables X and Y
$R(x_i)$	rank of X_i in vector $\boldsymbol{X}$
$C_{\boldsymbol{\theta}}(\cdot, \cdot)$	copula with parameter vector $\boldsymbol{\theta}$
$c_{\boldsymbol{\theta}}(\cdot, \cdot)$	copula density with parameter vector $\boldsymbol{\theta}$
$h_{\boldsymbol{\theta}}(\cdot, \cdot)$	h-function wit parameter vector $\boldsymbol{\theta}$
$x_i \vert x_j$	value of x_i given known value of x_j
$f(x_i \vert x_j)$	conditional density of X_i given X_j
$F(x_i \vert x_j)$	conditional distribution function of X_i given X_j
$E[\cdot]$, μ	expectation
$Var(\cdot)$, σ^2	variance
σ	standard deviation
$\bar{x}$	sample mean
γ	skewness
κ	kurtosis
h	distance between two time points

$\gamma(\cdot,\cdot)$	autocovariance function
$\rho(\cdot,\cdot)$	autocorrelation function
$\gamma(\cdot)$	autocovariance function of stationary time series
$\rho(\cdot)$	autocorrelation function of stationary time series
r	Pearson's correlation, coefficient of linear dependence
ρ	Spearman's coefficient of monotonic dependence
τ	Kendall's coefficient of monotonic dependence
λ_U	coefficient of upper tail dependence
λ_L	coefficient of lower tail dependence
$\hat{\cdot}$	estimation of $\cdot$
$\{\cdot\}$	stochastic process or sequence of $\cdot$ values
$\{\varepsilon_t\}$	white noise sequence
i.i.d.	identical independent distributed
$AR(p)$	autoregressive model of order p
$MA(q)$	moving average model of order q
$ARMA(p,q)$	autoregressive moving average model of order (p,q)
$ARCH(m)$	autoregressive conditionally heteroscedastic model of order m
$GARCH(m,r)$	generalized autoregressive conditionally heteroscedastic model of order (m,r)
$\hat{\varepsilon}_{t,raw}$	raw residuals
$\hat{\varepsilon}_t$	standardizes residuals
W	Shapiro-Wilk test statistic
χ^2	Jarque-Bera test statistic
$Q(\cdot)$	Ljung-Box test statistic
$Q^2(\cdot)$	Ljung-Box test statistic for squared values
TR^2	Engle's LM ARCH test statistic
$X \sim \cdot$	X follows distribution $\cdot$
$\mathcal{U}[a,b]$	uniform distribution on $[a,b]$
$\mathcal{N}(0,1)$	standard normal distribution
$\Phi(\cdot)$	distribution function of standard normal distribution
$\mathcal{N}(\mu,\sigma^2)$	normal distribution with parameters μ and σ^2
$\mathcal{SN}(\xi,\omega^2,\lambda)$	skewed normal distribution with parameters ξ, ω^2 and λ
$\mathcal{SN}(\xi,\omega^2,\lambda)^{stand}$	standard skewed normal distribution with parameters ξ, ω^2 and λ
t_ν	Student's t distribution with parameter ν
t_ν^{stand}	standard Student's t distribution with parameter ν
$t_{\nu,\lambda}$	skewed Student's t distribution with parameters ν and λ
$t_{\nu,\lambda}^{stand}$	standard skewed Student's t distribution with parameters ν and λ
$\mathcal{N}_d(\boldsymbol{\mu},\boldsymbol{\Sigma})$	multivariate normal distribution
$t_{\nu,\boldsymbol{\mu},\boldsymbol{\Sigma}}$	multivariate Student's t distribution
$\Gamma(\cdot)$	Gamma function
$B(\cdot,\cdot)$	Beta function
$D_k(\cdot)$	k-th Debye function
$\mathbb{R}$	real numbers
$KLIC$	Kullback-Leibler Information Criterion

Appendix B

R packages

`copula`	Fitting, sampling and log-likelihood estimation of some most common parametric copula families as Gauss, Student, Clayton, Gumbel, Plackett and Frank. Note, fitting of the Student copula is not stable. **Reference**: Yan (2007)
`timeSeries`	Handling with time series, evaluate returns. **Reference**: Wuertz and Chalabi (2009).
`mi`	Handling with missing values. **Reference**: Gelman, Hill, Yajima, Su, and Pittau (2009).
`fGarch`	Fitting and model checking of (ARMA-) GARCH models, ACF and PACF. Reference: Wuertz, with contribution from Michal Miklovic, Boudt, Chausse, et al. (2009).
`mlCopulaSelection`	Log-likelihood estimation of BB1 and BB7 copulas. **Reference**: Garcia and Gonzalez-Lopez (2006).
`fgac`	Log-likelihood estimation of BB1 and BB7 copulas. **Reference**: Gonzalez-Lopez (2009).
`CDVineMLE`	Maximum-Likelihoods estimation of D- and C-Vine (intern package of TU München). **Reference**: Almeida and Schepsmeier (2010).
`xtable`	Functions for converting an R table in a LaTex code. **Reference**: Dahl (2009)
`timeDate`	Handling with time series, kurtosis and skewness. **Reference**: Wuertz, with contributions from Martin Maechler, Byers, , et al. (2009).
`FinTS`	Handling with time series, Lagrange Multiplier (LM) test for autoregressive conditional heteroscedasticity (ARCH) that will be used in checking of (ARMA-) GARCH models. **Reference**: Graves (2009).
`QRMlib`	Fitting, sampling and log-likelihood functions of some most common parametric copula families. Note, fitting of the Student copula is more efficient as in the package `copula`. **Reference**: for S-Plus original; R port by Scott Ulman (2008).

Appendix C

Statistical distributions used

C.1 Univariate distributions

C.1.1 Normal distribution

Definition C.1 *A continuous random variable X is said to have a **Normal distribution**, denoted by $X \sim \mathcal{N}(\mu, \sigma^2)$, with location parameter $\mu \in \mathbb{R}$ and scale parameter $\sigma > 0$ if its density function has a form*

$$f_{\mathcal{N}(\mu,\sigma^2)}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad for \quad x \in \mathbb{R}\,.$$

For expectation and variance of $X \sim \mathcal{N}(\mu, \sigma^2)$, it holds

$$E[X] = \mu \quad \text{and} \quad \text{Var}(X) = \sigma^2\,.$$

The normal distribution with parameters $\mu = 0$ and $\sigma = 1$ is called *standard normal distribution* and will be defined by $\mathcal{N}(0, 1)$. It has a density function

$$f_{\mathcal{N}(0,1)}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,.$$

C.1.2 Skewed normal distribution

The *skewed normal distribution* was introduces by Azzalini (1985, 1986). It belongs to a parametric class of probability distributions and extends the normal distribution by an additional parameter λ that regulates the skewness, allowing for a continuous variation from normality to non-normality.

Definition C.2 *A continuous random variable X is said to follow a **skewed normal distribution**, denoted by $X \sim \mathcal{SN}(\xi, \omega^2, \lambda)$, with location parameter $\xi \in \mathbb{R}$, scale parameter $\omega > 0$ and skew parameter $\lambda \in \mathbb{R}$ if its density function can be expressed as*

$$f_{\mathcal{SN}(\xi,\omega^2,\lambda)}(x) = \frac{1}{\sqrt{\omega\pi}} e^{-\frac{(x-\xi)^2}{2\omega^2}} \int_{-\infty}^{\lambda\left(\frac{x-\xi}{\omega}\right)} e^{-\frac{-t^2}{2}}\, dt \quad for \quad x \in \mathbb{R}\,.$$

For expectation and variance of $X \sim \mathcal{SN}(\xi, \omega^2, \lambda)$, it holds

$$E[X] = \xi + \omega\lambda\sqrt{\frac{2}{\pi(1+\lambda^2)}} \quad \text{and} \quad Var(X) = \omega^2\left(1 - \frac{2\lambda^2}{\pi(1+\lambda^2)}\right) .$$

The skewed normal distribution is called *standard skewed normal* if $E[X] = 0$ and $\mathrm{Var}(x) = 1$ and will be denoted by $\mathcal{SN}(\xi, \omega^2, \lambda)^{stand}$. Figure C.1 (panel (a)) give an impression about differences between standard normal and skewed standard normal distribution plotted for three values of skew parameter, $\lambda \in \{\ 1.5, 0.5, 1.1\}$. As we ca see, the density curve is shifted to the left for $\lambda > 1$ and to the right for $\lambda < 1$, and for $\lambda \to 1$ it gets a symmetric form.

C.1.3 Student's t distribution

In some practical situations, observed data has some additional properties that can not be captured by the normal distribution. The Student's t distribution is more suitable for modeling such data. This distribution has heavier tails, i.e. it allows a higher probability for extreme values than the normal distribution.

Definition C.3 *A continuous random variable X is said to have a **Student's t distribution**, denoted by $X \sim t_\nu$, with the number of degrees of freedom $\nu > 0$ if its density function can be written in form*

$$f_{t_\nu}(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}(\nu\pi)^{-\frac{1}{2}}\left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} \quad \textit{for} \quad x \in \mathbb{R} ,$$

where $\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}dt$ is a Gamma function.

Parameter ν is also called a shape parameter. For $\nu \to \infty$ the t distribution collapses to the normal case. For expectation and variance of $X \sim t_\nu$, it holds

$$E[X] = 0 \text{ for } \nu > 0 \quad \text{and} \quad Var(X) = \frac{\nu}{\nu - 2} \text{ for } \nu > 2 .$$

Let $X \sim t_\nu$, $\nu > 2$, then $\tilde{X} = X\sqrt{\frac{\nu-2}{\nu}}$ has a *standard Student's t distribution*, i.e. $\tilde{X} \sim t_\nu^{stand}$, with unit variance, zero mean and probability function

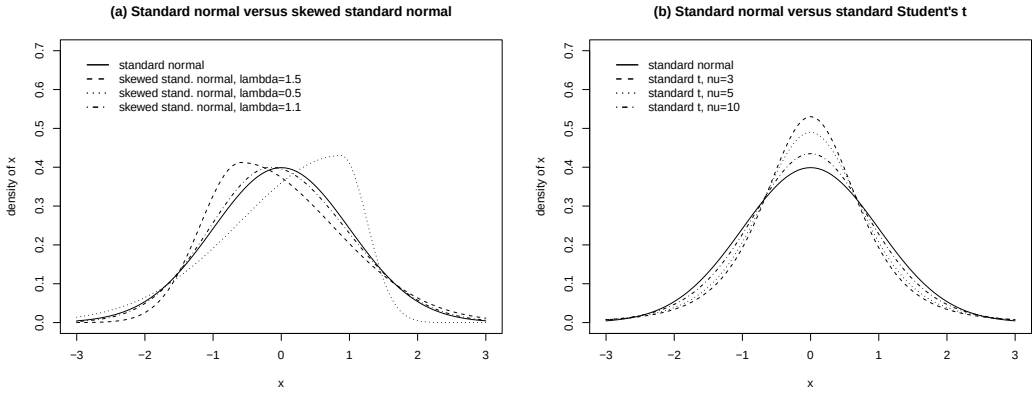
$$f_{t_\nu^{stand}}(\tilde{x}) = \frac{\Gamma\left(\frac{\nu+1}{\nu}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(\pi(\nu-1)\right)^{-\frac{1}{2}}\left(1 + \frac{\tilde{x}^2}{\nu-2}\right)^{-\frac{\nu+1}{2}} .$$

Panel (b) of Figure C.1 displays differences between standard normal and standard Student's t distributions plotted for three values of skew parameter, $\nu \in \{3, 5, 10\}$. As we can see, the density curves of t distribution are more spiky as in the standard normal case and have thicker tails.

C.1.4 Skewed Student's t distribution

The *skewed Student's t distribution* was first introduces by Hansen (1994). It has an additional parameter λ that allows to form asymmetry in the skewed t distribution. A presentation here follows Jondeau, Poon, and Rockinger (2007, pp. 160–162).

Figure C.1: Panel (a): Density of the univariate standard skewed normal distribution for $\lambda \in \{1.5, 0.5, 1.1\}$ compared to the standard normal distribution;
Panel (b): Density of the univariate standard Student's t distribution for $\nu \in \{3, 5, 10\}$ compared to the standard normal distribution



Definition C.4 *A continuous random variable X is said to have a* ***skewed Student's t distribution****, denoted by $X \sim t_{\nu,\lambda}$, with shape parameter $\nu \in (2, \infty)$ and skew parameter $\lambda \in (-1, 1)$ if its density function has a representation*

$$f_{t_{\nu,\lambda}}(x) = b\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(\pi(\nu-2)\right)^{-\frac{1}{2}}\left(1+\frac{\zeta^2}{\nu-2}\right)^{-\frac{\nu+1}{2}} \quad \textit{for} \quad x \in \mathbb{R}\,,$$

where $\Gamma(x) = \int_0^\infty t^{x-1}e^t \mathrm{d}t$ *is a Gamma function and*

$$\zeta = \begin{cases} \frac{bx+a}{1-\lambda} & \textit{if} \quad z < -\frac{a}{b}\,, \\ \frac{bx+a}{1+\lambda} & \textit{if} \quad z \geq -\frac{a}{b}\,. \end{cases}$$

For $\lambda \to 0$ the skewed t distribution reduces to the symmetric Student's t case. And for $\nu \to \infty$ it converges to the skewed normal distribution.

The constant terms a and b are defined as following to standardize the skewed t distribution (i.e for $X \sim t^{stand}_{\nu,\lambda}$ it follows $E[X] = 0$ and $\mathrm{Var}(X) = 1$)

$$a = 4\lambda c\frac{\nu-2}{\nu-1} \quad \text{and} \quad b = 1 + 3\lambda^2 - a^2$$

with

$$c = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(\pi(\nu-2)\right)^{-\frac{1}{2}}\,.$$

C.2 Multivariate distributions

This section presents two most common multivariate distributions, Normal (or Gaussian) and Student's t that we use for construction of elliptical copulas in Section 3.5.4. Multivariate distributions are generalizations of univariate distribution to higher dimensions. We focus our presentation mainly on a bivariate

case.

In this chapter, we denote by d a dimension of a continuous multidimensional random vector $\boldsymbol{X} = (X_1, \ldots, X_d)^t$. Let $\boldsymbol{\mu} \in \mathbb{R}^d$ states a mean vector of $\boldsymbol{X}$

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_d \end{pmatrix} = \begin{pmatrix} \mu(X_1) \\ \vdots \\ \mu(X_d) \end{pmatrix}$$

and $\boldsymbol{\Sigma} \in \mathbb{R}^{d \times d}$ (with inverse matrix $\boldsymbol{\Sigma}^{-1}$) a corresponding variance-covariance matrix

$$\boldsymbol{\Sigma} = (\mathrm{Cov}(X_i, X_j))_{i,j=1\ldots d}$$

Particularly, in bivariate case

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} ,$$

where

$$\sigma_1^2 = \mathrm{Var}(X_1) = \mathrm{Cov}(X_1, X_1) \quad \text{and} \quad \sigma_2^2 = \mathrm{Var}(X_2) = \mathrm{Cov}(X_2, X_2) ,$$

and r is a correlation between X_1 and X_2, i.e.

$$r = \mathrm{Cor}(X_1, X_2) = \frac{\mathrm{Cov}(X_1, X_2)}{\sqrt{\mathrm{Var}(X_1)\,\mathrm{Var}(X_2)}} = \frac{\sigma_{12}}{\sigma_1\sigma_2} .$$

Further, we denote by $\boldsymbol{x} = (x_1, \ldots, x_d)^t$ a realization of a random vector $\boldsymbol{X}$. We call a multivariate distribution *central* if $\boldsymbol{\mu} = \boldsymbol{0}$. Otherwise, it is said to be *non-central*.

C.2.1 Multivariate normal distribution

Most of the following discussions are taken from Johnson and Wichern (2007, pp. 149–152) "*Applied Multivariate Statistical Analysis*".

Definition C.5 *A continuous d-dimensional random vector $\boldsymbol{X} = (X_1, \ldots, X_d)^t$ is said to have a **multivariate (or d-variate) Normal distribution**, denoted by $\boldsymbol{X} \sim \mathcal{N}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with location vector $\boldsymbol{\mu} \in \mathbb{R}^d$ and symmetric positive definit variance-covariance matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{d \times d}$ if its joint probability density function is given by*

$$f_{\mathcal{N}_d(\boldsymbol{\mu},\boldsymbol{\Sigma})}(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^d}} |\det \boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left(-\frac{(\boldsymbol{x}-\boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})}{2} \right) .$$

It is obviously that for $\boldsymbol{X} \sim \mathcal{N}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ holds $E[\boldsymbol{X}] = \boldsymbol{\mu}$ and $\mathrm{Var}(\boldsymbol{X}) = \boldsymbol{\Sigma}$. In a bivariate case $\boldsymbol{X} = (X_1, X_2)$ we get following expression for density function

$$\begin{aligned} f_{\mathcal{N}_2(\boldsymbol{\mu},\boldsymbol{\Sigma})}(x_1, x_2) &= \frac{1}{2\pi} \frac{1}{\sqrt{\sigma_1\sigma_2(1-r^2)}} \\ &\cdot \exp\left\{ -\frac{1}{2(1-r^2)} \left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_1}} \right)^2 + \left(\frac{x_2-\mu_2}{\sqrt{\sigma_2}} \right)^2 - 2r \left(\frac{x_1-\mu_1}{\sqrt{\sigma_1}} \right) \left(\frac{x_2-\mu_2}{\sqrt{\sigma_2}} \right) \right] \right\} . \end{aligned} \tag{C.1}$$

If the random variables X_1 and X_2 are uncorrelated with $r = 0$, then the joint density $f_{\mathcal{N}_2(\boldsymbol{\mu},\boldsymbol{\Sigma})}$ can be written as a product of two univariate normal densities

$$\begin{aligned} f_{\mathcal{N}_2(\boldsymbol{\mu},\boldsymbol{\Sigma})}(x_1, x_2) &= \frac{1}{2\pi}\frac{1}{\sqrt{\sigma_1\sigma_2}} \exp\left\{-\frac{1}{2}\left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_1}}\right)^2 + \left(\frac{x_2-\mu_2}{\sqrt{\sigma_2}}\right)^2\right]\right\} \\ &= \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}\right) \cdot \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(x_2-\mu_2)^2}{2\sigma_2^2}\right) \\ &= f_{X_1}(x_1) \cdot f_{X_2}(x_2) \ , \end{aligned}$$

According to the definition (3.4), the last equation means the independence of X_1 and X_2. So it can be stated that for bivariate normal case, the independence is equivalent to the uncorrelateness.

A surface in the xy-plane of the bivariate standard normal probability distribution function (pdf) is displayed in Figure C.2 on the upper panel. The lower panel represents a corresponding contour plots. A contour line of a function of two variables is a curve along which the function has a constant value. As we can see, the bivariate standard normal distribution has its maximum at the origin. If marginal random variables X_1 and X_2 are independent ($r = 0$) then the surfaces of constant $f_{\mathcal{N}_2(\boldsymbol{\mu},\boldsymbol{\Sigma})}(x_1, x_2)$ are concentric circles around the origin. In dependence case ($r \neq 0$) the surfaces of constant $f_{\mathcal{N}_2(\boldsymbol{\mu},\boldsymbol{\Sigma})}(x_1, x_2)$ form ellipses around the origin. For positive dependence, the major axes of ellipses has a positive slope and vice versa for negative dependence.

C.2.2 Multivariate Student's t Distribution

In this chapter we refer to the book of Kotz and Nadarajah (2004, pp. 1–2) "*Multivariate t Distributions and Their Applications*". In contrast to the multivariate Normal distribution, we denote here by the symbol $\boldsymbol{\Sigma}$ a correlation matrix of $\boldsymbol{X}$, i.e

$$\boldsymbol{\Sigma} = (\mathrm{Cor}(X_i, X_j))_{i,j=1\ldots d} = (r_{ij})_{i,j=1\ldots d} \ .$$

Definition C.6 *A continuous d-dimensional random vector $\boldsymbol{X} = (X_1, \ldots, X_d)^t$ is said to have a **multivariate (or d-variate) t distribution**, denoted by $\boldsymbol{X} \sim t_{\nu,\boldsymbol{\mu},\boldsymbol{\Sigma}}$, with degrees of freedom $\nu > 0$, location vector $\boldsymbol{\mu} \in \mathbb{R}^d$ and symmetric positive definite dispersion matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{d\times d}$ if its joint probability density function is given by*

$$f_{t_{\nu,\boldsymbol{\mu},\boldsymbol{\Sigma}}}(\boldsymbol{x}) = \frac{1}{\sqrt{(\pi\nu)^d}} |\det\boldsymbol{\Sigma}|^{-\frac{1}{2}} \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{1}{\nu}(\boldsymbol{x}-\boldsymbol{\mu})^t\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})^t\right)^{-\frac{\nu+d}{2}} \ .$$

As in the univariate case, the shape parameter ν describes the peakedness of the multivariate distribution. For $d = 1$, $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = 1$ it reduces to univariate Student's t distribution with degrees of freedom ν. For $\boldsymbol{X} \sim t_{\nu,\boldsymbol{\mu},\boldsymbol{\Sigma}}$ it holds

$$E[\boldsymbol{X}] = \boldsymbol{\nu} \quad \text{for } \mu > 1$$

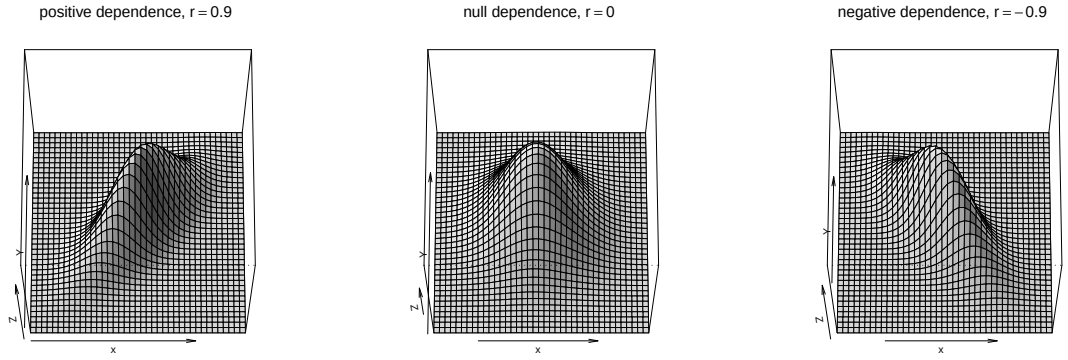
and

$$\mathrm{Var}(\boldsymbol{X}) = \frac{\nu}{\nu-2}\boldsymbol{\Sigma} \quad \text{for } \nu > 2 \ .$$

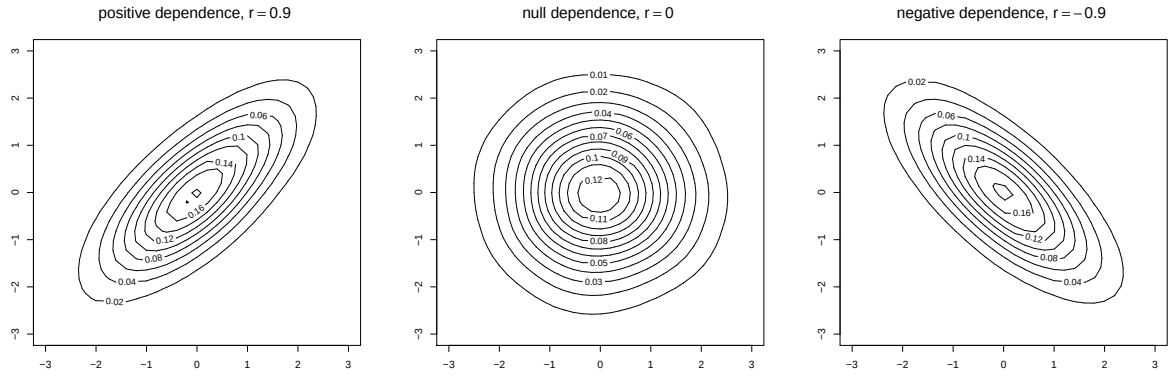
In the bivariate case $\boldsymbol{X} = (X_1, X_2)^t$ with central t distribution, the expression for $\boldsymbol{\Sigma}$ becomes a form

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}$$

Figure C.2: Visualization of a bivariate standard normal distribution with parameters $\boldsymbol{\mu} = (0, 0)^t$, $\sigma_1 = \sigma_2 = 1$ and three different values of correlation coefficient $r = -0.9,\ 0,\ 0.9$



(a) Theoretical 3D plots of bivariate density function



(b) Theoretical contour plots of bivariate density function

with $r = \mathrm{Cor}(X_1, X_2)$ and the density function simplifies to

$$f_{t_{\nu,r}}(x_1, x_2) = \frac{1}{2\pi}(1 - r^2)^{-\frac{1}{2}} \left(1 + \frac{x_1^2 - 2rx_1x_2 + x_2^2}{\nu(1 - r^2)}\right)^{-\frac{\nu+2}{2}} .$$

Appendix D

R code of Vuong and Clarke tests for goodness-of-fit testing for copulas

D.1 Vuong Test

```
#=============================================================================
#  -------------------------- VUONG-TEST  ---------------------------------
#=============================================================================
# Author: Natalia Djunushalieva, TU München, March 2010
# For more details see Quang H. Vuong, 1989, Econometrica
# "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses"
#-----------------------------------------------------------------------------
vuong.test<-function(loglik.model1,loglik.model2,alpha=0.05,p1=0,p2=0,
  correction="Schwarz",print.result=TRUE,name.model1=NULL,name.model2=NULL)
{
  #---------------------------------------------------------------------------
  # INPUT: PARAMETER DESCRIPTION
  # loglik.model1   - numerical vector, individual log-likelihoods of model 1
  # loglik.model2   - numerical vector, individual log-likelihoods of model 2
  # alpha           - numerical, significance level of the test
  # p1              - numerical, number of parameters in model 1
  # p2              - numerical, number of parameters in model 2
  # correction      - character, correction due to Schwarz or due to Akaike
  #                   ("Schwarz" or "Akaike")
  # print.result    - logical, should test results be printed?
  # name.model1     - character, model 1 denotation
  # name.model2     - character, model 2 denotation
  #---------------------------------------------------------------------------
  # OUPUT: PARAMETER DESCRIPTION
  # result          - numerical, favored model (1=model1, 2=model2 or 0=non)
  # nu              - numerical, value of test statistic
  # pvalue          - numerical, p-value of test statistic nu
  # kutosis         - numerical, kurtosis of diff=loglik.model1-loglik.model2
  #---------------------------------------------------------------------------
```

```
# REQUIRED PACKAGES: timeDate (for kurtosis() function)
#-------------------------------------------------------------------------------

# load required packages - timeDate etc.
if(!is.element(c("package:timeDate"),search())){library("timeDate")}

model.spec<-""
result<-NA

if((is.null(name.model1)+ is.null(name.model2))==0)
{
  model.spec1<-paste("(1) ",name.model1,sep="")
  model.spec2<-paste("(2) ",name.model2,sep="")
  model.spec<-paste(model.spec1,model.spec2,sep=" ~ ")
}

#cat(paste("VUONG TEST: ",model.spec,sep=""), "\n")
if(print.result)cat("H0: model (1) is eqivalent to model (2)", "\n")

 n<-length(loglik.model1)

if(correction=="Schwarz")
{
  correction.term<-(p1-p2)*log(n)/(2*n) # for individual log-likelihoods
}
else if(correction=="Akaike")
{
  correction.term<-(p1-p2)/n  # for individual log-likelihoods
}

# Calculate test statistic
m.i<-loglik.model1-loglik.model2-correction.term
kurt.ratios<-kurtosis(loglik.model1-loglik.model2)
nu<-(sqrt(n)*mean(m.i))/(sqrt((n-1)/n*var(m.i)))

if(abs(nu)<qnorm(1-alpha/2))
{
  decision<-"Decision: non of the models is favored"
  result<-0
}
if(nu >=qnorm(1-alpha/2) )
{
  decision<-"Decision: favor model 1"
  result<-1
}
if(nu <= -qnorm(1-alpha/2) )
{
  decision<-"Decision: favor model 2"
  result<-2
}
if(print.result)cat(decision,"\n")
```

```
  pvalue<-2*pnorm(-abs(nu))
  result<-data.frame(result,round(nu,digits=3),
    round(pvalue,digits=3),round(kurt.ratios,digits=3))
  names(result)<-c("model","nu","p.value","kurtosis")
  rownames(result)<-NULL

  if(print.result)print(result)
  if(print.result)cat("\n")

  return(result)
  rm(n,m.i,kurt.ratios,nu,pvalue,result,decision)
}  # end of vuong.test()
```

D.2 Clarke Test

```
#============================================================================
#  --------------------------- CLARKE-TEST  ----------------------------------
#============================================================================
# Author: Natalia Djunushalieva, TU München, March 2010
# For more details see Kevin A. Clarke, 2007, Political Analysis
# "A Simple Distribution-Free Test for Nonnested Model Selection"
#----------------------------------------------------------------------------
clarke.test<-function(loglik.model1,loglik.model2,alpha=0.05,p1=0,p2=0,
  correction="Schwarz",print.result=TRUE,name.model1=NULL,name.model2=NULL)
{
  #--------------------------------------------------------------------------
  # INPUT: PARAMETER DESCRIPTION
  # loglik.model1   - numerical vector, individual log-likelihoods of model 1
  # loglik.model2   - numerical vector, individual log-likelihoods of model 2
  # alpha           - numerical, significance level of the test
  # p1              - numerical, number of parameters in model 1
  # p2              - numerical, number of parameters in model 2
  # correction      - character, correction due to Schwarz or due to Akaike
  #                   ("Schwarz" or "Akaike")
  # print.result    - logical, should test results be printed?
  # name.model1     - character, model 1 denotation
  # name.model2     - character, model 2 denotation
  #--------------------------------------------------------------------------
  # OUPUT: PARAMETER DESCRIPTION
  # result          - numerical, favored model (1=model1, 2=model2 or 0=non)
  # nu              - numerical, value of test statistic
  # pvalue          - numerical, p-value of test statistic nu
  # kutosis         - numerical, kurtosis of diff=loglik.model1-loglik.model2
  #--------------------------------------------------------------------------
  # REQUIRED PACKAGES: timeDate (for kurtosis() function)
  #--------------------------------------------------------------------------

  # load required packages - timeDate etc.
```

```
if(!is.element(c("package:timeDate"),search())){library("timeDate")}

model.spec<-""
result<-NA

if((is.null(name.model1)+ is.null(name.model2))==0)
{
  model.spec1<-paste("(1) ",name.model1,sep="")
  model.spec2<-paste("(2) ",name.model2,sep="")
  model.spec<-paste(model.spec1,model.spec2,sep=" ~ ")
}

#cat(paste("CLARKE TEST: ",model.spec,sep=""), "\n")
if(print.result)cat("H0: model (1) is eqivalent to model (2)", "\n")

n<-length(loglik.model1)

if(correction=="Schwarz")
{
  correction.term<-(p1-p2)*log(n)/(2*n) # for individual log-likelihoods
}
else if(correction=="Akaike")
{
  correction.term<-(p1-p2)/n  # for individual log-likelihoods
}

# Calculate test statistic
m.i<-loglik.model1-loglik.model2-correction.term
kurt.ratios<-kurtosis(loglik.model1-loglik.model2)
B<-sum(m.i > 0)

# Calculate critical value
decision<-"Decision: non of the models is favored"
result<-0
if(B>=n/2)
{
  if(print.result)cat("Perform upper tail test", "\n")
  cAlphaPlus<-1+qbinom(p=(1-alpha),size=n,prob=0.5)
  #cat(paste("cAlphaPlus = ",cAlphaPlus), "\n")
  pvalue<-1-pbinom(B - 1, n, 0.5)
  if(pvalue<=alpha)
  {
    decision<-"Decision: favor model 1"
    result<-1
  }
}
if(B<n/2)
{
  if(print.result)cat("Perform lower tail test", "\n")
  cAlphaMinus<-qbinom(p=alpha,size=n,prob=0.5)
  #cat(paste("cAlphaMinus = ",cAlphaMinus), "\n")
```

```
    pvalue<-pbinom(B, n, 0.5)
    if(pvalue<=alpha)
    {
      decision<-"Decision: favor model 2"
      result<-2
    }
  }
  if(print.result)cat(decision,"\n")

  result<-data.frame(result,round(B,digits=3),round(pvalue,digits=3),
   round(kurt.ratios,digits=3))
  names(result)<-c("model","B","p.value","kurtosis")
  rownames(result)<-NULL

  if(print.result)print(result)
  if(print.result)cat("\n")

  return(result)
  rm(n,m.i,kurt.ratios,B,cAlphaMinus,cAlphaPlus,result,decision,pvalue)
}  # end of clarke.test()
```

Appendix E

Determining data for D-vine trees

In this appendix, we collect formulas for calculating conditional variables of the 10-dimensional D-vine which sequential estimation we have discussed in Section 5.4. As mentioned in Section 3.6.1 on the page 54, the conditional distribution of an arbitrary order can be computed using the h-function

$$F(x|\nu) = h_{\boldsymbol{\theta}_{x,\nu}}(x,\nu) = \frac{\partial C_{\boldsymbol{\theta}_{x,\nu}}\left(F_x(x), F_\nu(\nu)\right)}{\partial F_\nu(\nu)} \ ,$$

if x and ν are uniformly distributed. Here, ν is a conditioning variable, $C_{\boldsymbol{\theta}_{x,\nu}}\left(F_x(x), F_\nu(\nu)\right)$ is a copula that describes dependence between x and ν and $\boldsymbol{\theta}_{x,\nu}$ is its (estimated) parameter vector. Note, the variables x and ν may be itself conditional. We will use following h-function property

$$h_{\boldsymbol{\theta}_{x,\nu}}(x,\nu) = h_{\boldsymbol{\theta}_{\nu,x}}(x,\nu) \ .$$

It holds, because we can interchange the variables x and ν in the copula, i.e.

$$C_{\boldsymbol{\theta}_{x,\nu}}\left(F_x(x), F_\nu(\nu)\right) = C_{\boldsymbol{\theta}_{\nu,x}}\left(F_\nu(\nu), F_x(x)\right) \ .$$

and it does not impact the calculation of $F(x|\nu)$

$$F(x|\nu) = h_{\boldsymbol{\theta}_{x,\nu}}(x,\nu) = \frac{\partial C_{\boldsymbol{\theta}_{x,\nu}}\left(F_x(x), F_\nu(\nu)\right)}{\partial F_\nu(\nu)} = \frac{\partial C_{\boldsymbol{\theta}_{\nu,x}}\left(F_\nu(\nu), F_x(x)\right)}{\partial F_\nu(\nu)} = h_{\boldsymbol{\theta}_{\nu,x}}(x,\nu) \ .$$

To simplify matters, we substitute the notations $h_{\boldsymbol{\theta}_{x,\nu}}$ and $C_{\boldsymbol{\theta}_{x,\nu}}$ by its truncated forms $h_{x,\nu}$ and $C_{x,\nu}$ respectively.

Determining data for 2nd tree

Edge $2,9 \xrightarrow{C_{2,5|9}} 9,5$:

$C_{2,5|9}\left(F(x_2|x_9), F(x_5|x_9)\right)$
$F(x_2|x_9) = h_{\boldsymbol{\theta}_{2,9}}\left(x_2, x_9\right)$
$F(x_5|x_9) = h_{\boldsymbol{\theta}_{5,9}}\left(x_5, x_9\right) = h_{\boldsymbol{\theta}_{9,5}}\left(x_5, x_9\right)$

Edge $9,5 \xrightarrow{C_{9,4|5}} 5,4$:

$C_{9,4|5}\left(F(x_9|x_5), F(x_4|x_5)\right)$
$F(x_9|x_5) = h_{9,5}\left(x_9, x_5\right)$
$F(x_4|x_5) = h_{4,5}\left(x_4, x_5\right) = h_{5,4}\left(x_4, x_5\right)$

Edge $5,4 \xrightarrow{C_{5,10|4}} 4,10$:

$C_{5,10|4}\left(F(x_5|x_4), F(x_{10}|x_4)\right)$
$F(x_5|x_4) = h_{5,4}\left(x_5, x_4\right)$
$F(x_{10}|x_4) = h_{10,4}\left(x_{10}, x_4\right) = h_{4,10}\left(x_{10}, x_4\right)$

Edge $4,10 \xrightarrow{C_{4,6|10}} 10,6$:

$C_{4,6|10}\left(F(x_4|x_{10}), F(x_6|x_{10})\right)$
$F(x_4|x_{10}) = h_{4,10}\left(x_4, x_{10}\right)$
$F(x_6|x_{10}) = h_{6,10}\left(x_6, x_{10}\right) = h_{10,6}\left(x_6, x_{10}\right)$

Edge $10,6 \xrightarrow{C_{10,1|6}} 6,1$:

$C_{10,1|6}\left(F(x_{10}|x_6), F(x_1|x_6)\right)$
$F(x_{10}|x_6) = h_{10,6}\left(x_{10}, x_6\right)$
$F(x_1|x_6) = h_{1,6}\left(x_1, x_6\right) = h_{6,1}\left(x_1, x_6\right)$

Edge $6,1 \xrightarrow{C_{6,8|1}} 1,8$:

$C_{6,8|1}\left(F(x_6|x_1), F(x_8|x_1)\right)$
$F(x_6|x_1) = h_{6,1}\left(x_6, x_1\right)$
$F(x_8|x_1) = h_{8,1}\left(x_8, x_1\right) = h_{1,8}\left(x_8, x_1\right)$

Edge $1,8 \xrightarrow{C_{1,7|8}} 8,7$:

$C_{1,7|8}\left(F(x_1|x_8), F(x_7|x_8)\right)$
$F(x_1|x_8) = h_{1,8}\left(x_1, x_8\right)$
$F(x_7|x_8) = h_{7,8}\left(x_7, x_8\right) = h_{8,7}\left(x_7, x_8\right)$

Edge $8,7 \xrightarrow{C_{8,3|7}} 7,3$:

$C_{8,3|7}\left(F(x_8|x_7), F(x_3|x_7)\right)$
$F(x_8|x_7) = h_{8,7}\left(x_8, x_7\right)$
$F(x_3|x_7) = h_{3,7}\left(x_3, x_7\right) = h_{7,3}\left(x_3, x_7\right)$

Determining data for 3rd tree

Edge $2,5|9 \xrightarrow{C_{2,4|9,5}} 9,4|5$:

$C_{2,4|9,5}\left(F(x_2|x_9,x_5), F(x_4|x_9,x_5)\right)$
$F(x_2|x_9,x_5) = h_{2,5|9}\left(F(x_2|x_9), F(x_5|x_9)\right)$
$F(x_4|x_9,x_5) = h_{4,9|5}\left(F(x_4|x_5), F(x_9|x_5)\right) = h_{9,4|5}\left(F(x_4|x_5), F(x_9|x_5)\right)$

Edge $9,4|5 \xrightarrow{C_{9,10|5,4}} 5,10|4$:

$C_{9,10|5,4}\left(F(x_9|x_5,x_4), F(x_{10}|x_5,x_4)\right)$
$F(x_9|x_5,x_4) = h_{9,4|5}\left(F(x_9|x_5), F(x_4|x_5)\right)$
$F(x_{10}|x_5,x_4) = h_{10,5|4}\left(F(x_{10}|x_4), F(x_5|x_4)\right) = h_{5,10|4}\left(F(x_{10}|x_4), F(x_5|x_4)\right)$

Edge $5,10|4 \xrightarrow{C_{5,6|4,10}} 4,6|10$:

$C_{5,6|4,10}\left(F(x_5|x_4,x_{10}), F(x_6|x_4,x_{10})\right)$
$F(x_5|x_4,x_{10}) = h_{5,10|4}\left(F(x_5|x_4), F(x_{10}|x_4)\right)$
$F(x_6|x_4,x_{10}) = h_{6,4|10}\left(F(x_6|x_{10}), F(x_4|x_{10})\right) = h_{4,6|10}\left(F(x_6|x_{10}), F(x_4|x_{10})\right)$

Edge $4,6|10 \xrightarrow{C_{4,1|10,6}} 10,1|6$:

$C_{4,1|10,6}\left(F(x_4|x_{10},x_6), F(x_1|x_{10},x_6)\right)$
$F(x_4|x_{10},x_6) = h_{4,6|10}\left(F(x_4|x_{10}), F(x_6|x_{10})\right)$
$F(x_1|x_{10},x_6) = h_{1,10|6}\left(F(x_1|x_6), F(x_{10}|x_6)\right) = h_{10,1|6}\left(F(x_1|x_6), F(x_{10}|x_6)\right)$

Edge $10,1|6 \xrightarrow{C_{10,8|6,1}} 6,8|1$:

$C_{10,8|6,1}\left(F(x_{10}|x_6,x_1), F(x_8|x_6,x_1)\right)$
$F(x_{10}|x_6,x_1) = h_{10,6|1}\left(F(x_{10}|x_6), F(x_1|x_6)\right)$
$F(x_8|x_6,x_1) = h_{8,6|1}\left(F(x_8|x_1), F(x_6|x_1)\right) = h_{6,8|1}\left(F(x_8|x_1), F(x_6|x_1)\right)$

Edge $6,8|1 \xrightarrow{C_{6,7|1,8}} 1,7|8$:

$C_{6,7|1,8}\left(F(x_6|x_1,x_8), F(x_7|x_1,x_8)\right)$
$F(x_6|x_1,x_8) = h_{6,8|1}\left(F(x_6|x_1), F(x_8|x_1)\right)$
$F(x_7|x_1,x_8) = h_{7,1|8}\left(F(x_7|x_8), F(x_1|x_8)\right) = h_{1,7|8}\left(F(x_7|x_8), F(x_1|x_8)\right)$

Edge $1,7|8 \xrightarrow{C_{1,3|8,7}} 8,3|7$:

$C_{1,3|8,7}\left(F(x_1|x_8,x_7), F(x_3|x_8,x_7)\right)$
$F(x_1|x_8,x_7) = h_{1,7|8}\left(F(x_1|x_8), F(x_7|x_8)\right)$
$F(x_3|x_8,x_7) = h_{3,8|7}\left(F(x_3|x_7), F(x_8|x_7)\right) = h_{8,3|7}\left(F(x_3|x_7), F(x_8|x_7)\right)$

Determining data for 4th tree

Edge $2,4|9,5 \xrightarrow{C_{2,10|9,5,4}} 9,10|5,4$:

$C_{2,10|9,5,4}\left(F(x_2|x_9,x_5,x_4), F(x_{10}|x_9,x_5,x_4)\right)$
$F(x_2|x_9,x_5,x_4) = h_{2,4|9,5}\left(F(x_2|x_9,x_5), F(x_4|x_9,x_5)\right)$
$F(x_{10}|x_9,x_5,x_4) = h_{10,9|5,4}\left(F(x_{10}|x_5,x_4), F(x_9|x_5,x_4)\right) = h_{9,10|5,4}\left(F(x_{10}|x_5,x_4), F(x_9|x_5,x_4)\right)$

Edge $9,10|5,4 \xrightarrow{C_{9,6|5,4,10}} 5,6|4,10$:

$C_{9,6|5,4,10}\left(F(x_9|x_5,x_4,x_{10}), F(x_6|x_5,x_4,x_{10})\right)$
$F(x_9|x_5,x_4,x_{10}) = h_{9,10|5,4}\left(F(x_9|x_5,x_4), F(x_{10}|x_5,x_4)\right)$
$F(x_6|x_5,x_4,x_{10}) = h_{6,5|4,10}\left(F(x_6|x_4,x_{10}), F(x_5|x_4,x_{10})\right) = h_{5,6|4,10}\left(F(x_6|x_4,x_{10}), F(x_5|x_4,x_{10})\right)$

Edge $5,6|4,10 \xrightarrow{C_{5,1|4,10,6}} 4,1|10,6$:

$C_{5,1|4,10,6}\left(F(x_5|x_4,x_{10},x_6), F(x_1|x_4,x_{10},x_6)\right)$
$F(x_5|x_4,x_{10},x_6) = h_{5,6|4,10}\left(F(x_5|x_4,x_{10}), F(x_6|x_4,x_{10})\right)$
$F(x_1|x_4,x_{10},x_6) = h_{1,4|10,6}\left(F(x_1|x_{10},x_6), F(x_4|x_{10},x_6)\right) = h_{4,1|10,6}\left(F(x_1|x_{10},x_6), F(x_4|x_{10},x_6)\right)$

Edge $4,1|10,6 \xrightarrow{C_{4,8|10,6,1}} 10,8|6,1$:

$C_{4,8|10,6,1}\left(F(x_4|x_{10},x_6,x_1), F(x_8|x_{10},x_6,x_1)\right)$
$F(x_4|x_{10},x_6,x_1) = h_{4,1|10,6}\left(F(x_4|x_{10},x_6), F(x_1|x_{10},x_6)\right)$
$F(x_8|x_{10},x_6,x_1) = h_{8,10|6,1}\left(F(x_8|x_6,x_1), F(x_{10}|x_6,x_1)\right) = h_{10,8|6,1}\left(F(x_8|x_6,x_1), F(x_{10}|x_6,x_1)\right)$

Edge $10,8|6,1 \xrightarrow{C_{10,7|6,1,8}} 6,7|1,8$:

$C_{10,7|6,1,8}\left(F(x_{10}|x_6,x_1,x_8), F(x_7|x_6,x_1,x_8)\right)$
$F(x_{10}|x_6,x_1,x_8) = h_{10,8|6,1}\left(F(x_{10}|x_6,x_1), F(x_8|x_6,x_1)\right)$
$F(x_7|x_6,x_1,x_8) = h_{7,6|1,8}\left(F(x_7|x_1,x_8), F(x_6|x_1,x_8)\right) = h_{6,7|1,8}\left(F(x_7|x_1,x_8), F(x_6|x_1,x_8)\right)$

Edge $6,7|1,8 \xrightarrow{C_{6,3|1,8,7}} 1,3|8,7$:

$C_{6,3|1,8,7}\left(F(x_6|x_1,x_8,x_7),F(x_3|x_1,x_8,x_7)\right)$
$F(x_6|x_1,x_8,x_7) = h_{6,7|1,8}\left(F(x_6|x_1,x_8),F(x_7|x_1,x_8)\right)$
$F(x_3|x_1,x_8,x_7) = h_{3,1|8,7}\left(F(x_3|x_8,x_7),F(x_1|x_8,x_7)\right) = h_{1,3|8,7}\left(F(x_3|x_8,x_7),F(x_1|x_8,x_7)\right)$

Determining data for 5th tree

Edge $2,10|9,5,4 \xrightarrow{C_{2,6|9,5,4,10}} 9,6|5,4,10$:

$C_{2,6|9,5,4,10}\left(F(x_2|x_9,x_5,x_4,x_{10}),F(x_6|x_9,x_5,x_4,x_{10})\right)$
$F(x_2|x_9,x_5,x_4,x_{10}) = h_{2,10|9,5,4}\left(F(x_2|x_9,x_5,x_4),F(x_{10}|x_9,x_5,x_4)\right)$
$F(x_6|x_9,x_5,x_4,x_{10}) = h_{6,9|5,4,10}\left(F(x_6|x_5,x_4,x_{10}),F(x_9|x_5,x_4,x_{10})\right)$
$= h_{9,6|5,4,10}\left(F(x_6|x_5,x_4,x_{10}),F(x_9|x_5,x_4,x_{10})\right)$

Edge $9,6|5,4,10 \xrightarrow{C_{9,1|5,4,10,6}} 5,1|4,10,6$:

$C_{9,1|5,4,10,6}\left(F(x_9|x_5,x_4,x_{10},x_6),F(x_1|x_5,x_4,x_{10},x_6)\right)$
$F(x_9|x_5,x_4,x_{10},x_6) = h_{9,6|5,4,10}\left(F(x_9|x_5,x_4,x_{10}),F(x_6|x_5,x_4,x_{10})\right)$
$F(x_1|x_5,x_4,x_{10},x_6) = h_{1,5|4,10,6}\left(F(x_1|x_4,x_{10},x_6),F(x_5|x_4,x_{10},x_6)\right)$
$= h_{5,1|4,10,6}\left(F(x_1|x_4,x_{10},x_6),F(x_5|x_4,x_{10},x_6)\right)$

Edge $5,1|4,10,6 \xrightarrow{C_{5,8|4,10,6,1}} 4,8|10,6,1$:

$C_{5,8|4,10,6,1}\left(F(x_5|x_4,x_{10},x_6,x_1),F(x_8|x_4,x_{10},x_6,x_1)\right)$
$F(x_5|x_4,x_{10},x_6,x_1) = h_{5,1|4,10,6}\left(F(x_5|x_4,x_{10},x_6),F(x_1|x_4,x_{10},x_6)\right)$
$F(x_8|x_4,x_{10},x_6,x_1) = h_{8,4|10,6,1}\left(F(x_8|x_{10},x_6,x_1),F(x_4|x_{10},x_6,x_1)\right)$
$= h_{4,8|10,6,1}\left(F(x_8|x_{10},x_6,x_1),F(x_4|x_{10},x_6,x_1)\right)$

Edge $4,8|10,6,1 \xrightarrow{C_{4,7|10,6,1,8}} 10,7|6,1,8$:

$C_{4,7|10,6,1,8}\left(F(x_4|x_{10},x_6,x_1,x_8),F(x_7|x_{10},x_6,x_1,x_8)\right)$
$F(x_4|x_{10},x_6,x_1,x_8) = h_{4,8|10,6,1}\left(F(x_4|x_{10},x_6,x_1),F(x_8|x_{10},x_6,x_1)\right)$
$F(x_7|x_{10},x_6,x_1,x_8) = h_{7,10|6,1,8}\left(F(x_7|x_6,x_1,x_8),F(x_{10}|x_6,x_1,x_8)\right)$
$= h_{10,7|6,1,8}\left(F(x_7|x_6,x_1,x_8),F(x_{10}|x_6,x_1,x_8)\right)$

Edge $10,7|6,1,8 \xrightarrow{C_{10,3|6,1,8,7}} 6,3|1,8,7$:

$C_{10,3|6,1,8,7}\left(F(x_{10}|x_6,x_1,x_8,x_7),F(x_3|x_6,x_1,x_8,x_7)\right)$
$F(x_{10}|x_6,x_1,x_8,x_7) = h_{10,7|6,1,8}\left(F(x_{10}|x_6,x_1,x_8),F(x_7|x_6,x_1,x_8)\right)$
$F(x_3|x_6,x_1,x_8,x_7) = h_{3,6|1,8,7}\left(F(x_3|x_1,x_8,x_7),F(x_6|x_1,x_8,x_7)\right)$
$= h_{6,3|1,8,7}\left(F(x_3|x_1,x_8,x_7),F(x_6|x_1,x_8,x_7)\right)$

Determining data for 6th tree

Edge $2,6|9,5,4,10 \xrightarrow{C_{2,1|9,5,4,10,6}} 9,1|5,4,10,6$:

$C_{2,1|9,5,4,10,6}\left(F(x_2|x_9,x_5,x_4,x_{10},x_6),F(x_1|x_9,x_5,x_4,x_{10},x_6)\right)$
$F(x_2|x_9,x_5,x_4,x_{10},x_6) = h_{2,6|9,5,4,10}\left(F(x_2|x_9,x_5,x_4,x_{10}),F(x_6|x_9,x_5,x_4,x_{10})\right)$
$F(x_1|x_9,x_5,x_4,x_{10},x_6) = h_{1,9|5,4,10,6}\left(F(x_1|x_5,x_4,x_{10},x_6),F(x_9|x_5,x_4,x_{10},x_6)\right)$
$= h_{9,1|5,4,10,6}\left(F(x_1|x_5,x_4,x_{10},x_6),F(x_9|x_5,x_4,x_{10},x_6)\right)$

Edge $9,1|5,4,10,6 \xrightarrow{C_{9,8|5,4,10,6,1}} 5,8|4,10,6,1$:

$C_{9,8|5,4,10,6,1}\left(F(x_9|x_5,x_4,x_{10},x_6,x_1),F(x_8|x_5,x_4,x_{10},x_6,x_1)\right)$
$F(x_9|x_5,x_4,x_{10},x_6,x_1) = h_{9,1|5,4,10,6}\left(F(x_9|x_5,x_4,x_{10},x_6),F(x_1|x_5,x_4,x_{10},x_6)\right)$
$F(x_8|x_5,x_4,x_{10},x_6,x_1) = h_{8,5|4,10,6,1}\left(F(x_8|x_4,x_{10},x_6,x_1),F(x_5|x_4,x_{10},x_6,x_1)\right)$
$= h_{5,8|4,10,6,1}\left(F(x_8|x_4,x_{10},x_6,x_1),F(x_5|x_4,x_{10},x_6,x_1)\right)$

Edge $5,8|4,10,6,1 \xrightarrow{C_{5,7|4,10,6,1,8}} 4,7|10,6,1,8$:

$C_{5,7|4,10,6,1,8}\left(F(x_5|x_4,x_{10},x_6,x_1,x_8),F(x_7|x_4,x_{10},x_6,x_1,x_8)\right)$
$F(x_5|x_4,x_{10},x_6,x_1,x_8) = h_{5,8|4,10,6,1}\left(F(x_5|x_4,x_{10},x_6,x_1),F(x_8|x_4,x_{10},x_6,x_1)\right)$
$F(x_7|x_4,x_{10},x_6,x_1,x_8) = h_{7,4|10,6,1,8}\left(F(x_7|x_{10},x_6,x_1,x_8),F(x_4|x_{10},x_6,x_1,x_8)\right)$
$h_{4,7|10,6,1,8}\left(F(x_7|x_{10},x_6,x_1,x_8),F(x_4|x_{10},x_6,x_1,x_8)\right)$

Edge $4,7|10,6,1,8 \xrightarrow{C_{4,3|10,6,1,8,7}} 10,3|6,1,8,7$:

$C_{4,3|10,6,1,8,7}\left(F(x_4|x_{10},x_6,x_1,x_8,x_7),F(x_3|x_{10},x_6,x_1,x_8,x_7)\right)$
$F(x_4|x_{10},x_6,x_1,x_8,x_7) = h_{4,7|10,6,1,8}\left(F(x_4|x_{10},x_6,x_1,x_8),F(x_7|x_{10},x_6,x_1,x_8)\right)$
$F(x_3|x_{10},x_6,x_1,x_8,x_7) = h_{3,10|6,1,8,7}\left(F(x_3|x_6,x_1,x_8,x_7),F(x_{10}|x_6,x_1,x_8,x_7)\right)$
$h_{10,3|6,1,8,7}\left(F(x_3|x_6,x_1,x_8,x_7),F(x_{10}|x_6,x_1,x_8,x_7)\right)$

Determining data for 7th tree

Edge $2,1|9,5,4,10,6 \xrightarrow{C_{2,8|9,5,4,10,6,1}} 9,8|5,4,10,6,1$:

$C_{2,8|9,5,4,10,6,1}\left(F(x_2|x_9,x_5,x_4,x_{10},x_6,x_1), F(x_8|x_9,x_5,x_4,x_{10},x_6,x_1)\right)$
$F(x_2|x_9,x_5,x_4,x_{10},x_6,x_1) = h_{2,1|9,5,4,10,6}\left(F(x_2|x_9,x_5,x_4,x_{10},x_6), F(x_1|x_9,x_5,x_4,x_{10},x_6)\right)$
$F(x_8|x_9,x_5,x_4,x_{10},x_6,x_1) = h_{8,9|5,4,10,6,1}\left(F(x_8|x_5,x_4,x_{10},x_6,x_1), F(x_9|x_5,x_4,x_{10},x_6,x_1)\right)$
$= h_{9,8|5,4,10,6,1}\left(F(x_8|x_5,x_4,x_{10},x_6,x_1), F(x_9|x_5,x_4,x_{10},x_6x_1)\right)$

Edge $9,8|5,4,10,6,1 \xrightarrow{C_{9,7|5,4,10,6,1,8}} 5,7|4,10,6,1,8$:

$C_{9,7|5,4,10,6,1,8}\left(F(x_9|x_5,x_4,x_{10},x_6,x_1,x_8), F(x_7|x_5,x_4,x_{10},x_6,x_1,x_8)\right)$
$F(x_9|x_5,x_4,x_{10},x_6,x_1,x_8) = h_{9,8|5,4,10,6,1}\left(F(x_9|x_5,x_4,x_{10},x_6,x_1), F(x_8|x_5,x_4,x_{10},x_6,x_1)\right)$
$F(x_7|x_5,x_4,x_{10},x_6,x_1,x_8) = h_{7,5|4,10,6,1,8}\left(F(x_7|x_4,x_{10},x_6,x_1,x_8), F(x_5|x_4,x_{10},x_6,x_1,x_8)\right)$
$= h_{5,7|4,10,6,1,8}\left(F(x_7|x_4,x_{10},x_6,x_1,x_8), F(x_5|x_4,x_{10},x_6,x_1,x_8)\right)$

Edge $5,7|4,10,6,1,8 \xrightarrow{C_{5,3|4,10,6,1,8,7}} 4,3|10,6,1,8,7$:

$C_{5,3|4,10,6,1,8,7}\left(F(x_5|x_4,x_{10},x_6,x_1,x_8,x_7), F(x_3|x_4,x_{10},x_6,x_1,x_8,x_7)\right)$
$F(x_5|x_4,x_{10},x_6,x_1,x_8,x_7) = h_{5,7|4,10,6,1,8}\left(F(x_5|x_4,x_{10},x_6,x_1,x_8), F(x_7|x_4,x_{10},x_6,x_1,x_8)\right)$
$F(x_3|x_4,x_{10},x_6,x_1,x_8,x_7) = h_{3,4|10,6,1,8,7}\left(F(x_3|x_{10},x_6,x_1,x_8,x_7), F(x_4|x_{10},x_6,x_1,x_8,x_7)\right)$
$= h_{4,3|10,6,1,8,7}\left(F(x_3|x_{10},x_6,x_1,x_8,x_7), F(x_4|x_{10},x_6,x_1,x_8,x_7)\right)$

Determining data for 8th tree

Edge $2,8|9,5,4,10,6,1 \xrightarrow{C_{2,7|9,5,4,10,6,1,8}} 9,7|5,4,10,6,1,8$:

$C_{2,7|9,5,4,10,6,1,8}\left(F(x_2|x_9,x_5,x_4,x_{10},x_6,x_1,x_8), F(x_7|x_9,x_5,x_4,x_{10},x_6,x_1,x_8)\right)$
$F(x_2|x_9,x_5,x_4,x_{10},x_6,x_1,x_8)$
$= h_{2,8|9,5,4,10,6,1}\left(F(x_2|x_9,x_5,x_4,x_{10},x_6,x_1), F(x_8|x_9,x_5,x_4,x_{10},x_6,x_1)\right)$
$F(x_7|x_9,x_5,x_4,x_{10},x_6,x_1,x_8)$
$= h_{7,9|5,4,10,6,1,8}\left(F(x_7|x_5,x_4,x_{10},x_6,x_1,x_8), F(x_9|x_5,x_4,x_{10},x_6,x_1,x_8)\right)$
$= h_{9,7|5,4,10,6,1,8}\left(F(x_7|x_5,x_4,x_{10},x_6,x_1,x_8), F(x_9|x_5,x_4,x_{10},x_6,x_1,x_8)\right)$

Edge $9,7|5,4,10,6,1,8 \xrightarrow{C_{9,3|5,4,10,6,1,8,7}} 5,3|4,10,6,1,8,7$:

$$C_{9,3|5,4,10,6,1,8,7}\left(F(x_9|x_5,x_4,x_{10},x_6,x_1,x_8,x_7),F(x_3|x_5,x_4,x_{10},x_6,x_1,x_8,x_7)\right)$$
$$F(x_9|x_5,x_4,x_{10},x_6,x_1,x_8,x_7)$$
$$= h_{9,7|5,4,10,6,1,8}\left(F(x_9|x_5,x_4,x_{10},x_6,x_1,x_8),F(x_7|x_5,x_4,x_{10},x_6,x_1,x_8)\right)$$
$$F(x_3|x_5,x_4,x_{10},x_6,x_1,x_8,x_7)$$
$$= h_{3,5|4,10,6,1,8,7}\left(F(x_3|x_4,x_{10},x_6,x_1,x_8,x_7),F(x_5|x_4,x_{10},x_6,x_1,x_8,x_7)\right)$$
$$= h_{5,3|4,10,6,1,8,7}\left(F(x_3|x_4,x_{10},x_6,x_1,x_8,x_7),F(x_5|x_4,x_{10},x_6,x_1,x_8,x_7)\right)$$

Determining data for 9th tree

Edge $2,7|9,5,4,10,6,1,8 \xrightarrow{C_{2,3|9,5,4,10,6,1,8,7}} 9,3|5,4,10,6,1,8,7$:

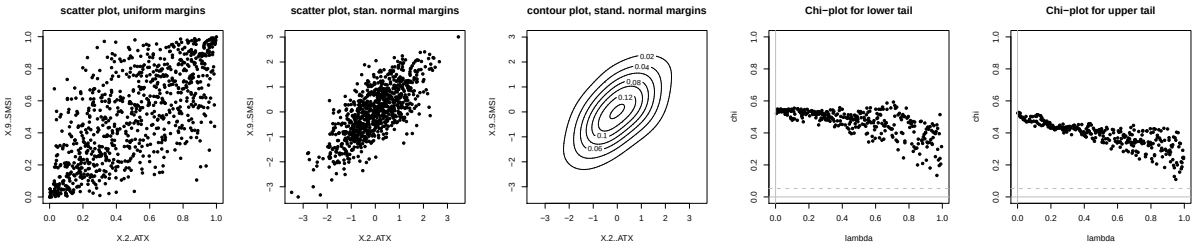
$$C_{2,3|9,5,4,10,6,1,8,7}\left(F(x_2|x_9,x_5,x_4,x_{10},x_6,x_1,x_8,x_7),F(x_3|x_9,x_5,x_4,x_{10},x_6,x_1,x_8,x_7)\right)$$
$$F(x_2|x_9,x_5,x_4,x_{10},x_6,x_1,x_8,x_7)$$
$$= h_{2,7|9,5,4,10,6,1,8}\left(F(x_2|x_9,x_5,x_4,x_{10},x_6,x_1,x_8),F(x_7|x_9,x_5,x_4,x_{10},x_6,x_1,x_8)\right)$$
$$F(x_3|x_9,x_5,x_4,x_{10},x_6,x_1,x_8,x_7)$$
$$= h_{3,9|5,4,10,6,1,8,7}\left(F(x_3|x_5,x_4,x_{10},x_6,x_1,x_8,x_7),F(x_9|x_5,x_4,x_{10},x_6,x_1,x_8,x_7)\right)$$
$$= h_{9,3|5,4,10,6,1,8,7}\left(F(x_3|x_5,x_4,x_{10},x_6,x_1,x_8,x_7),F(x_9|x_5,x_4,x_{10},x_6,x_1,x_8,x_7)\right)$$

Appendix F

D-vine sequential estimation

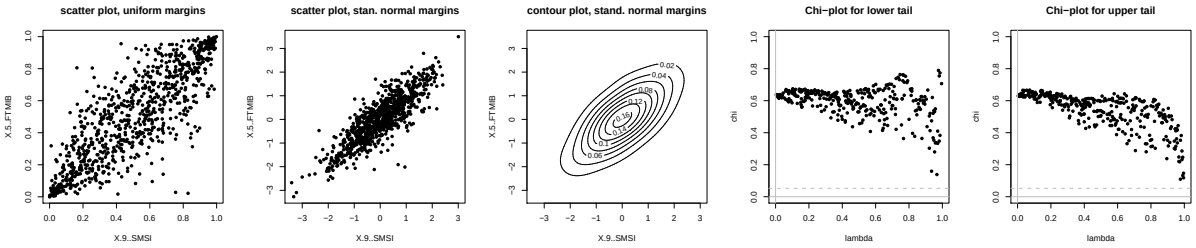
This appendix contains results of estimating procedures for each building block of the D-vine from Chapter 5. Such estimation technique of the D-vine parameters is called *sequential estimation*, i.e. pair-copula parameter will be estimated individually but not jointly. In many practical situation, D-vine estimations obtained in a such way are adequate. The log-likelihood function increases *slightly* when estimating all parameters jointly in comparison to the log-likelihood function estimated sequentially. The set of tools for determining an appropriate pair-copula was described in Section 5.4 on page 95. In following, we motivate our decision on the results of score test based on the Vuong statistic and justify it by coefficients of lower and upper tail-dependence as well as scatter, contour and Chi-plots. Moreover, we compare whether the tail-dependences of estimated Student copula describe co-movements in the data better than the tail-dependences of estimated BB1 copula and vice versa. Note, for small $\hat{\tau}$'s (< 0.3) we do not distinguish between several copula families and choose a Gauss copula as a most simple one.

Figure F.1: 1ST D-VINE TREE: determining copula family for building block $C_{2,9}$ ATX $\sim$ SMSI



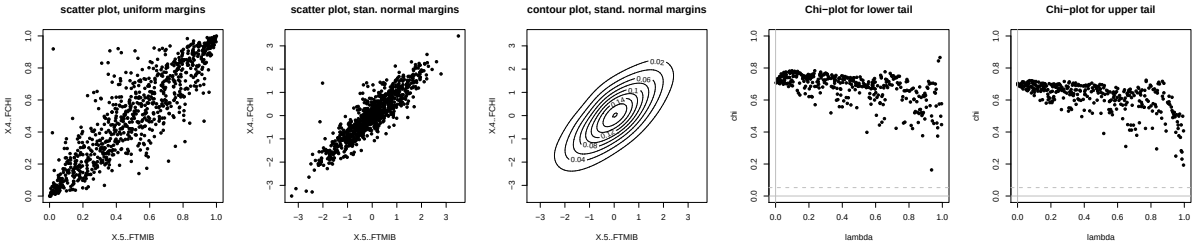
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	4	***5***	3	4	-4	-3	-5	-4
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.41$		$\hat{\lambda}_U^{t} = 0.19$		$\hat{\lambda}_U^{BB1} = 0.43$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.38$		$\hat{\lambda}_L^{t} = 0.19$		$\hat{\lambda}_L^{BB1} = 0.50$			
Chi-plots	nearly symmetrical tail dependences							
Empirical Kendall's $\hat{\tau}$	0.52^{***}							
Chosen copula	*BB1* with parameters $\hat{\theta} = 0.66$, $\hat{\delta} = 1.53$							

Figure F.2: 1ST D-VINE TREE: determining copula family for building block $C_{9,5}$ SMSI $\sim$ FTMIB



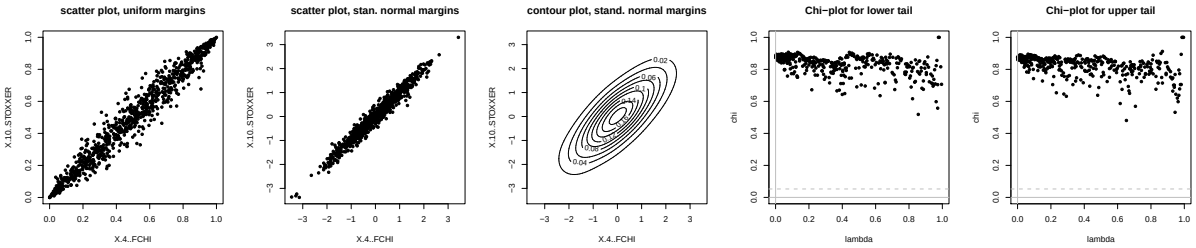
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	4	***6***	2	2	-4	0	-7	-3
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.51$		$\hat{\lambda}_U^{t} = 0.51$		$\hat{\lambda}_U^{BB1} = 0.61$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.62$		$\hat{\lambda}_L^{t} = 0.51$		$\hat{\lambda}_L^{BB1} = 0.64$			
Chi-plots	asymmetrical tail dependences							
Empirical Kendall's $\hat{\tau}$	0.66^{***}							
Chosen copula	*BB1* with parameters $\hat{\theta} = 0.73$, $\hat{\delta} = 2.10$							

Figure F.3: 1ST D-VINE TREE: determining copula family for building block $C_{5,4}$ FTMIB $\sim$ FCHI



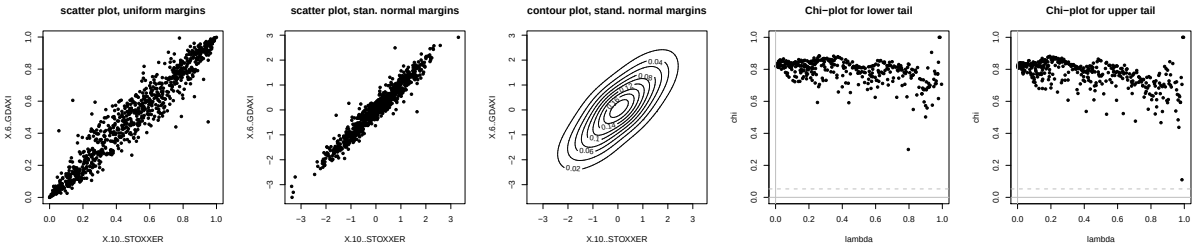
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	***5***	***5***	1	1	-3	0	-7	-2
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.68$		$\hat{\lambda}_U^{t} = 0.63$		$\hat{\lambda}_U^{BB1} = 0.69$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.62$		$\hat{\lambda}_L^{t} = 0.63$		$\hat{\lambda}_L^{BB1} = 0.72$			
Chi-plots	nearly symmetrical strong tail dependences							
Empirical Kendall's $\hat{\tau}$	0.73^{***}							
Chosen copula	*Student* with parameters $\hat{r} = 0.91$, $\hat{\nu} = 4.69$							

Figure F.4: 1ST D-VINE TREE: determining copula family for building block $C_{4,10}$
FCHI $\sim$ STOXXER



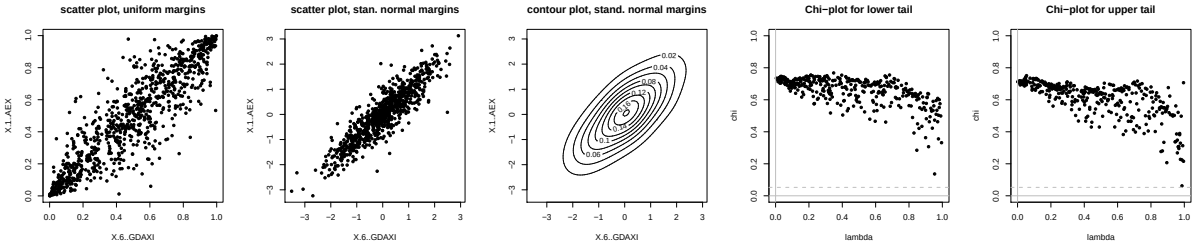
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	***5***	3	0	4	-4	-1	-6	-1
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.92$		$\hat{\lambda}_U^{t} = 0.80$		$\hat{\lambda}_U^{BB1} = 0.87$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.82$		$\hat{\lambda}_L^{t} = 0.80$		$\hat{\lambda}_L^{BB1} = 0.87$			
Chi-plots	symmetrical strong tail dependences							
Empirical Kendall's $\hat{\tau}$	0.88***							
Chosen copula	*Student* with parameters $\hat{} = 0.98$, $\hat{\nu} = 6.91$							

Figure F.5: 1ST D-VINE TREE: determining copula family for building block $C_{10,6}$
STOXXER $\sim$ GDAXI



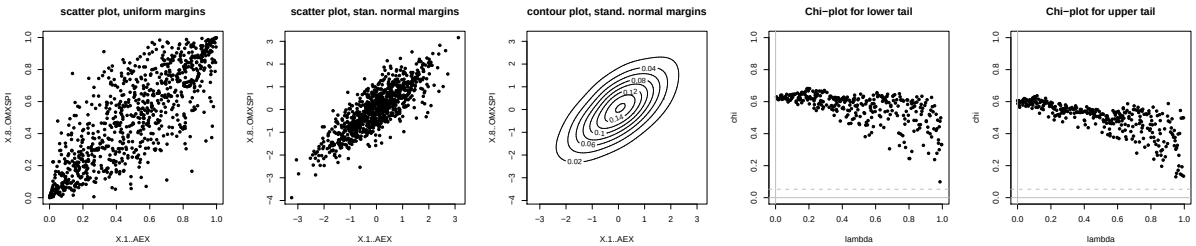
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	***6***	4	-1	1	-2	0	-7	-1
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.78$		$\hat{\lambda}_U^{t} = 0.76$		$\hat{\lambda}_U^{BB1} = 0.83$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.68$		$\hat{\lambda}_L^{t} = 0.76$		$\hat{\lambda}_L^{BB1} = 0.83$			
Chi-plots	symmetrical strong tail dependences							
Empirical Kendall's $\hat{\tau}$	0.85***							
Chosen copula	*Student* with parameters $\hat{r} = 0.97$, $\hat{\nu} = 5.63$							

Figure F.6: 1ST D-VINE TREE: determining copula family for building block $C_{6,1}$
GDAXI $\sim$ AEX



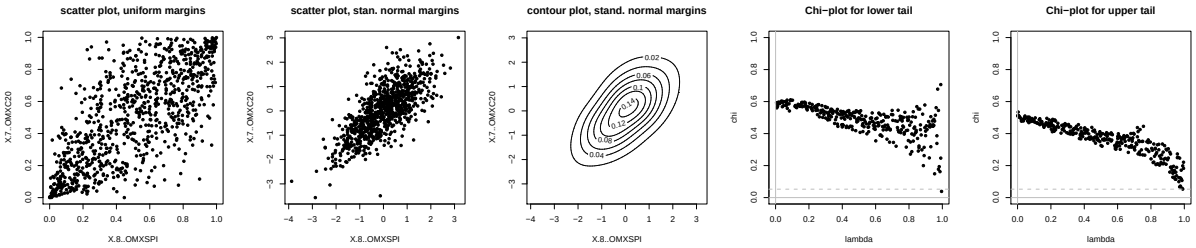
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	***6***	4	-1	2	-3	1	-7	-2
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.62$		$\hat{\lambda}_U^{t} = 0.60$		$\hat{\lambda}_U^{BB1} = 0.69$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.65$		$\hat{\lambda}_L^{t} = 0.60$		$\hat{\lambda}_L^{BB1} = 0.68$			
Chi-plots	symmetrical strong tail dependences							
Empirical Kendall's $\hat{\tau}$	0.72^{***}							
Chosen copula	*Student* with parameters $\hat{r} = 0.90$, $\hat{\nu} = 5.13$							

Figure F.7: 1ST D-VINE TREE: determining copula family for building block $C_{1,8}$
AEX $\sim$ OMXSPI



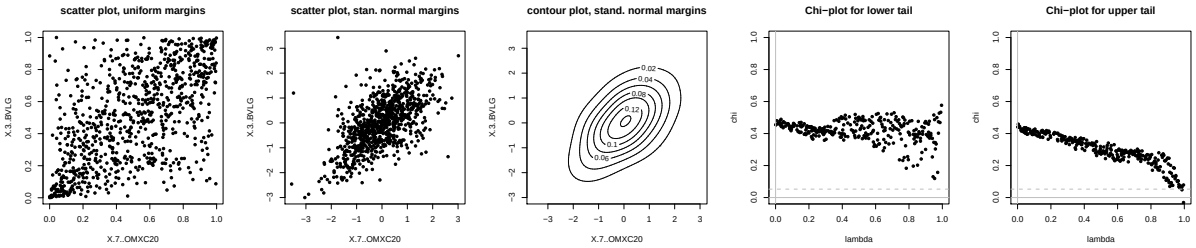
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	4	***6***	3	3	-4	-2	-6	-4
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.58$		$\hat{\lambda}_U^{t} = 0.44$		$\hat{\lambda}_U^{BB1} = 0.55$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.62$		$\hat{\lambda}_L^{t} = 0.44$		$\hat{\lambda}_L^{BB1} = 0.62$			
Chi-plots	nearly symmetrical middle tail dependences							
Empirical Kendall's $\hat{\tau}$	0.62^{***}							
Chosen copula	*BB1* with parameters $\hat{\theta} = 0.77$, $\hat{\delta} = 1.87$							

Figure F.8: 1ST D-VINE TREE: determining copula family for building block $C_{8,7}$ OMXSPI $\sim$ OMXC20



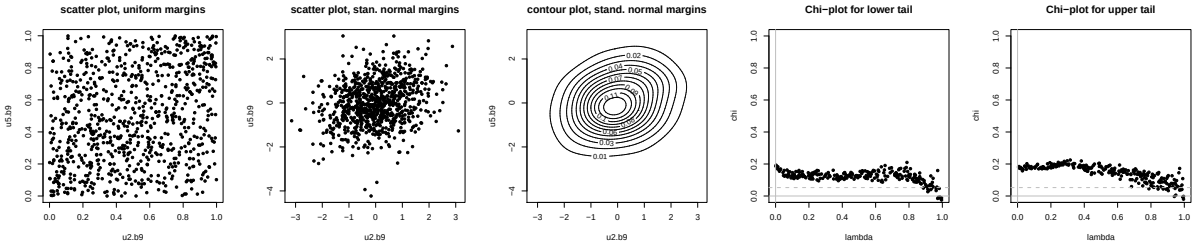
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	4	***5***	1	4	-2	-2	-4	-6
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.31$		$\hat{\lambda}_U^{t} = 0.20$		$\hat{\lambda}_U^{BB1} = 0.39$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.58$		$\hat{\lambda}_L^{t} = 0.20$		$\hat{\lambda}_L^{BB1} = 0.55$			
Chi-plots	not symmetrical tail dependences							
Empirical Kendall's $\hat{\tau}$	0.53***							
Chosen copula	*BB1* with parameters $\hat{\theta} = 0.81$, $\hat{\delta} = 1.46$							

Figure F.9: 1ST D-VINE TREE: determining copula family for building block $C_{7,3}$ OMXC20 $\sim$ BVLG



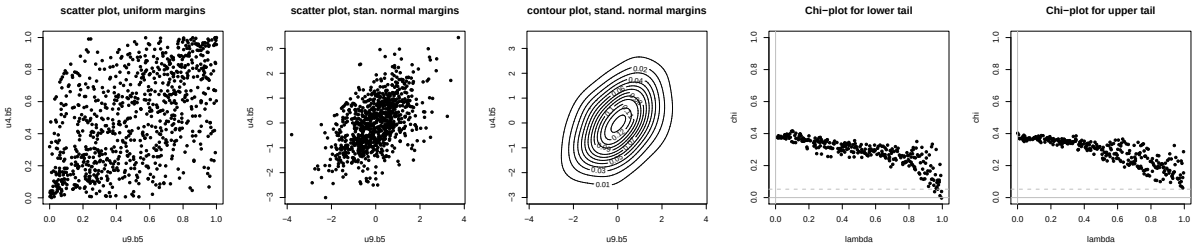
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	4	***5***	3	-2	-3	1	-1	-7
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.21$		$\hat{\lambda}_U^{t} = 0.30$		$\hat{\lambda}_U^{BB1} = 0.27$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.48$		$\hat{\lambda}_L^{t} = 0.30$		$\hat{\lambda}_L^{BB1} = 0.48$			
Chi-plots	not symmetrical tail dependences							
Empirical Kendall's $\hat{\tau}$	0.43***							
Chosen copula	*BB1* with parameters $\hat{\theta} = 0.75$, $\hat{\delta} = 1.26$							

Figure F.10: 2nd D-VINE TREE: determining copula family for building block $C_{2,5|9}$



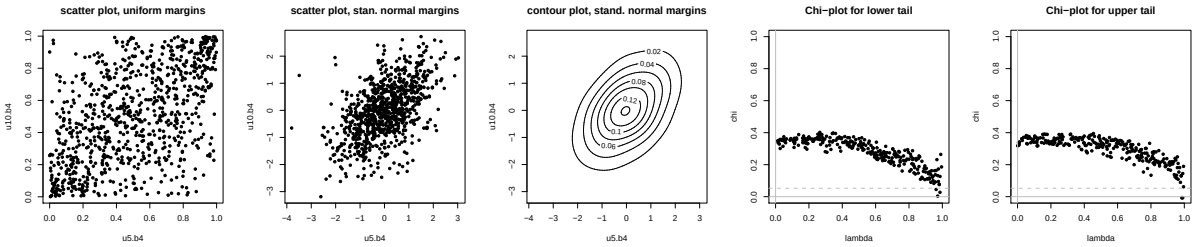
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	0	0	0	***1***	0	0	-1	0
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.14$		$\hat{\lambda}_U^{t} = 0.0006$		$\hat{\lambda}_U^{BB1} = 0.15$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.21$		$\hat{\lambda}_L^{t} = 0.0006$		$\hat{\lambda}_L^{BB1} = 0.002$			
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	0.17^{***}							
Chosen copula	*Gauss* with parameter $\hat{r} = 0.26$							

Figure F.11: 2nd D-VINE TREE: determining copula family for building block $C_{9,4|5}$



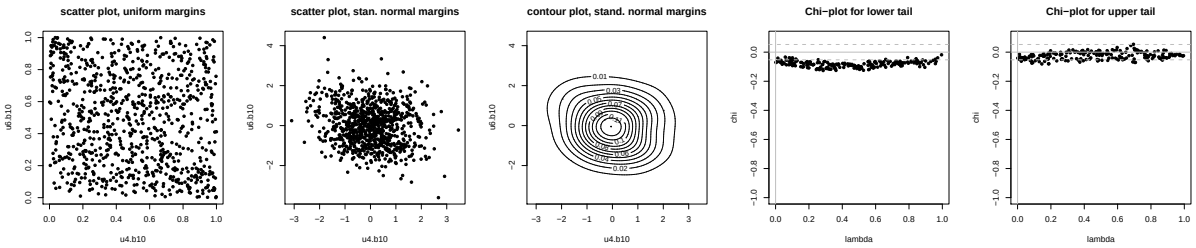
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	1	***2***	0	1	1	1	-7	1
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.24$		$\hat{\lambda}_U^{t} = 0.09$		$\hat{\lambda}_U^{BB1} = 0.29$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.31$		$\hat{\lambda}_L^{t} = 0.09$		$\hat{\lambda}_L^{BB1} = 0.25$			
Chi-plots	nearly symmetrical middle tail dependences							
Empirical Kendall's $\hat{\tau}$	0.37^{***}							
Chosen copula	*BB1* with parameter $\hat{\theta} = 0.37$, $\hat{\delta} = 1.30$							

Figure F.12: 2nd D-VINE TREE: determining copula family for building block $C_{5,10|4}$



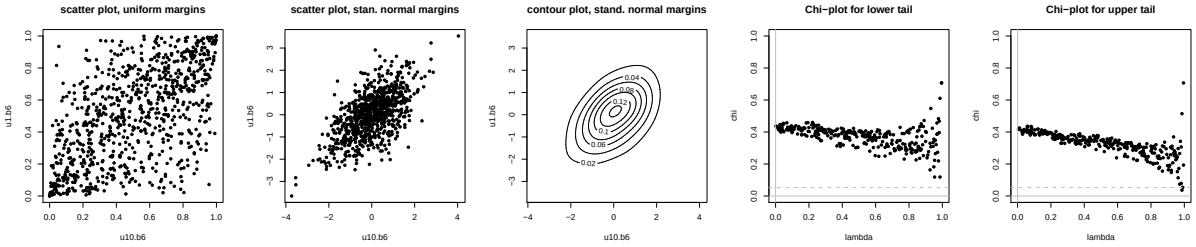
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	***3***	1	-1	1	1	1	-7	1
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.27$		$\hat{\lambda}_U^{t} = 0.12$		$\hat{\lambda}_U^{BB1} = 0.35$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.17$		$\hat{\lambda}_L^{t} = 0.12$		$\hat{\lambda}_L^{BB1} = 0.09$			
Chi-plots	symmetrical weak lower tail dependences							
Empirical Kendall's $\hat{\tau}$	0.36^{***}							
Chosen copula	*Student* with parameter $\hat{r} = 0.53$, $\hat{\nu} = 8.40$							

Figure F.13: 2nd D-VINE TREE: determining copula family for building block $C_{4,6|10}$



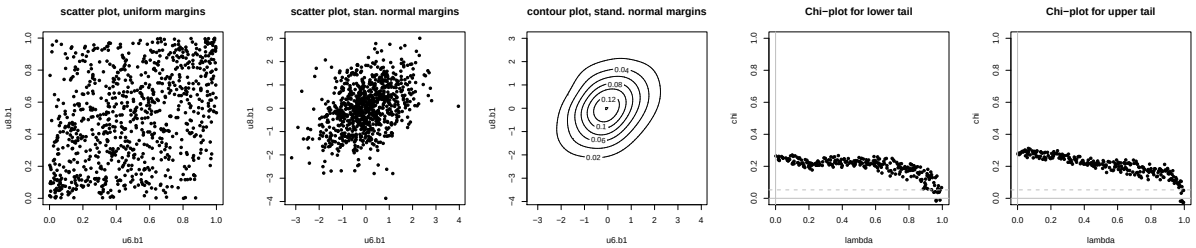
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	0	NA	NA	***2***	-2	0	NA	NA
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.07$		$\hat{\lambda}_U^{t} = 0.004$					
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0$		$\hat{\lambda}_L^{t} = 0.004$					
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	-0.09^{***}							
Chosen copula	*Gauss* with parameter $\hat{r} = -0.14$							

Figure F.14: 2nd D-VINE TREE: determining copula family for building block $C_{10,1|6}$



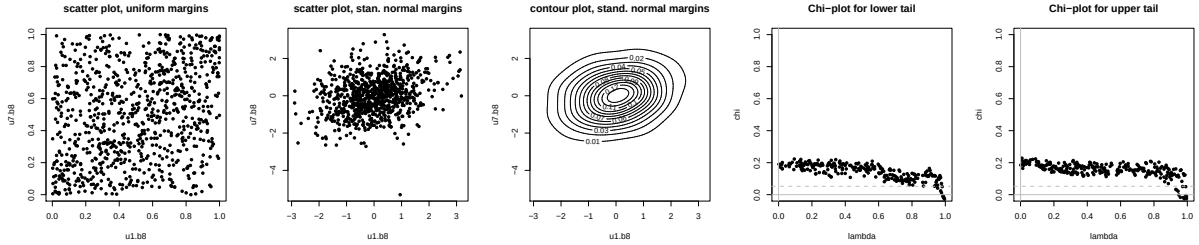
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	*4*	*4*	*4*	*4*	-5	-3	-4	-4
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.34$		$\hat{\lambda}_U^{t} = 0.18$		$\hat{\lambda}_U^{BB1} = 0.34$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.51$		$\hat{\lambda}_L^{t} = 0.18$		$\hat{\lambda}_L^{BB1} = 0.36$			
Chi-plots	not symmetrical tail dependences							
Empirical Kendall's $\hat{\tau}$	0.42^{***}							
Chosen copula	*BB1* with parameters $\hat{\theta} = 0.49$, $\hat{\delta} = 1.37$							

Figure F.15: 2nd D-VINE TREE: determining copula family for building block $C_{6,8|1}$



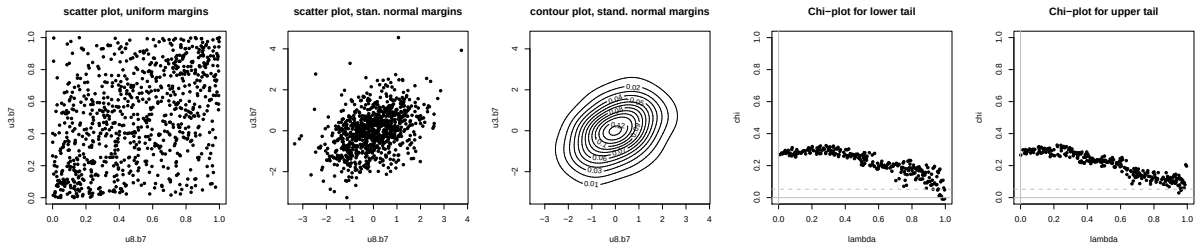
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	1	*2*	-1	1	1	1	-5	0
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.21$		$\hat{\lambda}_U^{t} = 0.013$		$\hat{\lambda}_U^{BB1} = 0.21$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.17$		$\hat{\lambda}_L^{t} = 0.013$		$\hat{\lambda}_L^{BB1} = 0.06$			
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	0.26^{***}							
Chosen copula	*Gauss* with parameters $\hat{r} = 0.39$							

Figure F.16: 2nd D-VINE TREE: determining copula family for building block $C_{1,7|8}$



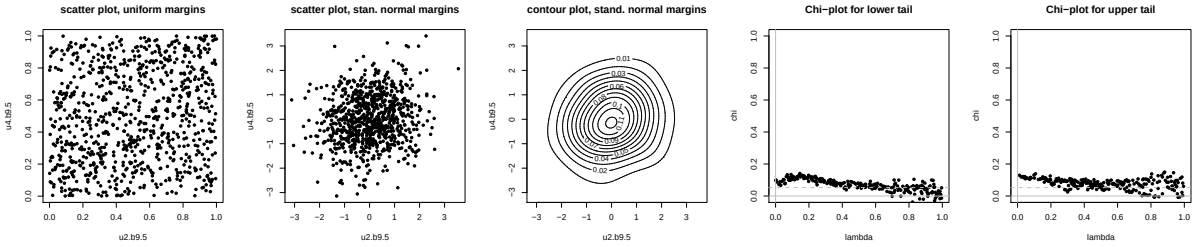
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	***1***	0	0	***1***	-1	***1***	-2	0
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.13$		$\hat{\lambda}_U^{t} = 0.013$		$\hat{\lambda}_U^{BB1} = 0.16$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.13$		$\hat{\lambda}_L^{t} = 0.013$		$\hat{\lambda}_L^{BB1} = 0.008$			
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	0.18***							
Chosen copula	*Gauss* with parameters $\hat{r} = 0.28$							

Figure F.17: 2nd D-VINE TREE: determining copula family for building block $C_{8,3|7}$



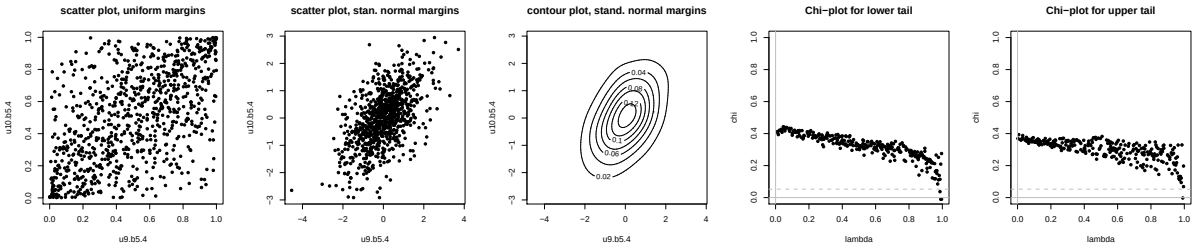
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	***3***	2	-2	1	1	2	-6	-2
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.14$		$\hat{\lambda}_U^{t} = 0.02$		$\hat{\lambda}_U^{BB1} = 0.20$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.14$		$\hat{\lambda}_L^{t} = 0.02$		$\hat{\lambda}_L^{BB1} = 0.12$			
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	0.28***							
Chosen copula	*Gauss* with parameters $\hat{r} = 0.43$							

Figure F.18: 3rd D-VINE TREE: determining copula family for building block $C_{2,4|9,5}$



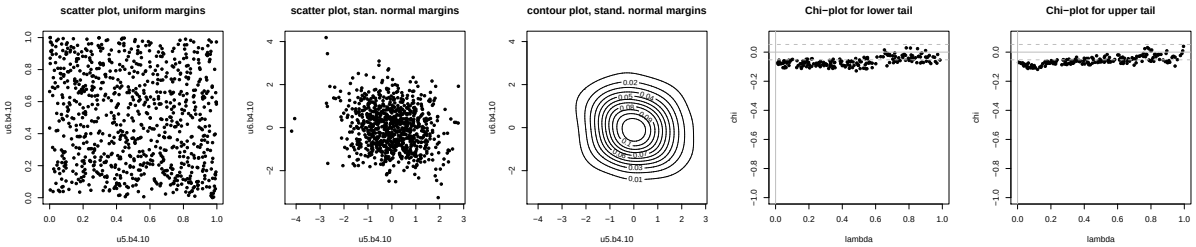
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	-1	-1	-1	***1***	0	0	0	2
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.17$		$\hat{\lambda}_U^{t} = 0$		$\hat{\lambda}_U^{BB1} = 0.09$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.03$		$\hat{\lambda}_L^{t} = 0$		$\hat{\lambda}_L^{BB1} = 0$			
Chi-plots	no lower and weak upper tail dependence							
Empirical Kendall's $\hat{\tau}$	0.10^{***}							
Chosen copula	*Gauss* with parameter $\hat{r} = 0.15$							

Figure F.19: 3rd D-VINE TREE: determining copula family for building block $C_{9,10|5,4}$



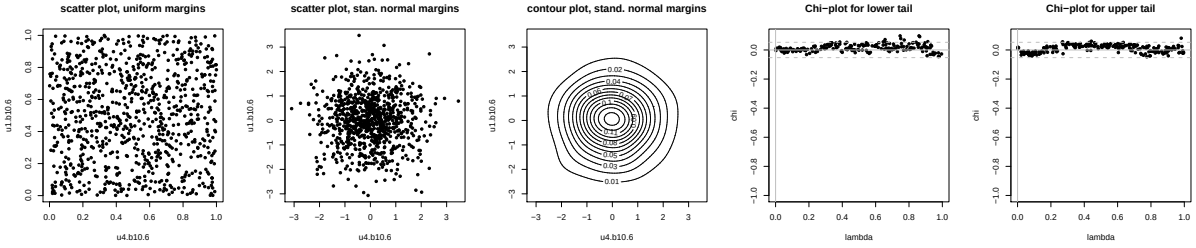
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	***3***	***3***	1	2	-3	2	-7	-1
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.34$		$\hat{\lambda}_U^{t} = 0.08$		$\hat{\lambda}_U^{BB1} = 0.29$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.27$		$\hat{\lambda}_L^{t} = 0.08$		$\hat{\lambda}_L^{BB1} = 0.15$			
Chi-plots	nearly symmetrical middle tail dependences							
Empirical Kendall's $\hat{\tau}$	0.38^{***}							
Chosen copula	*BB1* with parameter $\hat{\theta} = 0.31$, $\hat{\delta} = 1.36$							

Figure F.20: 3rd D-VINE TREE: determining copula family for building block $C_{5,6|4,10}$



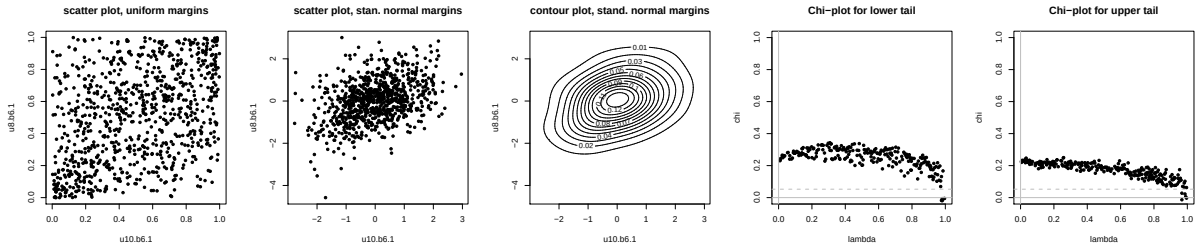
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	0	NA	NA	0	-1	***1***	NA	NA
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.07$		$\hat{\lambda}_U^{t} = 0.003$					
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.03$		$\hat{\lambda}_L^{t} = 0.003$					
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	-0.10^{***}							
Chosen copula	*Gauss* with parameter $\hat{r} = -0.15$							

Figure F.21: 3rd D-VINE TREE: determining copula family for building block $C_{4,1|10,6}$



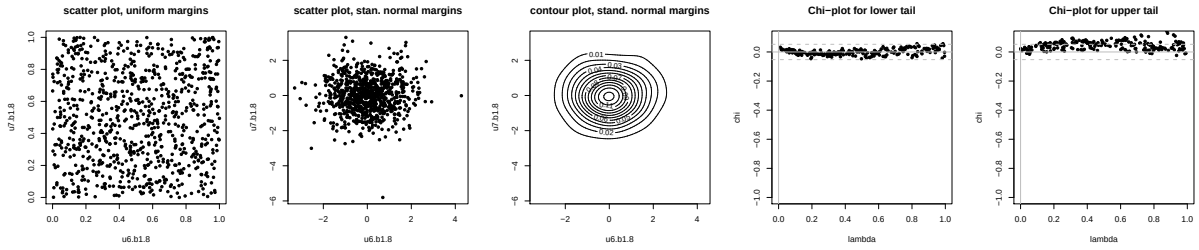
Chi-plots	no tail dependences	**Empirical Kendall's $\hat{\tau}$**	0.004, p-value = 0.86
Chosen copula	*Independent* copula		

Figure F.22: 3rd D-VINE TREE: determining copula family for building block $C_{10,8|6,1}$



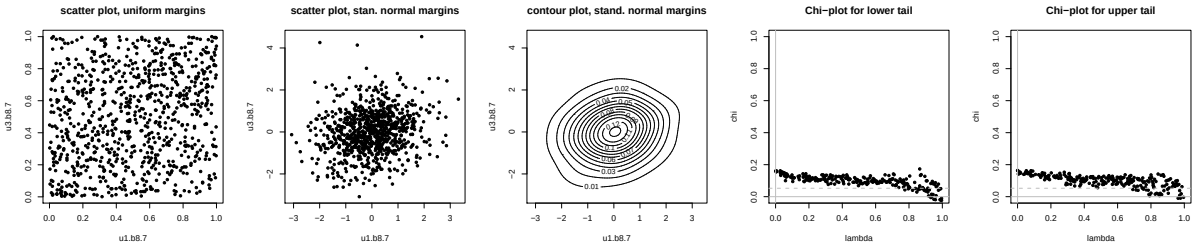
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	1	1	1	1	-1	***2***	0	-5
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.21$		$\hat{\lambda}_U^{t} = 0.016$		$\hat{\lambda}_U^{BB1} = 0.16$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.21$		$\hat{\lambda}_L^{t} = 0.016$		$\hat{\lambda}_L^{BB1} = 0.17$			
Chi-plots	approximately no tail dependences							
Empirical Kendall's $\hat{\tau}$	0.26^{***}							
Chosen copula	*Gauss* with parameter $\hat{r} = 0.39$							

Figure F.23: 3rd D-VINE TREE: determining copula family for building block $C_{6,7|1,8}$



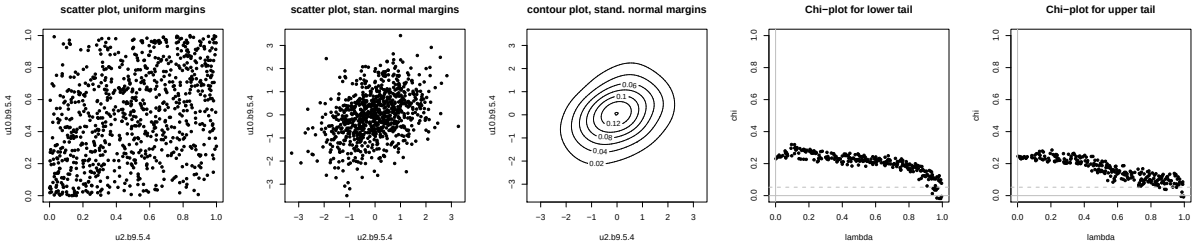
Chi-plots	no tail dependences	**Empirical Kendall's $\hat{\tau}$**	0.015, p-value = 0.5
Chosen copula	*Independent* copula		

Figure F.24: 3rd D-VINE TREE: determining copula family for building block $C_{1,3|8,7}$



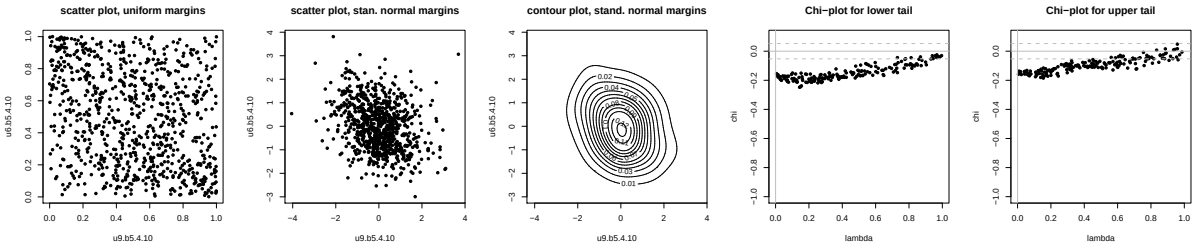
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	0	0	0	0	0	0	0	0
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.14$		$\hat{\lambda}_U^{t} = 0.002$		$\hat{\lambda}_U^{BB1} = 0.07$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.10$		$\hat{\lambda}_L^{t} = 0.002$		$\hat{\lambda}_L^{BB1} = 0.007$			
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	0.13^{***}							
Chosen copula	*Gauss* with parameter $\hat{r} = 0.19$							

Figure F.25: 4th D-VINE TREE: determining copula family for building block $C_{2,10|9,5,4}$



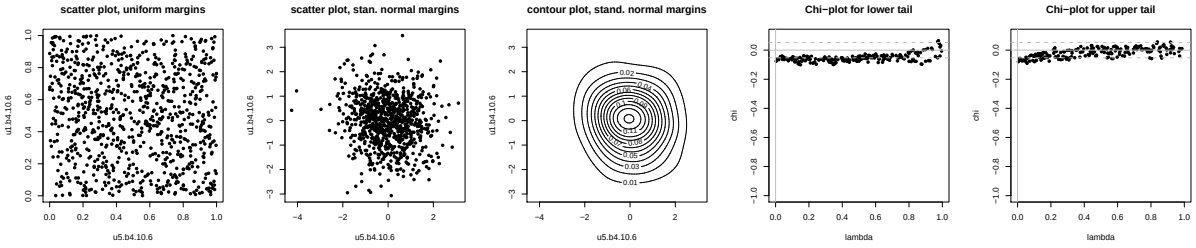
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	0	1	-1	***3***	1	1	-1	-4
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.13$		$\hat{\lambda}_U^{t} = 0$		$\hat{\lambda}_U^{BB1} = 0.16$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.13$		$\hat{\lambda}_L^{t} = 0$		$\hat{\lambda}_L^{BB1} = 0.06$			
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	0.25^{***}							
Chosen copula	*Gauss* with parameter $\hat{r} = 0.38$							

Figure F.26: 4th D-VINE TREE: determining copula family for building block $C_{9,6|5,4,10}$



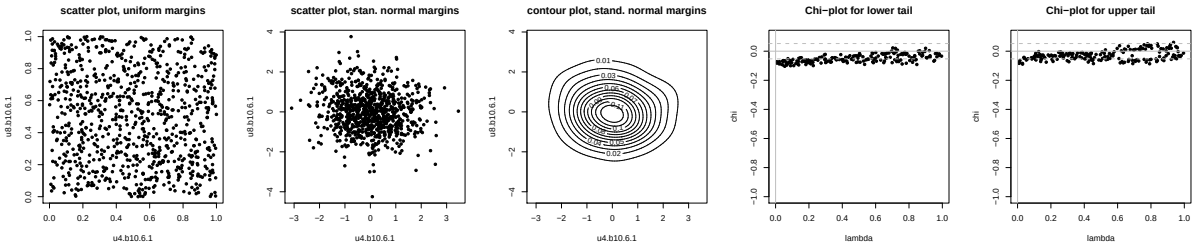
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	0	NA	NA	0	-1	***1***	NA	NA
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.07$		$\hat{\lambda}_U^{t} = 0.001$					
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0$		$\hat{\lambda}_L^{t} = 0.001$					
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	-0.18^{***}							
Chosen copula	*Gauss* with parameter $\hat{r} = -0.28$							

Figure F.27: 4th D-VINE TREE: determining copula family for building block $C_{5,1|4,10,6}$



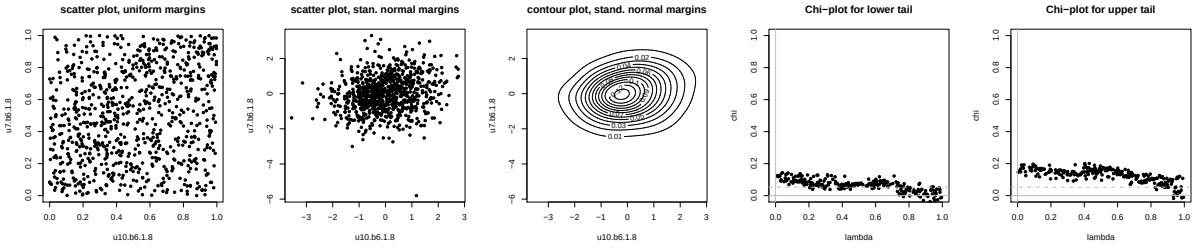
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	0	NA	NA	0	-1	***1***	NA	NA
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.07$		$\hat{\lambda}_U^{t} = 0.0003$					
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0$		$\hat{\lambda}_L^{t} = 0.0003$					
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	-0.07, p-value $= 0.003$							
Chosen copula	*Gauss* with parameter $\hat{r} = -0.105$							

Figure F.28: 4th D-VINE TREE: determining copula family for building block $C_{4,8|10,6,1}$



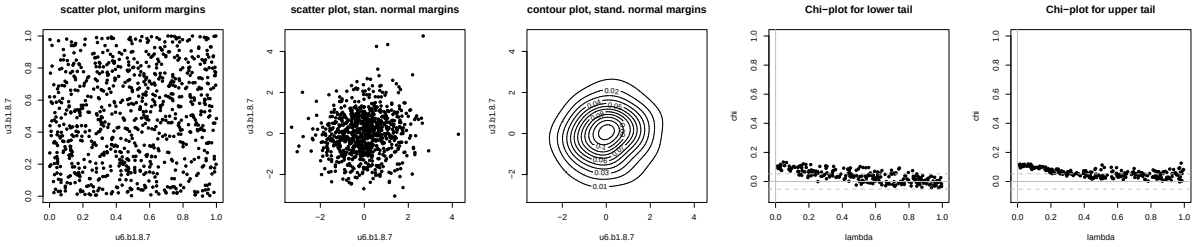
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	-3	NA	NA	***1***	***1***	***1***	NA	NA
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0$		$\hat{\lambda}_U^{t} = 0$					
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.03$		$\hat{\lambda}_L^{t} = 0$					
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	-0.06, p-value $= 0.004$							
Chosen copula	*Gauss* with parameter $\hat{r} = -0.10$							

Figure F.29: 4th D-VINE TREE: determining copula family for building block $C_{10,7|6,1,8}$



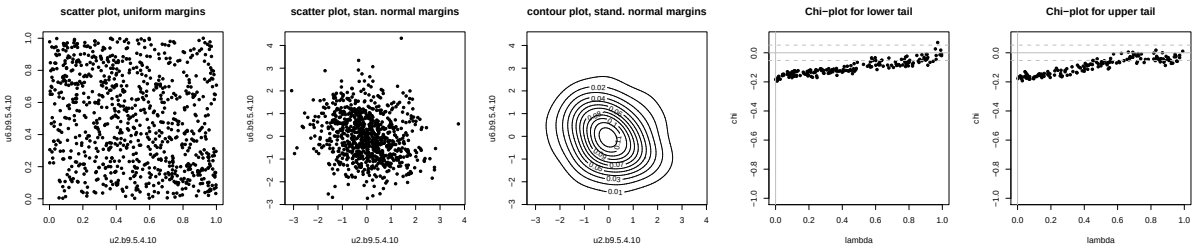
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	-2	0	-2	2	2	2	-5	***3***
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.10$		$\hat{\lambda}_U^{t} = 0$		$\hat{\lambda}_U^{BB1} = 0.14$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.03$		$\hat{\lambda}_L^{t} = 0$		$\hat{\lambda}_L^{BB1} = 0$			
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	0.13^{***}							
Chosen copula	*Gauss* with parameter $\hat{r} = 0.19$							

Figure F.30: 4th D-VINE TREE: determining copula family for building block $C_{6,3|1,8,7}$



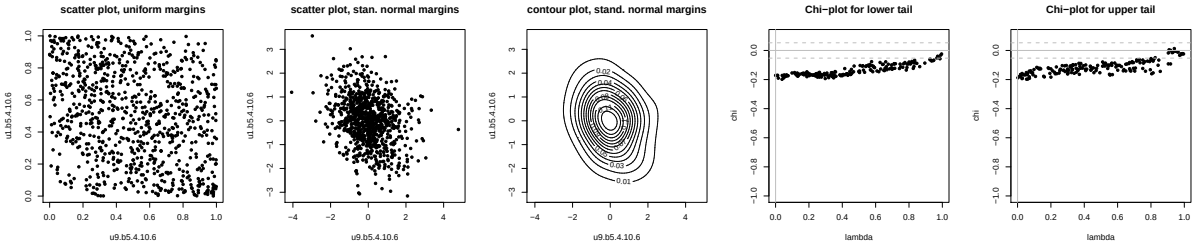
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	-4	-1	-1	2	2	2	-3	***3***
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.06$		$\hat{\lambda}_U^{t} = 0$		$\hat{\lambda}_U^{BB1} = 0.05$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.03$		$\hat{\lambda}_L^{t} = 0$		$\hat{\lambda}_L^{BB1} = 0$			
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	0.08, p-value = 0.0004							
Chosen copula	*Gauss* with parameter $\hat{r} = 0.12$							

Figure F.31: 5th D-VINE TREE: determining copula family for building block $C_{2,6|9,5,4,10}$



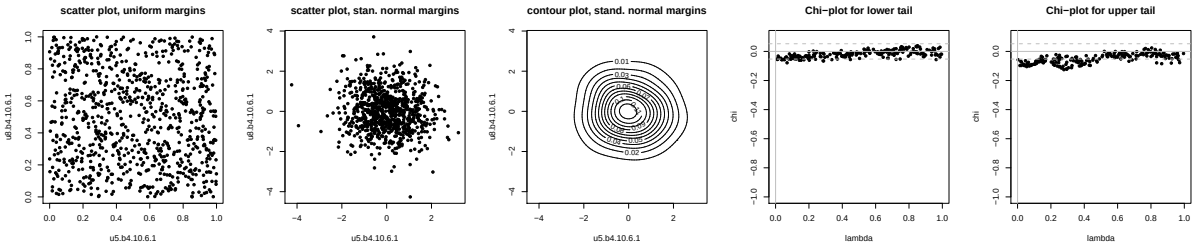
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	-2	NA	NA	-2	***2***	***2***	NA	NA
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0$		$\hat{\lambda}_U^{t} = 0$					
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0$		$\hat{\lambda}_L^{t} = 0$					
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	-0.16^{***}							
Chosen copula	*Gauss* with parameter $\hat{r} = -0.24$							

Figure F.32: 5th D-VINE TREE: determining copula family for building block $C_{9,1|5,4,10,6}$



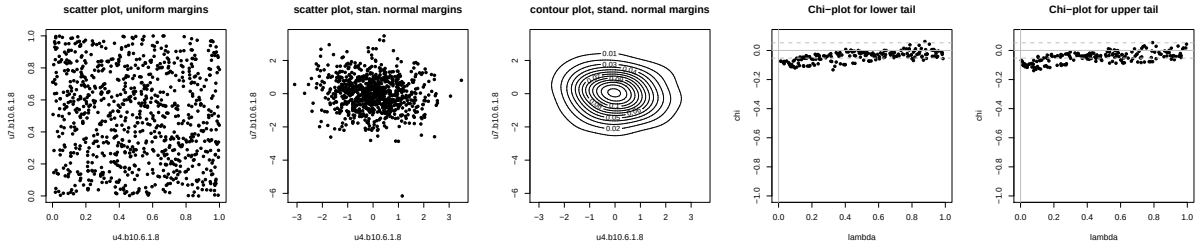
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	-1	NA	NA	***3***	-1	-1	NA	NA
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0$		$\hat{\lambda}_U^{t} = 0$					
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0$		$\hat{\lambda}_L^{t} = 0$					
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	-0.18^{***}							
Chosen copula	*Gauss* with parameter $\hat{r} = -0.27$							

Figure F.33: 5th D-VINE TREE: determining copula family for building block $C_{5,8|4,10,6,1}$



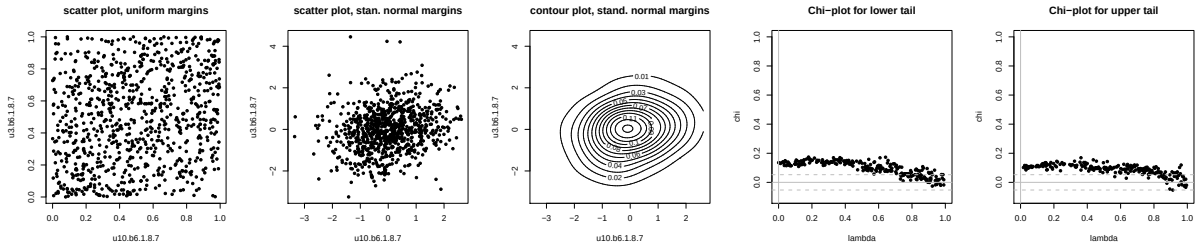
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	0	NA	NA	0	0	0	NA	NA
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.03$		$\hat{\lambda}_U^{t} = 0$					
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.03$		$\hat{\lambda}_L^{t} = 0$					
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	-0.06, p-value$=0.009$							
Chosen copula	*Gauss* with parameter $\hat{r} = -0.09$							

Figure F.34: 5th D-VINE TREE: determining copula family for building block $C_{4,7|10,6,1,8}$



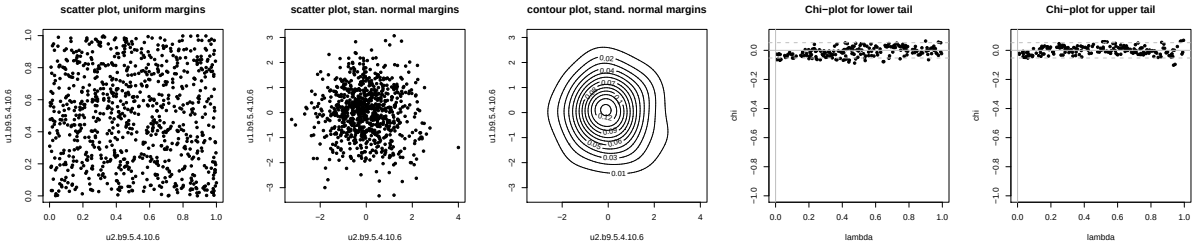
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	0	NA	NA	0	-1	***1***	NA	NA
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0$		$\hat{\lambda}_U^{t} = 0$					
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.06$		$\hat{\lambda}_L^{t} = 0$					
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	-0.07, p-value$=0.001$							
Chosen copula	*Gauss* with parameter $\hat{r} = -0.11$							

Figure F.35: 5th D-VINE TREE: determining copula family for building block $C_{10,3|6,1,8,7}$



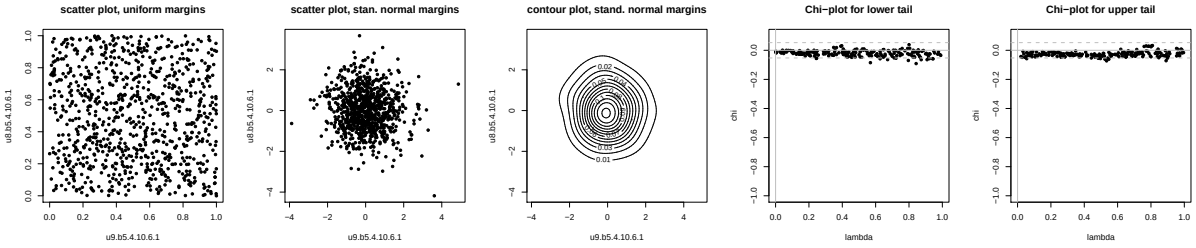
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	-1	-4	-4	***4***	***4***	***4***	0	-3
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.07$		$\hat{\lambda}_U^{t} = 0$		$\hat{\lambda}_U^{BB1} = 0.06$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.07$		$\hat{\lambda}_L^{t} = 0$		$\hat{\lambda}_L^{BB1} = 0.01$			
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	0.13^{***}							
Chosen copula	*Gauss* with parameter $\hat{r} = 0.19$							

Figure F.36: 6th D-VINE TREE: determining copula family for building block $C_{2,1|9,5,4,10,6}$



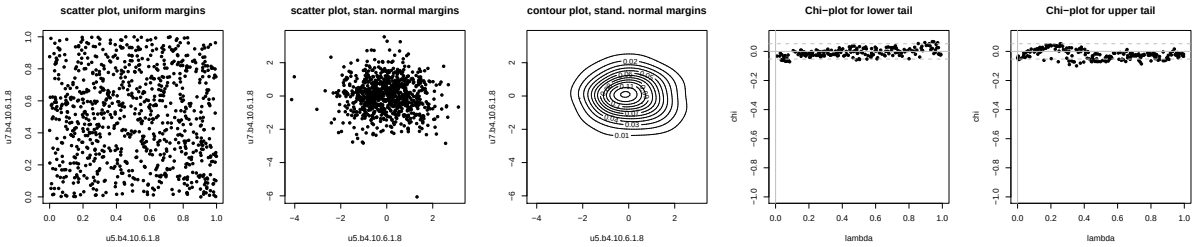
Chi-plots	no tail dependences	**Empirical Kendall's $\hat{\tau}$**	-0.023, p-value $= 0.3$
Chosen copula	*Independent* copula		

Figure F.37: 6th D-VINE TREE: determining copula family for building block $C_{9,8|5,4,10,6,1}$



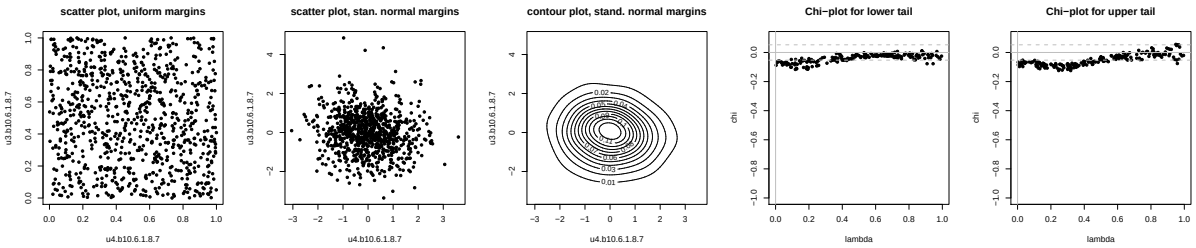
Chi-plots	no tail dependences	**Empirical Kendall's $\hat{\tau}$**	-0.03, p-value $= 0.15$
Chosen copula	*Independent* copula		

Figure F.38: 6th D-VINE TREE: determining copula family for building block $C_{5,7|4,10,6,1,8}$



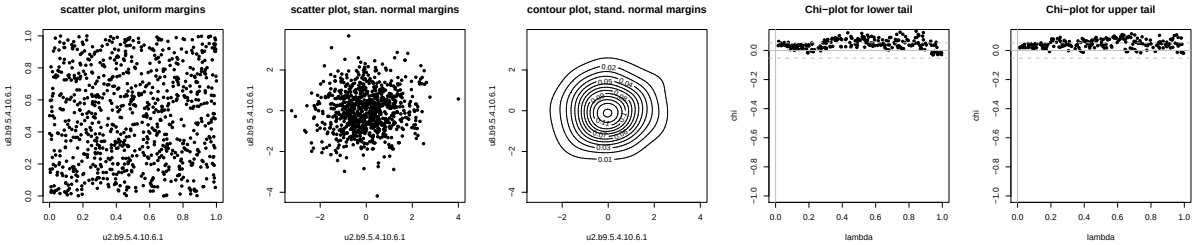
Chi-plots	no tail dependences	**Empirical Kendall's $\hat{\tau}$**	-0.025, p-value $= 0.25$
Chosen copula	*Independent* copula		

Figure F.39: 6th D-VINE TREE: determining copula family for building block $C_{4,3|10,6,1,8,7}$



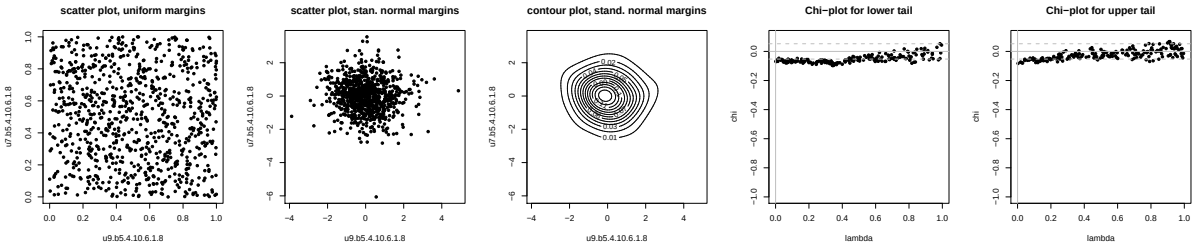
	Student	**BB1**	**BB7**	**Gauss**	**Frank**	**Plackett**	**Clayton**	**Gumbel**
Score Vuong test	-3	NA	NA	***1***	***1***	***1***	NA	NA
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.06$		$\hat{\lambda}_U^{t} = 0$					
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.03$		$\hat{\lambda}_L^{t} = 0$					
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	-0.074, p-value $= 0.001$							
Chosen copula	*Gauss* with parameter $\hat{r} = -0.099$							

Figure F.40: 7th D-VINE TREE: determining copula family for building block $C_{2,8|9,5,4,10,6,1}$



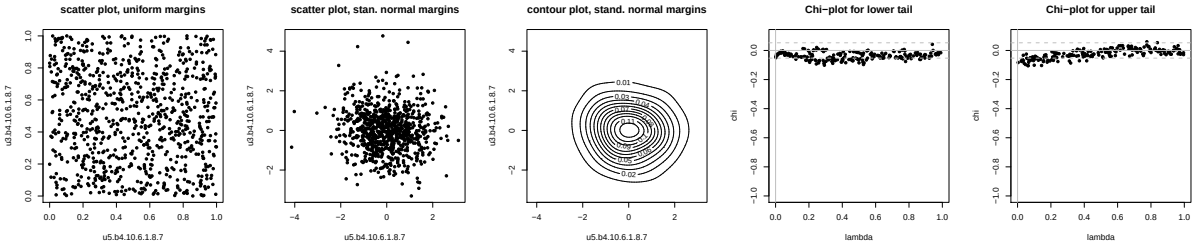
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	0	-1	-1	***2***	-1	-1	0	***2***
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.06$		$\hat{\lambda}_U^{t} = 0$		$\hat{\lambda}_U^{BB1} = 0.04$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.10$		$\hat{\lambda}_L^{t} = 0$		$\hat{\lambda}_L^{BB1} = 0$			
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	0.05, p-value = 0.02							
Chosen copula	*Gauss* with parameter $\hat{r} = 0.07$							

Figure F.41: 7th D-VINE TREE: determining copula family for building block $C_{9,7|5,4,10,6,1,8}$



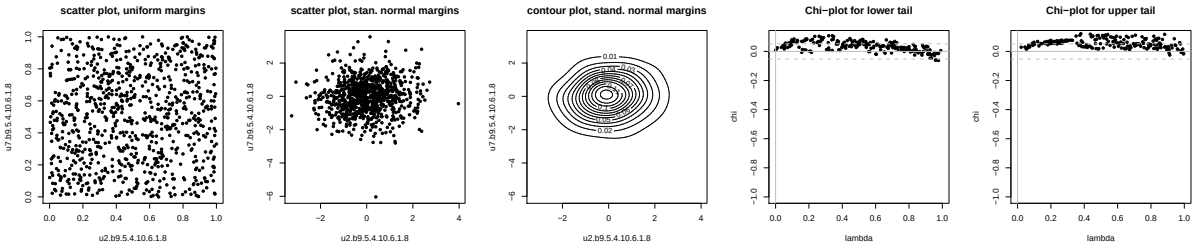
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	-3	NA	NA	***1***	***1***	***1***	NA	NA
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0$		$\hat{\lambda}_U^{t} = 0$					
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.03$		$\hat{\lambda}_L^{t} = 0$					
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	-0.05, p-value = 0.02							
Chosen copula	*Gauss* with parameter $\hat{r} = -0.08$							

Figure F.42: 7th D-VINE TREE: determining copula family for building block $C_{5,3|4,10,6,1,8,7}$



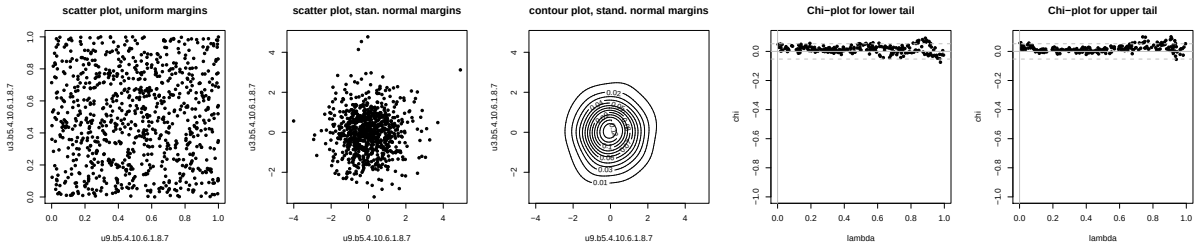
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	-3	NA	NA	***1***	***1***	***1***	NA	NA
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.03$		$\hat{\lambda}_U^{t} = 0$					
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0$		$\hat{\lambda}_L^{t} = 0$					
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	-0.05, p-value $= 0.02$							
Chosen copula	*Gauss* with parameter $\hat{r} = -0.08$							

Figure F.43: 8th D-VINE TREE: determining copula family for building block $C_{2,7|9,5,4,10,6,1,8}$



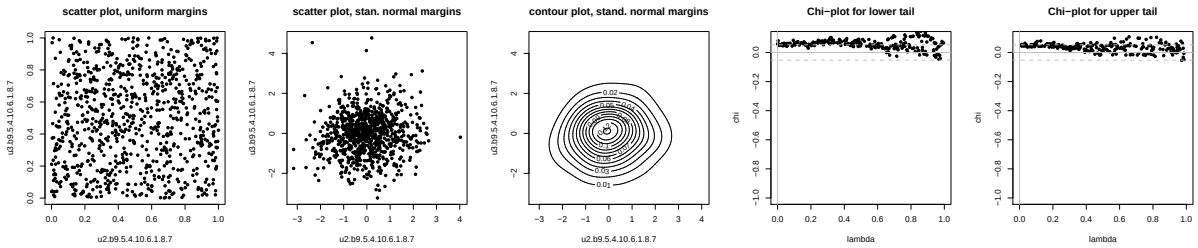
	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	-4	-1	-1	1	1	1	0	***3***
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.17$		$\hat{\lambda}_U^{t} = 0$		$\hat{\lambda}_U^{BB1} = 0.06$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.03$		$\hat{\lambda}_L^{t} = 0$		$\hat{\lambda}_L^{BB1} = 0$			
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	0.06, p-value $= 0.01$							
Chosen copula	*Gauss* with parameter $\hat{r} = 0.09$							

Figure F.44: 8th D-VINE TREE: determining copula family for building block $C_{9,3|5,4,10,6,1,8,7}$



Chi-plots	no tail dependences	**Empirical Kendall's $\hat{\tau}$**	0.11, p-value = 0.6
Chosen copula	*Independent* copula		

Figure F.45: 9th D-VINE TREE: determining copula family for building block $C_{2,3|9,5,4,10,6,1,8,7}$



	Student	BB1	BB7	Gauss	Frank	Plackett	Clayton	Gumbel
Score Vuong test	-3	-1	-1	***3***	1	1	0	0
Upper tail dependence	$\hat{\lambda}_U^{emp} = 0.06$		$\hat{\lambda}_U^{t} = 0$		$\hat{\lambda}_U^{BB1} = 0.02$			
Lower tail dependence	$\hat{\lambda}_L^{emp} = 0.06$		$\hat{\lambda}_L^{t} = 0$		$\hat{\lambda}_L^{BB1} = 0$			
Chi-plots	no tail dependences							
Empirical Kendall's $\hat{\tau}$	0.06, p-value = 0.01							
Chosen copula	*Gauss* with parameter $\hat{r} = 0.09$							