University of Cologne
Institute of Econometrics and Statistics

Detecting Structural Breaks in Factor Copula Models and in Vectors of Dependence Measures

Dissertation zur Erlangung des Doktorgrades der Wirtschafts- und Sozialwissenschaftlichen Fakultät der Universität zu Köln

2019

M. Sc. Florian Stark

aus

Euskirchen

Referent: Univ.-Prof. Dr. Dominik Wied
Korreferent: Prof. i.R. Dr. Josef G. Steinebach
Tag der Promotion: 26.03.2019

Danksagung

An dieser Stelle möchte ich meinen besonderen Dank nachstehenden Personen entgegenbringen, ohne deren Mithilfe die Anfertigung dieser Promotionsschrift niemals zustande gekommen wäre:

Mein Dank gilt zunächst Prof. Dr. Dominik Wied, für die Betreuung dieser Arbeit, der freundlichen Hilfe und der vielfältigen Ideengebung über die ganze Zeit.

Ich danke Prof. i.R. Dr. Josef G. Steinebach für die hilfsbereite und wissenschaftliche Betreuung als Zweitgutachter.

Ferner danke ich Prof. Dr. Hans Manner für die dauerhafte Betreuung meiner Arbeiten. Die netten sowie hilfreichen Telefonate trugen stets für ein gutes Gelingen bei.

Zudem bedanke ich mich bei der Deutschen Forschungsgemeinschaft (DFG) für die Finanzierung und Förderung dieses Projekts.

Vielen Dank auch an die gesamte Arbeitsgruppe für die freundschaftliche Arbeitsatmosphäre, viele wertvolle Anregungen und stete Hilfsbereitschaft, die wesentlich zum Gelingen dieser Arbeit beigetragen haben.

Besonders herzlich möchte ich mich bei meinen Eltern, meiner Oma, Dieter, Cari und meinen Freunden für die uneingeschränkte, liebevolle und vielseitige Unterstützung während meines Studiums bedanken, ohne die diese Arbeit so nicht möglich gewesen wäre.

Contents

1. INTRODUCTION AND MOTIVATION

The following work is based on the Papers ”Testing for structural breaks in factor copula models“, Manner, Stark, and Wied (2019) published in the *Journal of Econometrics* and the working Paper ”A monitoring procedure for detecting structural breaks in factor copula models“, Manner, Stark, and Wied (2018) such as my sole working Paper “On the applicability of a nonparametric test for constant copula-based dependence measures: Dating break points and analyzing different dependence measure sets”, Stark (2018). The first two papers are joint works with Hans Manner and Dominik Wied. Especially, I focused on the theoretical details and the implementation of the procedures in Matlab such as the realization of the simulation studies and the empirical applications. The joint works are presented in Section 3 and Section 4, where my sole work is presented in Section 5.
Further, during my Ph.D. studies I participated in the work of the working paper ”A non-parametric CUSUM-type test for testing relevant change in copulas“, Kutzker, Stark, and Wied (2018). All the work was supported by the *Deutsche Forschungsgemeinschaft* (DFG).

Dependence models based on copula functions have been an important topic for researchers and practitioner in the last 20 years (see Patton, 2012 and Fan and Patton, 2014 for reviews). These models offer an elegant approach for modelling multivariate distributions that has proven to be useful in many fields such as risk management, asset allocation or option pricing. Multivariate GARCH models (e.g. Engle, 2002 or Bauwens, Laurent, and Rombouts, 2006) or multivariate stochastic volatility model (Yu and Meyer, 2006) are the traditional way to model multivariate asset prices, but these models typically come with the drawback that they rely on the multivariate normal distribution, which contrasts stylized facts about the distribution of asset prices, in particular regarding the dependence structure. A number of parametric copula models exist that can capture the tail dependence and asymmetric dependence structure present in financial time series. More recently there have been two key advances in the literature on parametric copula modelling. First, the need for time-varying dependence has been recognized and a number of modelling approaches have been proposed. Patton (2006) extended Sklar's theorem for conditional distributions and proposed a simple observation driven model for the evolution of the copula parameter over time. Dias and Embrechts (2004) test for structural breaks at unknown dates using a sup LR statistic, whereas Garcia and Tsafack (2011), Stöber and Czado (2014) or Chollete, Heinen, and Valdesogo (2009) rely on Markov switching models assuming regime dependent parameters. A model that assumes a smooth evolution over time is proposed by Hafner and Reznikova (2010). A state space approach in which the copula parameter is driven by a latent process was advocated by Hafner and Manner (2012), whereas Creal, Koopman, and Lucas (2013) suggest a generalized autoregressive score model for time varying dependence.

A second innovation in the copula literature has been the availability of parametric models that are applicable in higher dimensional settings. Besides the obvious choice of elliptical copulas, typically Gaussian and Student copulas, three main approaches can be found in the literature. Within the class of Archimedean copulas hierarchical models have been studied by Savu and Trede (2010) and Okhrin, Okhrin, and Schmid (2013). However, in larger dimensions these models are still rather restrictive. A more popular approach is the class of vine copulas studied in Bedford and Cooke (2002), Aas, Czado, Frigessi, and Bakken (2009), Stöber and Czado (2011), Stöber, Joe, and Czado (2013) or Brechmann and Czado (2013). A time varying vine copula model has been proposed by Almeida, Czado, and Manner (2016). Finally, Oh and Patton (2017) and Krupskii and Joe (2013) introduced the class of factor copula models. Factor copulas are the copulas implied by a latent factor model, where the difference to traditional factor models is the fact that one is only interested in the copula implied by the factor structure, discarding its marginal information. The advantage of these models is that they can be used in relatively high dimensional applications and nevertheless capture the dependence structure by a low number of parameters. However, the estimation of this model is complicated by the fact that the factors are not observable. Several approaches have been proposed to tackle this problem. Oh and Patton (2013) suggest a simulated method of moments estimator, an approach that we adapt in this work. Krupskii and Joe (2013) propose maximum likelihood estimation by numerically integrating out the latent factor. This approach has the drawback that it is only applicable when the number of factors is relatively small. Murry, Dunson, Carin, and Lucas (2013) estimate a Gaussian Factor copula model with Bayesian methods. Factor copula models that allow for time-varying parameters have been proposed by Creal and Tsay (2015), who allow for stochastic autoregressive factor loading estimated with a Bayesian approach. An alternative approach can be found in Oh and Patton (2018) where the dynamics of the factor loadings are driven by a generalized autoregressive score model. This model is estimated using a multi stage maximum likelihood approach.

The aim of this work is to propose a different approach to allow for time-variation in factor copula models by testing for and dating breakpoints at unknown points in time. Several tests for constant dependencies have recently been developed, see e.g. Bücher and Ruppert (2013) for the case of copulas or Dehling, Vogel, Wendler, and Wied (2017) for the case of Kendall's tau. The main motivation for such tests is that dependencies usually increase in times of crises. Therefore, they can be applied to detect and quantify contagion between different financial markets or to construct optimal portfolios in portfolio management.

For the estimation of the model parameters, we rely on the simulated method of moments (SMM), which is different to standard method of moments applications, since the theoretical moment-counterparts are not available analytically and therefore need to be simulated. This complicates the derivation of results regarding the consistency and asymptotic distribution of the estimators. The reason is that the objective function is not continuous and furthermore not differentiable in the parameters and standard asymptotic approaches cannot be used here.

We first propose a new retro perspective fluctuation test, where successively parameter estimators are compared to the parameter estimates of the full sample and we then analyse the behaviour of the test under the null hypothesis of no parameter change (Section 3). In contrast to formerly proposed non-parametric tests for constant copulas by e.g. Bücher, Kojadinovic, Rohmer, and Segers (2014), our test is of parametric nature. The asymptotic distribution of the test statistic is non-trivial. Due to the non-smoothness of the objective function, we cannot make use of a Taylor expansion approach to derive the distribution under the null. To tackle this issue we propose a new construction principle inspired by Newey and McFadden (1994). These new functional limit theorems hold in general for SMM estimation and are therefore of broader interest. As the asymptotic distribution depends on unknown quantities we propose a bootstrap to estimate these.

In the context of retro perspective testing, i.e. no real time testing, we propose two possible tests, namely a fluctuation test based on parameter estimates and a test directly based on the dependence measure vector used to estimate the model. We analyze size and power properties of our tests in Monte Carlo simulation in various situations and compare our tests with the copula constancy test proposed by Bücher et al. (2014). While the Bücher et al. (2014) test has better properties for low dimensions, our test performs better in high dimensions. This reflects the fact that the drawback of having to estimate the model with simulated methods is more and more compensated with increasing dimensions. If the number of dimensions is kept fixed, one simply has more data for estimating the model, while, on the other hand, in a nonparametric copula constancy test, the complexity of the estimated objects increase. Finally, we provide an application to a set of stock returns from the Eurostoxx50.

After dealing with a retro-perspective approach for detecting changes in the copula parameters we are interested in deriving a procedure for real time applicability. There are many papers which deal with monitoring procedures for detecting structural changes, for instance Hoga and Wied (2017), who construct a sequential monitoring procedure for changes in the tail index and extreme quantiles of beta-mixing random variables, which can be based on a large class of tail index estimators. Furthermore, Pape, Wied, and Galeano (2017) propose a model-independent multivariate sequential procedure to monitor changes in the vector of componentwise unconditional variances in a sequence of p-variate random vectors, where Galeano and Wied (2013) developed a monitoring procedure to test for the constancy of the correlation coefficient of a sequence of random variables. Here the basic idea is that a historical sample is available and the goal is to monitor for changes in the correlation as new data become available. All these proposed monitoring procedures have in common that they are all of non parametric kind. A parametric approach for detecting structural breaks is shown for example in Chu, Stinchcombe, and White (1996), who construct real-time CUSUM based monitoring procedures to detect changes in the parameters of linear regression models, where they assume parameter constancy for an initial period. Also, Kurozumi (2017) propose a monitoring test for parameter change in linear regression models with endogenous regressors. In this article they consider a CUSUM-type test based on the instrumental variable (IV) estimation, as the IV method is standard for models with endogenous regressors.

The second aim of this work is to construct a new parametric monitoring procedure, based on moving sums (MOSUM), for the parameters in factor copula models, where rolling window parameter estimates are compared to a parameter estimate of an historical data sample, where we can assume constant parameter values (Section 4). By using rolling window parameter estimates, based on moving sums, new data has more impact on the estimated parameter, yielding higher power of our procedure. Further, a similar non-parametric monitoring procedure based on the dependence measure vector used in the SMM procedure is proposed. The pre proposed retro perspective parameter test and the non-parametric dependence measure test are useful to test the assumption of no parameter or dependence measure change within

the historical data set. We then analyze size and power properties of our procedures in single and multi break situations in Monte Carlo simulations. Finally, we use the parameter monitoring procedure in a real-data application during the last financial crisis.
Given the growing need for managing financial risk, risk prediction plays an increasing role in banking and finance. The value-at-Risk (VaR), is the most prominent measure of financial market risk. Despite it has been criticized as being theoretically not efficient and numerically problematic compare Dowd and Blake (2006), it is still the most widely used risk measure in practice. The number of methods for such calculations continues to increase. The theoretical and computational complexity of VaR models for calculating capital requirements is also increasing. Some examples include the use of extreme value theory see McNeil and Frey (2000), quantile regression methods see Manganelli and Engle (2004) and Markov switching techniques see Gray (1996) and Klaassen (2002). In the empirical application of the monitoring procedure in this work we propose a online procedure for evaluating the value at risk for the next time step, by simulating from the considered factor model.
Again note, that Section 3 and 4 are joint works with Hans Manner and Dominik Wied.

Lastly, we investigate the non-parametric dependence measure test, which compares different vectors of dependence measures jointly estimated using the whole sample information to successively estimated counterparts, where the dependence vectors consists of Spearman's rank correlation and quantile dependencies. The test is proposed in Section 3, where here the focus lies on the parameter test for detecting structural breaks in factor copula models. In Section 5 we want to pay more attention to the non-parametric dependence measure vector based test. The test is constructed to analyze the hypothesis of no dependence change in a pre-specified vector of dependence measures. We consider residual data from pre-estimated marginal time series models, namely ARCH and GARCH models such that the test is of non-parametric nature once we determined the residuals.

The asymptotic distribution of the test statistic is mainly obtained in Lemma 5 in the appendix and follows from a combination of the asymptotic behavior of the sequential copula process (cf. Bücher et al., 2014) and results from Bücher and Kojadinovic (2016), as well as Remillard (2017) to give a convergence result for the usage of residual data determined by pre-estimated time series models. For this reason it is important to use dependence measures which can be expressed in terms of the copula. Due to the fact that the asymptotic distribution is not known in closed form we have to estimate critical values by an i.i.d. bootstrap procedure. We extend previous simulation studies from Section 3 by analyzing size and power properties of the test for different skewed and fat tailed distributions for different settings of the used vector of dependence measures. We also propose a heuristic procedure to be able to make a statement for equality of two estimated break point locations, scaled to the uniform interval, using different dependence settings. Here, the (pivot) confidence intervals for both break point estimates have to be determined using a (percentile) bootstrap procedure and we consider two estimated break points as equal if they both lie in the intersection of the two confidence intervals. Finally, we use the test in a real-data application on daily log-returns of ten large financial firms during the last financial crisis, in which we use the test on the whole period and in a rolling window of a fixed window size.

The rest of this work is structured as follows. Section 2 gives an overview of the basic concepts of copula theory, dependence modeling, copula families and models such as parameter estimation with the SMM. Section 3 presents the retro perspective parameter and dependence measure test, where Section 4 deals with the online testing procedures to test for structural breaks in the copula parameters or the dependence measure vector in real time. Section 5 includes the investigation of the non-parametric dependence measure test for different settings of the considered dependence measure vector and a heuristic procedure to test for common breaks. A conclusion and an outlook are given in Section 6. All proofs can be found in the appendix in Section 7.

2. COPULA THEORY

In this section we want to introduce basic concepts of copula theory. Simply said: A copula is a function that couples marginal distribution functions to their multivariate distribution. We introduce the copula theory in a functional sense, without using random variables at all, the transition from the functional consideration to the usage of random variables is then straight forward. We define a random variable X as a quantity, whose values x are described by a known or unknown right-continuous probability distributon function $F(x) = P(X \leq x)$. In the same sense this definition can be used for the multivariate case, where we have random variables $(X_1, \dots, X_N)$, whose values $x_1, \dots, x_N$ are described by a joint distribution function $F(x_1, \dots, x_N) = P(X_1 \leq x_1, \dots, X_N \leq x_N)$.

2.1. Basic Concepts

This section is about definitions and properties of N-dimensional copulas. The main goal is to achieve the existence of a Copula from Sklar's Theorem. The whole subsection is based on the book from Nelson (2006). First we want to introduce some notations. Let $\bar{\mathbb{R}} := [-\infty, \infty]$ and with this define the N-space $\bar{\mathbb{R}}^N := \bar{\mathbb{R}} \times \cdots \times \bar{\mathbb{R}}$, where the operator " $\times$ " is the Cartesian product and N the dimension. We use vector notations for points in the space $\bar{\mathbb{R}}^N$, i.e. $\boldsymbol{a} = (a_1, \dots, a_N)$ and $\boldsymbol{a} \leq \boldsymbol{b}$ means $a_i \leq b_i$ for all i. With $[\boldsymbol{a}, \boldsymbol{b}]$ we denote the N-box $[a_1, b_1] \times \dots, \times [a_N, b_N]$. The vertices of the N-box are denoted as $\boldsymbol{c} = (c_1, \dots, c_N)$, where c_i can be equal to a_i or b_i. Let $\boldsymbol{I} = [0, 1]$ and define the unit N-cube $\boldsymbol{I}^N := \boldsymbol{I} \times \cdots \times \boldsymbol{I}$. We define the N-place real function H as a function, whose domain D_H is a subset of $\bar{\mathbb{R}}^N$ and whose range R_H is a subset of $\bar{\mathbb{R}}$.

Definition 1 (H-Volume). Let $A_1, \dots, A_N$ be non-empty subsets of $\bar{\mathbb{R}}^N$ and let H be an N-place real function with domain $D_H = A_1 \times \cdots \times A_N$. Let $B = [\boldsymbol{a}, \boldsymbol{b}]$ be an N-box and all of whose vertices are in D_H, then the H-volume of B is defined by

$$V_H(B) := \sum_{c \in B} \mathsf{sgn}(\boldsymbol{c}) H(\boldsymbol{c})$$

with

$$\mathsf{sgn}(\boldsymbol{c}) := \begin{cases} 1, & \text{if } c_i = a_i \text{ for an even number of i's} \\ -1, & \text{if } c_i = a_i \text{ for an odd number of i's.} \end{cases}$$

For example consider the two dimensional case where we have the domain $D_H = A_1 \times A_2$ and $B = [x_1, x_2] \times [y_1, y_2]$. Then the H-Volume is given by

$$V_H(B) = H(x_2, y_2) - H(x_1, y_2) - H(x_2, y_1) + H(x_1, y_1),$$

here $\boldsymbol{a} = (x_1, y_1)$, $\boldsymbol{b} = (x_2, y_2)$ and all vertices $\boldsymbol{c} \in B$ are given by $\{(x_1, y_1), (x_1, y_2), (x_2, y_1), (x_2, y_2)\}$. In the following some properties of the N-place real function H:

Definition 2 (N-increasing). An N-place real function H is N-increasing if $V_H(B) \geq 0$ for all N-boxes B whose vertices lie in the domain D_H.

Definition 3 (Grounded). We say that an N-place real function H is grounded, if $H(\tilde{\boldsymbol{c}}) = 0$ for all $\tilde{\boldsymbol{c}} \in D_H = A_1 \times \cdots \times A_N$, such that $\tilde{c}_i = a_i$ for at least one k and each A_i has at least one element a_i.

Definition 4 (Margins). Consider the domain $D_H = A_1 \times \cdots \times A_N$ and each A_i is nonempty and has a greatest element b_i, then we say that H has margins and the one-dimensional margins of H are the function H_i given by

$$H_i(x) = H(b_1, \ldots, b_{i-1}, x, b_{i+1}, \ldots, b_N),$$

with domain $D_{H_i} = A_i$ for all $x \in A_i$.

Note that higher dimensional margins are defined by fixing fewer places in H. With these definitions we derive the following Lemma, where the proof can be found in B.Schweizer and A.Sklar (1983).

Lemma 1 (Lipschitz Continuity for H). Let $A_1, \ldots, A_n$ be nonempty subsets of $\bar{\mathbb{R}}$ and H is a grounded N-increasing function with domain $D_H = A_1 \times \cdots \times A_N$ and one dimensional margins H_i. Let $\boldsymbol{x} = (x_1, \ldots, x_N)$ and $\boldsymbol{y} = (y_1, \ldots, y_N)$ be arbitrary points in D_H then we have

$$|H(\boldsymbol{x}) - H(\boldsymbol{y})| \leq \sum_{i=1}^{N} |H_i(x_i) - H_i(y_i)| \, .$$

We can later show that N- dimensional copulas are uniformly continuous by using Lemma 1. With these definitions in hand we can now define N-dimensional subcopulas and copulas, where subcopulas are a subclass of grounded N-increasing functions with margins and copulas are subcopulas with domain $\boldsymbol{I}^N$.

Definition 5 (Subcopulas). An N-dimensional subcopula is a function $\tilde{C}$ with the following properties:

1) $D_{\tilde{C}} = A_1 \times \cdots \times A_N$, where each A_i is a subset of $\boldsymbol{I}$ containing 0 and 1.

2) $\tilde{C}$ is grounded and N-increasing

3) $\tilde{C}$ has one dimensional margins $\tilde{C}_i(u) = u$, for $i = 1, \ldots, N$ and all $u \in A_i$

Because $\tilde{C}$ is grounded and N-increasing such as $\tilde{C}_i(u) = u \in \boldsymbol{I} = [0, 1]$ it follows that for every $\boldsymbol{u} \in D_{\tilde{C}}$, $0 \leq \tilde{C}(\boldsymbol{u}) \leq 1$. Now we can define copulas as a special case of subcopulas, whose domain is $\boldsymbol{I}^N$.

Definition 6 (Copulas). A Copula $C : \boldsymbol{I}^N \longrightarrow \boldsymbol{I}$ is a function with the following properties:

1) For every $\boldsymbol{u} \in \boldsymbol{I}^N$, $C(\boldsymbol{u}) = 0$ if at least one coordinate of $\boldsymbol{u}$ is 0 and if all coordinates of $\boldsymbol{u}$ are set to 1 except u_i, then $C(\boldsymbol{u}) = u_i$.

2) For every $\boldsymbol{a}$ and $\boldsymbol{b}$ in $\boldsymbol{I}^N$ such that $\boldsymbol{a} \leq \boldsymbol{b}$, C is N-increasing with $V_C([\boldsymbol{a}, \boldsymbol{b}]) \geq 0$.

With Definition 6 and Lemma 1 in Hand we can easily follow the following theorem for subcopulas $\tilde{C}$ and as a special case also for copulas C.

Theorem 2 (Lipschitz Continuity $\tilde{C}$). Let $\tilde{C}$ be an N-dimensional subcopula. Then for every $\boldsymbol{u}$ and $\boldsymbol{v}$ in $D_{\tilde{C}}$, we have

$$\left|\tilde{C}(\boldsymbol{u}) - \tilde{C}(\boldsymbol{v})\right| \leq \sum_{i=1}^{N} \left|\tilde{C}_i(u_i) - \tilde{C}_i(v_i)\right| = \sum_{i=1}^{N} |u_i - v_i| \, .$$

Because all norms are equivalent in $\mathbb{R}^N$, the subcopula $\tilde{C}$ fulfills a Lipschitz condition and thus we can establish the continuity of $\tilde{C}$. From Theorem 2 we can directly follow the continuity of copulas C by considering the domain $\boldsymbol{I}^N$. Before we want to state Sklar's theorem we first define N-dimensional distribution functions.

Definition 7 (Distribution Function). An N-dimensional distribution function is a function F with domain $\mathbb{R}^N$ such that F is N-increasing and $F(\boldsymbol{t}) = 0$ for all $\boldsymbol{t} \in \mathbb{R}^N$ if $t_i = -\infty$ for at least one i and $F(\infty, \ldots, \infty) = 1$.

Note, the N-dimensional distribution function F has the same characteristics (N-increasing, grounded and $0 \leq F \leq 1$) as the defined copula function. Furthermore there exist margins F_i for $i = 1, \ldots, N$, by setting all inputs of F to infinity except the i's input. We can now state Sklar's theorem.

Theorem 3 (Sklar's Theorem). Let F be an N-dimensional distribution function with margins $F_1, \ldots, F_N$, then there exists an N-dimensional copula C such that for all $\boldsymbol{x} = (x_1, \ldots, x_N) \in \mathbb{R}^N$

$$F(x_1, \ldots, x_N) = C(F_1(x_1), \ldots, F_N(x_N)).$$

Furthermore, if $F_1, \ldots, F_N$ are continuous, then C is unique, otherwise C is uniquely determined on the Cartesian product of the ranges of the margins. Conversely, if C is an N-dimensional copula and $F_1, \dots, F_N$ are marginal distribution functions, then the function $F(x_1, \dots, x_N) = C(F_1(x_1), \dots, F_N(x_N))$ is an N-dimensional distribution function with margins $F_1, \dots, F_N$.

A proof of Sklar's theorem can be found for example in Nelson (2006), where the theorem is first proven for the two dimensional case and then expanded to the N-dimensional case. If the inverses $F_1^{-1}, \dots, F_N^{-1}$ of $F_1, \dots, F_N$ exist then for any $\boldsymbol{u} \in \boldsymbol{I}^N$ we derive the copula function by

$$C(u_1, \dots, u_N) = F(F_1^{-1}(u_1), \dots, F_N^{-1}(u_N)).$$

Lastly, we want to state some important properties of copula functions.

Remark 1 (Properties).

1) Let $C(\cdot)$ be an N-dimensional copula function, then $\mathsf{max}(u_1 + \cdots + u_N - N + 1, 0) \leq C(\boldsymbol{u}) \leq \mathsf{min}(u_1, \dots, u_N)$, for all $\boldsymbol{u} = (u_1, \dots, u_N) \in \boldsymbol{I}^N$ (lower/upper Fréchet-Hoeffding boundary).

2) Let $(X_1, \dots, X_N)$ be a vector of continuous random variables with copula $C(\cdot)$, then $X_1, \dots, X_N$ are independent if and only if $C(u_1, \dots, u_N) = \prod_{i=1}^N u_i$ for $\boldsymbol{u} = (u_1, \dots, u_N) \in \boldsymbol{I}^N$.

3) Let $(X_1, \dots, X_N)$ be a vector of continuous random variables with copula $C(\cdot)$. If $\alpha_1, \dots, \alpha_N$ are strightly increasing functions on the range of $X_1, \dots, X_N$, then $\alpha_1(X_1), \dots, \alpha_N(X_N)$ has the same copula $C(\cdot)$.

2.2. Empirical Versions and asymptotic behavior

In this section we want to introduce the empirical version of the copula and study the asymptotic behavior of copula processes, which we later need to derive the asymptotics of our considered parameter process. We start by defining the empirical version of the copula function $C(\boldsymbol{u}) = C(u_1, \ldots, u_N)$, following Bücher et al. (2014). Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_T$ be random vectors of dimension $N \times 1$, then for $1 \leq k \leq l \leq T$ and $\boldsymbol{u} \in \boldsymbol{I}^N$, we define the empirical copula of the sample $\boldsymbol{x}_k, \ldots, \boldsymbol{x}_l$ as

$$\hat{C}_{k:l}(\boldsymbol{u}) = \frac{1}{l-k+1} \sum_{t=k}^{l} \mathbb{1}\{\hat{\boldsymbol{F}}_t^{k:l} \leq \boldsymbol{u}\}, \tag{2.1}$$

where

$$\hat{\boldsymbol{F}}_t^{k:l} = (\hat{F}_{1t}^{k:l}, \ldots, \hat{F}_{Nt}^{k:l}), \text{ for } t \in \{k, \ldots, l\}$$

and empirical distribution function

$$\hat{F}_{it}^{k:l}(x_{ti}) = \frac{1}{l-k+1} \sum_{j=k}^{l} \mathbb{1}\{x_{ij} \leq x_{it}\} \text{ for } t \in \{k, \ldots, l\} \text{ and } i \in \{1, \ldots, N\}.$$

Note that the empirical distribution function is computed using only information of the subsample $\boldsymbol{x}_k, \ldots, \boldsymbol{x}_l$ and not using the whole information $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T$.

Define the two-sided sequential empirical copula process $\mathbb{C}_T(n, m, \boldsymbol{u})$ as

$$\begin{aligned}\mathbb{C}_T(n, m, \boldsymbol{u}) =& \sqrt{T} \lambda_T(n, m) \left(\hat{C}_{1+\lfloor nT \rfloor:\lfloor mT \rfloor}(\boldsymbol{u}) - C(\boldsymbol{u})\right) \\ =& \frac{1}{\sqrt{T}} \sum_{t=\lfloor nT \rfloor+1}^{\lfloor mT \rfloor} \left(\mathbb{1}\{\hat{\boldsymbol{F}}_t^{\lfloor nT \rfloor+1:\lfloor mT \rfloor} \leq \boldsymbol{u}\} - C(\boldsymbol{u})\right),\end{aligned} \tag{2.2}$$

where $n < m$ and $n, m \in [0, 1]$, identifying $k = 1 + \lfloor nT \rfloor$ and $l = \lfloor mT \rfloor$, such as $\lambda_T(n, m) := \frac{\lfloor mT \rfloor - \lfloor nT \rfloor}{T}$.

We assume that the partial derivatives of $C(\cdot)$ exist, defined by $\dot{C}_j(\cdot) = \frac{\partial C}{\partial u_j}$, where $\dot{C}_j$ is continuous on $V_j = \{\boldsymbol{u} \in \boldsymbol{I}^N | u_j \in (0, 1)\}$ for every $j \in \{1, \ldots, N\}$ and that $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_T$ are drawn from a strictly stationary sequence $(\boldsymbol{X}_j)_{j \in \mathbb{Z}}$ with continuous margins and whose strong mixing coefficients satisfy $\alpha_r = \mathcal{O}(r^{-\alpha})$ and $\alpha > 1$.

Let $(\boldsymbol{U}_j)_{j \in \mathbb{Z}}$ be a strictly stationary sequence obtained from $(\boldsymbol{X}_j)_{j \in \mathbb{Z}}$ by using the probability integral transform $U_{i\cdot} = F_i(X_{i\cdot})$ for $i \in \{1, \ldots, N\}$, then we know from Bücher et al. (2014)

that the empirical process $\mathbb{C}_T$ converges in distribution to a special Gaussian process

$$\mathbb{C}_C = \mathbb{B}_C(n, m, \boldsymbol{u}) - \sum_{i=1}^{N} \dot{C}_j(\boldsymbol{u})\mathbb{B}_C(n, m, \boldsymbol{u}^{(j)}), \tag{2.3}$$

where $\mathbb{B}_C(n, m, \boldsymbol{u}) = \mathbb{Z}_C(m, \boldsymbol{u}) - \mathbb{Z}_C(n, \boldsymbol{u})$, $\mathbb{Z}_C$ a tight centered Gaussian process with covariance function

$$\mathsf{Cov}\left[\mathbb{Z}_C(n, \boldsymbol{u}), \mathbb{Z}_C(m, \boldsymbol{v})\right] = \mathsf{min}\{n, m\} \sum_{k \in \mathbb{Z}} \mathsf{Cov}\left[\mathbb{1}\{\boldsymbol{U}_0 \leq u\}, \mathbb{1}\{\boldsymbol{U}_k \leq v\}\right]$$

and $\boldsymbol{u}^{(j)}$ a vector in $\boldsymbol{I}^N$ defined by $u_i^{(j)} = u_j$ if $i = j$ and 1 otherwise. Note that $\mathbb{Z}_C(0, \boldsymbol{u}) = 0$ for all $\boldsymbol{u} \in \boldsymbol{I}^N$. The convergence result is later extended for the usage of residual data from pre estimated marginal time series models.

2.3. Dependence modeling with Copulas

In this section we present dependence measures which can be expressed by terms of the copula $C(\cdot)$. We are focusing on the two dimensional case where we have two random variables X and Y. This section follows Embrechts, Lindskog, and McNeil (2001), McNeil, R.Frey, and P.Embrechts (2010) and Oh and Patton (2013).

The most widely used dependence measure between two random variables X and Y is Pearson's linear correlation coefficient defined as

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

The main problem is, that it only captures linear dependence and does not capture dependence in higher orders, for example let $X \sim N(0, 1)$ and $Y = X^2$, obviously X and Y are perfectly dependent, but $\rho(X, Y) = 0$. Additionally it is only invariant under linear increasing transformations, but e.g. $\rho(e^X, Y) \neq \rho(X, Y)$. These disadvantages demonstrate why it might be useful to focus on pure dependence measures with better properties.

In this section we will present the dependence measures Spearmans's rank correlation, quantile dependence, Kendall's tau, and tail dependence, where we focus on the first two in later considerations. For continuous random variables X and Y, with distribution functions F_X and F_Y, whose copula is $C_{ij}(\cdot)$, Spearman's rank correlation coefficient in terms of the copula is defined as

$$\begin{aligned}\rho_S(X,Y) =& 12\int_0^1\int_0^1 C_{ij}(u,v)dudv - 3\\ =& 12\mathbb{E}(F_X(X)F_Y(Y)) - 3 = \rho(F_X(X), F_Y(Y)),\end{aligned} \tag{2.4}$$

where $u = F_X(x)$ and $v = F_Y(y)$. Simply said: ρ_S is the linear correlation coefficient of the underlying marginal distribution functions of X and Y.

For an empirical version of ρ_S we use the standard estimates for the expectation (mean) and the marginal distribution function (empirical distribution function) and receive

$$\hat{\rho}_S(X,Y) = \frac{12}{T}\sum_{t=1}^{T}\hat{F}_X(x_t)\hat{F}_Y(y_t), \tag{2.5}$$

for realizations $x_1,\ldots,x_T$ of X and $y_1,\ldots,y_T$ of Y, where

$$\hat{F}_Z(z) = \frac{1}{T+1}\sum_{t=1}^{T}\mathbb{1}\{z_t \le z\}$$

is the rescaled empirical distribution function of Z evaluated for realizations $z_1,\ldots,z_T$ at point z.

Let (X,Y) and (X^*,Y^*) be independent vectors of continuous random variables with joint distribution function F_{XY}, then Kendall's Tau between the variables X and Y is the probability of concordance minus the probability of discordance between (X,Y) and (X^*,Y^*) and can be expressed by terms of the copula

$$\begin{aligned}\tau(X,Y) =& P((X-X^*)(Y-Y^*)>0) - P((X-X^*)(Y-Y^*)<0)\\ =& 2P((X-X^*)(Y-Y^*)>0) - 1 = 4\mathbb{E}(C_{ij}(U,V)) - 1,\end{aligned} \tag{2.6}$$

where $U = F_X(X)$ and $V = F_Y(Y)$.

For the empirical version of τ we use the standard estimates for the expectation, the copula (empirical copula) and the marginal distribution function and receive

$$\hat{\tau}(X,Y) = \frac{4}{T}\sum_{t=1}^{T}\hat{C}_{ij}(\hat{F}_X(x_t), \hat{F}_Y(y_t)), \tag{2.7}$$

for realizations $x_1, \dots, x_N$ of X and $y_1, \dots, y_N$ of Y, where

$\hat{C}_{ij}(u,v) = \frac{1}{T}\sum_{t=1}^{T} \mathbb{1}\{\hat{F}_X(x_t) \leq u, \hat{F}_Y(y_t) \leq v\}$ is the bivariate empirical copula.

Note that $\rho_S, \tau \in [-1,1]$ and that $\rho_S = -1, \tau = -1$ and $\rho_S = 1, \tau = 1$ is equivalent to countermonotonicity and comonotonicity of X and Y. If X and Y are independent then $\rho_S = 0, \tau = 0$.

The quantile dependence between two continuous random variables X and Y for a specific quantile q, is the conditional probability that X is higher than the quantile value $Q_X(q)$ given that Y exceeds the corresponding quantile $Q_Y(q)$ for $q \in (0.5, 1)$ (upper quantile dependence) and the conditional probability that X is lower than the quantile value $Q_X(q)$ given that Y is smaler then the corresponding quantile $Q_Y(q)$ for $q \in (0, 0.5]$ (lower quantile dependence), i. e.

$$\lambda_q(X,Y) = \begin{cases} P(F_X(x) \leq q | F_Y(y) \leq q), & \text{for } q \in (0, 0.5] \\ P(F_X(x) > q | F_Y(y) > q), & \text{for } q \in (0.5, 1). \end{cases}$$

The quantile dependence $\lambda_q(X,Y)$ can be expressed in terms of the copula $C_{ij}(u,v)$ as

$$\lambda_q(X,Y) = \begin{cases} \frac{C_{ij}(q,q)}{q}, & \text{for } q \in (0, 0.5] \\ \frac{1-2q+C_{ij}(q,q)}{1-q}, & \text{for } q \in (0.5, 1). \end{cases} \tag{2.8}$$

We can estimate $\lambda_q(X, Y)$ in a fintite sample setting $x_1, \ldots, x_T$ and $y_1, \ldots, y_T$ as

$$\hat{\lambda}_q(X, Y) = \begin{cases} \frac{\hat{C}_{ij}(q,q)}{q}, \text{ for } q \in (0, 0.5] \\ \frac{1-2q+\hat{C}_{ij}(q,q)}{1-q}, \text{ for } q \in (0.5, 1), \end{cases} \tag{2.9}$$

where $\hat{C}_{ij}(q, q) = \frac{1}{T} \sum_{t=1}^{T} \mathbb{1}\{\hat{F}_X(x_t) \leq q, \hat{F}_Y(y_t) \leq q\}$.

The upper and lower tail dependence coefficients are the limit cases of the quantile dependence coefficient $\lambda_q(X, Y)$, where $q \longrightarrow 0$ or $q \longrightarrow 1$. These limit cases play an important role in the study of extreme events, noting that in the finite sample case q should be chosen close to 0 or 1. For a more detailed discussion see for example Joe (1997) and Nelson (2006).

2.4. Multivariate Copula Families

In this section we want to define some common families of copula functions, where the copula is known in closed form, considering elliptical copulas (Gaussian and Student's t) and the class of Archimedean copulas (Clayton, Gumbel and Frank), following Cherubini, Luciano, and Vecchiato (2004).

Starting with the class of elliptical copulas, whose contour plots are symmetric ellipses, we first introduce the multivariate Gaussian copula, for $u = (u_1, \ldots, u_N) \in \boldsymbol{I}^N$ defined as

$$C_{\Sigma}^{Ga}(\boldsymbol{u}) = \Phi_{\Sigma}\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_N)\right),$$

where Φ_{Σ} is the standardized multivariate normal distribution with covariance matrix Σ and Φ^{-1} the inverse of the standard univariate normal distribution function Φ. If the marginals are standard normal, we know from Sklar's theorem, that the Gaussian copula generates the joint standard Gaussian distribution. The density of the Gaussian copula can be easily derived by using the representation

$$\frac{1}{(2\pi)^{\frac{N}{2}}\mathsf{det}(\Sigma)^{\frac{1}{2}}}\mathsf{exp}\left(-\frac{1}{2}\boldsymbol{x}'\Sigma^{-1}\boldsymbol{x}\right) = c_{\Sigma}^{Ga}(\Phi(x_1),\ldots,\Phi(x_N))\prod_{i=1}^{N}\frac{1}{\sqrt{2\pi}}\mathsf{exp}\left(-\frac{1}{2}x_i^2\right)$$

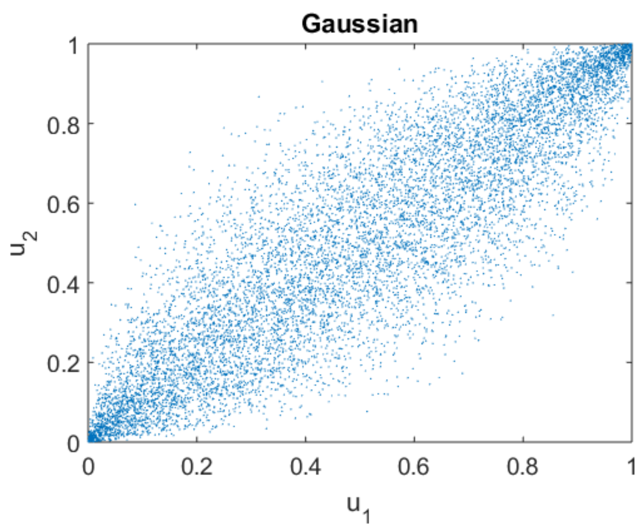
$$\Leftrightarrow c_{\Sigma}^{Ga}(\Phi(x_1),\ldots,\Phi(x_N)) = \frac{\frac{1}{(2\pi)^{\frac{N}{2}}\mathsf{det}(\Sigma)^{\frac{1}{2}}}\mathsf{exp}\left(-\frac{1}{2}\boldsymbol{x}'\Sigma^{-1}\boldsymbol{x}\right)}{\prod_{i=1}^{N}\frac{1}{\sqrt{2\pi}}\mathsf{exp}\left(-\frac{1}{2}x_i^2\right)}$$

and defining $u_i = \Phi(x_i)$, so that $x_i = \Phi^{-1}(u_i)$, we receive

$$c_{\Sigma}^{Ga}(\Phi(x_1),\ldots,\Phi(x_N)) = \frac{1}{\mathsf{det}(\Sigma)^{\frac{1}{2}}}\mathsf{exp}\left(-\frac{1}{2}\Phi^{-1}(\boldsymbol{u})'(\Sigma^{-1}-I)\Phi^{-1}(\boldsymbol{u})\right),$$

where $\Phi^{-1}(\boldsymbol{u}) := (\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_2))'$ and I the identity matrix of dimension $N \times N$.

Figure 2.1: Realisations of the Gaussian Copula



Note: Realisations of the Gaussian copula $C_{\Sigma}^{Ga}(u_1, u_2)$, with $Cov[u_1, u_2] = 0.9$.

The next copula we want to take a look at, in the class of elliptical copulas, is the multivariate

Student's t copula. For $\boldsymbol{u} \in \boldsymbol{I}^N$ we have

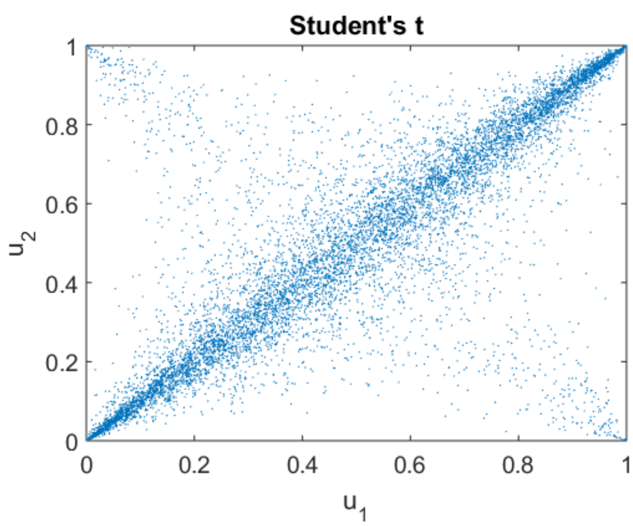
$$C_{\Sigma}^{St}(\boldsymbol{u}) = \int_{-\infty}^{t_{\nu}^{-1}(u_1)} \int_{-\infty}^{t_{\nu}^{-1}(u_2)} \cdots \int_{-\infty}^{t_{\nu}^{-1}(u_N)} \frac{\Gamma\left(\frac{\nu+N}{2}\right)\sqrt{\det(\Sigma)}}{\Gamma\left(\frac{\nu}{2}\right)(\nu\pi)} \left(1 + \frac{1}{\nu}\boldsymbol{x}'\Sigma^{-1}\boldsymbol{x}\right)^{-\frac{\nu+N}{2}} dx_1 dx_2 \ldots dx_N,$$

where t_{ν}^{-1} is the inverse of the univariate Student's t c.d.f with ν degrees of freedom, Σ the covariance matrix of the standardized multivariate Student's t distribution, $\Gamma(\cdot)$ the Gamma function and $\boldsymbol{x} := (x_1, \ldots, x_N)'$. The density of the multivariate Student's t copula is given by

$$c_{\Sigma}^{St}(\boldsymbol{u}) = \frac{\Gamma\left(\frac{\nu+N}{2}\right)\left(\Gamma\left(\frac{\nu}{2}\right)\right)^{N-1}}{\sqrt{\det(\Sigma)}\left(\Gamma\left(\frac{\nu+N}{2}\right)\right)^{N}} \left(1 + \frac{1}{\nu}t_{\nu}^{-1}(\boldsymbol{u})'\Sigma^{-1}t_{\nu}^{-1}(\boldsymbol{u})\right)^{-\frac{\nu+N}{2}} \prod_{i=1}^{N}\left(1 + \frac{t_{\nu}^{-1}(u_i)}{\nu}\right)^{\frac{\nu+1}{2}},$$

where $t_{\nu}^{-1}(\boldsymbol{u}) := (t_{\nu}^{-1}(u_1), \ldots, t_{\nu}^{-1}(u_N))'$.

Figure 2.2: Realisations of the Student Copula



Note: Realisations of the Student copula $C_{\Sigma}^{St}(u_1, u_2)$, with $Cov[u_1, u_2] = 0.9$ and $\nu = 0.5$.

Both, the multivariate Gaussian copula and the multivariate Student's t copula, are symmetric and the correlation matrix Σ is a symmetric, positive definite matrix with $diag(\Sigma) = (1, \ldots, 1)'$.

In contrast to the Gaussian copula, the Student's t copula allows for joint heavy tails with $\nu \in (2, \infty)$. The biggest disadvantage of these copulas is that they both are symmetric and hence unable to capture negative concurrent asymmetric extremal events.

The next family of copulas we want to define is the class of Archimedean copulas. Basically the copulas within this class are constructed via a generator function

$$\varphi(u) : [0, 1] \longrightarrow [0, \infty],$$

where φ is a strightly decreasing, convex and continuous function with monotonic generalized inverse $\varphi^{-1}(q) = \{u|\varphi(u) \leq q\}$ on $[0, \infty]$ and properties $\varphi(1) = 0$ and $\lim_{u \to 0} \varphi(u) \longrightarrow \infty$. Then we know from Kimberling (1974), that the function

$$C(\boldsymbol{u}) = \varphi^{-1}\left(\sum_{i=1}^{N} \varphi(u_i)\right)$$

is a copula for $\boldsymbol{u} \in \boldsymbol{I}^N$, where $\varphi(\cdot)$ is a valid generator function with several properties discussed in detail in McNeil and Nešlehová (2009). In the following, three examples of Archimedean copulas.

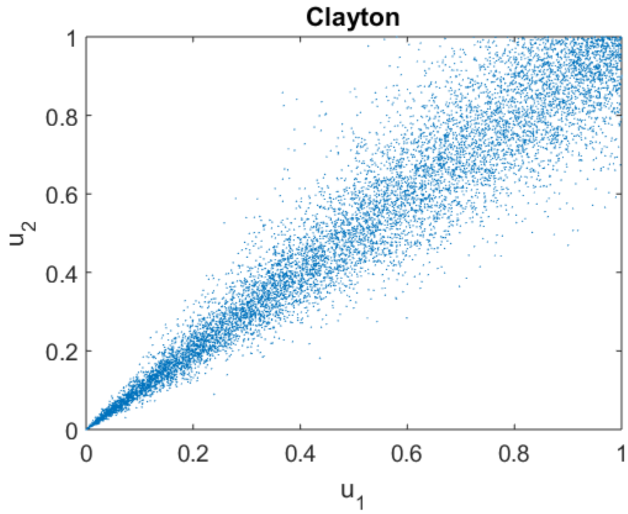
The Clayton copula is generated with the generator function $\varphi(u) = u^{-\alpha} - 1$, with parameter $\alpha > 0$ and is given by

$$C_\alpha^{Cl}(\boldsymbol{u}) = \left[\sum_{i=1}^{N} u_i^{-\alpha} - N + 1\right]^{-\frac{1}{\alpha}}.$$

The Clayton copula is asymmetric and has zero upper and positive lower tail dependence $2^{-\frac{1}{\alpha}}$. Note that the copula parameter α can be expressed by Kendall's τ (McNeil et al., 2010) and can therefore be estimated by the generalized method of moments.

The Gumbel copula is generated with the generator function $\varphi(u) = (-\log(u))^\alpha$, with

Figure 2.3: Realisations of the Clayton Copula



Note: Realisations of the Clayton copula $C_\alpha^{Cl}(u_1, u_2)$, with $\alpha = 10$.

parameter $\alpha > 0$ and is given by

$$C_\alpha^{Gu}(\boldsymbol{u}) = \exp\left(-\left(\sum_{i=1}^{N}(-\log\, u_i)^\alpha\right)^{\frac{1}{\alpha}}\right).$$

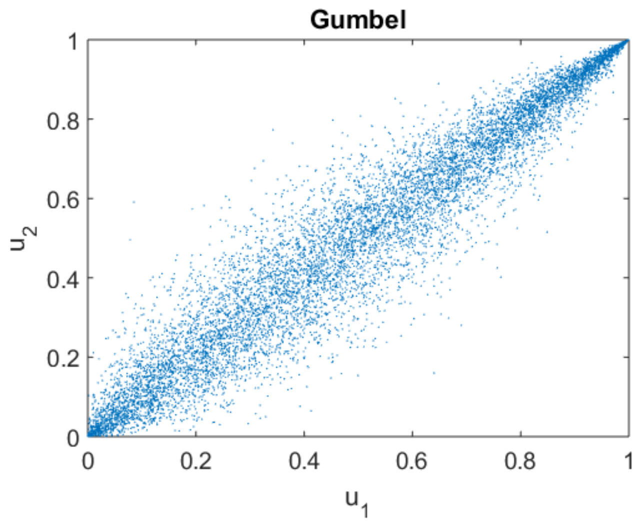
The Gumbel copula is asymmetric and has zero lower and positive upper tail dependence $2{-}2^{\frac{1}{\alpha}}$.

The Frank copula is generated with the generator function $\varphi(u) = \log\left(\frac{\exp(-\alpha u)-1}{\exp(-\alpha)-1}\right)$, with parameter $\alpha > 0$ and is given by

$$C_\alpha^{Fr}(\boldsymbol{u}) = -\frac{1}{\alpha}\log\left(1 + \frac{\prod_{i=1}^{N}(\exp(-\alpha u_i) - 1)}{(\exp(-\alpha) - 1)^{N-1}}\right).$$

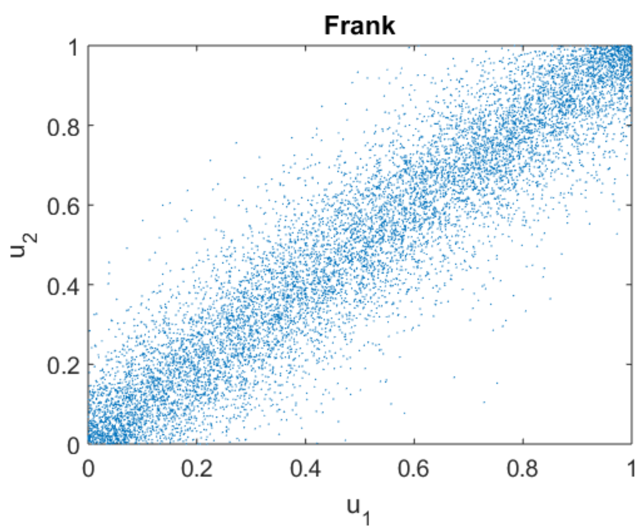
The Frank copula, in contrast to the Clayton or Gumble copula, is symmetric with zero tail dependence. Compared to the class of elliptical copulas the copulas within the Archimedean class allow for asymmetric effects.

Figure 2.4: Realisations of the Gumbel Copula



Note: Realisations of the Gumbel copula $C_{\alpha}^{Gu}(u_1, u_2)$, with $\alpha = 6$.

Figure 2.5: Realisations of the Frank Copula



Note: Realisations of the Gaussian copula $C_{\alpha}^{Fr}(u_1, u_2)$, with $\alpha = 15$.

2.5. Vine Copulas

This Section follows Aas et al. (2009), Brechmann and Schlepsmeier (2013), Czado (2010) and introduces R-vine, C-vine and D-vine copula construction. Considerable efforts have been undertaken to increase the flexibility of multivariate copula models beyond the scope of elliptical and Archimedean copulas which were discussed in the last section. Despite the factor copula models to model multivariate dependence we first take a look at another well known approach the vine copulas. With the help of Sklar's theorem we are able to separate the modeling of the marginal distributions and the joint distribution behavior. The main problem here is how to choose the copula model to model the multivariate dependence structure for higher dimensions. The literature is full of bivariate copula models, which are well investigated see for example Joe (1996) or Nelson (2006). For the multivariate case in higher dimension ($N > 2$) the variety of copula models is limited. Some examples of multivariate copula models were given in the last Section 2.4. The main problem here is that these copula families lack of flexibility to accurately model the dependence structure among larger numbers of variables. Generalizations of these models can offer some improvement, but become a rather intricate structure and other limitations arise for example parameter restrictions. Vine copulas give an approach to overcome these problems. Following Brechmann and Schlepsmeier (2013), "vines are a flexible graphical model for describing multivariate copulas built up using a cascade of bivariate copulas, so called pair-copulas first introduced by Joe (1996) and developed in more detail in Bedford and Cooke (2002) and in Kurowicka and Cooke (2006)“. With the help of Sklar's theorem and pair copula construction we are able to decompose multivariate densities into marginal densities and bivariate conditional/unconditional copula densities. The great advantage to standard multivariate models is that every bivariate copula can be chosen independently from each other and we can use the wide range of bivariate copula models. The great flexibility in the choice of bivariate copula models gives us the chance to account for asymmetries and tail dependence. We want to start with a three variables example to show the basic concepts of vine copula construction. Consider a multivariate density function with $N = 3$ variables, i.e. $f(x_1, x_2, x_3)$. We can write

$$f(x_1, x_2, x_3) = f_{3|1,2}(x_3|x_1, x_2) f_{2|1}(x_2|x_1) f_1(x_1), \tag{2.10}$$

where $f_{p|q}$ denotes the conditional density of the variables expressed by the number string p conditioned on the variables expressed by the number string q. Rewriting the densities with the belonging copula densities using Sklar's theorem leads to

$$\begin{aligned}
f_{2|1}(x_2|x_1) &= \frac{f_{12}(x_1, x_2)}{f_1(x_1)} = \frac{c_{12}(F_1(x_1), F_2(x_2)) f_1(x_1) f_2(x_2)}{f_1(x_1)} \\
&= c_{12}(F_1(x_1), F_2(x_2)) f_2(x_2) \\
f_{3|12}(x_3|x_1, x_2) &= \frac{f_{123}(x_1, x_2, x_3)}{f_{12}(x_1, x_2)} = \frac{f_{23|1}(x_2, x_3|x_1) f_1(x_1)}{f_{2|1}(x_2|x_1) f_1(x_1)} \\
&= \frac{c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)) f_{2|1}(x_2|x_1) f_{3|1}(x_3|x_1)}{f_{2|1}(x_2|x_1)} \\
&= c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)) c_{13}(F_1(x_1), F_3(x_3)) f_3(x_3).
\end{aligned}$$

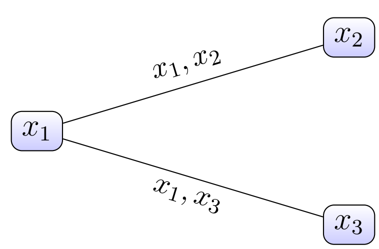
This leads to the bivariate copula decomposition

$$f(x_1, x_2, x_3) = f_1(x_1) f_2(x_2) f_3(x_3) c_{23|1}\left(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)\right) c_{13}\left(F_1(x_1), F_3(x_3)\right) c_{12}\left(F_1(x_1), F_2(x_2)\right).$$

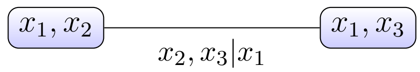
Thus, we decomposed the joint density function $f(x_1, x_2, x_3)$ in the bivariate conditional copula density $c_{23|1}$ and bivariate unconditional copula densities c_{12} and c_{13} such as the marginal densities f_1, f_2 and f_3. Note that the decomposition of the joint density in equation (2.10) is not unique. The decomposition can be classified by using a tree structure to arrange the $\frac{N(N-1)}{2}$ pair copulas in $N - 1$ linked trees called vines, for example see Kurowicka and Cooke (2006) and Kurowicka and Joe (2011). The first vine structure we want to introduce is the regular vine (R-vine) construction principle. In the first step we have a node for each variable $x_1, \ldots, x_N$ and we construct a tree (an undirected acyclic connected graph) with $N - 1$ edges, where every edge corresponds to an unconditional pair copula density for the considered node connection. The tree structure is not unique and can be determined by using minimal spanning tree algorithms from the field of graph theory, where the edge weights can for example be chosen as the values of Spearman's rank or Kendall's tau of the considered variables. In the second tree every edge of the first tree becomes a node and two nodes are connected with an edge if the corresponding edges shared a node in the first tree. Every edge in the second tree corresponds to a conditional bivariate copula density conditioned on the shared node of the first tree. Further trees are constructed in the same way. Every edge in tree T_i becomes a node in tree T_{i+1} and every edge in tree T_{i+1} corresponds to a conditional bivariate copula density conditioned on the corresponding shared node variables of the previous tree T_i.

The second vine structure we consider is the canonical (C-vine) tree representation. In the first tree each of the N variables again has a representative node. We choose a root node and connect all other nodes to this node. Each pair of nodes connected with an edge is than modeled using unconditional bivariate copulas. In the second tree the edges (i, j) connecting the nodes i and j in the first tree become a node. We now condition all previous connections with the last root node variable. This can be expressed with the edges in the second tree. "In general, a root node is chosen in each tree and all pairwise dependencies with respect to this node are modeled conditioned on all previous root nodes", following Brechmann and Schlepsmeier (2013). For the considered three variables example we have the following two C-vine trees T_1 and T_2.

T_1:



T_2:

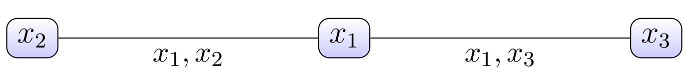


The edges of the trees T_1 and T_2 correspond to the bivariate unconditional/conditional copula densities in the multivariate density decomposition. The decomposition of a multivariate density in terms of bivariate copula densities and marginal densities for the C-vine structure with ordered root nodes $1, \ldots, N$ can in general be written as

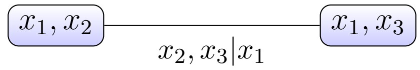
$$f(x_1, \ldots, x_N) = \prod_{k=1}^{N} f_k(x_k) \prod_{i=1}^{N-1} \prod_{j=1}^{N-i} c_{i,i+j|1:(i-1)} \left(F(x_i|x_1, \ldots, x_{i-1}), F(x_{i+j}|x_1, \ldots, x_{i-1})\right),$$

where f_k for $k = 1, \ldots, N$ denotes the marginal densities and $c_{i,i+j|1:(i-1)}$ denotes bivariate copula densities of the variables x_i and x_{i+j} conditioned on the variables $x_1, \ldots, x_{i-1}$. Note, a copula density is unconditioned if $i = 1$. Similar to the C-vine decomposition we can construct a drawable vine (D-vine) decomposition. For this we choose an order of the variables and connect two ordered nodes i and j with an edge (i, j) in the first tree. In the second tree every pair of nodes from the first tree becomes a node. Following Brechmann and Schlepsmeier (2013) "conditional dependence of the first and third given the second variable (the pair $(x_1, x_3|x_2)$), the second and fourth given the third (the pair $(x_2, x_4|x_3)$), and so on, is modeled“, indicated with an edge in the second tree. In the following trees the pairwise dependencies of two variables a and b is modeled in the same way. Different to the C-vine star structure the D-vine results in a path structure. For the considered example we have the following two D-vine trees T_1 and T_2, where the variable (node) ordering is chosen as x_2, x_1, x_3.

T_1:



T_2:



Obviously, for this simple example the C-vine and D-vine trees are the same but in general, the corresponding trees are different, see for example Czado (2010). The decomposition of a multivariate density in terms of bivariate copula and marginal densities for the D-vine structure with ordered root nodes $1, \ldots, N$ can in general be written as

$$f(x_1, \ldots, x_N) = \prod_{k=1}^{N} f_k(x_k) \prod_{i=1}^{N-1} \prod_{j=1}^{N-i} c_{j,j+i|(j+1):(j+i-1)} \left(F(x_j|x_{j+1}, \ldots, x_{j+i-1}), F(x_{j+i}|x_{j+1}, \ldots, x_{j+i-1})\right).$$

Note, the node order for the closed C-vine and D-vine representations can be changed without loss of generality.

To fit a vine copula one first has to choose a vine tree structure. The structure may be given by the data or has to be chosen manually. For example for a C-vine structure we can follow the optimal C-vine structure selection by Czado, Schepsmeier, and Min (2012), where for a D-vine structure the traveling salesman problem for D-vines can be applied. In both cases the edge weights in the corresponding graphs can for example be selected as the values of Kendall's tau or Spearman's rho of the considered variables. After determining the tree structure one selects and estimates adequate bivariate copula models. Copula selection can be done via Goodness-of-fit tests, Independence test, AIC/BIC-criterion or graphical tools like contour plots, where one can choose for example from elliptical, one-parametric Archimedean copulas introduced in the previous Section 2.4 or two-parametric Archimedean copulas. The estimation of the copulas can then be done for example by maximum likelihood estimation.

2.6. Marginal Distributions and Factor Copula Models

In this section we want to introduce the used marginal time series models and the class of factor copula models, which is the key copula model we focus on in this work.

We are interested in modeling the dependence between the $1 \times T$ dimensional random variables $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_N$, following Oh and Patton (2013) and Oh and Patton (2017), where we assume the following model for all time points $t = 1, \ldots, T$

$$\begin{aligned} Y_{it} &= \mu_{it}(\phi_i) + \sigma_{it}(\phi_i)\eta_{it}, \\ \boldsymbol{\eta}_t &= [\eta_{1t}, \ldots, \eta_{Nt}]' \sim F_{\boldsymbol{\eta}} = C(F_1(\eta_1), \ldots, F_N(\eta_N); \theta), \end{aligned} \tag{2.11}$$

with time varying conditional mean $\mu_{it}(\phi_i)$ and standard deviation terms $\sigma_{it}(\phi_i)$ for $i = 1, \ldots, N$ and $t = 1, \ldots, T$, $\phi = [\phi_1, \ldots, \phi_N]'$ the conditional mean and standard deviation parameter vector, θ the parameter vector of the copula model and residuals $\eta_i = [\eta_{i1}, \ldots, \eta_{iT}]$ with marginal distributions F_i for $i = 1, \ldots, N$. With this model we can filter out the time varying conditional mean and variance from each time series of the variables $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_N$. The cross sectional dependence structure over the dimension N is implied by the assumed dependence structure of the standardized residuals and is modeled with the copula C.

In a first stage we estimate the data parameter vector ϕ of the conditional mean μ_{it} and conditional standard deviation σ_{it} terms. For example we choose a AR-GARCH model and the estimator is denoted as $\hat{\phi} = [\hat{\phi}_1, \ldots, \hat{\phi}_N]$, consisting of all parameter estimates of the AR(P)-GARCH(P,Q) model defined below. For a conditional mean term we use the AR(P) process

$$Y_{it} = \phi_{i0} + \sum_{p=1}^{P} \phi_{i0} Y_{i(t-p)} + \sigma_{it}\eta_{it}$$

and for a conditional variance term we use the GARCH(P,Q) process

$$\sigma_{it}^2 = \omega_i + \sum_{p=1}^{P} \alpha_{ip}\sigma_{i(t-p)}^2\eta_{i(t-p)}^2 + \sum_{q=1}^{Q} \beta_{iq}\sigma_{i(t-q)}^2.$$

After estimating the conditional mean and conditional standard deviation terms, we can compute standardized residuals for $i = 1, \dots, N$ and $t = 1, \dots, T$

$$\hat{\eta}_{it} = \frac{Y_{it} - \mu_{it}(\hat{\phi}_i)}{\sigma_{it}(\hat{\phi}_i)}.$$

Note, the asymptotic distribution of the copula parameter estimate $\hat{\theta}$ is independent of the conditional dynamics of the marginal models, shown by Chen and Fan (2006), such that any misspecification in the conditional mean and variance models will asymptotically not have an impact on the estimation of the copula parameter θ.

For the residuals $\boldsymbol{\eta}_t$ we know from Sklar's theorem that

$$\boldsymbol{\eta}_t = [\eta_{1t}, \dots, \eta_{Nt}]' \sim F_{\boldsymbol{\eta}} = C(F_1(\eta_1), \dots, F_N(\eta_N); \theta),$$

where $F_1, \dots, F_N$ are the marginal distributions of $\eta_1, \dots, \eta_N$. Note, we are only interested in the copula C and discard the marginal distributions $F_1, \dots, F_N$. The idea of the factor copula model is that the copula C is not only a copula of the vector $\boldsymbol{\eta}_t$, but also a set of N latent variables X_{it} for $i = 1, \dots, N$, which are linear functions of the factors. In the following we are interested in the implied copula of the factor model. First we want to introduce the simple one factor copula model, using only one common factor Z_t and N idiosyncratic factors q_{it}

$$\begin{aligned}
X_{it} &= Z_t + q_{it}, \qquad i = 1, \dots, N \\
\boldsymbol{X}_t &= [X_{1t}, \dots, X_{Nt}]' \sim F_{\boldsymbol{X}} = C(F_{X_1}(\theta), \dots, F_{X_N}(\theta); \theta) \\
Z_t &\sim F_Z(\gamma), \qquad q_{it} \overset{\text{i.i.d.}}{\sim} F_q(\alpha),\ q_{it} \text{ is independent of } Z_t \text{ for all } i,
\end{aligned} \tag{2.12}$$

where $\theta = (\gamma, \alpha)'$. Note, the copula C is the copula we use to model the cross sectional dependence of the residuals $\boldsymbol{\eta_t}$. In general the marginal distributions F_{X_i} and F_i are different. We use the structure of the vector $\boldsymbol{X}_t$ only for the implied copula. The above defined dependence structure is called equidependence structure, as the dependence between any two variables X_{it} and X_{jt} for $i \neq j$ is the same. The advantage of these models is that the dependence structure is independent of the dimension N and is captured by just a few parameters. Unfortunately, only in rare cases, for example if F_Z and F_q are normal distributions, the implied copula from the factor structure is known in closed form and is the multivariate normal copula defined above in Section 2.4. In general there is no closed form copula function given. An economic interpretation of the model could be as follows: If we consider data from an asset portfolio, the common factor Z could represent the state of economy (market factor) where the idiosyncratic terms q_i represent the state of the individual firms for $i = 1, \ldots, N$.

The simple one factor copula model, defined above, can be extended by adding loadings on the common factor Z_t

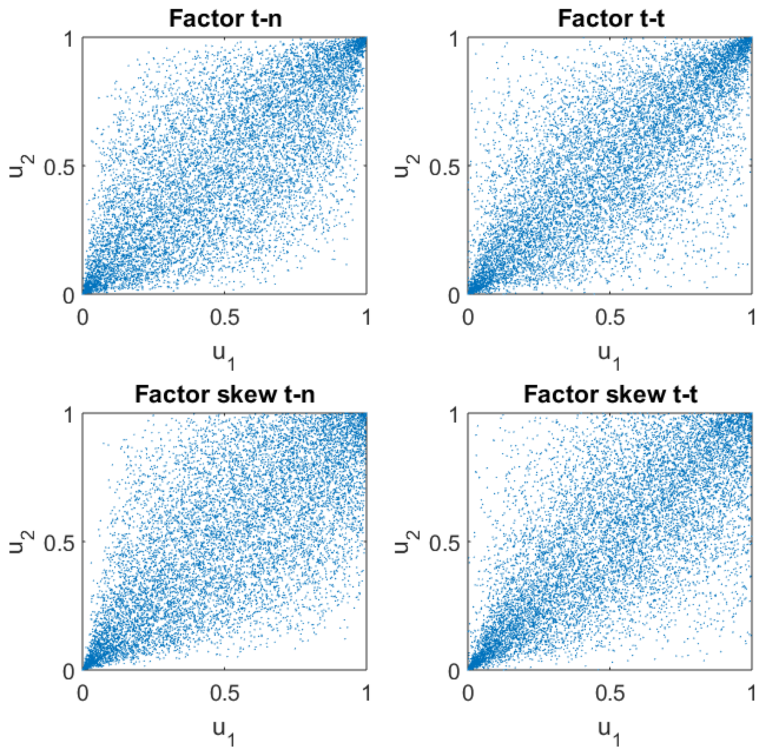
$$
\begin{aligned}
X_{it} &= \beta_i Z_t + q_{it}, \qquad i = 1, \ldots, N \\
\boldsymbol{X}_t &= [X_{1t}, \ldots, X_{Nt}]' \sim F_{\boldsymbol{X}} = C(F_{X_1}(\theta), \ldots, F_{X_N}(\theta); \theta) \\
Z_t &\sim F_Z(\gamma), \qquad q_{it} \overset{\text{i.i.d.}}{\sim} F_q(\alpha),\ q_{it} \text{ is independent of } Z_t \text{ for all } i,
\end{aligned} \tag{2.13}
$$

where $\theta = (\beta_1, \ldots, \beta_N, \gamma, \alpha)'$. An economic interpretation of such models could be as follows: If we consider data from an asset portfolio, than different stocks react differently to changes on the the common (market) factor Z.

Figure 2.6 shows scatter plots of the bivariate $t(\nu){-}N(0,1)$, $t(\nu){-}t(\nu)$, Skewed $t(\nu,\lambda){-}N(0,1)$ and Skewed $t(\nu,\lambda) - t(\nu)$ factor copulas with $\beta_i = 2\ \forall i, \nu = 4$ and $\lambda = -0.5$. The scatter plots reveal that all copulas have more mass in the tails and that Skewed $t(\nu,\lambda) - N(0,1)$ and Skewed $t(\nu,\lambda) - t(\nu)$ have asymmetric properties. The Skewed $t(\nu,\lambda)$ distribution first introduced by Hansen (1994) is described in the Appendix.

The main problem of the model decribed in (2.13) is, that the estimation of the parameters β_i could be extremly costly in high dimensional settings, where N is large. To overcome

Figure 2.6: Realisations of Factor Copulas



Note: Realisations of the Factor copula with factor loading $\beta_i = 2$ for all i and different distributions F_Z and F_q for Z and q.

this issue we can split the variables in different groups $S(i)$, which have the same factor loading parameter $\beta_{S(i)}$. Precisely, $S(i)$ sorts the variable with index i to its pre determined group. For example from an economic perspective it makes sense to sort the assets within the portfolio in their belonging industry sectors. Following Oh and Patton (2017) the block equidependence model is given as

$$
\begin{aligned}
X_{it} &= \beta_{S(i)} Z_t + q_{it}, \qquad i = 1, \dots, N \\
\boldsymbol{X}_t &= [X_{1t}, \dots, X_{Nt}]' \sim F_{\boldsymbol{X}} = C(F_{X_1}(\theta), \dots, F_{X_N}(\theta); \theta) \\
Z_t &\sim F_Z(\gamma), \qquad q_{it} \overset{\text{i.i.d.}}{\sim} F_q(\alpha),\ q_{it} \text{ is independent of } Z_t \text{ for all } i,
\end{aligned} \tag{2.14}
$$

where $\theta = (\beta_1, \dots, \beta_{G^*}, \gamma, \alpha)'$, with G^* being the amount of groups.

Another extension is to use more than only one common factor, i.e we can use K common factors and consider the following model

$$
\begin{aligned}
X_{it} &= \sum_{k=1}^{K} \beta_{S(i)k} Z_{kt} + q_{it}, \qquad i = 1, \dots, N \\
\boldsymbol{X}_t &= [X_{1t}, \dots, X_{Nt}]' \sim F_{\boldsymbol{X}} = C(F_{X_1}(\theta), \dots, F_{X_N}(\theta); \theta) \\
Z_{kt} &\sim F_{Z_k}(\gamma_k), \qquad q_{it} \overset{\text{i.i.d.}}{\sim} F_q(\alpha),
\end{aligned}
\tag{2.15}
$$

where q_{it} is independent of Z_{kt} for all i, k and Z_{kt} is independent of Z_{lt} for all $k \neq l$, with $\theta = (\mathsf{vec}\left([\beta_{1k}, \dots, \beta_{G^*k}]_{k=1}^{K}\right), \gamma, \alpha)'$, with G^* being the amount of groups.

For an economic application of such models we could use a joint market factor with different loadings and additionally add an industry factor for each industry group. An application of this model is later used in our real data application in Section 3.8.

2.7. Parameter estimation with Simulated Method of Moments (SMM)

In general, the factor copula model density function is not known in closed form, this prevents the application of direct Maximum Likelihood estimation. To overcome this issue we use the SMM proposed by Oh and Patton (2013). This section introduces the SMM procedure and its properties, which is the main estimation procedure for the estimation of the copula parameters in this work. Note, the SMM is not restricted to factor copula models and could also be used for other copula models.

In comparison to the method of moments or the generalized method of moments, where the to be estimated parameters are known functions of the moments, the SMM is used if this mapping from the parameters to the moments is unknown. The basic idea is to simulate data from the underlying copula model we want to fit to the data and minimize the difference of the moment vectors computed using the data and the moment vectors using the simulated data. The moment vectors we consider are vectors of dependence measures or their linear combinations, which makes sense if one is interested in analyzing the dependence structure of random variables. A specific structure of the moment vectors is explained later in this section, using the dependence measures introduced in Section 2.3.

We consider sample residuals $\hat{\boldsymbol{\eta}} = [\hat{\eta}_1, \ldots, \hat{\eta}_N]'$ from pre estimated time series models of size $N \times T$ with probability transformed series $\hat{\boldsymbol{u}} = [\hat{u}_1, \ldots, \hat{u}_N]'$ in $[0,1]^N$, using the marginal empirical distribution function, where the dependence structure is modeled with the copula $C(\cdot;\theta)$. Let $\hat{m}_T$ denote the $m^* \times 1$ vector, consisting of the copula based dependence measures introduced in Section 2.3 or linear combinations thereof, computed using $\hat{\boldsymbol{u}}$. Analogously, let $\tilde{m}_S(\theta)$ be the simulated counterpart using simulated data $\tilde{\boldsymbol{u}} = [\tilde{u}_1, \ldots, \tilde{u}_N]'$ from $\tilde{\boldsymbol{\eta}} = [\tilde{\eta}_1, \ldots, \tilde{\eta}_N]'$ of size $N \times S$ for a specific parameter vector θ of the copula with dimension $p \times 1$, where S is the number of simulations from the assumed copula model. We define the vector of difference as

$$g_{T,S}(\theta) := \hat{m}_T - \tilde{m}_S(\theta). \tag{2.16}$$

Then we can define the objective function of the SMM as the weighted sum squared difference between the empirical and simulated moments

$$Q_{T,S}(\theta) := g_{T,S}(\theta)' \hat{W}_T g_{T,S}(\theta), \tag{2.17}$$

where $\hat{W}$ is a positive definite weighting matrix of dimension $m^* \times m^*$. The task is then, find θ in a compact parameter space Θ such that $Q_{T,S}(\theta)$ is mimimized. The SMM estimator of the true paramater vector θ_0 is defined as

$$\hat{\theta}_{T,S} := \underset{\theta^* \in \Theta}{\arg\min}\; Q_{T,S}(\theta^*). \tag{2.18}$$

For an optimal choice of the weighting matrix $\hat{W}_T$ one can use the inverse of the asymptotic covariance matrix of $\hat{m}_T$ or simply the $m^* \times m^*$ identity matrix.

Oh and Patton (2013) showed, that under some assumptions (defined similar later in this work) the SMM estimator is consistent

$$\hat{\theta}_{T,S} \xrightarrow{a.s.} \theta_0 \tag{2.19}$$

and asymptotically normal distributed

$$\sqrt{T}(\hat{\theta}_{T,S} - \theta_0) \xrightarrow{d} \left(1 - \frac{1}{\sqrt{k}}\right) N\left(0, \Omega\right), \tag{2.20}$$

for $\frac{S}{T} \rightarrow k \in (0, \infty]$ and $T, S \rightarrow \infty$. With

$$\Omega = \left(G'WG\right)^{-1} G'W\Sigma WG \left(G'WG\right)^{-1}, \tag{2.21}$$

where Σ is the asymptotic covariance matrix of $\hat{m}_T$ and $G = \frac{\partial g_0(\theta)}{\partial \theta}_{|\theta=\theta_0}$ the derivative matrix of the asymptotic limit version $g_0(\theta)$ of $g_{T,S}(\theta)$ evaluated at $\theta = \theta_0$. Consequently one of the assumptions to derive the asymptotic normality needs to be the differentiability of $g_0(\theta)$ at θ_0. Note, in finite sample settings $g_{T,S}$ is in general not differentiable. To overcome this problem, Oh and Patton (2013) propose a two sided numerical derivative with appropriately chosen step size $\varepsilon_{T,S}$. The step size $\varepsilon_{T,S}$ is chosen larger than usually used step sizes when dealing with numerical derivatives. Oh and Patton (2013) suggest to use a step size $\varepsilon_{T,S} = 0.1$ leading to the most accurate estimates. The estimator for the k-th column of the $m \times p$ numerical derivative matrix $\hat{G}_{T,S}$ is given by

$$\hat{G}_{T,S,k} = \frac{g_{T,S}(\hat{\theta}_{T,S} + \varepsilon_{T,S}e_k) - g_{T,S}(\hat{\theta}_{T,S} - \varepsilon_{T,S}e_k)}{2\varepsilon_{T,S}}, \tag{2.22}$$

where e_k denotes the k-th unit vector of dimension $p \times 1$.

As mentioned the moment vectors $\hat{m}_T$ and $\tilde{m}_S(\cdot)$ consist of averaged pairwise dependence vectors, where the pairwise copula based dependence measures $\hat{\rho}_S, \hat{\tau}$ and $\hat{\lambda}_q$ introduced in Section 2.3 are used. Let δ_{ij}^d denote the d-th selected pairwise dependence measure ($\hat{\rho}_S, \hat{\tau}$ or $\hat{\lambda}_q$) between random variables i and j. If we have N variables we can define the $N \times N$

pairwise dependence matrix

$$D^d = \begin{bmatrix} 1 & \delta_{12}^d & \dots & \delta_{1N}^d \\ \delta_{21}^d & 1 & \dots & \delta_{2N}^d \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{N1} & \delta_{N2}^d & \dots & 1 \end{bmatrix}.$$

For the estimation of equidependence models, we take the average above all elements of D^d. Using the fact that D^d is symmetric, we consider the averaged d-th dependence measure

$$\bar{\delta}^d = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \delta_{ij}^d. \tag{2.23}$$

We stack all $\bar{\delta}^d$ in the moment (dependence) vector, depending on how many different dependence measures we use. For the estimation of block equidependence multi factor models, where we have our variables grouped into G^* groups with k_g members in group g, such that $N = \sum_{g=1}^{G^*} k_g$, we use the fact that all variables in the same group have an equidependence property and any pair of variables (i, j) in groups (r, s) has the same dependence as any other pair (i^*, j^*) in the same two groups (r, s), following Oh and Patton (2017). Having this in mind, we can now decompose the $N \times N$ matrix D^d into a matrix of submatrices D_{rs}^d of size $k_r \times k_s$ for $r, s \in \{1, \dots, G^*\}$

$$D^d = \begin{bmatrix} D_{11}^d & D_{12}^d & \dots & D_{1G^*}^d \\ D_{21}^d & D_{22}^d & \dots & D_{2G^*}^d \\ \vdots & \vdots & \ddots & \vdots \\ D_{G^*1}^d & D_{G^*2}^d & \dots & D_{G^*G^*}^d \end{bmatrix}.$$

Because D^d_{rs} is symmetric, we can transform D^d in the $G^* \times G^*$ matrix

$$D^d = \begin{bmatrix} \delta^{d*}_{11} & \delta^{d*}_{12} & \dots & \delta^{d*}_{1G^*} \\ \delta^{d*}_{21} & \delta^{d*}_{22} & \dots & \delta^{d*}_{2G^*} \\ \vdots & \vdots & \ddots & \vdots \\ \delta^{d*}_{G^*1} & \delta^{d*}_{G^*2} & \dots & \delta^{d*}_{G^*G^*} \end{bmatrix}$$

using $\delta^{d*}_{rr} := \frac{2}{k_r(k_r-1)} \sum_{i=1}^{k_r} \sum_{j=i+1}^{k_r} \hat{\delta}^d_{ij}$ for matrices D^d_{rr} and $\delta^{d*}_{rs} := \frac{1}{k_r k_s} \sum_{i=1}^{k_r} \sum_{j=1}^{k_s} \hat{\delta}^d_{ij}$ for matrices D^d_{rs} with $r \neq s$. Lastly create the vector

$$[\bar{\delta}^{d*}_1, \dots, \bar{\delta}^{d*}_{G^*}]' \tag{2.24}$$

,where $\bar{\delta}^{d*}_g := \frac{1}{G^*} \sum_{j=1}^{G^*} \delta^{d*}_{gj}$ for $g = 1, \dots, G^*$.

Let M denote the number of used dependence measures ($d \in \{1 \dots, M\}$) for a group g then we get a total number of $m^* = M \cdot G^*$ dependence measures.

3. TESTING FOR STRUCTURAL BREAKS IN FACTOR COPULAS

In this section we describe our results of the retro perspective parameter and dependence measure testing. Factor copula models and estimation by the SMM are reviewed in Section 3.1. Our null hypothesis and test statistic can be found in Section 3.2, whereas in Section 3.3 the asymptotic behaviour of the test is analysed. Our bootstrap algorithm is presented in Section 3.4. The broader applicability of our results is shortly explained in Section 3.5, whereas Section 3.6 discusses an important assumption that is made. All the proofs are included in Section 7.1.1 in the Appendix. The content in this section is a joint work with Hans Manner and Dominik Wied, supported by the DFG. The work belongs to the paper "Testing for structural breaks in factor copula models" Manner et al. (2019) published in the *Journal of Econometrics*.

3.1. Factor Copula Models and their Estimation

We consider the same model setup as in Oh and Patton (2013) and Oh and Patton (2017) with the difference that we allow underlying dependence parameter to be time-varying. The dynamics of the marginal distributions are determined by a parameter vector ϕ_0 and each variable can have time varying conditional mean $\mu_{it}(\phi_0)$ and standard deviation $\sigma_{it}(\phi_0)$ for $i = 1, \ldots, N$. The dependence of the joint distribution of the residuals η_t, captured by the parametric copula $C(., \theta_t)$, depends on the unknown parameters θ_t for $t = 1, \ldots, T$. The data-generating process is given by

$$[Y_{1t}, \ldots, Y_{Nt}]' =: \boldsymbol{Y}_t = \boldsymbol{\mu}_t(\phi_0) + \boldsymbol{\sigma}_t(\phi_0)\boldsymbol{\eta}_t,$$

with conditional mean $\boldsymbol{\mu}_t(\phi_0) := [\mu_{1t}(\phi_0), \ldots, \mu_{Nt}(\phi_0)]'$, conditional standard deviation $\boldsymbol{\sigma}_t(\phi_0) := \mathsf{diag}\{\sigma_{1t}(\phi_0), \ldots, \sigma_{Nt}(\phi_0)\}$ and $[\eta_{1t}, \ldots, \eta_{Nt}] =: \boldsymbol{\eta}_t \sim \boldsymbol{F}_\eta = C(F_1(\eta_1), \ldots, F_N(\eta_N); \theta_t)$, with marginal distributions F_i, where $\boldsymbol{\mu}_t$ and $\boldsymbol{\sigma}_t$ are $\mathcal{F}_{t-1}$-measurable and independent of $\boldsymbol{\eta}_t$. $\mathcal{F}_{t-1}$ is the sigma field containing information from the past $\{\boldsymbol{Y}_{t-1}, \boldsymbol{Y}_{t-2}, \dots\}$. Note that the $r \times 1$ vector ϕ_0 is $\sqrt{T}$ consistently estimable, which is fulfilled by many time series models, e.g. ARMA and GARCH models and the estimator is denoted as $\hat{\phi}$. The marginal distributions of the residuals $F_i(.)$ for $i = 1, \dots, N$ are estimated by the empirical distribution function $\hat{F}_i$. Using the residual information $\{\hat{\boldsymbol{\eta}}_t := \boldsymbol{\sigma}_t^{-1}(\hat{\boldsymbol{\phi}})[\boldsymbol{Y}_t - \boldsymbol{\mu}_t(\hat{\phi})]\}_{t=1}^T$ from the data, we are interested in estimating the $p \times 1$ vectors $\theta_t \in \Theta$ of the copula $C(., \theta_t)$ for all t. The copula we are interested in is the factor copula that is implied by the following factor structure

$$[X_{1t}, \dots, X_{Nt}]' =: X_t = \boldsymbol{\beta}_t \boldsymbol{Z}_t + \boldsymbol{q}_t, \tag{3.1}$$

with $X_{it} = \sum_{k=1}^{K} \beta_{ik}^t Z_{kt} + q_{it}$, where $\boldsymbol{q}_t := [q_{1t}, \dots, q_{Nt}]'$, $q_{it} \overset{i.i.d.}{\sim} F_q(\alpha_t)$ and $Z_{kt} \overset{i.i.d.}{\sim} F_{Z_k}(\gamma_{kt})$ for $i = 1, \dots, N$, $t = 1, \dots, T$ and $k = 1, \dots, K$. Note that Z_{kt} and q_{it} are independent $\forall i, k, t$ and the copula for $\boldsymbol{X}_t$ is given by

$$\boldsymbol{X}_t \sim \boldsymbol{F}_{\boldsymbol{X}_t} = C(G_{1t}(x_{1t}; \theta_t), \dots, G_{Nt}(x_{Nt}; \theta_t); \theta_t),$$

with marginal distributions $G_{it}(., \theta_t)$ and $\theta_t = \left[\{\{\beta_{ik}^t\}_{i=1}^N\}_{k=1}^K, \alpha_t', \gamma_{1t}', \dots, \gamma_{Kt}'\right]'$. Note that the marginal distributions of the factor model $G_{it}(., \theta_t)$ are not of interest and are discarded as one is only interested in the copula implied by this model. We assume that this implied copula governs the dependence of $\boldsymbol{Y}_t$.

In principle, the copula implied by (3.1) offers many possibilities regarding the type and heterogeneity of the dependence. Through the choice of appropriate distributions F_{Z_k} of the common factors and F_q of the idiosyncratic errors one has a lot of flexibility concerning the asymmetry and tail dependence properties of the copula; see Oh and Patton (2017) for details. Furthermore, by imposing the restriction of common factor loadings for specific groups of variables, e.g. those belonging to the same industry, one can reduce the number of parameters in higher dimensional applications.

As the notation suggests, we allow θ_t to be time-varying, having a piecewise constant model in mind. We directly consider the recursive estimation of the model for increasing sample sizes. For this, we denote $s \in (0,1]$ the fraction of the sample considered and we are interested in the recursively estimated parameter $\hat{\theta}_{sT,S}$ of $\theta_{\lfloor sT \rfloor} = \theta_t$. Note that the full sample estimator is recovered for $s = 1$. For the estimation we use the SMM estimator, introduced in Section 2.7, defined as

$$\hat{\theta}_{sT,S} := \underset{\theta \in \Theta}{\arg\min}\; Q_{sT,S}(\theta), \tag{3.2}$$

where the objective function is defined as $Q_{sT,S}(\theta) := g_{sT,S}(\theta)'\hat{W}_{sT} g_{sT,S}(\theta)$ with $g_{sT,S}(\theta) := \hat{m}_{sT} - \tilde{m}_S(\theta)$ and $\hat{W}_{sT}$ a $k \times k$ positive definite weight matrix. The $k \times 1$ vectors $\hat{m}_{sT}$ consist of appropriately chosen dependence measures that are potentially averaged from the pairwise measures $\hat{m}^{ij}_{sT}$, computed from the residuals $\{\hat{\eta}_t\}_{t=1}^{\lfloor sT \rfloor}$. As the dependence measures implied by the model are typically not available in closed form they have to be obtained by simulation. Hence, $\tilde{m}_S(\theta)$ is the corresponding vector of dependence measures computed from $\{\tilde{\eta}_l\}_{l=1}^{S}$, using S simulations from $\boldsymbol{F}_{\boldsymbol{X}_t}$. For the dependence measures of the pair (η_i, η_j) we need to consider copula based dependence measures that do not depend on the marginal distribution of the data. Following Oh and Patton (2013) we consider Spearman's rank correlation ρ^{ij} and quantile dependence λ^{ij}_q. These are defined in Section 2.3. The sample counterparts based on recursive samples are defined as

$$\hat{\rho}^{ij} := \frac{12}{\lfloor sT \rfloor} \sum_{t=1}^{\lfloor sT \rfloor} \hat{F}^s_i(\hat{\eta}_{it}) \hat{F}^s_j(\hat{\eta}_{jt}) - 3$$

$$\hat{\lambda}^{ij}_q := \begin{cases} \frac{\hat{C}^s_{ij}(q,q)}{q}, & q \in (0, 0.5] \\ \frac{1-2q+\hat{C}^s_{ij}(q,q)}{1-q}, & q \in (0.5, 1) \end{cases},$$

where $\hat{F}^s_i(y) := \frac{1}{\lfloor sT \rfloor} \sum_{t=1}^{\lfloor sT \rfloor} \mathbb{1}\{\hat{\eta}_{it} \leq y\}$ and $\hat{C}^s_{ij}(u,v) := \frac{1}{\lfloor sT \rfloor} \sum_{t=1}^{\lfloor sT \rfloor} \mathbb{1}\{\hat{F}^s_i(\hat{\eta}_{it}) \leq u, \hat{F}^s_j(\hat{\eta}_{jt}) \leq v\}$.

This means that $\hat{F}_j^s$ denotes the marginal empirical distribution function of the j-th component calculated from data up to time point $[sT]$. Hence, we are using *sequential ranks*. The sample moments for the simulated data $\{\tilde{\eta}_l\}_{l=1}^S$ are defined analogously and are denoted by $\tilde{\rho}^{ij}$ and $\tilde{\lambda}_q^{ij}$.

Depending on the precise model specification the pairwise dependence measures can be averaged for pairs that are assumed to have the same factor loading as is the case in equidependence or block equidependence models; see Oh and Patton (2017). This reduces the number of moment conditions accordingly.

3.2. Null Hypothesis and Test Statistics

The null hypothesis we are interested in is a constant copula parameter vector against the alternative of a single breakpoint at an unknown point in time,

$$H_0 : \theta_1 = \theta_2 = \cdots = \theta_T \qquad H_1 : \theta_1 = \theta_2 = \cdots = \theta_t \neq \theta_{t+1} = \cdots = \theta_T \text{ for some } t = \{1, \ldots, T-1\}.$$

The test statistic we propose is based on the difference between the recursive estimates of the parameter vector and its full sample analogue. Formally, it is defined as

$$\begin{aligned} P := P_{T,S} := \sup_{s\in[\varepsilon,1]} P_{sT,S} &:= \sup_{s\in[\varepsilon,1]} s^2 T(\hat{\theta}_{sT,S} - \hat{\theta}_{T,S})'(\hat{\theta}_{sT,S} - \hat{\theta}_{T,S}) \\ &\simeq \max_{\lfloor \varepsilon T \rfloor \leq t \leq T} \left(\frac{t}{T}\right)^2 T(\hat{\theta}_{t,S} - \hat{\theta}_{T,S})'(\hat{\theta}_{t,S} - \hat{\theta}_{T,S}), \end{aligned} \tag{3.3}$$

where $\hat{\theta}_{sT,S}$ is the recursive SMM estimator defined above that used the information up to time $t = \lfloor sT \rfloor$, T the sample size of the data, S the number of simulations in the SMM and $\varepsilon > 0$ a trimming parameter. Note, analytically ε has to be chosen strictly greater than zero and thus $s \in [\varepsilon, 1]$ to apply the required limit theorems for our proof of the asymptotic distributions. In the finite sample case ε should be chosen large enough so that the model parameters can be estimated in a reasonable way using $\lfloor \varepsilon T \rfloor$ observations.

Large values of the test statistic (3.3) indicate that the successively estimated parameter vector fluctuates too much over time compared to the full sample estimator, indicating instability.

The test statistic could also be applied to a subset of the parameter vector θ. For example, one may only be interested in testing the stability of the factor loadings assuming constant shape parameters. Another possibility is to consider a block-equidependence model and test for changing factor loadings only for a specific sector such as the financial sector during a financial crisis.

We consider an alternative test statistic that is based on the same principle as (3.3), but is based directly on the moment conditions used to estimate the model.

$$
\begin{aligned}
M := M_T := \sup_{s\in[\varepsilon,1]} M_{sT} &:= \sup_{s\in[\varepsilon,1]} s^2 T(\hat{m}_{sT} - \hat{m}_T)'(\hat{m}_{sT} - \hat{m}_T) \\
&\simeq \max_{\lfloor \varepsilon T \rfloor \leq t \leq T} \left(\frac{t}{T}\right)^2 T(\hat{m}_{sT} - \hat{m}_T)'(\hat{m}_{sT} - \hat{m}_T).
\end{aligned} \tag{3.4}
$$

This statistic is of nonparametric nature and has the advantage that is does not require recursive estimation of the model, which is computationally quite demanding. The disadvantage is that it does not allow testing the constancy of a subset of the parameters, but only can detect breaks in the whole copula. One may, however, consider an appropriate subset of the moment conditions and test for, e.g., breaks in the lower tail quantile dependence. The asymptotic distribution of M comes as a by product when deriving the asymptotic distribution of P. The corresponding asymptotic results can be found in the next subsection.

3.3. Asymptotic Analysis

For deriving analytical results for the asymptotic distribution of our test statistic we need the following assumptions. The first two ensure that the estimated rank correlation and quantile dependencies converge to their respective population counterparts.

***Assumption* 1.** i) The distribution function of the innovations F_η and the joint distribution function of the factors $F_X(\theta)$ are continuous.

ii) Every bivariate marginal copula $C_{ij}(u_i, u_j; \theta)$ of $\mathbf{C}(u; \theta)$ has continuous partial derivatives with respect to $u_i \in (0,1)$ and $u_j \in (0,1)$.

The assumption is similar to Assumption 1 in Oh and Patton (2013), but the assumption on the copula is relaxed in the sense that the restriction of u_i and v_i is relaxed to the open interval $(0,1)$.

***Assumption* 2.** Define $\gamma_{0t} := \sigma_t^{-1}(\hat{\phi})\dot{\mu}_t(\hat{\phi})$ and $\gamma_{1kt} := \sigma_t^{-1}(\hat{\phi})\dot{\sigma}_{kt}(\hat{\phi})$, where $\dot{\mu}_t(\phi) := \frac{\partial \mu_t(\phi)}{\partial \phi'}$ and $\dot{\sigma}_{kt}(\phi) := \frac{\partial [\sigma_t(\phi)]_{\text{k-th column}}}{\partial \phi'}$ for $k = 1, \ldots, N$. Define

$$d_t = \eta_t - \hat{\eta}_t - \left(\gamma_{0t} + \sum_{k=1}^{N} \eta_{kt}\gamma_{1kt}\right)(\hat{\phi} - \phi_0),$$

where η_{kt} is the k-th row of η_t and γ_{0t} and γ_{1kt} are $\mathcal{F}_{t-1}$-measurable and $\mathcal{F}_{t-1}$ containing information from the past as well as possible information from exogenous variables.

i) $\frac{1}{T}\sum_{t=1}^{\lfloor sT \rfloor} \gamma_{0t} \xrightarrow{p} s\Gamma_0$ and $\frac{1}{T}\sum_{t=1}^{\lfloor sT \rfloor} \gamma_{1kt} \xrightarrow{p} s\Gamma_{1k}$, uniformly in $s \in [\varepsilon, 1]$, $\varepsilon > 0$, where Γ_0 and Γ_{1k} are deterministic for $k = 1, \ldots, N$.

ii) $\frac{1}{T}\sum_{t=1}^{T} E(\|\gamma_{0t}\|)$, $\frac{1}{T}\sum_{t=1}^{T} E(\|\gamma_{0t}\|^2)$, $\frac{1}{T}\sum_{t=1}^{T} E(\|\gamma_{1kt}\|)$ and $\frac{1}{T}\sum_{t=1}^{T} E(\|\gamma_{1kt}\|^2)$ are bounded for $k = 1, \ldots, N$.

iii) There exists a sequence of positive terms $r_t > 0$ with $\sum_{i=1}^{\infty} r_t < \infty$, such that the sequence $\max_{1 \le t \le T} \frac{\|d_t\|}{r_t}$ is tight.

iv) $\max_{1 \le t \le T} \frac{\|\gamma_{0t}\|}{\sqrt{T}} = o_p(1)$ and $\max_{1 \le t \le T} \frac{|\eta_{kt}|\|\gamma_{1kt}\|}{\sqrt{T}} = o_p(1)$ for $k = 1, \ldots, N$.

v) $(\alpha_T(s), \sqrt{T}(\hat{\phi} - \phi_0))$ weakly converges to a continuous Gaussian process in $\mathcal{D}([0,1]^N) \times \mathbb{R}^r$, where $\mathcal{D}$ is the space of all càdlàg-functions on $[0,1]^N$, with

$$\alpha_T(s) := \frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor sT \rfloor} \left\{ \prod_{k=1}^{N} \mathbb{1}\{U_{kt} \leq u_k\} - \mathbf{C}(u;\theta) \right\}.$$

vi) $\frac{\partial F_\eta}{\partial \eta_k}$ and $\eta_k \frac{\partial F_\eta}{\partial \eta_k}$ are bounded and continuous on $\overline{\mathbb{R}}^N = [-\infty, \infty]^N$ for $k = 1, \ldots, N$.

vii) For $\mathbf{u} \in [0,1]^N$ and $\hat{\mathbf{F}}^s(\hat{\eta}_t) = (\hat{F}_1^s(\hat{\eta}_{1t}), \ldots, \hat{F}_N^s(\hat{\eta}_{Nt}))$, the sequential empirical copula process

$$\frac{1}{\sqrt{T}} \left[\sum_{t=1}^{\lfloor sT \rfloor} \mathbb{1}\{\hat{\mathbf{F}}^s(\hat{\eta}_t) \leq \mathbf{u}\} - C(\mathbf{u}) \right] \tag{3.5}$$

converges in distribution to some limit process $A^*(s, \mathbf{u})$.

Parts i) to vi) of this assumption are similar to Assumption 2 in Oh and Patton (2013), only part (i) is more restrictive. We need this because we consider successively estimated parameters. Part vii) ensures that the empirical copula process of the residuals has some well defined limit. Given the literature on this topic, the assumption is plausible, which is discussed in more detail in Subsection 3.6.

The next assumption is needed for consistency of the successively estimated parameters. It is the same as Assumption 3 in Oh and Patton (2013) with the difference that part (iv) is adapted to our situation.

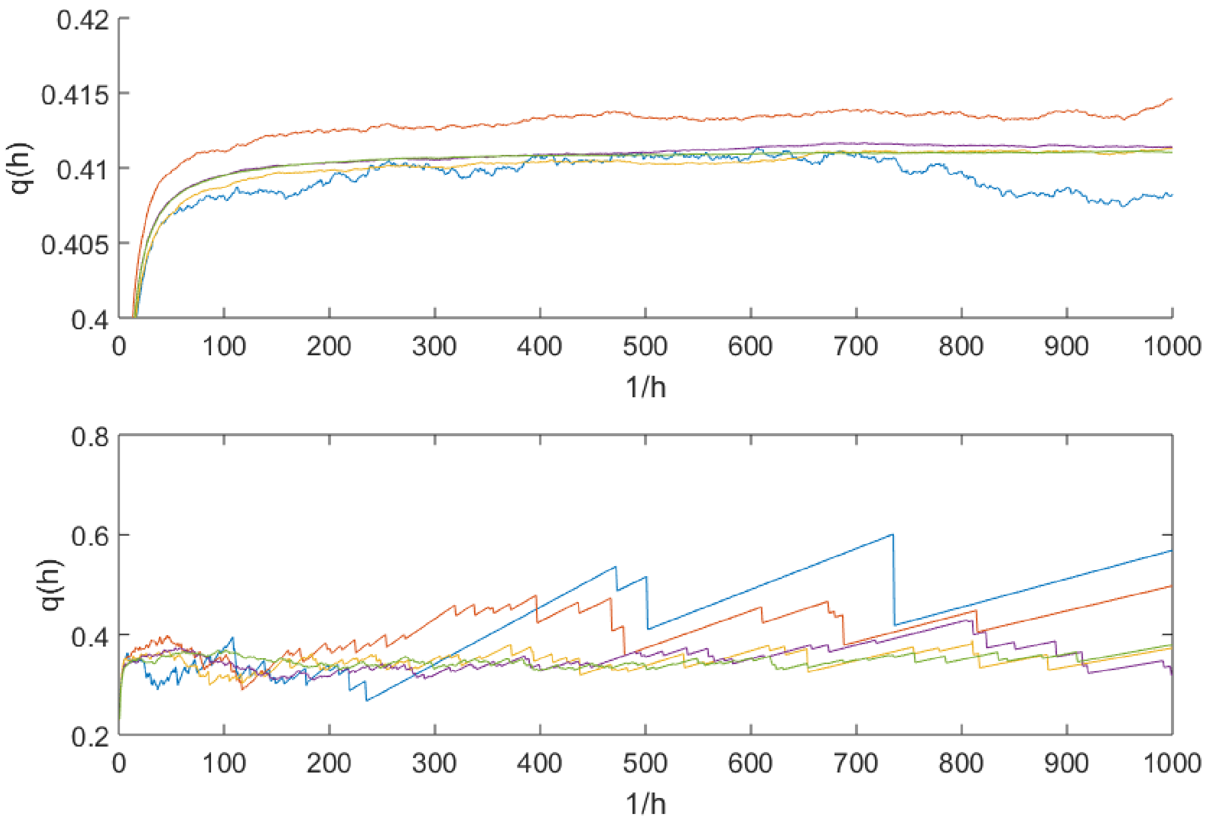
Assumption 3. i) $g_0(\theta) = 0$ only for $\theta = \theta_0$.

ii) The space Θ of all θ is compact.

iii) Every bivariate marginal copula $C_{ij}(u_i, u_j; \theta)$ of $\mathbf{C}(u;\theta)$ is Lipschitz-continuous for $(u_i, u_j) \in (0,1) \times (0,1)$ on Θ.

iv) The sequential weighting matrix $\hat{W}_{sT}$ is $O_p(1)$ and $\sup_{s \in [\varepsilon, 1]} \|\hat{W}_{sT} - W\| \xrightarrow{p} 0$ for $\varepsilon > 0$, where W is probability limit of W_{sT}.

Figure 3.7: Quotient $q(h)$



Note: Quotient $q(h)$ for $h = \frac{1}{i}$ for $i = 1, \ldots, 1000$, $\theta_2 = 1.0$ and $N = 10$ such as $T = \{250(\text{blue}), 500(\text{orange}), 1000(\text{yellow}), 2000(\text{purple}), 4000(\text{green})\}$. Results for $\hat{m}_{ij} = \hat{\rho}_{ij}$ (upper panel) and $\hat{m}^{ij} = \hat{\lambda}^{ij}_{0.1}$ (lower panel) using Model 3.7.

v) It holds for the moment simulating function $\tilde{m}_S(\theta)$ that, for $\theta_1, \theta_2 \in \Theta$,

$$|\tilde{m}_S(\theta_1) - \tilde{m}_S(\theta_2)| \leq C_S \|\theta_1 - \theta_2\|$$

with a random variable C_S that is independent of $\theta_1 - \theta_2$ and that fulfills $E(C_S^{2+\delta}) < \infty$ for some $\delta > 0$.

We checked Assumption 3 v) for the case of $\hat{m}_{ij} = \hat{\rho}_{ij}$ and $\hat{m}^{ij} = \hat{\lambda}^{ij}_{0.1}$ using Model 3.7. We considered $\theta_1 = \theta_2 + h$ where $h = \frac{1}{i}$ for $i = 1, \ldots, 1000$, $\theta_2 = 1.0$ and $N = 10$. We varied $T = \{250, 500, 1000, 2000, 4000\}$ and the Results can be seen in Figure 3.7.

Figure 3.7 reveals that the quotient $q(h) := \frac{|\tilde{m}_S(\theta_1)-\tilde{m}_S(\theta_2)|}{|\theta_1-\theta_2|}$ seems to be bounded for increasing sample size T independently of the parameter difference $\frac{1}{i}$.

Finally, we need an assumption for distributional results, which is the same as Assumption 4 in Oh and Patton (2013) with a difference in part iii).

Assumption 4. i) θ_0 is an interior point of Θ.

ii) $g_0(\theta)$ is differentiable at θ_0 with derivative G such that $G'WG$ is non singular.

iii) $\forall s \in [\varepsilon, 1], \varepsilon > 0 : \quad g_{sT,S}(\hat{\theta}_{sT,S})'\hat{W}_{sT}g_{sT,S}(\hat{\theta}_{sT,S}) = \inf_{\theta\in\Theta} g_{sT,S}(\theta)'\hat{W}_{sT}g_{sT,S}(\theta)+o_p^*((s^2T)^{-1})$, where $o_p^*((s^2T)^{-1})$ (instead of $o_p((s^2T)^{-1})$) indicates that the remainder term on the right hand side tends to zero *and* is non-negative.

With these assumptions, we can formulate our main theorem:

Theorem 4. Under the null hypothesis $H_0 : \theta_1 = \theta_2 = \cdots = \theta_T$ and if Assumptions 1-4 hold, we obtain for $\varepsilon > 0$

$$s\sqrt{T}\left(\hat{\theta}_{sT,S} - \theta_0\right) \overset{d}{\Longrightarrow} A^*(s)$$

as $T, S \to \infty$ in the space of càdlàg functions on the interval $[\varepsilon, 1]$ and $\frac{S}{T} \to k \in (0, \infty)$ or $\frac{S}{T} \to \infty$. Here, $A^*(s) = (G'WG)^{-1} G'W(A(s) - \frac{s}{\sqrt{k}}A(1))$, $A(s)$ is a Gaussian process defined in the proof of Lemma 11 in the appendix and θ_0 the value of all θ_t under the null.

With Theorem 4 we obtain the asymptotic distribution under the null of our parameter test statistic.

Corollary 1. Under the null hypothesis $H_0 : \theta_1 = \theta_2 = \cdots = \theta_T$ and if Assumptions 1-4 hold, we obtain for our test statistic

$$P = \sup_{s\in[\varepsilon,1]} s^2T(\hat{\theta}_{sT,S} - \hat{\theta}_{T,S})'(\hat{\theta}_{sT,S} - \hat{\theta}_{T,S}) \xrightarrow{d} \sup_{s\in[\varepsilon,1]} (A^*(s) - sA^*(1))'(A^*(s) - sA^*(1))$$

as $T, S \to \infty$ and $\frac{S}{T} \to k \in (0, \infty)$ or $\frac{S}{T} \to \infty$.

The estimation of the change point location is embedded in calculating the test statistic and is given by $\lfloor \tilde{s}T \rfloor$, where $\tilde{s}$ is the maximum point of the quadratic left side of Corollary 1, i.e.

$$\tilde{s} = \underset{s\in[\varepsilon,1]}{\mathsf{argmax}}\ s^2T(\hat{\theta}_{sT,S} - \hat{\theta}_{T,S})'(\hat{\theta}_{sT,S} - \hat{\theta}_{T,S}).$$

For our nonparametric moment (dependence measure) test we derive the following asymptotic distribution:

Corollary 2. Under the null hypothesis $H_0 : \theta_1 = \theta_2 = \cdots = \theta_T$ and if Assumptions 1-2 hold, we obtain for our test statistic

$$M = \sup_{s\in[\varepsilon,1]} s^2T(\hat{m}_{sT} - \hat{m}_T)'(\hat{m}_{sT} - \hat{m}_T) \xrightarrow{d} \sup_{s\in[\varepsilon,1]} (A(s) - sA(1))'(A(s) - sA(1))$$

as $T, S \to \infty$ and $\frac{S}{T} \to k \in (0, \infty)$ or $\frac{S}{T} \to \infty$.

The location of the changepoint is estimated in the same fashion as for P. Note, the asymptotic distribution of the moment test, as well as the asymptotic distribution of the parameter test, are not known in closed form and depend on the underlying sample. For this reason we cannot compute or simulate the critical values directly and need a bootstrap procedure to overcome this issue.

3.4. Bootstrap Distribution

The bootstrap distributions of the test statistics P and M are obtained by calculating B versions of the moment process $\frac{t}{T}\sqrt{T}\left(\hat{m}_t^{(b)} - \hat{m}_T^{(b)}\right)$, which can be calculated fast and directly from the data. It is therefore not necessary to solve B minimization problems which would produce a high computational effort.

We estimate the distribution under the null by using an i.i.d. bootstrap with the following steps:

i) Sample with replacement from the standardized residuals $\{\hat{\eta}_i\}_{i=1}^T$ to obtain B bootstrap samples $\{\hat{\eta}_i^{(b)}\}_{i=1}^T$, for $b = 1, \ldots, B$.

ii) Use $\{\hat{\eta}_i^{(b)}\}_{i=1}^t$ to compute $\hat{m}_t^{(b)}$ for $b = 1, \ldots, B$ and $t = \varepsilon T, \ldots, T$ and $\{\hat{\eta}_i\}_{i=1}^T$ to obtain $\hat{m}_T$.

iii) Calculate the bootstrap analogue of the limiting distribution of Corollary 1.

$$K^{(b)} := \max_{t \in \{\varepsilon T, \ldots, T\}} \left(A_*^{(b)} \left(\frac{t}{T} \right) - \frac{t}{T} A_*^{(b)} (1) \right)' \left(A_*^{(b)} \left(\frac{t}{T} \right) - \frac{t}{T} A_*^{(b)} (1) \right),$$

with $A_*^{(b)} \left(\frac{t}{T} \right) := (\hat{G}' \hat{W}_T \hat{G})^{-1} \hat{G}' \hat{W}_T A^{(b)} (\frac{t}{T})$ and $A^{(b)}(\frac{t}{T}) = \frac{t}{T} \sqrt{T} \left(\hat{m}_t^{(b)} - \hat{m}_T \right)$, where $\hat{G}$ is the two sided numerical derivative estimator of G, evaluated at point $\hat{\theta}_{T,S}$, computed with the full sample $\{\hat{\eta}_i\}_{i=1}^T$. We can compute the k-th column of $\hat{G}$ by

$$\hat{G}^k = \frac{g_{T,S}(\hat{\theta}_{T,S} + e_k \epsilon_{T,S}) - g_{T,S}(\hat{\theta}_{T,S} - e_k \epsilon_{T,S})}{2\epsilon_{T,S}}, \quad k \in \{1, \ldots, p\},$$

where e_k is the k-th unit vector, whose dimension ist $p \times 1$ and $\epsilon_{T,S}$ has to be chosen in a way that it fulfills $\epsilon_{T,S} \to 0$ and $\mathsf{min}\{\sqrt{T}, \sqrt{S}\} \epsilon_{T,S} \to \infty$.

iv) Compute B versions of $K^{(b)}$ and determine the critical value K such that

$$\frac{1}{B} \sum_{b=1}^{B} \mathbb{1}\{K^{(b)} > K\} = 0.05.$$

For our simulation study and the empirical application we use a step size $\epsilon_{T,S} = 0.1$, which worked best following Oh and Patton (2013). Critical values of the moment based test M are obtained similarly by adapting step iii) of the algorithm.

The intuition for the validity of the bootstrap, beside the fact that we only use the natural estimators for the respective terms, is as follows: Under the null hypothesis, we draw with replacement from the empirical distribution function which is close to the true distribution function. Due to the structure of the limit distribution of the test statistic, we can directly generate realizations from this without having to care about a suitable centering. Under the alternative of one fixed break at time t, the bootstrap quantiles remain bounded because the bootstrap procedure mimics a stationary distribution. By randomly drawing from either the data before or after the break, we effectively draw from stationary distribution which takes the parameters before the break with probability t/T and the ones after the break with probability 1-t/T. Using the above described bootstrap procedure the simulation results indicate that the test results in a reasonable sized and powered testing procedure, cf. Table 1, Table 2 and Table 3 in Section 3.7. A formal proof of the bootstrap validity is left for future research.

3.5. Discussion on Broader Applicability

Although the focus of this work lies on factor copulas, our tests are not restricted to this case. For example, if the factor copula structure in equation (3.1) is replaced by another copula, say an Archimedean copula, the parameter test (3.3) can be performed in a similar way. To obtain a valid test (size control and consistency under fixed alternatives), it is necessary that the SMM procedure yields consistent parameter estimators of the model under the null hypothesis of constant parameters. Oh and Patton (2017) show that many commonly used copulas can be expressed as factor copulas, whereas in Oh and Patton (2013) the estimation of other types of copula models by SMM is considered. On the other hand, we would like to stress that we consider factor copula models as the main application of SMM estimation, at least in financial econometrics. Simpler models can be estimated by ML or GMM, for which the literature already provides change point tests (see e.g. Wied, 2013).

If the model is misspecified (i.e., that the simulated moments do not arise from the correct model), it cannot be expected that the test is valid. We investigate this case in the simulation section and we find that the tests are, in fact, correctly sized in the case of misspecification. On the other hand, the moment-based test (3.4) can be interpreted as a general constancy test for dependence measures such as Spearman's rho or quantile dependencies (compare e.g. Wied, Dehling, van Kampen, and McFadden, 2013). It is in fact only indirectly linked to the factor copula model, i.e., this test detects changes in the moments which are induced by changes in the parameters. Therefore, the presence of a factor copula model is not necessary for this test.

3.6. Discussion of Assumption 2.7

Bücher et al. (2014) derive a result similar to the one we have in Assumption 2.7. In particular, in their Proposition 3.3, the limit process is given by

$$\mathbb{B}(s, \mathbf{u}) - \sum_{j=1}^{N} \partial_j C(\mathbf{u}) \mathbb{B}(s, \mathbf{u}^{(j)}).$$

Here, $\mathbf{u}^{(j)} \in [0,1]^N$ is defined by $\mathbf{u}_i^{(j)} = \mathbf{u}_j$, if $i = j$ and 1 otherwise. Moreover $\mathbb{B}(s, \mathbf{u})$ is a tight centered continuous Gaussian process with $\mathbb{B}(0, \mathbf{u}) = 0$ and

$$\mathsf{Cov}(\mathbb{B}(s, \mathbf{u}), \mathbb{B}(t, \mathbf{v})) = \mathsf{min}(s, t) \mathsf{Cov}(\mathbb{1}(\mathbf{F}(\eta) \leq \mathbf{u}), \mathbb{1}(\mathbf{F}(\eta) \leq \mathbf{v})).$$

The difference between their setting and ours is that they do not consider residuals, but the original observations. However, we can transfer this result by combining a result by Remillard (2017) and Bücher and Kojadinovic (2016). Remillard (2017) considers the case of residuals. We cannot use his copula results directly, because he only considers the case of $\hat{F}_j^1$, i.e., the case, where the ranks are not calculated sequentially. Nevertheless, Theorem 1 in Remillard (2017) gives a convergence result for the residuals themselves and thus also for the residuals transformed by the (unknown) limit of the empirical distribution function of the residuals. Combined with Theorem 3.4 in Bücher and Kojadinovic (2016), under the additional assumption that the residuals are strictly stationary, we obtain that the process in (3.5) converges to

$$A^*(s, \mathbf{u}) := \mathbb{B}^*(s, \mathbf{u}) - \sum_{j=1}^{N} \partial_j C(\mathbf{u}) \mathbb{B}^*(s, \mathbf{u}^{(j)}),$$

where $\mathbb{B}^*(s, \mathbf{u}) = \mathbb{B}(s, \mathbf{u}) + s\mathbb{B}^{**}(\mathbf{u})$ and details about $\mathbb{B}^{**}(\mathbf{u})$ can be found in Theorem 1 in

Remillard (2017). In particular, it follows that the limit of the sequential empirical copula CUSUM process (where $C(\mathbf{u})$ is replaced by $\frac{1}{T}\sum_{t=1}^{\lfloor T \rfloor} \mathbb{1}\{\hat{\mathbf{F}}^{s}(\hat{\eta}_t) \leq \mathbf{u}\}$) does not depend on whether residuals or the original observations are used.

3.7. Monte Carlo Simulations

In order to study the behaviour of our tests in finite samples and the quality of the bootstrap approximations we perform a small set Monte Carlo simulations. To this end we consider the one factor copula model

$$[X_{1t}, \ldots, X_{Nt}]' =: X_t = \boldsymbol{\beta}_t Z_t + \boldsymbol{q}_t, \tag{3.7}$$

with $\boldsymbol{\beta}_t = (\beta_t, \ldots, \beta_t)'$ a vector of size N, $Z_t \sim \mathsf{Skew\ t}\,(\nu^{-1}, \lambda)^1$ and $q_t \overset{i.i.d.}{\sim} \mathsf{t}\,(\nu^{-1})$ for $t = 1, \ldots, T$. We fix $\nu^{-1} = 0.25$ and $\lambda = -0.5$, such that our model is parametrized by the single factor loading $\theta_t = \beta_t$.

For the estimation of the sequential parameters β_t for $t = \varepsilon T, \ldots, T$ in the test statistic we use the SMM approach with $S = 25 \cdot T$ simulations to match the simulated dependence measures with the dependence measures computed from the data. For this we use five dependence measures, namely Spearman's rank correlation and the 0.05, 0.10, 0.90, 0.95 quantile dependence measures, averaged across all pairs. Note that the burn-in period $\lfloor \varepsilon T \rfloor$ has to be chosen sufficiently large in order to obtain reasonable parameter estimates for $\theta_{\lfloor \varepsilon T \rfloor}$ in our test statistic. Unreported simulations suggested that for samples with less than 100 observations highly unreasonable estimates can occur that severely affect the behaviour of our test. We decided to use $\varepsilon = 0.2$. While this is a limitation of our test in the sense that breaks at the beginning of the sample cannot be identified, truncating the sample is common in some tests for structural breaks, see Andrews (1993) or Qu and Perron (2007). Furthermore, breaks at the beginning and the end of the sample are typically hard to detect in any case.

[1] As in Oh and Patton (2017) this refers to the skewed t-distribution by Hansen (1994).

We consider three tests in this simulation exercise, namely the parameter based fluctuation test (P) given in equation (3.3), the test based on the moment condition (M) given in (3.4) and the nonparametric test for copula constancy proposed by Bücher et al. (2014) abbreviated as BKRS. The change point detection in the latter test is sensitive to changes in the copula of the multivariate continuous observations and is included as a benchmark. We do note, however, that this test is purely nonparametric in contrast to our test P that is based explicitely on factor copula models. Critical values of our tests are computed using the bootstrap algorithm from Section 3.4 with $B = 1000$ bootstrap replications. The tests are performed at the $\alpha = 0.05$ significance level and we use 301 Monte Carlo replications. The computational complexity of the simulations was extremely high due to the fact that for each test $\hat{\theta}_{sT,S}$ needs to be estimated a large number of times using the computationally heavy SMM estimator and because critical values have to be bootstrapped. This explains why we had to restrict ourselves to a limited number of situations for a fairly simple model. Furthermore, numerical instabilities were present in more complex models when repeatedly estimating the model parameters. Such problems can be dealt within empirical applications, but further restrict the potential model complexity in simulations. The computations were implemented in Matlab, parallelized and performed using CHEOPS, a scientific High Performance Computer at the Regional Computing Center of the University of Cologne (RRZK) funded by the DFG.

Table 1: Size retro perspective Testing

	$\theta_0 = 0.5$	$N=5$	$N=10$	$N=20$	$\theta_0 = 1$	$N=5$	$N=10$	$N=20$	$\theta_0 = 2.0$	$N=5$	$N=10$	$N=20$
$T = 500$												
	P	0.102	0.079	0.056	P	0.066	0.056	0.053	P	0.059	0.059	0.056
	M	0.029	0.036	0.046	M	0.030	0.039	0.056	M	0.033	0.043	0.043
	BKRS	0.046	0.036	0.033	BKRS	0.049	0.053	0.049	BKRS	0.056	0.046	0.023
$T = 1000$												
	P	0.089	0.049	0.049	P	0.056	0.046	0.069	P	0.049	0.059	0.049
	M	0.046	0.033	0.056	M	0.049	0.043	0.076	M	0.043	0.056	0.069
	BKRS	0.043	0.046	0.056	BKRS	0.066	0.056	0.076	BKRS	0.069	0.060	0.046
$T = 1500$												
	P	0.073	0.059	0.043	P	0.056	0.069	0.066	P	0.066	0.069	0.066
	M	0.056	0.056	0.049	M	0.049	0.063	0.066	M	0.063	0.069	0.069
	BKRS	0.046	0.056	0.069	BKRS	0.053	0.069	0.066	BKRS	0.046	0.046	0.049

Note: Table 1 reports the rejection rate for $\theta_0 = 0.5$, $\theta_0 = 1$ and $\theta_0 = 2$ in the model (3.7) for the parameter Test (P) with $\varepsilon = 0.2$, the moment function test (M) and the nonparametric test of Bücher et al. (BKRS).

We begin by studying the size of the tests for three parameter values $\theta_0 = 0.5$, $\theta_0 = 1$ and $\theta_0 = 2$, sample sizes $T = 500, 1000, 1500$ and cross sectional dimensions $N = 5, 10, 20$. Results are presented in Table 1. All tests have acceptable size properties except the parameter based test for small dimensions and sample sizes in the case $\theta_0 = 0.5$. However, as N and T increase the size clearly tends to the nominal level of 5%.[2]

Furthermore, we study the size of the tests when the DGP is not a factor copula. Here we again consider $N = 5, 10, 20$, but only $T = 1000$. For the parameter based test this means that the model is misspecified. The first case is a Clayton copula with parameter $\theta = 1$, a model implying equidependence with kendall's τ equal to 1/3 and lower tail dependence. The second DGP is a (truncated) D-vine copula model (see Aas et al., 2009). One the first tree all pairs are connected with a Clayton copula with $\theta = 2$, the second tree has Gaussian copulas with $\rho = 0.5$ and the third tree survival Gumbel copulas with parameter $\gamma = 1.25$. All remaining trees have conditional independence, implying the truncation of the model. This model does not imply equidependence, but lower tail dependence is still present for all pairs. The parametric test P is based on the same one-factor copula model (3.7) for both cases. The results in Table 2 show that all tests have good size properties. The parameter based test P is slightly undersized for $N = 5$. From the two examples (Clayton and D-vine copula), it seems that the test is also reliable when the underlying model is misspecified.

To study the power of the test, we generate data with a break point at $\frac{T}{2}$ for all sample sizes, where the data is simulated with $\theta_t = 1$ for $t \in \{\varepsilon T, \ldots, \frac{T}{2}\}$, denoted by θ_0, whereas after the break we increase the parameter to $\theta_t = \{1.2, 1.4, 1.6, 1.8, 2.0\}$ for $t \in \{\frac{T}{2} + 1, \ldots, T\}$, denoted by θ_1. Here we consider $N = 5, 10, 20, 40$, but restrict the sample size to the cases $T = 500, 1000$. The results can be found in Table 3. Note that the first column of the table

[2]Note that a larger burn-in period εT leads to a slightly better size properties, in particular for small values of T and N, which can be explained by a lower degree of variation in the numerical minimization procedure.

Table 2: Size under alternative copula DGPs

	$N=5$	$N=10$	$N=20$	$N=5$	$N=10$	$N=20$
$T=1000$	Clayton copula			D-vine copula		
P	0.033	0.037	0.049	0.033	0.056	0.059
M	0.043	0.047	0.069	0.033	0.063	0.063
BKRS	0.057	0.049	0.043	0.043	0.053	0.073

Note: Table 2 reports the rejection rate under H_0 for alternative data generating processes for the parameter Test (P) with $\varepsilon = 0.2$ and assuming a one-factor copula, the moment function test (M) and the nonparametric test of Bücher et al. (BKRS), where a Clayton and a D-vine copula are used for the DGP.

reports the size of the tests again. For $N = 40$ the tests have good size properties for $T = 500$, but are slightly oversized for $T = 1000$. The BKRS test has the most severe size distortions in this case. We observe that all tests have good power that increases with θ_1 and sample size T. The parameter based test P and the moment test M have increasing power as N increases from 5 to 40, whereas the power of the BKRS test decreases for the higher dimensional case. For $N = 5, 10, 20$ the BKRS test has the highest power followed by the parameter based test. For $N = 40$, however, the P test performs better and even the M test has (mostly) more power than the nonparametric BKRS test. This indicates that the tests based on the factor copula model are preferable for higher dimensional situations. This can be explained by the fact that more available data improves the SMM estimation, while in a nonparametric copula constancy test the complexity of the estimated objects increase.

3.8. Empirical Application

In this section we apply our test to a financial dataset. We use daily stock return data over a time span ranging from July 2005 to May 2009 from the EURO STOXX 50 of the four largest industry sectors Finance, Energy, Telecom and Media and Consumer Retail and

Table 3: Power retro perspective Testing

	$\theta_0 = 1$	$\theta_1 = 1.2$	$\theta_1 = 1.4$	$\theta_1 = 1.6$	$\theta_1 = 1.8$	$\theta_1 = 2.0$
$N = 5$,$T = 500$						
P	0.066	0.272	0.551	0.833	0.963	0.993
M	0.030	0.173	0.452	0.771	0.940	0.987
BKRS	0.049	0.272	0.727	0.946	0.996	1.000
$N = 10$,$T = 500$						
P	0.056	0.266	0.658	0.887	0.993	1.000
M	0.039	0.236	0.558	0.877	0.983	0.997
BKRS	0.053	0.285	0.764	0.973	1.000	1.000
$N = 20$,$T = 500$						
P	0.053	0.299	0.704	0.907	1.000	1.000
M	0.056	0.259	0.628	0.900	0.993	1.000
BKRS	0.049	0.275	0.750	0.966	1.000	1.000
$N = 40$, $T = 500$						
P	0.043	0.302	0.691	0.910	0.996	1.000
M	0.059	0.282	0.635	0.920	0.993	1.000
BKRS	0.059	0.225	0.588	0.903	0.996	1.000
$N = 5$,$T = 1000$						
P	0.056	0.352	0.781	0.980	1.000	1.000
M	0.049	0.285	0.717	0.966	1.000	1.000
BKRS	0.066	0.481	0.946	1.000	1.000	1.000
$N = 10$,$T = 1000$						
P	0.046	0.415	0.874	0.993	1.000	1.000
M	0.043	0.352	0.801	0.993	1.000	1.000
BKRS	0.056	0.478	0.963	1.000	1.000	1.000
$N = 20$,$T = 1000$						
P	0.069	0.455	0.887	1.000	1.000	1.000
M	0.076	0.389	0.834	0.993	1.000	1.000
BKRS	0.076	0.465	0.943	0.996	1.000	1.000
$N = 40$, $T = 1000$						
P	0.076	0.4751	0.927	1.000	1.000	1.000
M	0.073	0.399	0.844	0.993	1.000	1.000
BKRS	0.093	0.398	0.880	0.993	0.996	1.000

Note: Table 3 reports the rejection rate for $\theta_0 = 1.0$ and $\theta_1 = 1.2, 1.4, 1.6, 1.8, 2$ in the model (3.7) for the parameter Test (P) with $\varepsilon = 0.2$, the moment function test (M) and the nonparametric test of Bücher et al. (BKRS).

we choose the subdivision in Table 4, implying $T = 1000$ and $N = 32$, with group sizes $k_1 = 13, k_2 = 8, k_3 = 5$ and $k_4 = 6$.

Table 4: Included Stocks by Industry

Finance	Allianz, Axa, Banco Bilbao, Banco Santander, BNP Paribas, Deutsche Bank, Deutsche Börse, Generali, ING Groep, Intesa, Münchener Rück, Société Générale, Unicredit
Energy	E.ON, ENEL, ENI, SUEZ, Iberdrola, Repsol, RWE, Total
Telecom and media	Deutsche Telekom, France Telecom, Telecom Italia, Telefonica, Vivendi
Consumer retail	Anheuser Busch, Carrefour, Danone, L'Oreal, LVMH, Unilever

To model the conditional mean and variance we estimate an AR(1)-GARCH(1,1) model for each return series and compute the standardized residuals,

$$r_{i,t} = \alpha_i + \beta_i r_{i,t-1} + \sigma_{i,t}\eta_{it},$$

$$\sigma_{it}^2 = \gamma_{i0} + \gamma_{i1}\sigma_{i,t-1}^2 + \gamma_{i2}\eta_{i,t-1}^2,$$

for $t = 1, \dots, 1000$. The marginal distribution of the residuals are estimated using the empirical CDF. Following Oh and Patton (2017) we specify the following block-equidependence five factor copula model:

$$[X_{1t}, \dots, X_{Nt}]' =: X_t = \begin{pmatrix} \boldsymbol{\beta}_{1t} \\ \boldsymbol{\beta}_{2t} \\ \boldsymbol{\beta}_{3t} \\ \boldsymbol{\beta}_{4t} \end{pmatrix} Z_{0t} + \begin{pmatrix} \boldsymbol{\beta}_{5t} Z_{1t} \\ \boldsymbol{\beta}_{6t} Z_{2t} \\ \boldsymbol{\beta}_{7t} Z_{3t} \\ \boldsymbol{\beta}_{8t} Z_{4t} \end{pmatrix} + \boldsymbol{q}_t, \tag{3.8}$$

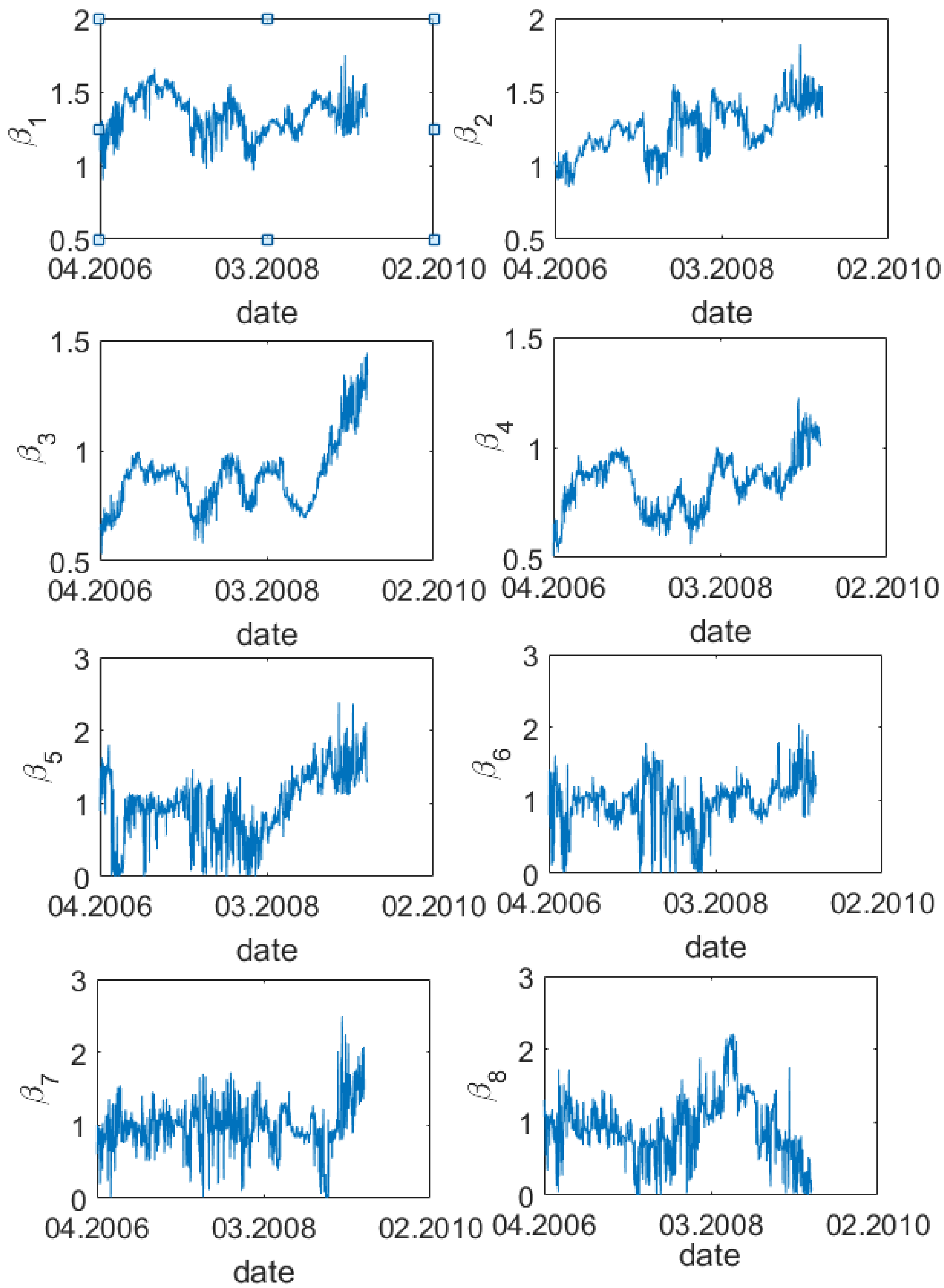
with $\boldsymbol{\beta}_{it} = (\beta_{it}, \dots, \beta_{it})'$ of size k_i for $i = 1, 2, 3, 4$, where $Z_{0t} \sim \mathsf{Skew\ t}\,(\nu^{-1}, \lambda)$ and $Z_{it} \sim \mathsf{t}\,(\nu^{-1})$ for $i = 1, 2, 3, 4$ and $\boldsymbol{q}_t \overset{i.i.d.}{\sim} \mathsf{t}\,(\nu^{-1})$ for $t = 1, \dots, T$. Thus, we have have one common factor with industry specific factor loadings $\boldsymbol{\beta}_{it}$ for $i = 1, \dots, 4$ and four industry specific factors with corresponding loadings $\boldsymbol{\beta}_{it}$ for $i = 5, \dots, 8$. We assume identical degrees of freedom for the common factors and the idiosyncratic errors implying a model with tail dependence strictly between zero and one.

For the estimation of the model we use the SMM approach described above with $S = 25 \cdot T$ simulations. The moment conditions are based on five dependence measures, namely Spearman's rank correlation and the 0.05, 0.10, 0.90, 0.95 quantile dependence. In the block equidependence model with four groups and five dependence measures this gives us a total number of $4 \times 5 = 20$ dependence measures.

The full sample estimates can be found in Table 5. For studying the time-variation we fix λ and ν at their full sample estimates to avoid numerical problems as these parameters are difficult to estimate for small samples. As a preliminary analysis, we estimate the model over a rolling window of 200 days. Figure 3.8 shows that there is some variation over time in the factor loadings with an apparent increase in most parameters towards the end of the sample. The results of the tests for a structural break in the factor copula parameters can be found in Table 6. The moment based test M finds a significant breakpoint on January 8, 2008. The BKRS test finds a similar break data (Jan. 17), but the statistic is only significant at the 10% level. The parameter test P applied to all factor loadings indicates a break slightly later on March 7, 2008. This is a little earlier than the peak of the financial crisis with Lehman Brothers filing bankruptcy on September 15. Some of the estimated parameters after the break are larger than before the break while other decrease. This makes the direct interpretation of the change in dependence difficult. We return to the implied dependence of the model before and after the break below.

As the dataset contains companies from different sectors we applied the P test to a number of subvectors of the factor loadings. To be precise, we tested for a break in the loading of the market factor alone and of the loadings on the market and group specific factors for each respective sector, while fixing the remaining model parameters at their full sample estimates.

Figure 3.8: Rolling Window Parameter Estimates



Note: Rolling window parameter estimation for the factor loadings β_i for a window of size 200.

For all subsets we find evidence of a structural break. However, the estimated break dates are later than for the full set of loading mainly around the peak of the financial crisis. The break for the loading corresponding to the energy sector is later in December 2008. Comparing the estimated parameters before and after the break, some of the loadings decrease after the estimated breakpoint. Part of the apparent discrepancies between the results for the full loading vector and the analysis on the subsets can be explained by the differences in estimated break dates coupled with the fact that the estimation uncertainty for the relatively small post-break period is quite large, which is due to the fact that factor copulas are difficult to estimate on such small samples.

A direct interpretation of the change in the factor loadings is difficult due to the complex interactive effect the different factors have on the overall dependence structure. Therefore, we computed (by simulation) the rank correlations implied by the different break models. The result can be found in Table 7. As we have a block-equidependence model the implied dependence for assets within each sector is the same, as is the dependence between assets from two sectors. The within sector dependence is given on the main diagonal of the presented matrices, while the between sector dependences are given by the off-diagonal elements.

The results based on the break in all factor loadings indicate increasing (Energy, Telecom) or almost stable (Finance, Consumer) within sector rank correlations and slightly increasing rank correlations between the sectors. The break for the market factor loadings implies a similar change in dependence, but a stronger increase between the Telecom and Energy sectors. For the sector specific breaks we note that the results for the finance sector indicate a slight decrease within the finance sector, but increased dependence with the other sectors, which can be interpreted as an indication of contagion from the finance sector to the other sectors. For the energy sector specific case we observe an increase both within the sector and across sectors. The case of the telecommunication sector indicates an increase within the sector, but mostly stable dependence with the other sectors. For the consumption sector-specific breaks

Table 5: Full sample Parameter Estimates of the Model (3.8)

	$\hat{\nu}$	$\hat{\lambda}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\beta}_8$
$\hat{\theta}_{T,S}$	9.263	-0.157	1.491	0.970	1.253	0.889	1.095	0.517	0.678	1.668
std	1.249	0.065	0.132	0.048	0.082	0.073	0.433	0.159	0.169	0.082

we again see stable within sector dependence, but increased dependence across the sectors. Overall, we conclude from this that dependence has indeed increased after the break but the increase in dependence is not as strong and clear cut as one might expect. Estimation uncertainty for the sub-periods may partly explain the mixed results.

In order to get a clearer picture of the evolution of the size and structure of the dependence with respect to the breakpoint we computed the (averaged) dependence measures that were used for estimation before and after the breakpoint indicated by the M test, see Table 8. The results indicate that the overall dependence measured by the rank correlation ρ increases. Similarly, the upper quantile dependence measures $\lambda_{0.9}$ and $\lambda_{0.95}$ increase after the break. Surprisingly, the lower quantile dependence stays approximately constant indicating that the dependence of the (left) tail risk for the data at hand has not increased after the estimated breakpoint while overall the diversification benefits have decreased.

Table 6: Breakpoint Test Results

		test stat	p-val	Break date	$\hat{\theta}_{pre}$	$\hat{\theta}_{post}$
θ_{factor}	$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \\ \beta_7 \\ \beta_8 \end{pmatrix}$	2008.50	0.000	07.03.2008	$\begin{pmatrix} 1.36 \\ 0.77 \\ 1.16 \\ 0.86 \\ 1.55 \\ 0.73 \\ 0.76 \\ 0.93 \end{pmatrix}$	$\begin{pmatrix} 1.32 \\ 1.16 \\ 1.28 \\ 1.00 \\ 1.22 \\ 1.10 \\ 1.31 \\ 0.56 \end{pmatrix}$
θ_{market}	$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$	49.71	0.000	15.08.2008	$\begin{pmatrix} 1.38 \\ 0.74 \\ 1.14 \\ 0.82 \end{pmatrix}$	$\begin{pmatrix} 1.24 \\ 1.41 \\ 1.30 \\ 0.93 \end{pmatrix}$
$\theta_{finance}$	$\begin{pmatrix} \beta_1 \\ \beta_5 \end{pmatrix}$	1297.10	0.000	02.09.2008	$\begin{pmatrix} 1.24 \\ 1.88 \end{pmatrix}$	$\begin{pmatrix} 1.35 \\ 1.16 \end{pmatrix}$
θ_{energy}	$\begin{pmatrix} \beta_2 \\ \beta_6 \end{pmatrix}$	323.98	0.001	05.12.2008	$\begin{pmatrix} 0.76 \\ 0.65 \end{pmatrix}$	$\begin{pmatrix} 1.41 \\ 1.33 \end{pmatrix}$
θ_{tele}	$\begin{pmatrix} \beta_3 \\ \beta_7 \end{pmatrix}$	907.39	0.000	02.09.2008	$\begin{pmatrix} 1.20 \\ 1.44 \end{pmatrix}$	$\begin{pmatrix} 1.52 \\ 0.75 \end{pmatrix}$
θ_{consum}	$\begin{pmatrix} \beta_4 \\ \beta_8 \end{pmatrix}$	583.23	0.000	02.07.2008	$\begin{pmatrix} 0.65 \\ 1.47 \end{pmatrix}$	$\begin{pmatrix} 0.95 \\ 1.20 \end{pmatrix}$
M		5.72	0.000	09.01.2008		
$BKRS$		8.25	0.065	17.01.2008		

Note: Table 6 reports tests for a structural break in the factor copula model (3.8). The penultimate row gives the results of the moment based test M. The last row gives the results of the nonparametric test of Bücher et al. (BKRS).The other rows show the results of the parameter based test P for the given subsets of the parameter vector while fixing the remaining parameter values at the full sample estimates. $\hat{\theta}_{pre}$ and $\hat{\theta}_{post}$ denote the parameter estimates before and after the estimated break dates, respectively. We use 1000 bootstrap replications.

Table 7: Implied Rank Correlations

	Pre-break				Post-break			
	Finance	Energy	Telecom	Consumer	Finance	Energy	Telecom	Consumer
Break all factors loadings								
Finance	0.79	0.30	0.38	0.30	0.74	0.37	0.37	0.40
Energy		0.52	0.34	0.27		0.70	0.36	0.38
Telecom			0.63	0.34			0.74	0.39
Consumer				0.59				0.55
Break market factor loadings								
Finance	0.73	0.36	0.44	0.25	0.71	0.47	0.44	0.27
Energy		0.44	0.36	0.21		0.67	0.54	0.32
Telecom			0.61	0.26			0.66	0.30
Consumer				0.75			0.76	
Break financial sector loadings								
Finance	0.82	0.32	0.35	0.20	0.74	0.41	0.45	0.27
Energy		0.53	0.45	0.26		0.53	0.45	0.26
Telecom			0.64	0.29			0.65	0.29
Consumer				0.76			0.76	
Break energy sector loadings								
Finance	0.75	0.36	0.48	0.28	0.75	0.43	0.48	0.28
Energy		0.49	0.37	0.22		0.77	0.44	0.26
Telecom			0.64	0.29			0.65	0.29
Consumer				0.76			0.76	
Break telecommunication sector loadings								
Finance	0.75	0.44	0.38	0.28	0.75	0.44	0.52	0.28
Energy		0.53	0.35	0.26		0.53	0.48	0.26
Telecom			0.75	0.23			0.72	0.31
Consumer				0.76				0.76
Break consumption sector loadings								
Finance	0.75	0.44	0.48	0.23	0.75	0.44	0.48	0.35
Energy		0.53	0.45	0.22		0.53	0.45	0.32
Telecom			0.64	0.24			0.65	0.36
Consumer				0.69				0.68

Note: Table 7 shows the model implied rank correlations before and after the estimated breakpoint corresponding to the subsets of factor loading allowed to break in Table 6 and using the corresponding break date and parameter estimates. The entries on the main diagonal are implied rank correlations between assets within the respective sector, the off-diagonal elements are the implied rank correlations between the sectors.

Table 8: Average Dependence Measures

	Full sample	Pre-break	Post-break
ρ_1	0.46	0.45	0.48
ρ_2	0.39	0.36	0.43
ρ_3	0.45	0.43	0.48
ρ_4	0.39	0.37	0.43
$\lambda^1_{0.05}$	0.31	0.29	0.27
$\lambda^2_{0.05}$	0.27	0.21	0.25
$\lambda^3_{0.05}$	0.28	0.27	0.25
$\lambda^4_{0.05}$	0.25	0.24	0.21
$\lambda^1_{0.1}$	0.40	0.37	0.36
$\lambda^2_{0.1}$	0.33	0.29	0.34
$\lambda^3_{0.1}$	0.38	0.35	0.34
$\lambda^4_{0.1}$	0.35	0.30	0.31
$\lambda^1_{0.9}$	0.34	0.30	0.39
$\lambda^2_{0.9}$	0.28	0.24	0.35
$\lambda^3_{0.9}$	0.31	0.28	0.36
$\lambda^4_{0.9}$	0.28	0.25	0.34
$\lambda^1_{0.95}$	0.27	0.19	0.33
$\lambda^2_{0.95}$	0.22	0.14	0.27
$\lambda^3_{0.95}$	0.23	0.15	0.29
$\lambda^4_{0.95}$	0.22	0.13	0.26

Note: Table 8 contains the (average) empirical moments used for the model estimator for the full sample and the subsamples implied by a structural break on Jan. 9, 2008 that was detected by the moment based structural break test. ρ_i denotes the rank correlation, whereas λ^i_q is the quantile q dependence measure for the sectors $i = 1, \ldots, 4$, i.e. finance, energy, telecom and media, consumer retail.

4. A MONITORING PROCEDURE FOR DETECTING STRUCTURAL BREAKS IN FACTOR COPULA MODELS

In this section we describe our results of the parameter and dependence measure monitoring procedure. Different to the retro perspective test presented in Section 3 the proposed procedures can be used in real time applications. The section is structured as follows: Section 4.1 presents the null hypothesis and monitoring procedure, whereas in Section 4.2 we study its asymptotic distribution under the setting of simulated method of moments estimation. Results from the Monte Carlo simulations can be found in Section 4.3. Section 4.4 presents our empirical application with an application for the evaluation of the Value at Risk (VaR) and the main proof can be found in Section 7.1.2 in the appendix. The work belongs to the working paper “A monitoring procedure for detecting structural breaks in factor copula models”, Manner et al. (2018).

4.1. Null Hypothesis and Testing Procedure

We are again interested in testing the hypothesis of no parameter change within the residual dependence model. The main idea is now to compare a parameter estimate from a historical data set of size $\lfloor mT \rfloor$ to sequential estimated parameters from a rolling data window of the same size, where T is the length of the monitored time period and m a value between $(0, 1]$. Since we are interested in sequentially monitoring whether or not the parameter θ_t changes in $t = \lfloor mT \rfloor + 1, \dots, T$, we assume that the parameters remain constant over the initial sample $t = 1, \dots, \lfloor mT \rfloor$, meaning that:

Assumption 5.

$$\theta_1 = \cdots = \theta_{\lfloor mT \rfloor}. \tag{4.1}$$

In practice, if a sufficient amount of initial data is available, this assumption can be tested by using the test for parmeter constancy in factor copulas proposed in Section 3.

We are interested in testing the null hypothesis

$$H_0: \quad \theta_1 = \dots = \theta_{\lfloor mT \rfloor} = \theta_{\lfloor mT \rfloor + 1} = \dots$$

versus the alternative

$$H_1: \quad \theta_1 = \dots = \theta_{\lfloor mT \rfloor} = \dots = \theta_{\lfloor mT \rfloor + k^* - 1} \neq \theta_{\lfloor mT \rfloor + k^*} = \theta_{\lfloor mT \rfloor + k^* + 1} = \dots,$$

by using the detector

$$D_T(s) := m^2 T (\hat{\theta}_{1+(s-m)T:sT} - \hat{\theta}_{1:mT})' (\hat{\theta}_{1+(s-m)T:sT} - \hat{\theta}_{1:mT}), \tag{4.2}$$

where $k^* \geq 1$ and $\lfloor mT \rfloor + k^*$ is the unknown change point and $\hat{\theta}_{t_1:t_2}$ a consistent estimator for θ that is based on the subsample ranging from t_1 to t_2. Note that for the sake of thrift, we use the same parameter m for the initial period and further rolling window periods. Furthermore we do not need a certain weighted deviation factor, due to the fact we consider a MOSUM-type test statistic in contrast to for example Pape et al. (2017), where an expanding window is used. We stop our monitoring procedure if the MOSUM-type detector defined in (4.2) exceeds the appropriately chosen constant critical value c for the first time k. This yields the stopping rule

$$\tau_T := \inf_k \left\{ k \leq T : D_T\left(\frac{k}{T}\right) > c \right\},$$

where τ_T is the stopping time of the monitoring procedure. Here c is chosen in a way that under H_0 the monitoring procedure holds the size level $\lim_{T,S \to \infty} P(\tau_T < \infty | H_0) = \alpha$, with $\alpha \in (0, 1)$. We write $\tau_T < \infty$ to indicate that the monitoring has been terminated during the testing period, meaning that the detector crossed the boundary value c at a time point $k \leq T$. On the other hand, we write $\tau_T = \infty$, if D_T does not cross the boundary value during the testing period.

Note that the detected stopping time τ_T is not meant to be an estimator of change point, as the actual change point is likely to be earlier. This is due to the fact the monitoring procedure needs a sufficient number of observations after a change point before it can be detected. In the next chapter we present a procedure for estimating the change point conditional on H_0 having been rejected.

4.2. Estimation and Asymptotics

In this section we describe our theoretical results. The estimation of the factor copula model by SMM is reviewed in Section 4.2.1, whereas the asymptotic behaviour of our monitoring procedures is studied in Section 4.2.2. A bootstrap algorithm to approximate the asymptotic distribution is presented in Section 4.2.3 and a procedure for detecting multiple breaks is described in Section 4.2.4.

4.2.1. SMM Estimation.

We are interested in estimating the parameter vector $\theta_{uT:vT}$ for the subsample ranging from $\lfloor uT \rfloor$ to $\lfloor vT \rfloor$, where $u < v$ and $u, v \in [\varepsilon, 1]$, with $\varepsilon > 0$. This is again achieved by using the SMM, where the estimator is defined as

$$\hat{\theta}_{uT:vT,S} := \operatorname*{arg\,min}_{\theta \in \Theta} Q_{uT:vT,S}(\theta),$$

where $Q_{uT:vT,S}(\theta) := g_{uT:vT,S}(\theta)' \hat{W}_{(uT:vT)} g_{uT:vT,S}(\theta)$ is the objective function, $g_{uT:vT,S}(\theta) := \hat{m}_{uT:vT} - \tilde{m}_S(\theta)$ and $\hat{W}_{(uT:vT)}$ a positive definite weight matrix with probability limit W, for simplicity one can chose the $k \times k$ identity matrix. The moment conditions (dependence vectors) $\hat{m}_{uT:vT}$ are $k \times 1$ vectors of appropriately chosen pairwise dependence measures $\hat{m}^{ij}_{uT:vT}$ (possibility averaged over equidependent pairs), computed from the residuals $\{\hat{\eta}_t\}_{t=\lfloor uT \rfloor}^{\lfloor vT \rfloor}$, whereas $\tilde{m}_S(\theta)$ can be thought of the corresponding vector of true dependence measures. The residuals are obtained in the same way as in Section 3.1. Note that the dependence measures implied by the factor copula model are typically not available in closed form and they have to be obtained by simulation. Therefore, the classical method of moments (MM) or generalized method of moments (GMM) cannot be used here. The true dependence measures are approximated using S simulations $\{\tilde{\eta}_t\}_{t=1}^S$ from $\boldsymbol{F}_{\boldsymbol{X}_t}$, and hence the objective function, the estimator, and consequently our detector defined in equation (4.2) depend on the number of simulations S. Following the simulation studies in Oh and Patton (2013), we chose $S = 25 \cdot (vT - uT)$ and we need to ensure that the sub-sample ranging from $\lfloor uT \rfloor$ to $\lfloor vT \rfloor$ is large enough to receive reasonable SMM estimates. In our simulation studies we find that our procedure still results in reasonable size and power properties by choosing $\lfloor uT \rfloor - \lfloor vT \rfloor = mT = 250$ data points. For the dependence measures of the pair (η_i, η_j), we use again Spearman's rank correlation ρ^{ij} and the quantile dependence λ_q^{ij}, defined in Section 2.3. The considered sample counterparts for the observations between $\lfloor uT \rfloor$ and $\lfloor vT \rfloor$ are defined as

$$\hat{\rho}^{ij} := \frac{12}{\lfloor vT - uT \rfloor} \sum_{t=\lfloor uT \rfloor}^{\lfloor vT \rfloor} \hat{F}_i^{uT:vT}(\hat{\eta}_{it}) \hat{F}_j^{uT:vT}(\hat{\eta}_{jt}) - 3$$

$$\hat{\lambda}_q^{ij} := \begin{cases} \frac{\hat{C}_{ij}^{uT:vT}(q,q)}{q}, & q \in (0, 0.5] \\ \frac{1-2q+\hat{C}_{ij}^{uT:vT}(q,q)}{1-q}, & q \in (0.5, 1) \end{cases},$$

where $\hat{F}_i^{uT:vT}(y) := \frac{1}{\lfloor vT - uT \rfloor} \sum_{t=\lfloor uT \rfloor}^{\lfloor vT \rfloor} \mathbb{1}\{\hat{\eta}_{it} \leq y\}$ and $\hat{C}_{ij}^{uT:vT}(u, v) := \frac{1}{\lfloor vT - uT \rfloor} \sum_{t=\lfloor uT \rfloor}^{\lfloor vT \rfloor} \mathbb{1}\{\hat{F}_i^{uT:vT}(\hat{\eta}_{it}) \leq u, \hat{F}_j^{uT:vT}(\hat{\eta}_{jt}) \leq v\}$. The simulated counterparts of these dependence measures based on the simulations $\{\tilde{\eta}_t\}_{t=1}^S$ are defined analogously and are denoted by $\tilde{\rho}^{ij}$ and $\tilde{\lambda}_q^{ij}$.

In summary, the SMM estimator minimizes the weighted difference between suitable sample dependence measures and their model counterparts obtained by simulation. Depending on the precise model specification, the pairwise dependence measures are averaged for groups, which have the same factor loadings. For more information on SMM estimation and a suitable way to average the pairwise dependence measures for equidependence or block equidependence models see Section 2.7 or in more detail Oh and Patton (2013) and Oh and Patton (2017).

4.2.2. Asymptotics. To derive the asymptotic distribution of our detector (4.2), we consider Assumption 5 and Assumptions 6-9 (given in the appendix) and follow similar steps as in Manner et al. (2019). The difference is that we replace the scale factor $s\sqrt{T}$ by $m\sqrt{T}$ and that we derive the following distributional limit for the process $s \mapsto m\sqrt{T}g_{1+(s-m)T:sT,S}(\theta)$, for $\frac{S}{T} \to k \in (0, \infty]$ and $T, S \to \infty$,

$$
\begin{aligned}
&m\sqrt{T}g_{1+(s-m)T:sT,S}(\theta) = m\sqrt{T}\left(\hat{m}_{1+(s-m)T:sT} - \tilde{m}_S(\theta)\right) \\
=&m\sqrt{T}\left(\hat{m}_{1+(s-m)T:sT} - m_0(\theta)\right) - m\sqrt{T}\left(\tilde{m}_S - m_0(\theta)\right) \\
=&m\sqrt{T}\left(\hat{m}_{1+(s-m)T:sT} - m_0(\theta)\right) - \sqrt{\frac{T}{S}}m\sqrt{S}\left(\tilde{m}_S - m_0(\theta)\right) \\
\xrightarrow{d}&A(s) - \frac{m}{\sqrt{k}}B.
\end{aligned}
$$

Here $A(s)$ is a Gaussian process defined in the proof of Theorem 5 in the appendix and $B := N(0, \Sigma_0)$ a centered Gaussian distribution with covariance matrix Σ_0, for details see Oh and Patton (2013). The limit result follows by using the independence of the moment process calculated from the data and the moment process corresponding to the simulated data. Note that the term $\frac{m}{\sqrt{k}}B$ cancels out in later considerations, e.g to determine the critical value c using the bootstrap procedure proposed in Section 4.2.3.

Theorem 5. Under the null hypothesis $H_0 : \theta_1 = \dots = \theta_{mT} = \theta_{mT+1} = \dots$ and under Assumption 5 and Assumptions 6-9 in the appendix, we obtain for $m \geq \varepsilon > 0$

$$
m\sqrt{T}\left(\hat{\theta}_{1+(s-m)T:sT,S} - \theta_0\right) \xrightarrow{d} A^*(s)
$$

as $T, S \to \infty$ in the space of càdlàg functions on the interval $[m, 1]$ and $\frac{S}{T} \to k \in (0, \infty]$. Here, $A^*(s) = \left(G'WG\right)^{-1}G'W(A(s) - \frac{m}{\sqrt{k}}B)$ and θ_0 is the (constant) value of θ_t under the null. Note that G is the derivative matrix of g_0, which is the probability limit of $g_{\cdot,S}$.

With Theorem 5 we obtain for $T, S \to \infty$

$$\begin{aligned}
&m\sqrt{T}(\hat{\theta}_{1+(s-m)T:sT,S} - \hat{\theta}_{1:mT,S}) \\
=&m\sqrt{T}(\hat{\theta}_{1+(s-m)T:sT,S} - \theta_0) - m\sqrt{T}(\hat{\theta}_{1:mT,S} - \theta_0) \\
\stackrel{d}{\Longrightarrow}&A^*(s) - A^*(m).
\end{aligned}$$

From this we can conclude the asymptotic behaviour of our detector under H_0, which we state in Corollary 3.

Corollary 3. Under the null hypothesis $H_0 : \theta_1 = \cdots = \theta_{mT} = \theta_{mT+1} = \ldots$ and if all mentioned Assumptions hold, we obtain for our detector

$$\begin{aligned}
D_{T,S}(s) =&m^2T(\hat{\theta}_{1+(s-m)T:sT,S} - \hat{\theta}_{1:mT,S})'(\hat{\theta}_{1+(s-m)T:sT,S} - \hat{\theta}_{1:mT,S}) \\
\stackrel{d}{\Longrightarrow}&(A^*(s) - A^*(m))'(A^*(s) - A^*(m)) =: Q(s)
\end{aligned}$$

as $T, S \to \infty$ and $\frac{S}{T} \to k \in (0, \infty]$.

Similarly, we may define a monitoring detector that is based directly on the moment conditions (dependence vector). This allows for monitoring of the corresponding dependence measures in a model-free way. Under the assumed factor copula model it can be used to monitor the stability of the model parameters. Furthermore, it has the additional advantage of being computationally much less demanding since no model parameters have to be estimated and it does not depend on any simulated quantities. The following corollary defines such a detector and describes its asymptotic behaviour.

Corollary 4. Under the null hypothesis $H_0 : \theta_1 = \cdots = \theta_{mT} = \theta_{mT+1} = \ldots$ and if all mentioned Assumptions hold, we obtain

$$\begin{aligned} M_T(s) :=& m^2 T(\hat{m}_{1+(s-m)T:sT} - \hat{m}_{1:mT})'(\hat{m}_{1+(s-m)T:sT} - \hat{m}_{1:mT}) \\ \xrightarrow{d}& (A(s) - A(m))'(A(s) - A(m)) =: R(s) \end{aligned}$$

as $T \to \infty$.

With the limit distribution of our detector $Q(s)$, we define the boundary value c in our monitoring procedure as the upper α-quantile of

$$\sup_{s\in[m,1]} Q(s) = \sup_{s\in[m,1]} (A^*(s) - A^*(m))'(A^*(s) - A^*(m)), \quad m \geq \varepsilon > 0. \tag{4.3}$$

Thus, $\lim_{T,S\to\infty} P(\tau_T < \infty | H_0) = \lim_{T,S\to\infty} P(\inf_k \{k \leq T : D_{T,S}(k) > c\} < \infty | H_0) = \alpha$.
In the same way the critical value of the moment monitoring procedure is determined as the upper α-quantile of $\sup_{s\in[m,1]} R(s)$.
For the estimation of the break point $mT + k^*$, once H_0 is rejected, we propose $mT + \hat{k}$, with

$$\hat{k} := \operatorname*{argmax}_{\lfloor\gamma(\tau_T - mT)\rfloor \leq i \leq \tau_T - mT} \frac{i^2}{\tau_T - mT} (\hat{\theta}_{1+mT:mT+i-1,S} - \hat{\theta}_{1+mT:\tau_T-1,S})'(\hat{\theta}_{1+mT:mT+i-1,S} - \hat{\theta}_{1+mT:\tau_T-1,S}), \tag{4.4}$$

where we only consider the information from $mT + 1$ to $\tau_m - 1$. Note that we need to trim a sufficient fraction $\lfloor\gamma(\tau_T - mT)\rfloor$ of the beginning, where $\gamma > 0$ to receive reasonable SMM estimates. In a similar way, the size of the rolling window mT should not be chosen too small. Note that the stopping time and the break point estimator for the moment monitoring procedure are defined analogously to the parameter monitoring procedure. As mentioned above, the moment based monitoring procedure is easy to implement and can be calculated fast, but in general it has lower power than the parametric procedure. Furthermore, as outlined in Manner et al. (2019), another disadvantage is that it does not allow testing the constancy of a subset of the parameters, but only can detect breaks in the whole copula. It may, however, be used to test for breaks in the dependence in selected regions of the support such as the lower tail. We leave this possibility for future research.

The limit distributions of $D_{T,S}$ and M_T are not known in closed form. To overcome this issue we have to simulate the critical values using an i.i.d. bootstrap procedure, which is described in the next section.

4.2.3. Bootstrap Distribution. First note that the limit result mainly consists of the limit distribution of the moment vectors, which can be computed relatively fast, compared to the detector that requires solving a minimization problem. This fact is used for the construction of the bootstrap. In order to approximate the limiting distribution under the null we use an i.i.d. bootstrap consisting of the following steps:

i) Sample with replacement from $\{\tilde{\eta}_i\}_{i=1}^T$ to obtain B bootstrap samples $\{\tilde{\eta}_i^{(b)}\}_{i=1}^T$, for $b = 1, \ldots, B$, where $\{\tilde{\eta}_i\}_{i=1}^T$ stacks the initial residual data $\{\hat{\eta}_i\}_{i=1}^{mT}$ and simulated data $\{\tilde{\eta}_i^*\}_{i=mT+1}^T$ from the assumed model, using the parameter estimate $\hat{\theta}_{1:mT,S}$ from the initial sample period.

ii) Use $\{\tilde{\eta}_i^{(b)}\}_{i=1+t-mT}^t$ to compute $\hat{m}_{1+t-mT:t}^{(b)}$ for $t = mT, \ldots, T$ and use $\{\tilde{\eta}_i^{(b)}\}_{i=1}^T$ to obtain $\hat{m}_{1:T}^{(b)}$, for $b = 1, \ldots, B$.

iii) Calculate the bootstrap version of the limiting distribution of our detector

$$K^{(b)} := \underset{t \in \{mT, \ldots, T\}}{max} \left(A^{*(b)} \left(\frac{t}{T} \right) - A^{*(b)} (m) \right)' \left(A^{*(b)} \left(\frac{t}{T} \right) - A^{*(b)} (m) \right),$$

with $A^{*(b)} \left(\frac{t}{T} \right) := (\hat{G}'\hat{W}_T\hat{G})^{-1}\hat{G}'\hat{W}_T A^{(b)}(\frac{t}{T})$ and $A^{(b)}(\frac{t}{T}) = m\sqrt{T} \left(\hat{m}_{1+t-mT:t}^{(b)} - \hat{m}_{1:T}^{(b)} \right)$, where $\hat{G}$ is the two sided numerical derivative estimator of G, evaluated at point $\hat{\theta}_{1:mT,S}$, computed with the historical sample $\{\hat{\eta}_i\}_{i=1}^{mT}$. We can compute the k-th column of $\hat{G}$ by

$$\hat{G}^k = \frac{g_{T,S}(\hat{\theta}_{1:mT,S} + e_k \varepsilon_{T,S}) - g_{T,S}(\hat{\theta}_{1:mT,S} - e_k \varepsilon_{T,S})}{2\varepsilon_{T,S}}, \quad k \in \{1, \ldots, p\},$$

where e_k is the k-th unit vector, whose dimension is $p \times 1$ and $\varepsilon_{T,S}$ has to be chosen in a way that it fulfils $\varepsilon_{T,S} \to 0$ and $\mathsf{min}\{\sqrt{T}, \sqrt{S}\}\varepsilon_{T,S} \to \infty$.

iv) Compute B versions of $K^{(b)}$ and determine the boundary value c such that

$$\frac{1}{B}\sum_{b=1}^{B} \mathbb{1}\{K^{(b)} > c\} \overset{!}{=} 0.05.$$

This bootstrap method is similar to the bootstrap used in Section 3, where iii) is adapted to the monitoring situation. The same intuitive argument holds for the validity of the bootstrap, which is only based on natural estimators of the respective terms. Furthermore, draws from the empirical distribution are close to draws from the population distribution and the structure of the limiting distribution allows for a direct computation without the need for centering. Under the alternative of a change point the bootstrap quantiles are bounded because the i.i.d. bootstrap destroys the temporal structure of the time series and thus mimicks a stationary distribution. Critical values of the moment based test M are obtained similarly by adapting step iii) of the algorithm. Our Monte Carlo simulations below confirm that the bootstrap indeed results in reasonably sized and powered tests.

4.2.4. Multiple Break Testing. In practice if one is interested in detecting multiple structural breaks in factor copula models in real time, we propose the following procedure that consists of steps applying the monitoring procedure proposed in this Section 4 and the retrospective change point test for factor copulas from Section 3. In particular, the retrospective test is used to test for the constant parameter Assumption 5 in the initial sample period and to detect the break point location once the monitoring procedure stops.

1) Compute the restrospective change point statistic $\underset{s\in[\varepsilon,m]}{\mathsf{sup}} P_{sT,S}$ from Section 3 for the initial mT observation. If a changepoint is detected go to step 2a). If no changepoint is detected go to step 2b).

2a) Estimate the breakpoint location and remove all pre-change observations. Restock the subsample to mT observations and return to step 1). If there are not enough observations left to restock the subsample to mT observations go to step 4).

2b) Take the sample as initial sample period. Apply the monitoring procedure to the residuals, i.e. compute $D_{T,S}(s)$ for $s \in (m, 1]$. Compute the bootstrap critical value c as described in Section 4.2.3. If a changepoint is detected go to step 3). If no changepoint is detected go to step 4).

3) Estimate the location of the changepoint. Then, remove the pre-change observations, use the first mT observations of the resulting dataset as the new initial sample and return to step 1). If there are not enough observations left to restock the subsample to mT observations go to step 4).

4) Terminate the procedure.

In the same way this procedure can be adapted for the moment monitoring procedure. Simulation results for single and multiple break testing, using the moment or the parameter monitoring procedure can be found in the next section. An obvious issue with this procedure is its multiple testing nature, in particular given that a pre-test has to be applied to the initial sample period to ensure that Assumption 5 holds. One should adapt the confidence levels accordingly and be aware of this when interpreting testing results.

4.3. Simulations

We now want to investigate the size and power and the estimation of the break point location of our monitoring procedure. We again consider the simple one factor copula model

$$[X_{1t}, \ldots, X_{Nt}]' =: X_t = \beta_t \boldsymbol{Z}_t + \boldsymbol{q}_t, \tag{4.5}$$

where $Z_t \sim \mathsf{Skew\ t}(\sigma^2, \nu^{-1}, \lambda)$ and $q_t \overset{i.i.d.}{\sim} \mathsf{t}(\nu^{-1})$ for $t = 1, \ldots, T$. We fix $\sigma^2 = 1$, $\nu^{-1} = 0.25$ and $\lambda = -0.5$, so that our model is parametrized by the factor loading parameter $\theta_t = \beta_t$. The sequential parameter estimates $\hat{\theta}_t = \hat{\theta}_{1-mT+t:t,S}$ for $t = mT, \ldots, T$ in the detector are computed using the SMM approach with $S = 25 \cdot mT$ simulations. For this we use five dependence measures, namely Spearman's rank correlation and the 0.05, 0.10, 0.90, 0.95 quantile dependence measures, averaged across all pairs. Critical values for the monitoring procedure are computed using $B = 1000$ bootstrap replications.

The nominal size of the tests is chosen to be 5% and we use 700 Monte Carlo replications for all settings.[3]

4.3.1. Size and Single Break Case. The size of the testing period is always fixed to be $T = 1500$.

Table 9: Size Monitoring Procedure

		$N = 10$	$N = 20$	$N = 30$
i)	$mT = 250$	0.051	0.052	0.051
	$mT = 400$	0.055	0.054	0.050
	$mT = 500$	0.057	0.047	0.054
ii)	$mT = 250$	0.062	0.064	0.061
	$mT = 400$	0.062	0.065	0.065
	$mT = 500$	0.061	0.054	0.055

Note: Empirical size for $\theta_0 = 1.0$, $T = 1500$ and 700 simulations, using i) the whole sample up to time point T and using ii) the initial data set and simulated data from $mT + 1$ up to T.

[3]The computational complexity of the simulations was extremely high due to the fact that for every monitoring procedure the parameter values need to be estimated a large number of times using the computationally heavy SMM estimator and because critical values have to be bootstrapped. This explains why we had to restrict ourselves to a limited number of situations for a fairly simple model. Furthermore, numerical instabilities were present in more complex models when repeatedly estimating the model parameters. Such problems can be dealt within empirical applications, but further restrict the potential model complexity in simulations. The computations were implemented in Matlab, parallelized and performed using CHEOPS, a scientific High Performance Computer at the Regional Computing Center of the University of Cologne (RRZK) funded by the DFG.

We begin with the case of testing against a single break. The rejection rate under the null are presented in Table 9 for $\theta_0 = 1$, for various combinations of the length of the initial sample mT and dimension N, where the critical values are calculated using one of the following two possibilities:

i) Calculate the critical value c using the whole, in general not known, data up to time point T. This mimics the situation that the test is used in a retrospective fashion, i.e. once all T observations are available.

ii) Calculate the critical value c using the initial data set together with simulated data from $mT + 1$ up to T, based on the estimated parameter vector $\hat{\theta}_{1:mT,S}$.

The test is slightly oversized for both settings. The empirical size is slightly higher for the second procedure ii), due to the fluctuation in the parameter estimation in the SMM procedure, but overall between 0.05 and 0.1.

To study the power of the procedure, we generate data with a break point at $\frac{T}{2}$, where the data is simulated with $\theta_t = 1$ for $t \in \{1, \ldots, \frac{T}{2}\}$, denoted as θ_0 and with $\theta_t = \{1.2, 1.4, 1.6, 1.8, 3.0\}$ for $t \in \{\frac{T}{2}+1, \ldots, T\}$, denoted as θ_1. With power we mean the probability that our monitoring procedure stops within the monitored testing period ($\tau_T < \infty$). The upper panel of Table 10 reveals that the power of the procedure increases with the size of the initial sample for the two possibilities i) and ii). The moment monitoring procedure based on M_T has similar size characteristics but lower power compared to the parameter-based procedure. This result is in line with the results for the retrospective test in Section 3.

The second and third panels of the table present the (average) relative stopping times and break point estimates using (4.4). The table reveals, that the averaged stopping time, given that a break has been detected, occurs with a significant delay after the true break point. It is closer to the true location $\frac{1}{2}$ for a smaller monitoring window, due to the greater impact of new data and, of course, for an increase of the step size between θ_0 and θ_1. If the step size is large enough ($\theta_1 = 3.0$) the monitoring procedure consistently stops shortly after the true break point.

The averaged estimated break point locations based on equation (4.4) are closer to the true break point. It always detects the break before the stopping time. For small shifts in θ it estimates the break too late, whereas for large shifts in θ breaks are estimated a little too early. It seems that a larger initial sample always results in slightly later stopping times, that can be explained due to the greater impact on the detector by new observed data in small rolling window sample sizes. However the usage of smaller window sizes imply lower power of the procedure. Note that the moment monitoring tends to result in later stopping times and break point estimates in all cases.

4.3.2. Multiple Breaks. For the analysis of multiple breaks we allow for breaks at $\frac{T}{3}$ and $\frac{2T}{3}$ with sample size $T = 1500$, and dimensions $N = 10$ and $N = 20$. The parameter varies from $\theta_0 = 1.0$ for $t \in \{1, \ldots, \frac{T}{3}\}$ to $\theta_1 = 1.5$ for $t \in \{\frac{T}{3} + 1, \ldots, \frac{2T}{3}\}$ and $\theta_2 = 0.8$ for $t \in \{\frac{2T}{3} + 1, \ldots, T\}$. The results using the procedure proposed in Section 4.2.4 can be found in Table 11. The tables report the averaged stopping times, averaged break point estimates and rejection rates for the first, second, and the joint first and second break events.

The rejection rates increase with the size of the initial sample period mT. Power increases in the dimension N, although this effect is only moderate for both tests. As before, the tests based on $D_{T,S}$ has larger power than the one based on M_T. We also note that the second break point is detected more frequently than the first one, that is of course due to the higher magnitude of the second break compared to the first break. Furthermore, if the monitoring procedure detects the first break point it is very likely that the second break point is detected as well, which can be seen by the almost identical rejection rates of rej_1 and rej_{all}. Again, the average stopping time is much later than the true break, but the estimated break point $\hat{k}$

Table 10: Power Monitoring Procedure

			$\theta_0 = 1.0$	$\theta_1 = 1.2$	$\theta_1 = 1.4$	$\theta_1 = 1.6$	$\theta_1 = 1.8$	$\theta_1 = 3.0$
rej	i)	$mT = 250$	0.059	0.375	0.787	0.973	1.000	1.000
		$mT = 400$	0.066	0.435	0.877	0.993	1.000	1.000
		$mT = 500$	0.066	0.465	0.910	1.000	1.000	1.000
	ii)	$mT = 250$	0.079	0.409	0.787	0.967	1.000	1.000
		$mT = 400$	0.069	0.468	0.860	0.990	1.000	1.000
		$mT = 500$	0.076	0.485	0.894	1.000	1.000	1.000
	m_T	$mT = 250$	0.053	0.193	0.482	0.780	0.944	1.000
		$mT = 400$	0.076	0.223	0.671	0.944	0.993	1.000
		$mT = 500$	0.063	0.306	0.738	0.960	1.000	1.000
$\frac{\tau_T}{T}$	i)	$mT = 250$		0.715	0.667	0.625	0.579	0.513
		$mT = 400$		0.751	0.689	0.629	0.588	0.523
		$mT = 500$		0.767	0.677	0.639	0.596	0.525
	ii)	$mT = 250$		0.698	0.660	0.619	0.581	0.518
		$mT = 400$		0.733	0.675	0.626	0.587	0.525
		$mT = 500$		0.759	0.699	0.638	0.595	0.527
	m_T	$mT = 250$		0.718	0.695	0.662	0.627	0.525
		$mT = 400$		0.738	0.725	0.672	0.625	0.528
		$mT = 500$		0.765	0.741	0.679	0.626	0.530
$\frac{\hat{k}}{T}$	i)	$mT = 250$		0.516	0.487	0.483	0.473	0.457
		$mT = 400$		0.544	0.508	0.493	0.487	0.479
		$mT = 500$		0.562	0.522	0.497	0.492	0.487
	ii)	$mT = 250$		0.511	0.484	0.479	0.471	0.464
		$mT = 400$		0.538	0.502	0.491	0.485	0.483
		$mT = 500$		0.562	0.518	0.497	0.491	0.489
	m_T	$mT = 250$		0.517	0.495	0.487	0.486	0.485
		$mT = 400$		0.541	0.518	0.500	0.495	0.492
		$mT = 500$		0.561	0.534	0.507	0.499	0.496

Note: Rejection frequency (rej), average stopping time $\frac{\tau_T}{T}$ and average breakpoint estimate $\frac{\hat{k}}{T}$ for $\theta_0 = 1$, $T = 1500$ $N = 10$ and 301 simulations for the parameter monitoring procedure, where critical values c computed with the two possibilities i) and ii) and for the moment monitoring procedure. Data was generated with a break in $\frac{T}{2}$ and post-break parameter θ_1.

is able to detect the breaks reasonably well. Thus, we can conclude that the procedure works fairly well for the case of multiple breaks and that both the power of detecting changes and estimating the break locations can be achieved in a fairly reliable manner.

Table 11: Multibreak Power Monitoring Procedure

		$\frac{\tau_T^1}{T}$	$\frac{\hat{k}^1}{T}$	rej_1	$\frac{\tau_T^2}{T}$	$\frac{\hat{k}^2}{T}$	rej_2	$(\frac{\tau_T^1}{T}\ \frac{\tau_T^2}{T})$	$(\frac{\hat{k}^1}{T}\ \frac{\hat{k}^2}{T})$	rej_{all}
Parameter based										
$N = 10$	$mT = 250$	0.458	0.336	0.800	0.805	0.667	0.851	(0.458 0.801)	(0.337 0.665)	0.777
	$mT = 400$	0.479	0.365	0.861	0.836	0.711	0.954	(0.479 0.836)	(0.365 0.716)	0.854
$N = 20$	$mT = 250$	0.457	0.338	0.810	0.799	0.661	0.864	(0.456 0.794)	(0.339 0.659)	0.787
	$mT = 400$	0.475	0.363	0.874	0.827	0.708	0.970	(0.475 0.827)	(0.362 0.713)	0.867
Moment based										
$N = 10$	$mT = 250$	0.493	0.343	0.570	0.806	0.663	0.618	(0.495 0.796)	(0.344 0.661)	0.558
	$mT = 400$	0.517	0.369	0.691	0.843	0.721	0.834	(0.517 0.840)	(0.369 0.729)	0.691
$N = 20$	$mT = 250$	0.491	0.338	0.591	0.808	0.660	0.671	(0.493 0.797)	(0.338 0.658)	0.588
	$mT = 400$	0.513	0.366	0.730	0.839	0.714	0.884	(0.513 0.834)	(0.366 0.721)	0.730

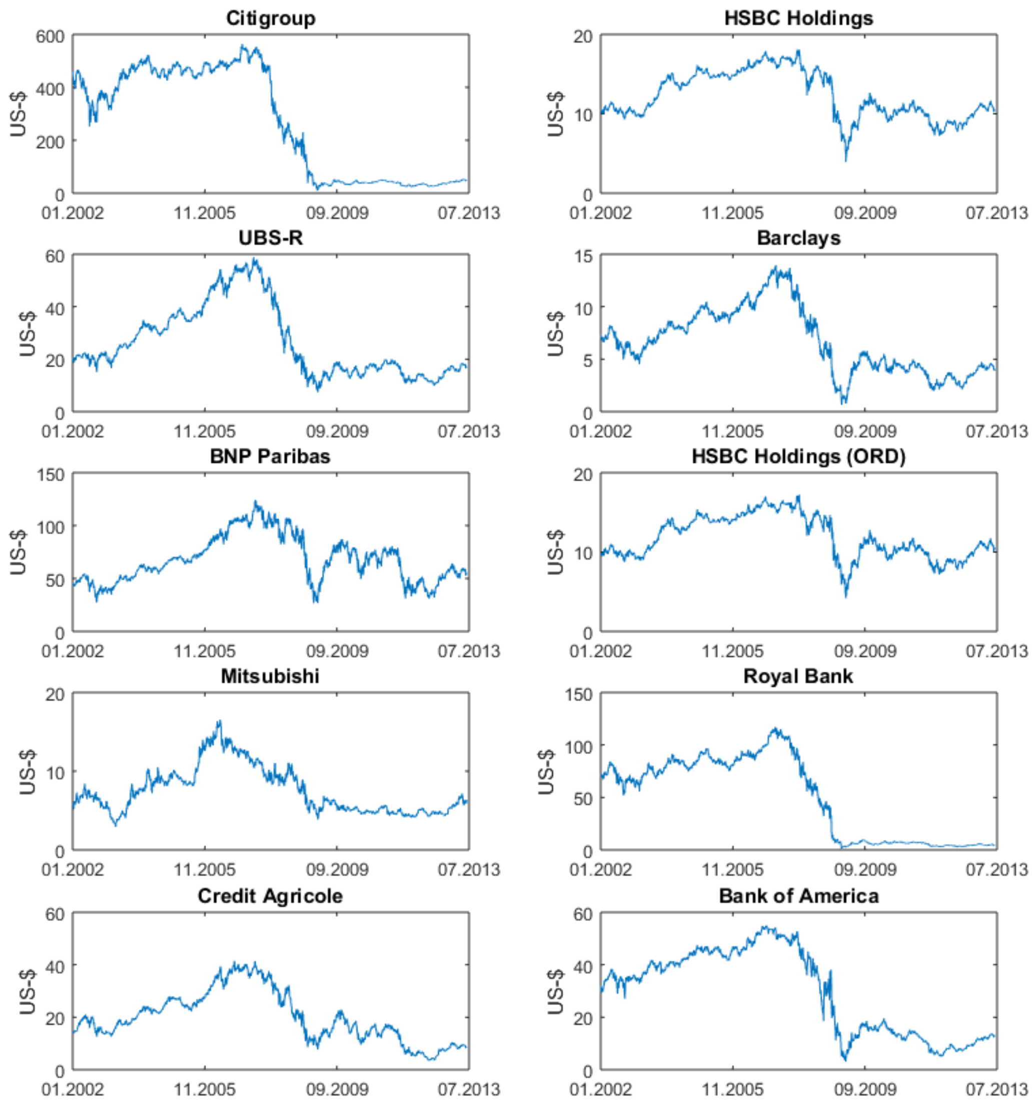
Note: Average detected break point location $\frac{\hat{k}^i}{T}$, stopping time $\frac{\tau_T^i}{T}$ and rejection frequency using 301 simulations for the parameter monitoring procedure. Data was generated with breaks at $\frac{T}{3}$ and $\frac{2T}{3}$, with $T = 1500$, $N = 10, 20$, $\theta_0 = 1.0$, $\theta_1 = 1.5$, $\theta_2 = 0.8$. Results are based on the parameter based detector $D_{T,S}$ (top panel) and the moment based detector M_T (bottom panel).

4.4. Empirical Application

In this section we apply our test to a real data set. We use daily log-returns of stock prices over a time span ranging from 29.01.2002 to 01.07.2013 of ten large firms, namely Citigroup, HSBC Holdings ($), UBS-R, Barclays, BNP Paribas, HSBC Holdings (ORD), Mitsubishi, Royal Bank, Credit Agricole and Bank of America. This implies a monitored period of size $T = 2980$ and $N = 10$. Figure 4.9 is a plot of the stock values in US-$ of the ten assets over the whole monitored period.

We use the same factor copula model as in (4.5) and we fix $\nu = 2.855$ and $\lambda = -0.0057$ for the monitoring procedure, i.e. we only monitor the factor loading parameter. These fixed values correspond to the parameter estimates from the initial sample period of size $mT = 400$.

Figure 4.9: Asset values of the Portfolio



Note: Asset values S_t^i in US-$ in our considered portfolio for data between 29.01.2002 and 01.07.2013, $T = 2980$ and $N = 10$.

For the conditional mean and variance we specify the following AR(1)-GARCH(1,1).

$$r_{i,t} = \alpha + \beta r_{i,t-1} + \sigma_{i,t}\eta_{it},$$

$$\sigma_{it}^2 = \gamma_0 + \gamma_1\sigma_{i,t-1}^2 + \gamma_2\eta_{i,t-1}^2,$$

for $t = 2, \ldots, 2980$, and $i = 1, \ldots, 10$. Note that for the monitoring procedure the parameters of the conditional mean and variance models are always re-estimated on the corresponding rolling window sample of size mT.

4.4.1. Monitoring Procedure. A rolling window parameter analysis of the whole data set with window size 400 can be seen in Figure 4.10, indicating parameter changes between 2006 and 2009.

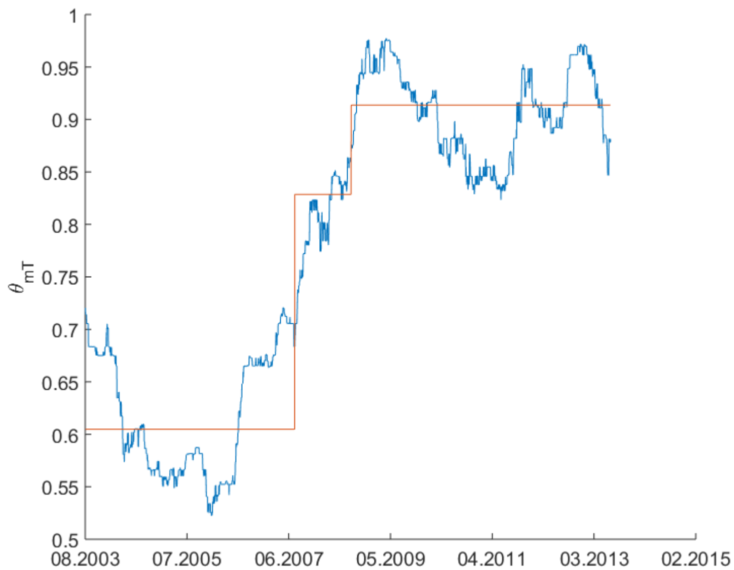
The results of the monitoring procedure of the whole considered period can be seen in Table 12. We choose the historical period $mT = 400$ from 29.01.2002 to 11.08.2003, where we first estimate the marginal AR(1)-GARCH(1,1) model to obtain the residuals. We use the retrospective constant parameter test from Section 3 to test the hypothesis of no parameter change within the historical data set and the null hypothesis cannot be rejected. Note that for the retrospective parameter test a burn-in period of 20 % of the behold data is used. We then apply our constructed monitoring procedure. The monitoring procedure stops at the 21.11.2008 and the estimated break point location is found at the 19.07.2007, where we used the retrospective parameter break point estimate with data from the end of the historical data set 12.08.2003 to the stopping time 21.11.2008.

Figure 4.11 is a plot of $D_{T,S}$ for every time point between $mT + 1$ (12.08.2003) and the stopping point, where $D_{T,S}$ exceeds the 0.95-quantile value of (4.3) equal to 4.4512.

We then cut of all the data in front of the estimated break point location (19.07.2007) and test for the null hypothesis of no parameter change in the period from 20.07.2007 to 29.01.2009 of size $mT = 400$, using again the retrospective parameter test and the null is rejected. The estimated break point is found at the 11.08.2008.

For the next subsample we try the period from 12.08.2008 to 23.02.2010 and get a retrospective test statistic value $S_{T,S}$ of 1.4442 and a quantile value of 2.5156, hence the null hypothesis

Figure 4.10: Rolling Window Parameter Estimate



Note: Rolling window estimate of θ_{mT} for $mT = 400$ and $N = 10$ between 11.08.2003 and 01.07.2013, with parameter values estimated from break to break. Each parameter value is associated to the end time point of the rolling window.

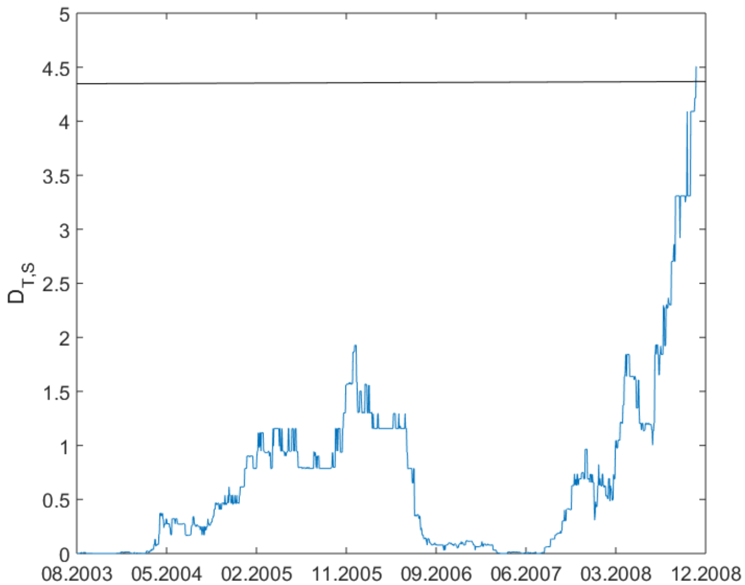
Table 12: Testing Results empirical Application

Monitored/Testing Period	τ_T	$\hat{k}$	T
29.01.2002-11.08.2003			400
12.08.2003-01.07.2013	21.11.2008	19.07.2007	2580
20.07.2007-29.01.2009		11.08.2008	400
12.08.2008-23.02.2010			400
24.02.2010-01.07.2013			874

Note: Stopping time τ_T, estimated break point location $\hat{k}$ and associated sample size T for monitored or tested periods using the monitoring procedure or the retrospective parameter test.

cannot be rejected and we choose this period as our new historical period and restart our monitoring procedure from 24.02.2010 to 01.07.2013. The detector $D_{T,S}$ does not cross the boundary value $c = 12.0020$ and the procedure stops at the end of the monitored period,

Figure 4.11: Value Detector $D_{T,S}(s)$



Note: $D_{T,S}(s)$ for $T = 2980$, $mT = 400$ and $N = 10$. Stopping date at 21.11.2008 and $c = 4.4512$.

without rejecting the null.

4.4.2. Value-at-Risk Predictions. Given the growing need for managing financial risk, risk prediction plays an increasing role in banking and finance. The value-at-Risk (VaR), is the most prominent measure of financial market risk. Despite it having been criticized as being theoretically not efficient and numerically problematic (see Dowd and Blake, 2006), it is still the most widely used risk measure in practice. The number of methods for such calculations continues to increase. The theoretical and computational complexity of VaR models for calculating capital requirements is also increasing. Some examples include the use of extreme value theory (see McNeil and Frey, 2000), quantile regression methods (see Manganelli and Engle, 2004), and Markov switching techniques (see Gray, 1996 and Klaassen, 2002).

First, we want to define the Value at Risk (VaR). We define the log-return of a single asset i to a time point t as $r_t^i = \ln(S_t^i) - \ln(S_{t-1}^i)$, where S_t^i is the stock value of asset i to a specific time point t. The change in the portfolio value over the time interval $[t-1, t]$ is then

$$\Delta V_t = \sum_{i=1}^{N} w_i r_t^i,$$

where w_i are portfolio weights. The (negative) α-quantile of the distribution of $\Delta V := \{\Delta V_t\}_{t=1}^{T}$ is the day t Value-at-Risk at level α.

Here we want to show that our monitoring procedure can help improve the day-ahead predictions of the VaR based on a factor copula model.

The VaR predictions based on the monitoring procedure for the factor copula model are computed as follows. In general, based on $\mathcal{F}_t$, the information available at time t, we want to predict the VaR for period $t+1$. The prediction of the VaR is always based on the following four steps.

1. Simulate M draws from the copula model $\tilde{\boldsymbol{u}}_{t+1} \sim C(\cdot, \hat{\theta}_t)$, where $\tilde{\boldsymbol{u}}_{t+1} = [\tilde{\boldsymbol{u}}_{1,t+1}, \ldots, \tilde{\boldsymbol{u}}_{N,t+1}]$ is an $M \times N$ matrix of simulated observation and $\hat{\theta}_t$ is an appropriate parameter estimate based on information up to time t.

2. Use the inverse marginal distribution function of the standardized residuals η to transform every component of $\tilde{\boldsymbol{u}}_{t+1}$ to $\tilde{\boldsymbol{\eta}}_{t+1} = \left[F_1^{-1}(\tilde{\boldsymbol{u}}_{1,t+1}), \ldots, F_N^{-1}(\tilde{\boldsymbol{u}}_{N,t+1})\right]$, where $F_i^{-1}(\cdot)$ is estimated by the inverse integrated kernel density estimator of the residuals $\hat{\boldsymbol{\eta}}$ with a sufficiently large number of evaluation points.

3. Compute the simulated returns $\tilde{r}_{t+1} := [\tilde{r}_{t+1}^1, \ldots, \tilde{r}_{t+1}^N]' = \boldsymbol{\mu}(\hat{\phi}_t) + \boldsymbol{\sigma}(\hat{\phi}_t)\tilde{\boldsymbol{\eta}}_{t+1}$, where $\hat{\phi}_t$ are the estimated parameters from models for the conditional mean and variance using information up to time t.

4. Form the portfolio of interest from the simulated returns and compute the appropriate quantile from the distribution of the portfolio to obtain the VaR prediction for time $t+1$.

This procedure for predicting the VaR is generic. The monitoring procedure for the copula parameter θ_t is used to determine the appropriate information set on which the parameter estimate in Step 1. is based. The basic idea is to use as much information as possible as long as no changepoint is detected. In case a changepoint is found only the most recent observations should be used to estimate θ_t. Recall that mT observations for which the dependence is assumed to be constant are available at the beginning of the sample. Further, denote $\hat{\theta}_{s:t}$ the estimator of the copula parameter based on the observations from time s to t. At each point in time t, compute $D_{T,S}(t)$.

i. Before a changepoint is detected, i.e. as long as $D_{T,S}(t) < c$ the draws from the copula in Step 1 above are based on $\hat{\theta}_{1:t}$

ii. Assume the monitoring procedure stops at time $t = \hat{\tau}$, i.e. when $D_{T,S}(t) > c$. Compute the breakpoint estimate $\hat{k}$ using (4.4). Use the estimate $\hat{\theta}_{\hat{k}:t}$ in Step 1 above. If $t - \hat{k} < 400$, i.e. if less than 400 observations are available use $\hat{\theta}_{t-400:t}$. In other words, after a breakpoint is identified use either all observations after the breakpoint estimate or the most recent 400 observations to estimate the copula parameter.[4]

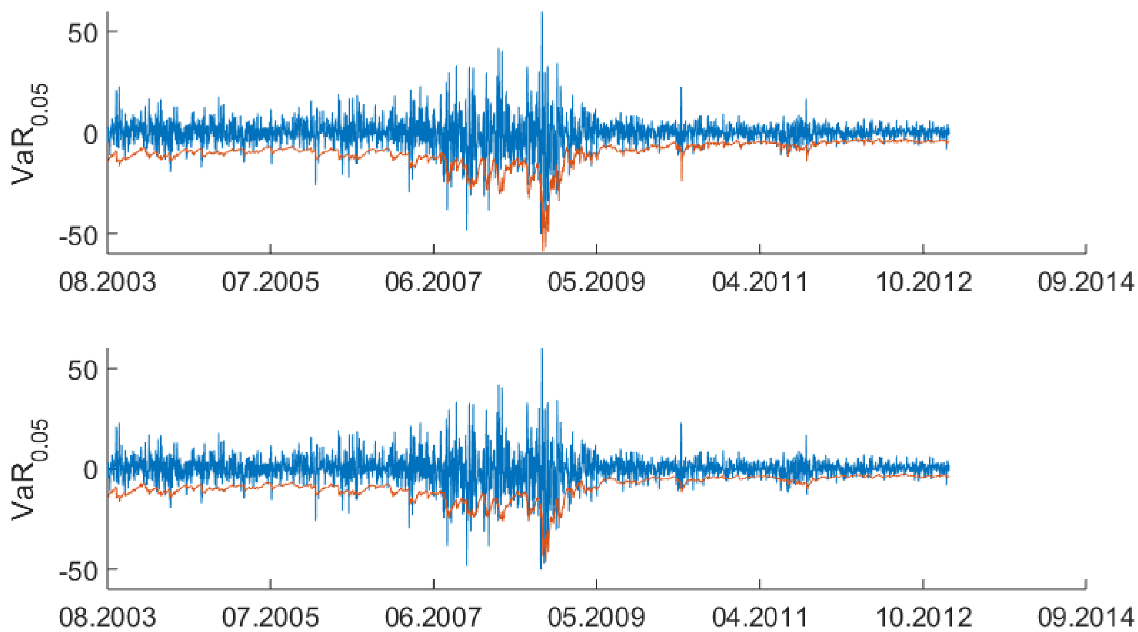
iii. If $t - \hat{k} \leq mT$ proceed as in Step ii. Otherwise use the window $[\hat{k}, \hat{k} + mT]$ as the new initial sample and apply the monitoring procedure. As long as no further breakpoint is detected the parameter estimate $\hat{\theta}_{\hat{k}:t}$ is used. When the monitoring procedure stops again return to Step ii.

The results for the online VaR evaluation based on $M = 1500$ simulations for each period and for $\alpha = 0.05$ can be seen in Figure 4.12. As an alternative, we consider the same model without the monitoring procedure. In that case the copula parameter is estimated using the

[4] The minimum number of observations required for model estimation depends on the complexity of the chosen model. However, for the type of model we found that one needs at least 400 observations to obtain reliable and numerically stable parameter estimates.

full sample available at time t using an expanding window. The model for the margins is an AR(1)-GARCH(1,1) in both cases.

Figure 4.12: Portfolio Returns and Value at Risk



Note: Portfolio returns ΔV_t and the $\alpha = 0.05$ predicted Value-at-Risk based on the monitoring procedure, allowing for structural breaks (upper panel) and without (lower panel) for the period between 29.01.2002 and 01.07.2013.

Visually, the online procedure tracks the 5 % VaR well. The empirical VaR exceedance rate is, in fact, 6.05% (156 exceedances in 2580 days) and therefore reasonably close to 5 %. In the model without structural breaks, where the parameters are estimated from the beginning of the sample on, the exceedance rate is higher with 7.71% (199 exceedances). With a binomial test (compare Berens, Wied, Weiß, and Ziggel, 2014), we test the null hypothesis of unconditional coverage, i.e.,

$$\mathbb{E}\left(\frac{1}{T}\sum_{t=1}^{T} I_t(0.05)\right) = \alpha = 0.05,$$

where α is the VaR coverage probability and

$$I_t(0.05) = \begin{cases} 0, \text{ if } \Delta V_t \geq -VaR_{0.05} \\ \\ 1, \text{ if } \Delta V_t < -VaR_{0.05}. \end{cases}$$

One expects 129 exceedances under H_0 and at the 1% significance level the critical value of the test is 158 exceedances. This implies that the null of unconditional coverage is rejected in the model without structural breaks, but not in the model with structural breaks.

5. On the Applicability of a Nonparametric Test for Constant Copula-Based Dependence Measures: Dating Breakpoints and Analyzing Different Dependence Measure Sets

In this section we want to investigate the non-parametric dependence measure test with test statistic (3.4) for different dependence measure settings proposed in Section 3. The section is structured as follows: Section 5.1 presents the null hypothesis, test statistic, a bootstrap procedure and a heuristic to test for common breaks. Results from the Monte Carlo simulations can be found in Section 5.2 and Section 5.3 presents the empirical application. The work belongs to the working paper "On the Applicability of a nonparametric Test for constant Copula-based Dependence Measures: Dating Breakpoints and analyzing different Dependence Measure Sets", Stark (2018).

5.1. Null Hypothesis and Test Statistic

The null hypothesis which is considered is a constant dependence measure vector against the alternative of a single breakpoint at an unknown point in time,

$$H_0 : m_1 = m_2 = \cdots = m_T \qquad H_1 : m_1 = \cdots = m_t \neq m_{t+1} = \cdots = m_T \text{ for some } t = \{1, \ldots, T-1\}.$$

The CUSUM-type test statistic is based on the maximum difference between the recursive estimates and the full sample estimate of the dependence measure vector. Formally, it is defined as

$$\begin{aligned} M := \sup_{s\in[\varepsilon,1]} M_{sT} &:= \sup_{s\in[\varepsilon,1]} \; s^2 T(\hat{m}_{sT} - \hat{m}_T)'(\hat{m}_{sT} - \hat{m}_T) \\ &\simeq \max_{\lfloor \varepsilon T\rfloor \le t\le T} \left(\frac{t}{T}\right)^2 T(\hat{m}_t - \hat{m}_T)'(\hat{m}_t - \hat{m}_T), \end{aligned} \tag{5.1}$$

where $\hat{m}_{sT}$ is the recursive dependence measure vector defined above that uses the information up to time point $t = \lfloor sT \rfloor$ and $\hat{m}_T$ the full sample dependence measure vector analogue such as $\varepsilon > 0$ a trimming parameter. In Section 3 it is noted that ε has to be chosen strictly greater than zero and in finite sample cases it should be chosen in a way that we have enough data information to receive reasonable dependence measure vector estimates. For the finite sample case in Section 3 we propose $\varepsilon = 0.2$ in the context of the copula parameter test for a better comparison of the two proposed tests here. Note that it is sufficient to use $\varepsilon = 0.1$ for the non parametric dependence measure test in the later considered simulation study and empirical application. The test rejects the null hypothesis of a constant dependence measure vector, if the sequential estimated dependence vectors fluctuate to much over time, which is measured by $\sup_{s\in[\varepsilon,1]} s^2T(\hat{m}_{sT} - \hat{m}_T)'(\hat{m}_{sT} - \hat{m}_T)$. The term s^2T puts less weight on deviations at the beginning of the observed period, due to the fact that $\hat{m}_{sT}$ fluctuates more for smaller sequential sample sizes sT. As mentioned, the statistic is of non-parametric nature and one can consider an appropriate subset of dependence measure settings and test for, e.g. breaks in the lower/upper quantile dependencies, Spearman's rank correlation or a combination of both quantile dependencies and rank correlation. An analysis of different selected dependence measure settings, where residual data is simulated from different fat-tailed and skewed copula distribution models can be found in the simulation section.

The analytical results for the asymptotic distribution of the test statistic M can be found in Corollary 2 in Section 3. Here the asymptotic results can be obtained by using two main assumptions (Assumption 1 and 2 from Section 3) where these ensure that the estimated rank correlation and quantile dependencies converge to their respective population counterparts, using residual data from pre-estimated marginal time series models. Then under the null hypothesis $H_0 : m_1 = m_2 = \cdots = m_T$ and if these assumptions hold, it follows that

$$M = \sup_{s\in[\varepsilon,1]} s^2T(\hat{m}_{sT} - \hat{m}_T)'(\hat{m}_{sT} - \hat{m}_T) \xrightarrow{d} \sup_{s\in[\varepsilon,1]} (A(s) - sA(1))'(A(s) - sA(1)), \tag{5.2}$$

as $T \to \infty$, where $A(s)$ is a Gaussian process defined in the proof of Lemma 11 from Section 3. The proof of the limit result in equation (5.2) follows by a steady transformation of the result from Section 3 in Lemma 11. We reject the null if

$$M > q_{1-\alpha}, \tag{5.3}$$

where $q_{1-\alpha}$ is the $(1-\alpha)$-quantile of $\sup\limits_{s\in[\varepsilon,1]} (A(s) - sA(1))'(A(s) - sA(1))$. If we reject the null hypothesis we speak of a structural break. The estimation of the change point location, once we detected a structural break, is embedded in calculating the test statistic and is given by $\hat{k} := \lfloor \hat{s}T \rfloor$, where $\hat{s}$ is the maximum point of the quadratic left side of (5.2), i.e.

$$\hat{s} = \operatorname*{argmax}_{s\in[\varepsilon,1]} s^2 T(\hat{m}_{sT} - \hat{m}_T)'(\hat{m}_{sT} - \hat{m}_T). \tag{5.4}$$

The distribution term $A(s)$ in the asymptotic distribution of the test statistic is in general not known in closed form and depends on the underlying sample. For this reason critical values cannot be computed or simulated directly. To overcome this issue a bootstrap procedure similar to the one in Section 3 is used. The bootstrap distribution of the test statistic M is obtained by calculating B versions of the process $\frac{t}{T}\sqrt{T}\left(\hat{m}_t^{(b)} - \hat{m}_T^{(b)}\right)$, which can be calculated fast and directly from the data, where $t = sT$. The following steps are used:

i) Sample with replacement from the standardized residuals $\{\hat{\eta}_i\}_{i=1}^T$ to obtain B bootstrap samples $\{\hat{\eta}_i^{(b)}\}_{i=1}^T$, for $b = 1, \dots, B$.

ii) Use $\{\hat{\eta}_i^{(b)}\}_{i=1}^t$ to compute $\hat{m}_t^{(b)}$ for $b = 1, \dots, B$ and $t = \varepsilon T, \dots, T$ and $\{\hat{\eta}_i\}_{i=1}^T$ to obtain $\hat{m}_T$.

iii) Calculate the bootstrap analogue of the limiting distribution in equation (5.2)

$$K^{(b)} := \max_{t\in\{\varepsilon T,\dots,T\}} \left(A^{(b)}\left(\frac{t}{T}\right) - \frac{t}{T}A^{(b)}(1)\right)' \left(A^{(b)}\left(\frac{t}{T}\right) - \frac{t}{T}A^{(b)}(1)\right),$$

with $A^{(b)}\left(\frac{t}{T}\right) := \frac{t}{T}\sqrt{T}\left(\hat{m}_t^{(b)} - \hat{m}_T\right)$.

iv) Compute B versions of $K^{(b)}$ and determine the critical value K such that

$$\frac{1}{B}\sum_{b=1}^{B}\mathbb{1}\{K^{(b)} > K\} = \alpha,$$

for a confidence level $\alpha \in (0,1)$, e.g. $\alpha = 0.05$.

In Section 3 an intuition for the validity of the bootstrap is given. A formal proof of the bootstrap validity is left for future research.

For a better comparison of the break point location estimates in empirical applications, determined by different dependence measure settings within the testing procedure, we propose a heuristic procedure. With this we are able to make a statement whether two estimated break point locations $\hat{s}_a$ and $\hat{s}_b$, with $a \neq b$, belong to the same class of break points, choosing different subsets of the used dependence measures. Here, the subscripts a and b denote the choice of the different vector of dependence measures $\hat{m}_a^{ij}$ and $\hat{m}_b^{ij}$. Note that we consider the break point location estimator defined in equation (5.4), which is a scalar in the uniform interval $(0,1]$. We determine (pivot) confidence intervals $\hat{K}_a := [\hat{K}_a^-, \hat{K}_a^+] := [2\hat{s}_a - \hat{c}^a_{1-\frac{\alpha}{2}}, 2\hat{s}_a - \hat{c}^a_{\frac{\alpha}{2}}]$ and $\hat{K}_b := [2\hat{s}_b - \hat{c}^b_{1-\frac{\alpha}{2}}, 2\hat{s}_b - \hat{c}^b_{\frac{\alpha}{2}}]$, where $\hat{c}^a_{(\cdot)}$ and $\hat{c}^b_{(\cdot)}$ are estimated quantiles of the bootstrap distribution of $\hat{s}_a$ and $\hat{s}_b$ which can be determined by using the following (percentile) bootstrap procedure.

We consider the residual sample $\{\boldsymbol{\eta}_t\}_{t=1}^T$ in which we detected two break point locations $\hat{s}_a$ and $\hat{s}_b$, using the dependence measure setting $\hat{m}_a^{ij}$ and $\hat{m}_b^{ij}$.

i) Split the sample in $\{\hat{\boldsymbol{\eta}}_t\}_{t=1}^{\hat{s}_a T}$ and $\{\hat{\boldsymbol{\eta}}_t\}_{t=\hat{s}_a T+1}^{T}$ for setting $\hat{m}_a^{ij}$ and $\{\hat{\boldsymbol{\eta}}_t\}_{t=1}^{\hat{s}_b T}$ and $\{\hat{\boldsymbol{\eta}}_t\}_{t=\hat{s}_b T+1}^{T}$ for setting $\hat{m}_b^{ij}$.

ii) Sample separately with replacement from $\{\hat{\boldsymbol{\eta}}_t\}_{t=1}^{\hat{s}_a T}$ and $\{\hat{\boldsymbol{\eta}}_t\}_{t=\hat{s}_a T+1}^{T}$ such as $\{\hat{\boldsymbol{\eta}}_t\}_{t=1}^{\hat{s}_b T}$ and $\{\hat{\boldsymbol{\eta}}_t\}_{t=\hat{s}_b T+1}^{T}$ to obtain B bootstrap samples $\{\hat{\boldsymbol{\eta}}_{t,a}^{(p)}\}_{t=1}^{T}$ and $\{\hat{\boldsymbol{\eta}}_{t,b}^{(p)}\}_{t=1}^{T}$, for $p = 1, \ldots, B$.

iii) Estimate the break point location $\hat{s}_a^{(p)}$ for each bootstrap sample $\{\hat{\boldsymbol{\eta}}_{t,a}^{(p)}\}_{t=1}^{T}$ and $\hat{s}_b^{(p)}$ for each bootstrap sample $\{\hat{\boldsymbol{\eta}}_{t,b}^{(p)}\}_{t=1}^{T}$for $p = 1, \ldots, B$, using (5.4).

iv) Compute the quantiles of the bootstrap distribution $\hat{c}^a_{\frac{\alpha}{2}}, \hat{c}^b_{\frac{\alpha}{2}}$ and $\hat{c}^a_{1-\frac{\alpha}{2}}, \hat{c}^b_{1-\frac{\alpha}{2}}$ of $\{\hat{s}_a^{(p)}\}_{p=1}^{B}$ and $\{\hat{s}_b^{(p)}\}_{p=1}^{B}$, where $\alpha \in (0, 1)$.

We say that two estimated break point locations $\hat{s}_a$ and $\hat{s}_b$ can be considered equal if both lie in the intersection of the two determined confidence intervals, i.e.

$$\hat{s}_a, \hat{s}_b \in \hat{K}_a \cap \hat{K}_b. \tag{5.5}$$

Note, this procedure is only plausible if we consider the same testing period for both dependence settings $\hat{m}_a^{ij}$ and $\hat{m}_b^{ij}$. Therefore, in the empirical application the procedure can only be applied for a break comparison in the full sample testing and cannot be used in the rolling window testing procedure, due to the fact that similar break point locations may belong to different tested periods. Further, notice that the estimation error between the estimated break point $\hat{s}$ and the true break point s_0 is approximately the same as the difference between the estimated break location $\hat{s}$ and the bootstrap break estimate $\hat{s}^{(p)}$, i.e.

$$\begin{aligned}
1 - \alpha &\approx P(\hat{c}_{\frac{\alpha}{2}} \leq \hat{s}^{(p)} \leq \hat{c}_{1-\frac{\alpha}{2}}) \\
&= P(\hat{s} - \hat{c}_{1-\frac{\alpha}{2}} \leq \hat{s} - \hat{s}^{(p)} \leq \hat{s} - \hat{c}_{\frac{\alpha}{2}}) \\
&\approx P(\hat{s} - \hat{c}_{1-\frac{\alpha}{2}} \leq s_0 - \hat{s} \leq \hat{s} - \hat{c}_{\frac{\alpha}{2}}) \\
&= P(2\hat{s} - \hat{c}_{1-\frac{\alpha}{2}} \leq s_0 \leq 2\hat{s} - \hat{c}_{\frac{\alpha}{2}}).
\end{aligned}$$

The construction of the confidence intervals is for example similar to the one in Hušková and Kirch (2008).

5.2. Simulations

In this section we want to analyze size and power properties of the proposed test for different dependence measure settings in Monte Carlo simulations. We consider the following dependence measure settings

$$\hat{m}_1^{ij} = (\hat{\rho}^{ij}\ \hat{\lambda}_{0.05}^{ij}\ \hat{\lambda}_{0.1}^{ij}\ \hat{\lambda}_{0.9}^{ij}\ \hat{\lambda}_{0.95}^{ij})'$$

$$\hat{m}_2^{ij} = (\hat{\lambda}_{0.05}^{ij}\ \hat{\lambda}_{0.1}^{ij}\ \hat{\lambda}_{0.9}^{ij}\ \hat{\lambda}_{0.95}^{ij})'$$

$$\hat{m}_3^{ij} = (\hat{\lambda}_{0.9}^{ij}\ \hat{\lambda}_{0.95}^{ij})'$$

$$\hat{m}_4^{ij} = (\hat{\lambda}_{0.05}^{ij}\ \hat{\lambda}_{0.1}^{ij}\ )'$$

$$\hat{m}_5^{ij} = \hat{\rho}^{ij},$$

where we average all pairwise dependence measures in an equidependence way, i.e. $\hat{m} = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{m}^{ij}$, due to $\hat{m}^{ij} = \hat{m}^{ji}$ for $i \neq j$ and $\hat{m}^{ii} = 1$.

For our investigations, we consider a level of 5% and repeat the test 301 times for every scenario. Due to the fact that we are mainly interested in comparing the different dependence settings, we fix the data size and cross sectional dimension to $T = 1000$ and $N = 10$. An analysis for different combinations of T and N in the case of $\hat{m}_1^{ij} = (\hat{\rho}^{ij}\ \hat{\lambda}_{0.05}^{ij}\ \hat{\lambda}_{0.1}^{ij}\ \hat{\lambda}_{0.9}^{ij}\ \hat{\lambda}_{0.95}^{ij})'$ can be found in Section 3. For the simulations we simulate data from different skewed and fat tailed distributions. We first consider our data $\boldsymbol{\eta}_t$ to be jointly distributed with a simple one factor copula model following Oh and Patton (2013) and Oh and Patton (2017), where the copula is implied by the following factor structure

$$\boldsymbol{\eta}_t = [\eta_{1t}, \ldots, \eta_{Nt}]' = \boldsymbol{\theta}_t Z + \boldsymbol{q}, \tag{5.6}$$

with $\boldsymbol{\theta}_t = (\theta_t, \ldots, \theta_t)'$ a parameter vector of size N, $Z \sim \mathsf{Skew\ t}\,(\nu^{-1}, \lambda)$[5] and $\boldsymbol{q} = [q_{1t}, \ldots, q_{Nt}]'$ with $q_{it} \overset{i.i.d.}{\sim} \mathsf{t}\,(\nu^{-1})$ for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. We fix $\nu^{-1} = 0.25$ and vary $\lambda = \{-0.5, 0, 0.5\}$, such that our model is parametrized by the single factor loading θ_t.

For the power analysis we generate data with a break point at $\frac{T}{2}$ for different sample sizes, where the data is simulated with $\theta_t = 1$ for the first $\frac{T}{2}$ data points, denoted by θ_0, whereas

[5] As in Oh and Patton (2017) this refers to the skewed t-distribution by Hansen (1994).

Table 13: Size and Power Factor Copula

$T = 1000, N = 10$	$\theta_0 = 1$	$\theta_1 = 1.1$	$\theta_1 = 1.2$	$\theta_1 = 1.3$	$\theta_1 = 1.4$	$\theta_1 = 1.5$
			$\lambda = -0.5$			
$\hat{m}_1^{ij}$	0.0465	0.1628	0.3887	0.6013	0.8007	0.9269
$\hat{m}_2^{ij}$	0.0498	0.1462	0.3189	0.5249	0.6944	0.8704
$\hat{m}_3^{ij}$	0.0365	0.0897	0.2259	0.4784	0.7043	0.8571
$\hat{m}_4^{ij}$	0.0498	0.1329	0.2724	0.4286	0.5781	0.7176
$\hat{m}_5^{ij}$	0.0532	0.2558	0.6645	0.9435	0.9934	1.0000
			$\lambda = 0$			
$\hat{m}_1^{ij}$	0.0532	0.1993	0.4485	0.7010	0.9003	0.9767
$\hat{m}_2^{ij}$	0.0532	0.1927	0.3787	0.6213	0.8538	0.9358
$\hat{m}_3^{ij}$	0.0432	0.1229	0.2625	0.4385	0.6478	0.8206
$\hat{m}_4^{ij}$	0.0565	0.1495	0.2857	0.4485	0.6146	0.8641
$\hat{m}_5^{ij}$	0.0598	0.2791	0.7176	0.9668	0.9967	1.0000
			$\lambda = 0.5$			
$\hat{m}_1^{ij}$	0.0565	0.1661	0.3322	0.5781	0.8538	0.9635
$\hat{m}_2^{ij}$	0.0764	0.1495	0.2890	0.4917	0.7342	0.9203
$\hat{m}_3^{ij}$	0.0731	0.1395	0.2658	0.3854	0.5781	0.7741
$\hat{m}_4^{ij}$	0.0332	0.1096	0.2558	0.4651	0.6611	0.8538
$\hat{m}_5^{ij}$	0.0498	0.2658	0.7043	0.9502	1.0000	1.0000

Note: Table 13 reports the rejection rate for $\theta_0 = 1.0$ and $\theta_1 = \{1.1, 1.2, 1.3, 1.4, 1.5\}$ for the different dependence measure combinations $\hat{m}_a^{ij}$ for $a = 1, \ldots, 5$ and $\lambda = \{-0.5, 0, 0.5\}$ in the DGP (5.6).

after the break we increase the parameter to $\theta_t = \{1.1, 1.2, 1.3, 1.4, 1.5\}$, denoted by θ_1. The results for the factor copula model (5.6) using $\lambda = \{-0.5, 0, 0.5\}$ can be seen in Table 13. Table 13 reveals that $\hat{m}_5^{ij} = \hat{\rho}^{ij}$ gains overall the highest power, directly followed by the setting $\hat{m}_1^{ij}$ considering Spearman's rho, lower and upper quantile dependence measures. The cases where only upper $\hat{m}_3^{ij}$ or lower quantiles $\hat{m}_4^{ij}$ are considered suffer from poor power properties compared to the other dependence settings. Considering both upper and lower quantile dependencies in $\hat{m}_2^{ij}$ has better power properties than the separated cases. For a clearer comparison of the different dependence vector settings, especially for the cases where only lower or upper quantile dependence measures are used, we consider two more data generating processes. First we consider residual data generated by a Clayton copula, where we vary the parameter from $\theta_0 = 2.5$ to $\theta_1 = \{3.0, 3.5, 4.0, 5.0, 5.5\}$. Second we consider residual data generated from a Gumbel copula, where we also vary $\theta_0 = 2.0$ to $\theta_1 = \{2.2, 2.4, 2.6, 2.8, 3.0\}$. In both cases again θ_1 denotes the parameter value after the break at $\frac{T}{2}$ and θ_0 the parameter value before the break. Note, with the choice of θ_0 and θ_1 the implied upper quantile dependence for the Clayton copula and implied lower quantile dependence for the Gumbel Copula is in the same magnitude for better comparability. Due to the heavy tailed behavior of the Clayton (strong lower quantile dependence) and Gumbel Copula (strong upper quantile dependence), the dependence structure in the lower (Clayton) and upper (Gumbel) cases just changes slightly by varying the parameter values after the break. This yields poor power properties of the test by only using lower quantile dependencies $\hat{m}_4^{ij}$, where the Clayton Copula is used as the data generating process (DGP) (cf. Table 14). On the other hand, only allowing for upper quantile dependencies $\hat{m}_3^{ij}$, where the Gumbel Copula is used as the DGP, the test suffers also from poor power properties (cf. Table 15). Similar results, only using upper or lower quantile dependencies can be seen for the factor copula model in the case of $\lambda = \{-0.5, 0.5\}$, see Table 13. A combination of lower and upper

Table 14: Size and Power Clayton Copula

$T = 1000, N = 10$	$\theta_0 = 2.5$	$\theta_1 = 3.0$	$\theta_1 = 3.5$	$\theta_1 = 4.0$	$\theta_1 = 4.5$	$\theta_1 = 5.0$
$\hat{m}_1^{ij}$	0.0532	0.1960	0.4518	0.7342	0.9336	0.9701
$\hat{m}_2^{ij}$	0.0565	0.1761	0.3488	0.6179	0.8372	0.8970
$\hat{m}_3^{ij}$	0.0631	0.1827	0.4219	0.7010	0.8738	0.9402
$\hat{m}_4^{ij}$	0.0565	0.1694	0.2924	0.3821	0.5382	0.6047
$\hat{m}_5^{ij}$	0.0332	0.3854	0.9468	1.0000	1.0000	1.0000

Note: Table 14 reports the rejection rate for $\theta_0 = 2.5$ and $\theta_1 = \{3.0, 3.5, 4.0, 4.5, 5.0\}$ for different dependence measure combinations, where the DGP is a Clayton Copula with parameter α (more mass on lower tail).

quantile dependence measures ($\hat{m}_2^{ij}$) again yields therefore better power properties. Yet, the dependence vector settings $\hat{m}_1^{ij}$ and $\hat{m}_5^{ij}$, where Spearman's rank correlation is included, imply again better power properties of the test. This can be explained due to the fact that quantile dependencies suffer from a poor data amount in the tails, i.e. for a data size of $T = 1000$ we consider only 100 or 50 data points by choosing the 0.1/0.9 or 0.05/0.95 quantiles. Consequently, a larger data size is required to gain better power properties, compared to the usage of Spearman's rho, where the final rank correlation coefficient is computed out of the whole data information and is a global dependence measure. Before taking a look at the simulation results one would expect that the usage of more dependence measures within the dependence measure vector increases the power of the test (5.3). However, it is not the case for the setting $\hat{m}_1^{ij}$. It seems that the dependence vector compounded of a mixture of quantile dependencies and rank correlation inherits the bad power properties of the quantile dependencies and the better performance, compared to the settings $\hat{m}_2^{ij}, \hat{m}_3^{ij}$ and $\hat{m}_4^{ij}$, is mainly driven by the usage of the rank correlation coefficient.

However, the usage of various dependence settings may give us different break point estimates, which can be seen in the empirical application, cf. Section 5.3. In times of clearer structural break, i.e. periods which can be assigned to events that effected the financial market, one can be more sure whether the detected break is plausible if several dependence settings result in the same break event. Another possibility is to divide the data in sensible subsets and separately test for structural breaks using different combinations of the considered dependence measures. For example, one could split the data in different industry sectors. To test for equality of two found break points use the confidence interval procedure explained in the previous Section.

Table 15: Size and Power Gumbel Copula

$T = 1000, N = 10$	$\theta_0 = 2.0$	$\theta_1 = 2.2$	$\theta_1 = 2.4$	$\theta_1 = 2.6$	$\theta_1 = 2.8$	$\theta_1 = 3.0$
$\hat{m}_1^{ij}$	0.0399	0.1628	0.4352	0.8671	0.9668	1.0000
$\hat{m}_2^{ij}$	0.0365	0.1329	0.3654	0.7874	0.9003	0.9834
$\hat{m}_3^{ij}$	0.0399	0.1229	0.2492	0.4618	0.5648	0.6678
$\hat{m}_4^{ij}$	0.0565	0.3522	0.6445	0.8571	0.9402	0.9734
$\hat{m}_5^{ij}$	0.0532	0.5282	0.9435	1.0000	1.0000	1.0000

Note: Table 15 reports the rejection rate for $\theta_0 = 2.0$ and $\theta_1 = \{2.2, 2.4, 2.6, 2.8, 3.0\}$ for different dependence measure combinations, where the DGP is a Gumbel Copula with parameter α (more mass on upper tail).

In the following we include a small simulation study for the confidence interval procedure, proposed in Section 5.1, using the dependence measure settings $\hat{m}_1^{ij}$ and $\hat{m}_3^{ij}$, where the break estimates of these settings are later compared in the full sample testing in the empirical application. We simulate a residual data set, using the DGP in (5.6) with $\lambda = -0.5$, where we constructed a break at $\frac{T}{2}$, i.e. $s_0 = 0.5$. We fix the cross sectional dimension to $N = 10$ and vary the sample size $T = \{500, 1000, 1500\}$ and the break size by $\theta_1 = \{1.5, 2.0, 2.5\}$. For all simulations we use $B = 500$ bootstrap replications. We present the coverage probability $P(0.5 \in \hat{K}_1)$, the coverage probability $P(0.5 \in \hat{K}_3)$ and the probability that the constructed break at $s_0 = 0.5$ lies in the intersection of $\hat{K}_1 \cap \hat{K}_3$, i.e. $P(0.5 \in \hat{K}_1 \cap \hat{K}_3)$, using again 301 Monte-Carlo simulations and a confidence level of 5 percent. Table 16 reveals that the coverage probability of $\hat{K}_1$ and $\hat{K}_3$ tends to $1-\alpha = 0.95$ for increasing sample size and break size. The probability that the actual break at $s_0 = 0.5$ lies in the interval $\hat{K}_1 \cap \hat{K}_3$ tends to $(1-\alpha)^2$. Note, that the practitioner can control the size level α^* of the common break test by considering $(1-\alpha)^2 = 1-\alpha^*$.

Table 16: Coverage Probability same Breakpoints

	$\left(P(0.5 \in \hat{K}_1)\ P(0.5 \in \hat{K}_3)\ P(0.5 \in \hat{K}_1 \cap \hat{K}_3)\right)$		
$B = 500, N = 10$	$T = 500$	$T = 1000$	$T = 1500$
$\theta_1 = 1.5$	(0.80 0.78 0.64)	(0.93 0.87 0.81)	(0.94 0.90 0.84)
$\theta_1 = 2.0$	(0.91 0.88 0.80)	(0.95 0.93 0.89)	(0.95 0.92 0.89)
$\theta_1 = 2.5$	(0.93 0.94 0.88)	(0.93 0.94 0.88)	(0.95 0.95 0.91)

Note: Table 16 reports the coverage probability of 301 simulated confidence intervals $\hat{K}_1$ and $\hat{K}_3$ for a break constructed at 0.5 such as the coverage probability of $0.5 \in \hat{K}_1 \cap \hat{K}_3$ where the factor copula model (5.6) is used as the DGP and $\alpha = 0.05$.

Next, we simulate two residual data sets, using again the DGP in (5.6), where we constructed break points, lying nearby at $s_0^{(1)} = \frac{6}{14} = 0.429$ in the first set and a break at $s_0^{(2)} = \frac{7}{14} = 0.5$ in the second set. Note, this simulation setting mimics the situation where we split our sample in subsets. For all simulations we consider the same scenarios as before and we present the coverage probability $P(0.429 \in \hat{K}_1)$, the coverage probability $P(0.5 \in \hat{K}_3)$ and the probability that both constructed breaks 0.429 and 0.5 lie in the intersection of $\hat{K}_1 \cap \hat{K}_3$, i.e. $P(0.429, 0.5 \in \hat{K}_1 \cap \hat{K}_3)$. The results can be seen in Table 17.

Here, the coverage probability of $\hat{K}_1$ for a break at 0.429 and $\hat{K}_3$ for a break at 0.5 tends to $1-\alpha = 0.95$ for increasing sample size and break size, where the results of $P(0.5 \in \hat{K}_3)$ are obviously the same as in Table 16. On the other hand, the probability that the two break points 0.429 and 0.5 lie in the interval $\hat{K}_1 \cap \hat{K}_3$ tends to zero for increasing sample size and break steps, where the convergence is much faster in the break step θ_1. Note that for example

Table 17: Coverage Probability different Breakpoints

	$\Big(P(0.429 \in \hat{K}_1)$	$P(0.5 \in \hat{K}_3)$	$P(0.429, 0.5 \in \hat{K}_1 \cap \hat{K}_3)\Big)$
$B = 500, N = 10$	$T = 500$	$T = 1000$	$T = 1500$
$\theta_1 = 1.5$	(0.81 0.78 0.46)	(0.91 0.87 0.52)	(0.92 0.90 0.31)
$\theta_1 = 2.0$	(0.93 0.88 0.21)	(0.95 0.93 0.01)	(0.96 0.92 0.00)
$\theta_1 = 2.5$	(0.95 0.94 0.01)	(0.95 0.94 0.00)	(0.95 0.95 0.00)

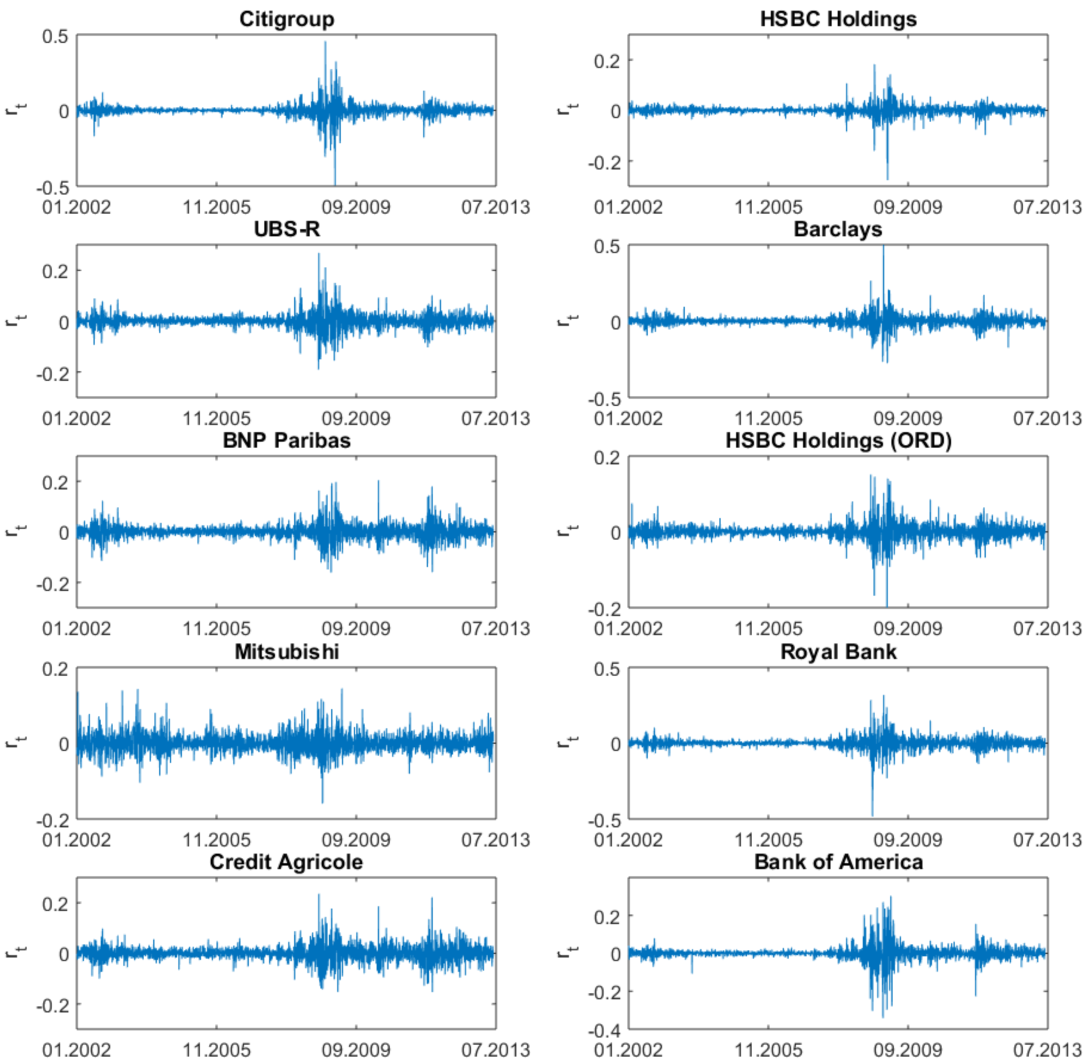
Note: Table 17 reports the coverage probability of 301 simulated confidence intervals $\hat{K}_1$ and $\hat{K}_3$ for a break constructed at 0.5 and 0.429 such as the coverage probability of $0.429, 0.5 \in \hat{K}_1 \cap \hat{K}_3$ where the factor copula model (5.6) is used as the DGP and $\alpha = 0.05$.

a break step of $\theta_1 = 1.5$ implies a rank correlation change before and after the break of 0.17, where a break change of $\theta_1 = 2.0$ implies a rank correlation change of 0.25. Thus we conclude, that our procedure to test for equality of two estimated break point locations results in a reasonable sized and powered testing procedure, if the break steps and the sample size are high enough.

5.3. Empirical Application

To analyze the applicability of the proposed test we apply the test to the same financial data set as in Section 4 and determine structural breaks in the vector of dependence measures. We are interested in the similarity and diversity of the estimated break point locations using different dependence measure settings analyzed in the simulation Section 5.2. For a better comparison of similar break dates we use the confidence interval procedure presented in Section 5.1 and analyzed in Section 5.2. Again we have a sample size of $T = 2980$ and cross sectional dimension $N = 10$. By taking a look at the time evolution of the log-returns of the portfolio in Figure 5.13, one can immediately see the strong fluctuations between 2002-2003, 2007-2008 and 2011-2012 in nearly all assets, indicating a joint behavior during these periods within the portfolio. To get a better understanding of the joint dependence behavior we calculated the pairwise averaged Spearman's rank correlation coefficient in a rolling window of size 150, the result can be seen in Figure 5.14.

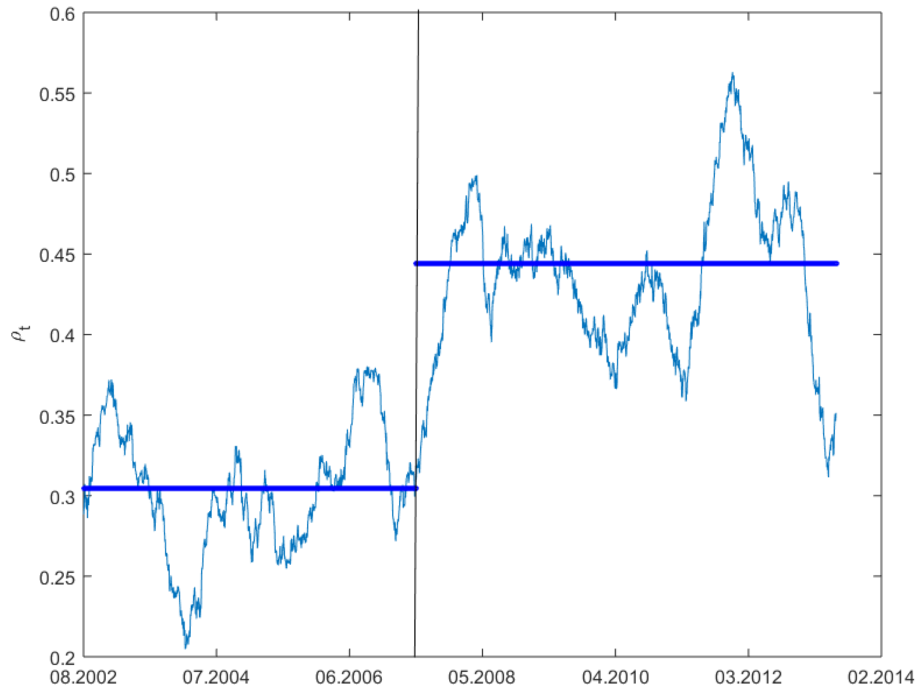
Figure 5.13: Log-returns of the Portfolio



Note: Daily log-returns of our considered portfolio between 29.01.2002 to 01.07.2013.

The strongest joint fluctuations and increase of the correlation coefficient appear between the time span of the heights of the last financial crisis between the beginning of 2007 and the end of 2008. The other strong fluctuations and increase of the correlation can be explained

Figure 5.14: Rolling Window Spearman's Rank Correlation Coefficient Approach 1)



Note: Pairwise averaged Spearman's rank correlation coefficient in a rolling window of size 150 with averaged estimated break point location (vertical line) at 17.07.2007 using approach 1) as well as the estimated rank correlation coefficient from break to break (thick blue line).

by the downturn in stock prices in stock exchanges across the United States, Canada, Asia and Europe in October 2002 and the euro crisis peek in 2011.

Due to the fact that we consider residual data, we first have to estimate a time series model for each log-return series $i = 1, \ldots, N$. Therefore we use an AR(1)-GARCH(1,1) model to model the conditional mean and variance

$$r_{i,t} = \alpha_i + \beta_i r_{i,t-1} + \sigma_{i,t}\eta_{it},$$

$$\sigma_{it}^2 = \gamma_{i0} + \gamma_{i1}\sigma_{i,t-1}^2 + \gamma_{i2}\sigma_{i,t-1}^2\eta_{i,t-1}^2,$$

for $t = 1, \ldots, 2980$.

Using the estimated parameters we are able to compute the standardized residuals. Note that the marginal distributions of the residuals are estimated using the empirical CDF. We apply the test to the residual data considering the five dependence vector settings from the simulation section. We consider two approaches, where approach 2) can be applied to test for multiple breaks:

1) We apply the test to the pre-determined residual data considering information between 1 and T.

2) We apply the test in a rolling window setting where we consider periods of size 400. If a breakpoint is detected in the period $[t_1, (t_1 - 1) + 400]$ we estimate the break point location $\hat{k}$ and $[t_1 + 1, t_1 + 400]$ is the next considered period, where $t_1 = \hat{k}$. If no break point is detected consider the next time step $t_1 + 1$. We start the procedure by setting $t_1 = 1$ and terminate the procedure if $t_1 + 400 > T$. The marginal models are re-estimated for every considered period.

The break detection results for approach 1) can be found in Table 18 and Figure 5.14 where the results for approach 2) can be found in Table 19 and Figure 5.15. First we take a look

Table 18: List of found Breakpoints Approach 1)

	$\hat{m}_1^{ij}$	$\hat{m}_2^{ij}$	$\hat{m}_3^{ij}$	$\hat{m}_4^{ij}$	$\hat{m}_5^{ij}$	avg
$\hat{s}_a T$	09.07.2007	09.07.2007	08.08.2007	09.07.2007	09.07.2007	17.07.2007
$\hat{K}_a^-$	21.12.2006	31.08.2006	07.03.2006	30.03.2006	22.02.2007	
$\hat{K}_a^+$	25.10.2007	22.11.2007	20.06.2008	28.02.2008	18.10.2007	

Note: Table 18 reports the found break points and confidence intervals $[\hat{K}_a^-,\hat{K}_a^+]$ for the five dependence vector setting within the test using approach 1) such as the averaged break point location.

at the results for approach 1), here nearly all dependence settings found the same break at

09.07.2007 (transformed to the uniform interval $\hat{s}_1 = 0.476$) except the setting choosing the upper quantiles, here the break is found close by at 08.08.2007 (transformed to the uniform interval $\hat{s}_3 = 0.484$). It seems obvious that the different estimated break point also belongs to the other four break events, which can be explained by the last financial crisis. Nevertheless we check the suspicion using the confidence interval procedure from Section 5.1 with the settings $\hat{m}_1^{ij}$ and $\hat{m}_3^{ij}$. We find that the two estimated break point locations $\hat{s}_1 = 0.476$ and $\hat{s}_3 = 0.484$ lie in the intersection of both confidence intervals and thus conclude that the estimated break point locations belong to the same break event. As we will see by using approach 2), there might be more break point locations, but approach 1) gives us the most significant break in our data set. By taking a look at Figure 5.14, where we plotted the pairwise averaged Spearman's rank correlation coefficient in a rolling window of size 150, we see a strong increase of the rank correlation coefficient after the break (indicated by the black line) from 0.31 up to 0.44 (indicated by the thick blue line), where the overall maximum change of the rolling window estimates is even higher between 0.27 and 0.49.

Table 19: List of found Breakpoints Approach 2)

year	$\hat{m}_1^{ij}$	$\hat{m}_2^{ij}$	$\hat{m}_3^{ij}$	$\hat{m}_4^{ij}$	$\hat{m}_5^{ij}$	avg
2002/2003	30.12.2002	08.01.2003	20.12.2002		23.12.2002	27.12.2002
2004	19.02.2004	26.02.2004		05.03.2004	04.03.2004	01.03.2004
2005						
2006				11.05.2006		11.05.2006
2007	24.07.2007	11.07.2007	17.07.2007	16.02.2007	09.07.2007	15.06.2007
2008	16.07.2008	08.08.2008	16.07.2008	17.07.2008		23.07.2008
2009						
2010	21.04.2010	15.06.2010	29.04.2010	10.06.2010		19.05.2010
2011	28.06.2011	28.06.2011	21.09.2011	14.06.2011	20.05.2011	05.07.2011
2012				14.08.2012		15.08.2012

Note: Table 19 reports the found break points for the five dependence vector setting within the test using approach 2) as well as the averaged break point locations.

Next, we take a look at the results for approach 2), where we tested the whole period sequentially in a rolling window of size 400, where we found some more breaks. Note that an obvious issue with this procedure is its multiple testing nature. In particular one should adapt the confidence levels accordingly and be aware of this when interpreting testing results. By taking a look at the breakpoints (cf. Table 19) most detections can be explained by well known financial market crashes from the last twenty years. First as already mentioned the downturn in stock prices in stock exchanges across the United States, Canada, Asia and Europe in October 2002, the start of the Iraqi war 2003/2004, start of the last financial crisis in 2007 as well as the bankruptcy of Lehman Brother's in 2008 and last the Euro crisis starting at the end of 2009 with its height in 2011. The break point estimates of the dependence settings $\hat{m}_1^{ij}, \hat{m}_2^{ij}$ and $\hat{m}_3^{ij}$ seem to be really closely related and belong to the above mentioned events, where the break event in 2004 seems not to be significant for the upper quantile setting. In contrast to the simulation study, where the setting $\hat{m}_5^{ij}$ gains the highest power over all other settings, the test with the setting $\hat{m}_5^{ij}$ detects only four significant break dates at 2002, 2004, 2007 and 2011, which are overall the most significant. A slightly different break result is given by the lower quantile setting $\hat{m}_4^{ij}$, where the tested periods in 2002/2003 seem not to be significant. On the other hand, breaks are detected in the mid of 2006 and 2012.

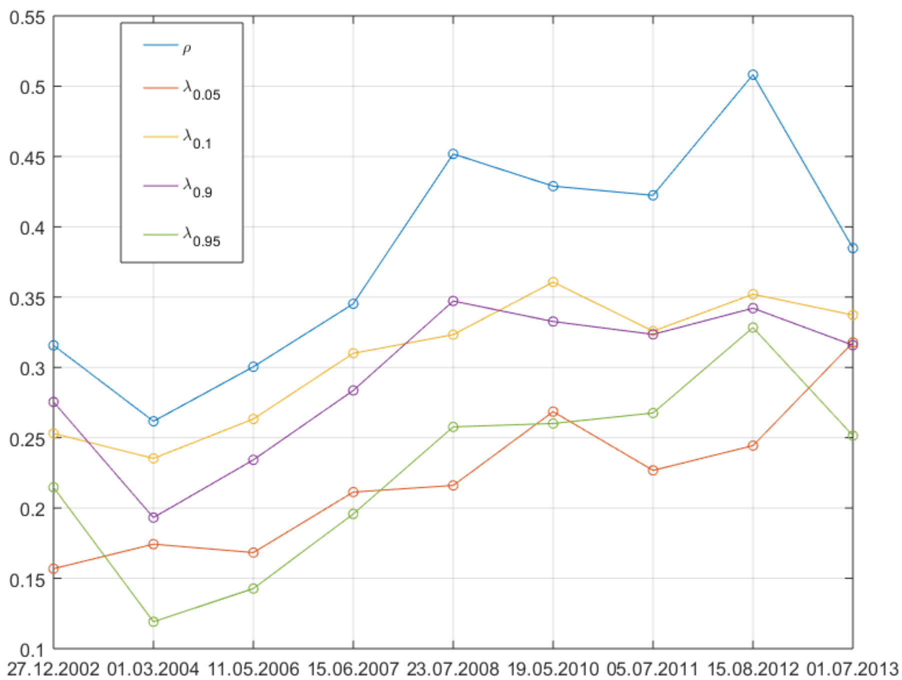
One advantage of the use of different dependence settings is that we are able to conclude that a detected break point is in a way more relevant in specific regions of the distribution, if it is detected by more than one setting. The break events in 2002, 2004, 2007, 2008, 2010 and 2011 are such break points and the breaks in 2002, 2004, 2008 and 2010 are detected by four settings where the break events in 2007 and 2011 are detected by even five settings, corresponding to the most significant breaks and can be explained by the pre-mentioned well known financial market crashes. The detected break point in 2007 is also in line with the detected break event using approach 1), corresponding to the highest dependence change in the considered testing period. On the other hand we also get a different break picture as in the case of using the setting $\hat{m}_4^{ij}$, motivating the usage of flexible dependence measure settings for a clearer and wider interpretation of the results. Due to the fact that the found break dates correspond over all to the same break events and being really close to each other in the most cases, we average the break dates over all settings (cf. last column in Table 19).

Figure 5.15: Rolling Window Spearman's Rank Correlation Coefficient Approach 2)

Note: Pairwise averaged Spearman's rank correlation coefficient in a rolling window of size 150 and averaged break point locations (averaged over five similar detected break points (black line), averaged over four similar detected break points (red line) and single detected break points (green line)) using approach 2) such as the estimated rank correlation coefficient from break to break (thick blue line).

We again plot the pairwise averaged Spearman's rank correlation coefficient in a rolling window and mark the averaged break point estimates together with the rank correlation estimates from break to break. Here five similar averaged break dates marked with the black line, four averaged detected break dates marked with the red line and single detected break points corresponding to the green line. The horizontal blue lines correspond to the values of the estimated rank correlation coefficient from break to break, see Figure 5.15. Noticeable are the high jumps of the rank correlation in the periods of the most significant breaks in 2007 and 2011, where the correlation jumps from 0.35 to 0.45 (2007) and 0.42 to 0.51 (2011) considering break to break estimates. Note that the overall increase, considering the rolling window rank correlation estimates, is even higher. Furthermore the upper quantile dependence measures increase strongly here (cf. Figure 5.16).

Figure 5.16: Dependence Measure Estimates from Break to Break



Note: Dependence measure estimates (Spearman's rank correlation and $(0.05, 0.1, 0.9, 0.95)$-quantile dependencies) from break to break.

In general, nearly all detected break events correspond in an increase of the considered dependence measures (cf. Figure 5.16), except the first detected break in 2002/2003, where the peak in this period of nearly all dependencies is reached. Also in the time span after the last financial crisis, we see a decrease of the rank correlation and the 0.9-quantile dependence. However, we see an overall increase of the dependencies within the portfolio. From a portfolio manager point of view the increase of the dependencies within our considered portfolio is disadvantageous and one is rather interested to decrease the dependencies by changing the portfolio to lower the risk of losses, which is known as the diversification effect.

Lastly, we are interested in the break behavior if we divide our considered portfolio in subsets of assets. For this we split our portfolio in two groups of five assets, where we collected the assets with the highest unconditional variance in the first group. For the group with the highest variance we consider Cititgroup, Barclays, Royal Bank, Credit Agricole and Bank of America (Group 1), where the other group consists of HSBC Holdings, UBS-R, BNP Paribas, HSBC Holdings (ORD) (Group 2). We apply the test separately for each group of assets, considering the five dependence vector settings, where we only consider approach 1). Table 20 shows the estimated break point locations and confidence intervals for Group 1 (upper panel) and Group 2 (lower panel).

Table 20: List of found Breakpoints Approach 1) splitted Sample

Group 1	$\hat{m}_1^{ij}$	$\hat{m}_2^{ij}$	$\hat{m}_3^{ij}$	$\hat{m}_4^{ij}$	$\hat{m}_5^{ij}$
$\hat{s}_a T$	09.07.2007	16.02.2007	26.11.2007	16.02.2007	09.07.2007
$\hat{K}_a^-$	24.11.2006	22.02.2006	09.12.2005	23.12.2005	14.03.2007
$\hat{K}_a^+$	27.09.2007	25.05.2007	05.12.2008	06.07.2007	26.09.2007
Group 2	$\hat{m}_1^{ij}$	$\hat{m}_2^{ij}$	$\hat{m}_3^{ij}$	$\hat{m}_4^{ij}$	$\hat{m}_5^{ij}$
$\hat{s}_a T$	09.07.2007	09.07.2007			31.05.2006
$\hat{K}_a^-$	23.11.2005	18.02.2005			13.07.2004
$\hat{K}_a^+$	19.06.2008	10.10.2008			14.03.2007

Note: Table 20 reports the found break points and confidence intervals [$\hat{K}_a^-$,$\hat{K}_a^+$] for Group 1 (upper panel) and Group 2 (lower panel).

Using the first group data all dependence measure settings find a significant break at the $\alpha = 0.05$ level, whereas in the second group only the breaks of the settings $\hat{m}_1^{ij}, \hat{m}_2^{ij}$ and $\hat{m}_5^{ij}$ are significant. The results are more varied than in Table 18. Most of the found breaks correspond to the summer of 2007 and the same break date is detected at 09.07.2007 using the settings $\hat{m}_1^{ij}$, $\hat{m}_5^{ij}$ for Group 1 and $\hat{m}_1^{ij}$, $\hat{m}_2^{ij}$ for Group 2. Considering the first group data the settings $\hat{m}_2^{ij}, \hat{m}_4^{ij}$ find an earlier break at 16.02.2007, while the setting $\hat{m}_3^{ij}$ detects a break at 26.11.2007. A more separated break is detected within the second group data at 31.05.2006 using the setting $\hat{m}_5^{ij}$, where only Spearman's rank correlation coefficient is used. Using the common break procedure we find that the break event at 31.05.2006 is clearly separated from the break event at 09.07.2007. The found break in 2006 may correspond to the early beginning of the last financial crisis in the summer of 2006. Further, the break event at 16.02.2007 is also separated to the one at the mid of 2007, whereas the separation is not that clear.

6. CONCLUSION AND OUTLOOK

In this work we proposed and investigated retro-perspective and online applicable monitoring procedures to test for constant parameters in factor copula models. Further, we developed and investigated a non parametric dependence measure test to test for constancy in a vector of copula based dependence measures like Spearman's rank correlation and quantile dependencies.

First, we proposed a new fluctuation tests for detecting structural breaks in factor copula models and analysed the behaviour under the null hypothesis of no parameter change. This is the work presented in Section 3, which is based on the published paper "Testing for structural Breaks in Factor Copulas", Manner et al. (2019). Due to the discontinuity of the SMM objective function this requires additional effort to derive a functional limit theorem for the model parameters. General approaches like Taylor expansions of first or second order are not applicable here. Further we proposed a non-parametric retrospective test to test for a constant vector of dependence measures. The presence of nuisance parameters in the asymptotic distribution of the two proposed test statistics requires a bootstrap approximation for parts of the asymptotic distribution. The proposed tests show good size and power properties in finite samples. An empirical application to a set of 32 stock returns indicates the presence of a breakpoint early in 2008, before the Lehman Brothers bankruptcy. Dependence has increased after this break providing evidence of a diversification breakdown and contagion among different stocks.

To apply the test in real time we further developed a new monitoring procedure for detecting structural breaks in factor copula models and analysed the behaviour under the null hypothesis of no change. Additionally, we extended the proposed non-parametric dependence measure test to online applicability. This is the work presented in Section 4, which is based on the paper "A Monitoring Procedure for detecting structural Breaks in Factor Copula Models", Manner et al. (2018). Again, the presence of nuisance parameters in the asymptotic distribution of the two proposed detectors requires a bootstrap approximation for parts of the asymptotic distribution. The case of detecting multiple breaks is also treated. In simulations, the proposed procedures show good size and power properties in single and multiple break settings in finite samples. An empirical application to a set of 10 stock returns of large financial firms indicates the presence of break points around July 2007 and August 2008, time points of the heights of the last financial crisis. The proposed online Value-at-Risk procedure shows the usefulness of the monitoring procedure in portfolio management. Here, we simulate residual data from the assumed factor copula model and transform the simulated returns in a second step to simulated log-returns, using the estimated time series models. After transforming the log-returns into general returns and a reasonable aggregation, we can calculate the VaR for the next time step. The proposed monitoring procedures gives us information about the periods of data we should use for our residual simulations. The real data example showed the usefulness to allow for structural breaks for the VaR evaluation especially in the time span of the last financial crisis. The presented works in Section 3 and 4 are joint works with Hans Manner and Dominik Wied. In these two sections we focused on the parametric test to test for a constant parameter vector of the considered factor copula model. In Section 5 we wanted to focus on the non-parametric dependence measure vector test first proposed in Section 3, which is based on the paper "On the Applicability of a nonparametric Test for constant Copula-based Dependence Measures: Dating Breakpoints and analyzing different Dependence Measure Sets", Stark (2018).

We investigated the test for a constant copula based dependence measure vector, considering pairwise averaged Spearman's rho and quantile dependencies in equidependence settings. Again, the asymptotic null distribution is not known in closed form and therefore estimated by an i.i.d. bootstrap procedure. A size and power anaylsis, using different dependence measure settings for different simulated fat and skewed distributed data, is considered. Here the best power properties are gained by considering solely Spearman's rank correlation and a combination of Spearman's rank correlation and quantile dependencies, where the simple setting using only the rank correlation coefficient works best. The settings using only upper or lower quantile dependencies suffer from poor power properties. Further we found that the use of upper quantile dependencies results in better power properties at present strongly left skewed data compared to lower quantile dependencies and on the other way around lower quantile dependencies result in better power properties by considering right skewed data compared to the usage of upper quantile dependencies. Considering jointly lower and upper quantile dependencies always results in better power properties as the separate considerations. The test is also applied to a real data application to show the usefulness of the flexibility by the choice of different dependence measure settings. We use historical data of ten large companies during the last financial crisis from 2002 to the mid of 2013. One advantage of the use of different dependence settings is that we are able to conclude that a detected break point is in a way more significant than another, if it is detected by more than one setting. On the other hand we also get a different break picture, motivating the usage of flexible dependence measure settings and the combination of rank correlation and quantile dependencies. Further we propose a heuristic procedure to be able to make a statement for equality of two estimated break point locations, transformed to the uniform interval, using different dependence vector settings.

Inspired by the motivation that dependence within a portfolio usually increases in times of financial crisis and the well known diversification strategy to optimize the portfolio, the tests can be used to detect and quantify contagion between different financial markets or to construct optimal portfolios in portfolio management.

In future research, our work could be extended in several interesting directions. First, one could extend the known model of factor copulas with time-varying exogenous regressors and discuss estimation methods. To be precise, the factor loadings should depend on a known function of observable regressors and a parameter vector κ. For example, in the context of returns on risky securities, regressors could be macroeconomic variables such as GDP or the central bank's interest rate. Also time-delayed market returns, volatility measures (such as the VIX) or deterministic trigonometric functions may also be useful. Challenging will be the estimation of the time-independent parameter vector κ. For this we will investigate to what extend the SMM procedure can be applied in this context. A theoretical demanding task will be to derive the consistency and asymptotic behavior of the derived parameter estimator $\hat{\kappa}$. An interesting question in this connection will be, how the regressors change the shape of the asymptotic distribution.

In a next step the new theoretical founding can be used to extend the test and monitoring procedures from Section 3 and Section 4 to the case of time-varying exogenous regressors and test for structural breaks in the parameter vector κ. Asymptotic distribution of the considered test statistic has to be derived and discussed. An extensive simulation study and empirical application should conclude the findings.

So far, factor copulas have always been considered to model cross sectional dependency. Another approach could be to model dependencies over time. We could examine how many time lags should be considered and which assumptions are required on the parameters. The next step would be the development of a factor copula VAR model, which models dependence of time and cross section.

7. APPENDIX

7.1. Proofs

This section contains all the proofs needed to proof Theorem 4 and Theorem 5.

7.1.1. Proof Theorem 4. Theorem 4 is proved in different steps. First, we provide a consistency result in Lemma 6. Then, Theorem 8, which is based at Theorem 7, yields a general convergence result for SMM estimators. Lemma 10, which is based at Lemma 9 provides stochastic equicontinuity for the objective function in a general SMM setting. Finally, Lemma 11 yields distribution results for the empirical moments in our specific problem. All these results are then used for proving Theorem 4. Note, that we use the abbreviation *c.s.* to indicate the usage of the Cauchy–Schwarz inequality.

Lemma 6. If $\hat{\theta}_{T,S} \xrightarrow{a.s.} \theta_0, \qquad T, S \to \infty$, then

$$\sup_{s\in[\varepsilon,1]} \|\hat{\theta}_{sT,S} - \theta_0\| \xrightarrow{p} 0, \qquad \forall\varepsilon > 0, \qquad T, S \to \infty.$$

Proof. Let $\delta > 0$, $\hat{\theta}_{T,S} \xrightarrow{a.s.} \theta_0$ and choose any $\varepsilon > 0$

$\Rightarrow \forall\gamma > 0$ there exists $T_0^*, S_0^* \in \mathbb{N}_+$, such that for all $T \geq T_0^*, S \geq S_0^*, \qquad \|\hat{\theta}_{T,S} - \theta_0\| < \gamma$

$\Rightarrow$ there exists $T_0, S_0 \in \mathbb{N}_+$ such that for all $T \geq T_0, S \geq S_0, \qquad \|\hat{\theta}_{T,S} - \theta_0\| < \delta$

Choose T, S with $\varepsilon T \geq T_0 \Leftrightarrow T \geq \frac{T_0}{\varepsilon}, S \geq S_0, \quad \forall\varepsilon > 0$ (in all cases $T \geq T_0$)

$\Rightarrow \forall s \in [\varepsilon, 1] : \|\hat{\theta}_{sT,S} - \theta_0\| < \delta$, for all $T \geq \frac{T_0}{\varepsilon}, S \geq S_0, \quad \forall\varepsilon > 0$

$\Rightarrow \sup_{s\in[\varepsilon,1]} \|\hat{\theta}_{sT,S} - \theta_0\| < \delta$, for all $T \geq \frac{T_0}{\varepsilon}, S \geq S_0, \quad \forall\varepsilon > 0$

$\Rightarrow \sup_{s\in[\varepsilon,1]} \|\hat{\theta}_{sT,S} - \theta_0\| \xrightarrow{p} 0, \qquad \forall\varepsilon > 0, \quad T, S \to \infty.$ □

Theorem 7. Under the null hypothesis $H_0 : \theta_1 = \theta_2 = \cdots = \theta_T$, suppose that

$\forall s \in [\varepsilon, 1], \varepsilon > 0 \quad Q_{sT,S}(\hat{\theta}_{sT,S}) = \sup_{\theta\in\Theta} Q_{sT,S}(\theta) - o_p^*((s^2T)^{-1}), \sup_{s\in[\varepsilon,1)} \|\hat{\theta}_{sT,S} - \theta_0\| \xrightarrow{p} 0,$

$T, S \to \infty$ and:

i) $Q_0(\theta)$ is maximized on $\theta_0 (= \theta_1 = \cdots = \theta_T)$

ii) $\theta_0 (= \theta_1 = \dots = \theta_T)$ is an interior point of Θ

iii) $Q_0(\theta)$ is twice differentiable at θ_0 with non singular second derivative $H = \nabla_{\theta\theta} Q_0(\theta_0)$

iv) $s\sqrt{T}\hat{D}_{sT}(\theta_0) \stackrel{d}{\Longrightarrow} A(s)$

v) $\forall \delta \to 0 \quad \sup\limits_{s\in[\varepsilon,1]} \sup\limits_{\|\theta-\theta_0\|<\delta} \left| \frac{\hat{R}_{sT}(\theta)}{1+s\sqrt{T}\|\theta-\theta_0\|} \right| \stackrel{p}{\longrightarrow} 0$
with $\hat{R}_{sT} = \frac{s\sqrt{T}[Q_{sT,S}(\theta)-Q_{sT,S}(\theta_0)-\hat{D}_{sT}(\theta-\theta_0)-(Q_0(\theta)-Q_0(\theta_0))]}{\|\theta-\theta_0\|}$

$\Rightarrow s\sqrt{T}(\hat{\theta}_{sT,S} - \theta_0) \stackrel{d}{\Longrightarrow} A^*(s) \qquad \forall s \in [\varepsilon, 1], \varepsilon > 0$ and $A^*(s) = H^{-1}A(s)$,
where $A(s)$ is a continuous Gaussian process.

Proof. For simplification set $Q := Q_0$ and $\hat{Q} := Q_{sT,S}$. We first show that $s\sqrt{T}\|\hat{\theta}_{sT,S} - \theta_0\| = O_p(1)$. With a Taylor-expansion of $Q(\theta)$ around θ_0 and knowing $\nabla_\theta Q(\theta_0) = 0$, due to condition i), we receive $Q(\theta) = Q(\theta_0) + \frac{1}{2}(\theta-\theta_0)'H(\theta-\theta_0) + o(\|\theta-\theta_0\|^3)$. We also know from condition i) and iii), that $\exists C > 0 : \ (\theta-\theta_0)'H(\theta-\theta_0) + o(\|\theta-\theta_0\|^3) \le -C\|\theta-\theta_0\|^2$
$\Rightarrow Q(\hat{\theta}_{sT,S}) \le Q(\theta_0) - C\|\hat{\theta}_{sT,S} - \theta_0\|^2$ and we obtain

$$
\begin{aligned}
0 &= \hat{Q}(\hat{\theta}_{sT,S}) - \hat{Q}(\theta_0) + o_p^*((s^2T)^{-1}) \\
&= Q(\hat{\theta}_{sT,S}) - Q(\theta_0) + \hat{D}'_{sT}(\hat{\theta}_{sT,S} - \theta_0) + \frac{1}{s\sqrt{T}}\|\hat{\theta}_{sT,S} - \theta_0\|\hat{R}_{sT}(\hat{\theta}_{sT,S}) + o_p^*((s^2T)^{-1}) \\
&\stackrel{c.s.}{\le} -C\|\hat{\theta}_{sT,S} - \theta_0\|^2 + \|\hat{D}'_{sT}\|\|\hat{\theta}_{sT,S} - \theta_0\| \\
&+ \|\hat{\theta}_{sT,S} - \theta_0\|(1 + s\sqrt{T}\|\hat{\theta}_{sT,S} - \theta_0\|)o_p(s^{-1}T^{-\frac{1}{2}}) + o_p^*((s^2T)^{-1}) \\
&= -(C + o_p(1))\|\hat{\theta}_{sT,S} - \theta_0\|^2 + \|\hat{\theta}_{sT,S} - \theta_0\|(\|\hat{D}'_{sT}\| + o_p(s^{-1}T^{-\frac{1}{2}})) + o_p^*((s^2T)^{-1}) \\
&\le -(C + o_p(1))\|\hat{\theta}_{sT,S} - \theta_0\|^2 + \|\hat{\theta}_{sT,S} - \theta_0\|O_p(s^{-1}T^{-\frac{1}{2}}) + o_p^*((s^2T)^{-1}) \\
&\Rightarrow \|\hat{\theta}_{sT,S} - \theta_0\|^2 \le \|\hat{\theta}_{sT,S} - \theta_0\|O_p(s^{-1}T^{-\frac{1}{2}}) + o_p^*((s^2T)^{-1}), \quad \forall s \in [\varepsilon, 1]. \qquad (1)
\end{aligned}
$$

Consider

$$\begin{aligned}\left(\|\hat{\theta}_{sT,S}-\theta_0\|+O_p(s^{-1}T^{-\frac{1}{2}})\right)^2 &= \|\hat{\theta}_{sT,S}-\theta_0\|^2+\|\hat{\theta}_{sT,S}-\theta_0\|O_p(s^{-1}T^{-\frac{1}{2}})+O_p(s^{-2}T^{-1})\\ &\overset{(1)}{\leq} \|\hat{\theta}_{sT,S}-\theta_0\|O_p(s^{-1}T^{-\frac{1}{2}})+o_p^*((s^2T)^{-1})+O_p(s^{-2}T^{-1})\\ &\leq O_p(s^{-2}T^{-1})\end{aligned}$$

$$\Rightarrow \left|\|\hat{\theta}_{sT,S}-\theta_0\|+O_p(s^{-1}T^{-\frac{1}{2}})\right| \leq O_p(s^{-1}T^{-\frac{1}{2}}), \quad \forall s\in[\varepsilon,1] \qquad (2)$$

and we get

$$\begin{aligned}\|\hat{\theta}_{sT,S}-\theta_0\| &= |\|\hat{\theta}_{sT,S}-\theta_0\|+O_p(s^{-1}T^{-\frac{1}{2}})-O_p(s^{-1}T^{-\frac{1}{2}})|\\ &\overset{c.s.}{\leq} |\|\hat{\theta}_{sT,S}-\theta_0\|+O_p(s^{-1}T^{-\frac{1}{2}})|+|-O_p(s^{-1}T^{-\frac{1}{2}})|\\ &\overset{(2)}{\leq} O_p(s^{-1}T^{-\frac{1}{2}})\end{aligned}$$

$$\Rightarrow s\sqrt{T}\|\hat{\theta}_{sT,S}-\theta_0\| = O_p(1), \quad \forall s\in[\varepsilon,1]. \qquad (3)$$

Note that for the numerator of the remainder Term $\hat{R}_{sT}$, without the factor $s\sqrt{T}$, we get with condition v) the scale

$$\begin{aligned}&o_p(1)(1+s\sqrt{T}\|\hat{\theta}_{sT,S}-\theta_0\|)\|\hat{\theta}_{sT,S}-\theta_0\|\frac{1}{s\sqrt{T}}\\ =&o_p\left(\frac{\|\hat{\theta}_{sT,S}-\theta_0\|}{s\sqrt{T}}+\|\hat{\theta}_{sT,S}-\theta_0\|^2\right)\\ \overset{(3)}{=}&o_p\left(O_p((s^2T)^{-1}+O_p((s^2T)^{-1})\right)\\ =&o_p((s^2T)^{-1}). \qquad (4)\end{aligned}$$

Now we can show the asymptotic behavior of $s\sqrt{T}(\hat{\theta}_{sT,S}-\theta_0)$. First let $\tilde{\theta}_{sT,S}=\theta_0-H^{-1}\hat{D}_{sT} \Rightarrow$

$\hat{D}_{sT} = -H(\tilde{\theta}_{sT,S} - \theta_0)$ (5) be the maximum of the approximation

$$\begin{aligned}
\hat{Q}(\theta) \approx& \hat{Q}(\theta_0) + \hat{D}'_{sT}(\theta - \theta_0) + Q(\theta) - Q(\theta_0) \\
\approx& \hat{Q}(\theta_0) + \hat{D}'_{sT}(\theta - \theta_0)' + \frac{1}{2}(\theta - \theta_0)H(\theta - \theta_0) \qquad (6)
\end{aligned}$$

and by construction $s\sqrt{T}$-consistent.

From the previous result (4), we know the rate of convergence of the remainder term of the approximation in (6). So we receive

$$\begin{aligned}
2[\hat{Q}(\hat{\theta}_{sT,S}) - \hat{Q}(\theta_0)] &= 2\hat{D}'_{sT}(\hat{\theta}_{sT,S} - \theta_0) + (\hat{\theta}_{sT,S} - \theta_0)'H(\hat{\theta}_{sT,S} - \theta_0) + o_p((s^2T)^{-1}) \\
&\overset{(5)}{=} (\hat{\theta}_{sT,S} - \theta_0)'H(\hat{\theta}_{sT,S} - \theta_0) - 2(\tilde{\theta}_{sT,S} - \theta_0)'H(\hat{\theta}_{sT,S} - \theta_0) + o_p((s^2T)^{-1})
\end{aligned}$$

and analogously for $\tilde{\theta}_{sT,S}$

$$\begin{aligned}
2[\hat{Q}(\tilde{\theta}_{sT,S}) - \hat{Q}(\theta_0)] &= 2\hat{D}'_{sT}(\tilde{\theta}_{sT,S} - \theta_0) + (\tilde{\theta}_{sT,S} - \theta_0)'H(\tilde{\theta}_{sT,S} - \theta_0) + o_p((s^2T)^{-1}) \\
&\overset{(5)}{=} -(\tilde{\theta}_{sT,S} - \theta_0)'H(\tilde{\theta}_{sT,S} - \theta_0) + o_p((s^2T)^{-1}).
\end{aligned}$$

Because $\hat{\theta}_{sT,S}, \tilde{\theta}_{sT,S} \in \Theta$, the rate of convergence of the remainder terms are known and $H = H(\theta_0)$ is negative definite and non singular

$$\begin{aligned}
\Rightarrow o_p^*((s^2T)^{-1}) =& 2[\hat{Q}(\hat{\theta}_{sT,S}) - \hat{Q}(\theta_0)] - 2[\hat{Q}(\tilde{\theta}_{sT,S}) - \hat{Q}(\theta_0)] \\
=& (\hat{\theta}_{sT,S} - \theta_0)'H(\hat{\theta}_{sT,S} - \theta_0) - 2(\tilde{\theta}_{sT,S} - \theta_0)'H(\hat{\theta}_{sT,S} - \theta_0) + (\tilde{\theta}_{sT,S} - \theta_0)'H(\tilde{\theta}_{sT,S} - \theta_0) \\
=& (\hat{\theta}_{sT,S} - \tilde{\theta}_{sT,S})'H(\hat{\theta}_{sT,S} - \tilde{\theta}_{sT,S}) \leq -C\|\hat{\theta}_{sT,S} - \tilde{\theta}_{sT,S}\|^2 \\
\Rightarrow& s\sqrt{T}\|\hat{\theta}_{sT,S} - \tilde{\theta}_{sT,S}\| = o_p^*(1). \qquad (7)
\end{aligned}$$

So we have $\forall s \in [\varepsilon, 1], \varepsilon > 0$

$$\begin{aligned}
&\|s\sqrt{T}(\hat{\theta}_{sT,S} - \theta_0) - (-s\sqrt{T}H^{-1}\hat{D}_{sT})\| \\
\overset{(5)}{=}&\|s\sqrt{T}(\hat{\theta}_{sT,S} - \theta_0) - s\sqrt{T}(\tilde{\theta}_{sT,S} - \theta_0)\| \\
=&\|s\sqrt{T}(\hat{\theta}_{sT,S} - \tilde{\theta}_{sT,S})\| \\
=&s\sqrt{T}\|(\hat{\theta}_{sT,S} - \tilde{\theta}_{sT,S})\| \overset{(7)}{=} o_p^*(1)
\end{aligned}$$

$$\Rightarrow s\sqrt{T}(\hat{\theta}_{sT,S} - \theta_0) \overset{p}{\Longrightarrow} -H^{-1}s\sqrt{T}\hat{D}_{sT} \underset{iv)}{\overset{d}{\Longrightarrow}} -H^{-1}A(s) = A^*(s).$$

$\square$

Theorem 8. Under the null hypothesis $H_0 : \theta_1 = \theta_2 = \cdots = \theta_T$, suppose that
$\forall s \in [\varepsilon, 1], \varepsilon > 0 : \quad g_{sT,S}(\hat{\theta}_{sT,S})'\hat{W}_{sT}g_{sT,S}(\hat{\theta}_{sT,S}) = \inf_{\theta\in\Theta} g_{sT,S}(\theta)'\hat{W}_{sT}g_{sT,S}(\theta) + o_p^*((s^2T)^{-1})$,
$\sup_{s\in[\varepsilon,1]} \|\hat{\theta}_{sT,S} - \theta_0\| \overset{p}{\longrightarrow} 0$, $\sup_{s\in[\varepsilon,1]} \|\hat{W}_{sT} - W\| \overset{p}{\longrightarrow} 0$, $\quad T, S \to \infty$ and:

i) There is a $\theta_0 (= \theta_1 = \cdots = \theta_T)$ such that $g_0(\theta_0) = 0$

ii) $\theta_0 (= \theta_1 = \cdots = \theta_T)$ is an interior point of Θ

iii) $g_0(\theta)$ is differentiable at θ_0 with derivative G such that $G'WG$ is non singular

iv) $s\sqrt{T}g_{sT,S}(\theta_0) \overset{d}{\Longrightarrow} A(s)$

v) $\forall \delta \to 0 \quad \sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} s\sqrt{T}\frac{\|g_{sT,S}(\theta)-g_{sT,S}(\theta_0)-g_0(\theta)\|}{1+s\sqrt{T}\|\theta-\theta_0\|} \overset{p}{\longrightarrow} 0$

$\Rightarrow s\sqrt{T}(\hat{\theta}_{sT,S} - \theta_0) \overset{d}{\Longrightarrow} A^*(s) \qquad \forall s \in [\varepsilon, 1], \varepsilon > 0$
and $A^*(s) = (G'WG)^{-1}G'WA(s)$,
where $A(s)$ is a continuous Gaussian process.

Proof. Theorem 8 follows by verifying the conditions of Theorem 7. Set $\hat{Q}(\theta) := Q_{sT}(\theta) := -\frac{1}{2}\hat{g}(\theta)'\hat{W}_{sT}\hat{g}(\theta) + \hat{\Delta}_{sT}(\theta)$ with $\hat{g}(\theta) := g_{sT,S}(\theta)$ and $Q(\theta) := Q_0(\theta) := -\frac{1}{2}g(\theta)'Wg(\theta)$ with

$g(\theta) := g_0(\theta)$. With a Taylor-expansion of $g(\theta)$ around θ_0

$$g(\theta) = g(\theta_0) + G(\theta - \theta_0) + o(\|\theta - \theta_0\|^2) = G(\theta - \theta_0) + o(\|\theta - \theta_0\|^2) \qquad (8),$$

we obtain

$$Q(\theta) = -\frac{1}{2}g(\theta)'Wg(\theta) \overset{(8)}{=} -\frac{1}{2}[G(\theta - \theta_0) + o(\|\theta - \theta_0\|^2)]'W[G(\theta - \theta_0) + o(\|\theta - \theta_0\|^2)]$$

and comparing this with a Taylor-expansion of $Q(\theta)$ around θ_0

$$Q(\theta) = Q(\theta_0) + \frac{1}{2}(\theta - \theta_0)'H(\theta - \theta_0) + o(\|\theta - \theta_0\|^3),$$

noting that $Q(\theta)$ is maximized at θ_0, it follows $H(\theta_0) = -G'WG$, where H is a non singular negative definite matrix. Because H is by construction a nonsingular negative definite matrix, $\exists$ neighborhood of θ_0, where $Q(\theta)$ has a unique maximum at θ_0 with $Q(\theta_0) = 0$.
$\Rightarrow$ Conditions i), ii) and iii) of Theorem 7 are satisfied. By choosing $\hat{D}_{sT} = -G'\hat{W}_{sT}g_{sT,S}(\theta_0)$ it follows, $\forall s \in [\varepsilon, 1]$,

$$s\sqrt{T}\hat{D}_{sT} = -s\sqrt{T}G'\hat{W}_{sT}g_{sT,S}(\theta_0) \underset{iv)}{\overset{d}{\Longrightarrow}} -G'WA(s),$$

thus condition iv) of Theorem 7 is fulfilled. Now we define

$$\hat{\varepsilon}(\theta) := \frac{\hat{g}(\theta) - \hat{g}(\theta_0) - g(\theta)}{1 + s\sqrt{T}\|\theta - \theta_0\|} \Leftrightarrow \hat{g}(\theta) = [1 + s\sqrt{T}\|\theta - \theta_0\|]\hat{\varepsilon}(\theta) + \hat{g}(\theta_0) + g(\theta) \qquad (9)$$

and we get

$$\begin{aligned}\hat{g}(\theta)'\hat{W}_{sT}\hat{g}(\theta) \overset{(9)}{=} &[1 + 2s\sqrt{T}\|\theta - \theta_0\| + s^2T\|\theta - \theta_0\|^2]\hat{\varepsilon}(\theta)'\hat{W}_{sT}\hat{\varepsilon}(\theta) \\ &+g(\theta)'\hat{W}_{sT}g(\theta) + \hat{g}(\theta_0)'\hat{W}_{sT}\hat{g}(\theta_0) + 2g(\theta)'\hat{W}_{sT}\hat{g}(\theta_0) \\ &+2[g(\theta) + \hat{g}(\theta_0)]'\hat{W}_{sT}\hat{\varepsilon}(\theta)[1 + s\sqrt{T}\|\theta - \theta_0\|] \qquad (10)\end{aligned}$$

Next we define the remainder term of $\hat{Q}(\theta)$

$$\hat{Q}(\theta) = -\frac{1}{2}\hat{g}(\theta)'\hat{W}_{sT}\hat{g}(\theta) + \hat{\Delta}_{sT}(\theta) = -\frac{1}{2}\hat{g}(\theta)'\hat{W}_{sT}\hat{g}(\theta) + \frac{1}{2}\hat{\varepsilon}(\theta)'\hat{W}_{sT}\hat{\varepsilon}(\theta) + \hat{g}(\theta_0)'\hat{W}_{sT}\hat{\varepsilon}(\theta).$$

The remainder term is just chosen in this way, that $\hat{Q}(\theta)$ is consistent with $-\frac{1}{2}\hat{g}(\theta)'\hat{W}_{sT}\hat{g}(\theta)$, which is shown in the next window and that we get the right rate of convergence, when checking condition v) of Theorem 7. First notice that by condition v) $\forall\delta > 0 \quad \sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} \|\hat{\varepsilon}(\theta)\| = o_p(s^{-1}T^{-\frac{1}{2}})$, furthermore

$\sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} \|\hat{g}(\theta_0)\| = o_p(s^{-1}T^{-\frac{1}{2}}) \quad , \quad \sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} \|\hat{W}_{sT}\| = O_p(1)$ and $\frac{\|g(\theta)-g(\theta_0)\|}{\|\theta-\theta_0\|} = O_p(1) \quad (11).$

$$\begin{aligned}
&\Rightarrow \forall\delta > 0 \quad \sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} \left|\hat{Q}(\theta) - (-\frac{1}{2}\hat{g}(\theta)'\hat{W}_{sT}\hat{g}(\theta))\right| \\
&= \sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} \left|\frac{1}{2}\hat{\varepsilon}(\theta)'\hat{W}_{sT}\hat{\varepsilon}(\theta) + \hat{g}(\theta_0)'\hat{W}_{sT}\hat{\varepsilon}(\theta)\right| \\
&\overset{c.s.}{\leq} \sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} \frac{1}{2}\|\hat{\varepsilon}(\theta)\|\|\hat{W}_{sT}\|\|\hat{\varepsilon}(\theta)\| + \|\hat{g}(\theta_0)\|\|\hat{W}_{sT}\|\|\hat{\varepsilon}(\theta)\| \\
&\overset{(11)}{=} O_p(1)(o_p(s^{-2}T^{-1}) + o_p(s^{-2}T^{-1})) = o_p(s^{-2}T^{-1}). \qquad (12)
\end{aligned}$$

With the consistency of $\hat{Q}(\theta)$ we can show the initial condition of Theorem 7

$$\begin{aligned}
&\forall s \in [\varepsilon,1], \varepsilon > 0 \quad \hat{g}(\hat{\theta}_{sT,S})'\hat{W}_{sT}\hat{g}(\hat{\theta}_{sT,S}) = \inf_{\theta\in\Theta} \hat{g}(\theta)'\hat{W}_{sT}\hat{g}(\theta) + o_p^*((s^2T)^{-1}) \\
\Leftrightarrow &\forall s \in [\varepsilon,1], \varepsilon > 0 \quad -\frac{1}{2}\hat{g}(\hat{\theta}_{sT,S})'\hat{W}_{sT}\hat{g}(\hat{\theta}_{sT,S}) = -\inf_{\theta\in\Theta} \frac{1}{2}\hat{g}(\theta)'\hat{W}_{sT}\hat{g}(\theta) - o_p^*((s^2T)^{-1}) \\
\overset{(12)}{\Leftrightarrow} &\forall s \in [\varepsilon,1], \varepsilon > 0 \quad \hat{Q}(\hat{\theta}_{sT,S}) = \sup_{\theta\in\Theta} \hat{Q}(\theta) - o_p^*((s^2T)^{-1}).
\end{aligned}$$

Finally we have to check condition v) of Theorem 7, for that we calculate

$$\begin{aligned}
&\left|\frac{\hat{R}_{sT}(\theta)}{1+s\sqrt{T}\|\theta-\theta_0\|}\right| \\
=&s\sqrt{T}\left|\frac{\hat{Q}(\theta)-\hat{Q}(\theta_0)-\hat{D}_{sT}(\theta-\theta_0)-(Q(\theta)-Q(\theta_0))}{\|\theta-\theta_0\|(1+s\sqrt{T}\|\theta-\theta_0\|)}\right| \\
=&s\sqrt{T}\left|\frac{-\frac{1}{2}\hat{g}(\theta)'\hat{W}_{sT}\hat{g}(\theta)+\frac{1}{2}\hat{\varepsilon}(\theta)'\hat{W}_{sT}\hat{\varepsilon}(\theta)+\hat{g}(\theta_0)'\hat{W}_{sT}\hat{\varepsilon}(\theta)+\frac{1}{2}\hat{g}(\theta_0)'\hat{W}_{sT}\hat{g}(\theta_0)}{\|\theta-\theta_0\|(1+s\sqrt{T}\|\theta-\theta_0\|)}\right. \\
&\left.+\frac{-\hat{D}_{sT}(\theta-\theta_0)-(Q(\theta)-Q(\theta_0))}{\|\theta-\theta_0\|(1+s\sqrt{T}\|\theta-\theta_0\|)}\right| \qquad (\hat{\varepsilon}(\theta_0)=0),
\end{aligned}$$

inserting (10) and $Q(\theta)=-\frac{1}{2}g(\theta)'Wg(\theta)$, sorting, triangle inequality, choosing $\hat{D}_{sT}=-G'\hat{W}_{sT}\hat{g}(\theta_0)$ and size up the resulting terms, leads to

$$\begin{aligned}
\leq&\frac{s\sqrt{T}[2s\sqrt{T}\|\theta-\theta_0\|+s^2T\|\theta-\theta_0\|^2]\left|\hat{\varepsilon}(\theta)'\hat{W}_{sT}\hat{\varepsilon}(\theta)\right|}{\|\theta-\theta_0\|(1+s\sqrt{T}\|\theta-\theta_0\|)} \qquad (=:r_1(\theta)) \\
&+\frac{s\sqrt{T}\left|(-g(\theta)+G(\theta-\theta_0))'\hat{W}_{sT}\hat{g}(\theta_0)\right|}{\|\theta-\theta_0\|(1+s\sqrt{T}\|\theta-\theta_0\|)} \qquad (=:r_2(\theta)) \\
&+\frac{s^2T\left|(g(\theta)+\hat{g}(\theta_0))'\hat{W}_{sT}\hat{\varepsilon}(\theta)\right|}{1+s\sqrt{T}\|\theta-\theta_0\|} \qquad (=:r_3(\theta)) \\
&+\frac{s\sqrt{T}\left|g(\theta)'\hat{W}_{sT}\hat{\varepsilon(\theta)}\right|}{\|\theta-\theta_0\|} \qquad (=:r_4(\theta)) \\
&+\frac{s\sqrt{T}\left|g(\theta)'[W-\hat{W}_{sT}]g(\theta)\right|}{\|\theta-\theta_0\|(1+s\sqrt{T}\|\theta-\theta_0\|)}. \qquad (=:r_5(\theta))
\end{aligned}$$

Now we have

$$\begin{aligned}
\forall\delta\to 0 \quad \sup_{s\in[\varepsilon,1]}\ \sup_{\|\theta-\theta_0\|<\delta}&\left|\frac{\hat{R}_{sT}(\theta)}{1+s\sqrt{T}\|\theta-\theta_0\|}\right| \\
\leq\sup_{s\in[\varepsilon,1]}\ \sup_{\|\theta-\theta_0\|<\delta}&\sum_{i=1}^{5}r_i(\theta)=o_p(1)
\end{aligned}$$

and we just have to check the convergence of the $r_i(\theta)$ terms for $i\in\{1,2,3,4,5\}$. For r_1, we have

$$
\begin{aligned}
\sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} r_1(\theta) &= \sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} \frac{s\sqrt{T}[2s\sqrt{T}\|\theta-\theta_0\| + s^2T\|\theta-\theta_0\|^2]\left|\hat{\varepsilon}(\theta)'\hat{W}_{sT}\hat{\varepsilon}(\theta)\right|}{\|\theta-\theta_0\|(1+s\sqrt{T}\|\theta-\theta_0\|)} \\
&\overset{c.s.}{\leq} \sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} \frac{s\sqrt{T}(s\sqrt{T}\|\theta-\theta_0\|(2+s\sqrt{T}\|\theta-\theta_0\|))\|\hat{\varepsilon}(\theta)\|^2\|\hat{W}_{sT}\|}{\|\theta-\theta_0\|(1+s\sqrt{T}\|\theta-\theta_0\|)} \\
&\leq \sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} cs^2T\|\hat{\varepsilon}(\theta)\|^2\|\hat{W}_{sT}\| \qquad \text{(for a constant } c>1) \\
&\overset{(11)}{=} o_p(1).
\end{aligned}
$$

For r_2, we obtain

$$
\begin{aligned}
\sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} r_2(\theta) &= \sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} \frac{s\sqrt{T}\left|(-g(\theta)+G(\theta-\theta_0))'\hat{W}_{sT}\hat{g}(\theta_0)\right|}{\|\theta-\theta_0\|(1+s\sqrt{T}\|\theta-\theta_0\|)} \\
&\overset{c.s.}{\leq} \sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} \frac{s\sqrt{T}o(\|\theta-\theta_0\|^2)\|\hat{W}_{sT}\|\|\hat{g}(\theta_0)\|}{\|\theta-\theta_0\|} \\
&\leq \sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} s\sqrt{T}o(\|\theta-\theta_0\|)\|\hat{W}_{sT}\|\|\hat{g}(\theta_0)\| \\
&\overset{(11)}{=} o_p(1).
\end{aligned}
$$

Considering r_3 yields

$$
\begin{aligned}
\sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} r_3(\theta) &= \sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} \frac{s^2T\left|(g(\theta)+\hat{g}(\theta_0))'\hat{W}_{sT}\hat{\varepsilon}(\theta)\right|}{1+s\sqrt{T}\|\theta-\theta_0\|} \\
&\overset{c.s.}{\leq} \sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} \left(s^2T\|\hat{g}(\theta_0)\| + sT^{\frac{1}{2}}\frac{\|g(\theta)\|}{\|\theta-\theta_0\|}\right)\|\hat{W}_{sT}\|\|\hat{\varepsilon}(\theta)\| \\
&\leq \sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} \left(s^2To_p(s^{-1}T^{-\frac{1}{2}}) + sT^{\frac{1}{2}}O_p(1)\right)\|\hat{W}_{sT}\|\|\hat{\varepsilon}(\theta)\| \\
&\leq \sup_{s\in[\varepsilon,1]} \sup_{\|\theta-\theta_0\|<\delta} sT^{\frac{1}{2}}O_p(1)\|\hat{W}_{sT}\|\|\hat{\varepsilon}(\theta)\| \\
&\overset{(11)}{=} o_p(1).
\end{aligned}
$$

For r_4, it holds

$$\begin{aligned}\sup_{s\in[\varepsilon,1]}\sup_{\|\theta-\theta_0\|<\delta} r_4(\theta) &= \sup_{s\in[\varepsilon,1]}\sup_{\|\theta-\theta_0\|<\delta} \frac{s\sqrt{T}\left|g(\theta)'\hat{W}_{sT}\varepsilon(\hat{\theta})\right|}{\|\theta-\theta_0\|}\\ &\overset{c.s.}{\leq} s\sqrt{T}O_p(1)\|\hat{W}_{sT}\|\|\hat{\varepsilon}(\theta)\|\\ &\overset{(11)}{=} o_p(1).\end{aligned}$$

Finally, for r_5,

$$\begin{aligned}\sup_{s\in[\varepsilon,1]}\sup_{\|\theta-\theta_0\|<\delta} r_5(\theta) &= \sup_{s\in[\varepsilon,1]}\sup_{\|\theta-\theta_0\|<\delta} \frac{s\sqrt{T}\left|g(\theta)'[W-\hat{W}_{sT}]g(\theta)\right|}{\|\theta-\theta_0\|(1+s\sqrt{T}\|\theta-\theta_0\|)}\\ &\overset{c.s.}{\leq} \sup_{s\in[\varepsilon,1]}\sup_{\|\theta-\theta_0\|<\delta} \frac{s\sqrt{T}\|g(\theta)\|^2\|W-\hat{W}_{sT}\|}{s\sqrt{T}\|\theta-\theta_0\|^2}\\ &= \sup_{s\in[\varepsilon,1]}\sup_{\|\theta-\theta_0\|<\delta}\left(\frac{\|g(\theta)\|}{\|\theta-\theta_0\|}\right)^2 o_p(1)\\ &= o_p(1).\end{aligned}$$

$\square$

Lemma 9. Under Assumption 1, 2, 3.ii) and 3.iii)

i) $g_{sT,S}(\theta)$ is stochastically Lipschitz-continuous $\forall s\in[\varepsilon,1], \varepsilon>0$, i.e.,

$$\exists B = O_p(1) \text{ such that } \forall \theta_1,\theta_2\in\Theta: \qquad \|g_{sT,S}(\theta_1)-g_{sT,S}(\theta_2)\| \leq B\|\theta_1-\theta_2\|$$

ii) $\exists\delta>0$ such that

$$\limsup_{T,S\to\infty} E\left(B^{2+\delta}\right) < \infty.$$

Proof. Without loss of generality suppose $g_{sT,S}(\theta)$ is a one-dimensional function, otherwise show the Lipschitz-continuity for every entry of the vector $g_{sT,S}(\theta)$.

i) We know $\tilde{m}_S(\theta) = m_0(\theta) + o_p(1) \quad (13)$, and from Assumption 3.iii), $m_0(\theta)$ is Lipschitz-continuous, due to combination of Lipschitz-continuous bivariate copulas $C_{ij}(\theta)$. Further from Assumption 3. v) we have

$$|\tilde{m}_S(\theta_1) - \tilde{m}_S(\theta_2)| \leq C_S \|\theta_1 - \theta_2\|. \qquad (14)$$

Now consider

$$\begin{aligned}|g_{sT,S}(\theta_1) - g_{sT,S}(\theta_2)| =& |\hat{m}_{sT} - \tilde{m}_S(\theta_1) - \hat{m}_{sT} + \tilde{m}_S(\theta_2)| \\ =& |\tilde{m}_S(\theta_2) - \tilde{m}_S(\theta_1)| = |\tilde{m}_S(\theta_1) - \tilde{m}_S(\theta_2)| \\ \overset{(14)}{\leq}& C_S \|\theta_1 - \theta_2\|.\end{aligned}$$

ii) For some $\delta > 0$

$$\Rightarrow \limsup_{T,S\to\infty} E\left(C_S^{2+\delta}\right) < \infty.$$

□

Lemma 10. Under Assumption 1, 2, 3.ii) and 3.iii), for $\frac{S}{T} \to \infty$ or $\frac{S}{T} \to k \in (0, \infty)$,

$$v_{sT,S}(\theta) = \sqrt{sT}[g_{sT,S}(\theta) - g_0(\theta)] \quad \text{is stochastically equicontinuous} \quad \forall s \in [\varepsilon, 1], \varepsilon > 0$$

Proof. By Lemma 9)i) $\{g_{sT,S}(\theta) : \theta \in \Theta\}$ is Lipschitz-continuous $\forall s \in [\varepsilon, 1], \varepsilon > 0$ and so a Type II class of functions in Andrews (1994). By Theorem 2 of Andrews $\{g_{sT,S}(\theta) : \theta \in \Theta\}$ satisfies Pollard's entropy condition with envelope

$$\max\{1, \quad \sup_{\theta\in\Theta} \|g_{sT,S}(\theta)\|, \quad B\}, \qquad \forall s \in [\varepsilon, 1], \varepsilon > 0.$$

$\Rightarrow$ Assumption A of Andrews (1994) is satisfied.

Furthermore $g_{sT,S}(\theta)$ is bounded and by Lemma 9)ii) it holds

$$\limsup_{T,S\to\infty} E\left(B^{2+\delta}\right) < \infty.$$

$\Rightarrow$ Assumption B of Andrews (1994) is satisfied. Then with Theorem 1 of Andrews (1994) and noting, that Assumption C is fulfilled by construction

$$v_{sT,S}(\theta) = \sqrt{sT}[g_{sT,S}(\theta) - g_0(\theta)] \quad \text{is stochastically equicontinuous} \quad \forall s \in [\varepsilon, 1], \varepsilon > 0.$$

□

Lemma 11. We consider the dependence measures Spearman's rho and quantile dependence measures, which are functions only depending on bivariate copulas.
Under the null and Assumption 1 and 2, for $T \to \infty$,

$$s\sqrt{T}(\hat{m}_{sT} - m_0(\theta_0)) \overset{d}{\Longrightarrow} A(s), \qquad T \to \infty, \quad \forall s \in [\varepsilon, 1], \varepsilon > 0$$

where $A(s)$ is defined in the proof and θ_0 the value of all θ_t under the null.

Proof. By Assumption 2.7 (15) the sequential empirical copula of the N-dimensional random vectors fulfills

$$\begin{aligned} \mathbb{C}_{sT} :=& s\sqrt{T}\left[\hat{C}^s(\mathbf{u}) - C(\mathbf{u})\right] \\ =& \frac{1}{\sqrt{T}}\left[\sum_{t=1}^{\lfloor sT \rfloor} \mathbb{1}\{\hat{\mathbf{F}}^s(\hat{\eta}_t) \leq \mathbf{u}\} - C(\mathbf{u})\right] \\ \underset{(15)}{\overset{d}{\Longrightarrow}}& A^*(s, \mathbf{u}), \qquad T \to \infty, \quad \forall s \in [\varepsilon, 1], \varepsilon > 0, \end{aligned}$$

Note that Spearman's rho between the i-th and j-th component is given by

$$12 \int_0^1 \int_0^1 C(1, \dots, 1, u_i, 1, \dots, 1, u_j, 1, \dots, 1) du_i du_j - 3$$

and that the quantile dependencies are projections of the N-dimensional copula onto one specific point divided by some prespecified constant. Define the function $m^{ij}(C)$ as the function which generates a vector of all considered dependence measures (Spearman's rho and/or quantile dependencies for different levels) between the i-th and j-th component out of the copula C. Without loss of generality consider the equidependence case (in the same way the argumentation holds for the block equidependence case, only that we average all intra- and inter-group dependence measures), then the function

$$\begin{aligned} m(C): D[0,1]^N &\to \mathbb{R}^k \\ C &\to m(C) = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} m^{ij*}(C) \end{aligned}$$

is continuous and we directly obtain

$$s\sqrt{T}(\hat{m}_{sT} - m_0(\theta)) = s\sqrt{T}\,[m(C^s) - m(C)] \stackrel{d}{\Longrightarrow} \frac{2}{N(N-1)} \left(\sum_{i,j} m^{ij}(A^*(s, \mathbf{u})) \right) =: A(s)$$

as $T \to \infty$ with $s \in [\varepsilon, 1], \varepsilon > 0$. Here, $m^{ij}(\cdot)$ is the same function as $m^{ij*}(\cdot)$ with the only difference that the formula for Spearman's rho between the i-th and j-th component is replaced by

$$12 \int_0^1 \int_0^1 C(1, \ldots, 1, u_i, 1, \ldots, 1, u_j, 1, \ldots, 1) du_i du_j.$$

$\square$

Proof of Theorem 4

The proof follows by checking the conditions of Theorem 8. The initial conditions of Theorem 8 follow by Assumption 4.iii) and Lemma 6.

i) $g_0(\theta_0) = 0$ follows directly by construction, because $g_0(\theta) = m_0(\theta_0) - m_0(\theta)$.

ii) $\theta_0(= \theta_1 = \cdots = \theta_T)$ is an interior point of Θ given by Assumption 4.i).

iii) $g_0(\theta)$ is differentiable at θ_0 with derivative G such that $G'WG$ is non singular, given by Assumption 4.ii).

iv) 1) If $\frac{S}{T} \to \infty$ as $T, S \to \infty$,

$$\begin{aligned} s\sqrt{T} g_{sT,S}(\theta_0) =& s\sqrt{T}(\hat{m}_{sT} - \tilde{m}_S(\theta_0)) \\ =& s\sqrt{T}(\hat{m}_{sT} - m_0(\theta_0)) + s\sqrt{T}(m_0(\theta_0) - \tilde{m}_S(\theta_0)) \\ =& s\sqrt{T}(\hat{m}_{sT} - m_0(\theta_0)) - \frac{\sqrt{T}}{\sqrt{S}} s\sqrt{S}(\tilde{m}_S(\theta_0) - m_0(\theta_0)) \\ \underset{Lemma\ 11}{\overset{d}{\Longrightarrow}}& A(s) \end{aligned}$$

2)If $\frac{S}{T} \to k \in (0, \infty)$ as $T, S \to \infty$,

$$\begin{aligned} s\sqrt{T} g_{sT,S}(\theta_0) =& s\sqrt{T}(\hat{m}_{sT} - \tilde{m}_S(\theta_0)) \\ =& s\sqrt{T}(\hat{m}_{sT} - m_0(\theta_0)) + s\sqrt{T}(m_0(\theta_0) - \tilde{m}_S(\theta_0)) \\ =& s\sqrt{T}(\hat{m}_{sT} - m_0(\theta_0)) - \frac{\sqrt{T}}{\sqrt{S}} s\sqrt{S}(\tilde{m}_S(\theta_0) - m_0(\theta_0)) \\ \underset{Lemma\ 11}{\overset{d}{\Longrightarrow}}& A(s) - \frac{s}{\sqrt{k}} A(1), \end{aligned}$$

combined we get

$$s\sqrt{T} g_{sT,S}(\theta_0) \overset{d}{\Longrightarrow} A(s) - \frac{s}{\sqrt{k}} A(1), \qquad T, S \to \infty, \quad \forall s \in [\varepsilon, 1], \varepsilon > 0.$$

v) We know by Lemma 10, that for $\frac{S}{T} \to \infty$ or $\frac{S}{T} \to k \in (0, \infty)$
$v_{sT,S}(\theta) = \sqrt{sT}[g_{sT,S}(\theta) - g_0(\theta)]$ is stochastically equicontinuous $\forall s \in [\varepsilon, 1], \varepsilon > 0$.

$$\begin{aligned} \Rightarrow \forall \varepsilon > 0,\ \eta > 0,\ \exists \delta > 0 : \limsup_{T \to \infty} P & \left[\sup_{s \in [\varepsilon, 1]} \sup_{\|\theta - \theta_0\| < \delta} \|v_{sT,S}(\theta) - v_{sT,S}(\theta_0)\| > \eta \right] \\ = \limsup_{T \to \infty} P & \left[\sup_{s \in [\varepsilon, 1]} \sup_{\|\theta - \theta_0\| < \delta} \sqrt{sT} \|g_{sT,S}(\theta) - g_{sT,S}(\theta_0) - g_0(\theta)\| > \eta \right] < \varepsilon. \quad (16) \end{aligned}$$

Furthermore the inequality

$$s\sqrt{T} \frac{\|g_{sT,S}(\theta) - g_{sT,S}(\theta_0) - g_0(\theta)\|}{1 + s\sqrt{T}\|\theta - \theta_0\|} \leq s\sqrt{T} \|g_{sT,S}(\theta) - g_{sT,S}(\theta_0) - g_0(\theta)\| \qquad (17)$$

is valid $\forall s \in [\varepsilon, 1]$.

Finally we obtain

$$\begin{aligned}
&\limsup_{T\to\infty} P\left[\sup_{s\in[\varepsilon,1]}\sup_{\|\theta-\theta_0\|<\delta} s\sqrt{T}\frac{\|g_{sT,S}(\theta)-g_{sT,S}(\theta_0)-g_0(\theta)\|}{1+s\sqrt{T}\|\theta-\theta_0\|}>\eta\right]\\
\leq&\limsup_{T\to\infty} P\left[\sup_{s\in[\varepsilon,1]}\sup_{\|\theta-\theta_0\|<\delta} \sqrt{sT}\frac{\|g_{sT,S}(\theta)-g_{sT,S}(\theta_0)-g_0(\theta)\|}{1+s\sqrt{T}\|\theta-\theta_0\|}>\eta\right]\\
\overset{(17)}{\leq}&\limsup_{T\to\infty} P\left[\sup_{s\in[\varepsilon,1]}\sup_{\|\theta-\theta_0\|<\delta} \sqrt{sT}\|g_{sT,S}(\theta)-g_{sT,S}(\theta_0)-g_0(\theta)\|>\eta\right]\overset{(16)}{<}\varepsilon.
\end{aligned}$$

Note that, for the first inequality, we use that $0 < s \leq \sqrt{s}\ \forall s \in [\varepsilon, 1], \varepsilon > 0$.

This completes the proof. $\square$

7.1.2. Proof Theorem 5. For the proof of Theorem 5 we first state our assumptions. Assumption 6 and Assumption 7 ensure that the estimated rank correlation and quantile dependencies converge to their respective population counterparts.

***Assumption* 6.** i) The distribution function of the innovations F_η and the joint distribution function of the factors $F_X(\theta)$ are continuous.

ii) Every bivariate marginal copula $C_{ij}(u_i, u_j; \theta)$ of $\mathbf{C}(u; \theta)$ has continuous partial derivatives with respect to $u_i \in (0,1)$ and $u_j \in (0,1)$.

The assumption is similar to Assumption 1 in Oh and Patton (2013), but the assumption on the copula is relaxed in the sense that the restriction of u_i and v_i is relaxed to the open interval $(0,1)$.

***Assumption* 7.** The first order derivatives of the functions $\phi \mapsto \mu_t(\phi)$ and $\phi \mapsto \sigma_t(\phi)$ exist and are given by $\dot{\mu}_t(\phi) := \frac{\partial \mu_t(\phi)}{\partial \phi'}$ and $\dot{\sigma}_{kt}(\phi) := \frac{\partial [\sigma_t(\phi)]_{\text{k-th column}}}{\partial \phi'}$ for $k = 1, \ldots, N$. Moreover, define $\gamma_{0t} := \sigma_t^{-1}(\hat{\phi})\dot{\mu}_t(\hat{\phi})$ and $\gamma_{1kt} := \sigma_t^{-1}(\hat{\phi})\dot{\sigma}_{kt}(\hat{\phi})$ such as

$$d_t := \eta_t - \hat{\eta}_t - \left(\gamma_{0t} + \sum_{k=1}^{N} \eta_{kt}\gamma_{1kt}\right)(\hat{\phi} - \phi_0),$$

with η_{kt} being the k-th row of η_t and γ_{0t} such as γ_{1kt} are $\mathcal{E}_{t-1}$-measurable, where $\mathcal{E}_{t-1}$ contains information from the past as well as possible information from exogenous variables.

i) $\frac{1}{T}\sum_{t=1}^{\lfloor sT \rfloor} \gamma_{0t} \xrightarrow{p} s\Gamma_0$ and $\frac{1}{T}\sum_{t=1}^{\lfloor sT \rfloor} \gamma_{1kt} \xrightarrow{p} s\Gamma_{1k}$, uniformly in $s \in [\varepsilon, 1]$, $\varepsilon > 0$, where Γ_0 and Γ_{1k} are deterministic for $k = 1, \ldots, N$.

ii) $\frac{1}{T}\sum_{t=1}^{T} E(\|\gamma_{0t}\|)$, $\frac{1}{T}\sum_{t=1}^{T} E(\|\gamma_{0t}\|^2)$, $\frac{1}{T}\sum_{t=1}^{T} E(\|\gamma_{1kt}\|)$ and $\frac{1}{T}\sum_{t=1}^{T} E(\|\gamma_{1kt}\|^2)$ are bounded for $k = 1, \ldots, N$.

iii) There exists a sequence of positive numbers $r_t > 0$ with $\sum_{i=1}^{\infty} r_t < \infty$, such that the sequence $\max_{1 \le t \le T} \frac{\|d_t\|}{r_t}$ is tight.

iv) $\max\limits_{1\leq t\leq T}\frac{\|\gamma_{0t}\|}{\sqrt{T}} = o_p(1)$ and $\max\limits_{1\leq t\leq T}\frac{|\eta_{kt}|\|\gamma_{1kt}\|}{\sqrt{T}} = o_p(1)$ for $k = 1, \ldots, N$.

v) $(\alpha_T(s, \mathbf{u}), \sqrt{T}(\hat{\phi} - \phi_0))$ weakly converges to a continuous Gaussian process in $\mathcal{D}((0,1] \times [0,1]^N) \times \mathbb{R}^r$, where $\mathcal{D}((0,1] \times [0,1]^N)$ is the space of all càdlàg-functions on $(0,1] \times [0,1]^N$, with

$$\alpha_T(s, \mathbf{u}) := \frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor sT \rfloor} \left\{ \prod_{k=1}^{N} \mathbb{1}\{U_{kt} \leq u_k\} - \mathbf{C}(\mathbf{u}; \theta) \right\}.$$

vi) $\frac{\partial F_\eta}{\partial \eta_k}$ and $\eta_k \frac{\partial F_\eta}{\partial \eta_k}$ are bounded and continuous on $\overline{\mathbb{R}}^N = [-\infty, \infty]^N$ for $k = 1, \ldots, N$.

vii) For $\mathbf{u} \in [0,1]^N$, $s \in [m, 1]$ and $\hat{\mathbf{F}}^{1+(s-m)T:st}(\hat{\eta}_t) = (\hat{F}_1^{1+(s-m)T:st}(\hat{\eta}_{1t}), \ldots, \hat{F}_N^{1+(s-m)T:st}(\hat{\eta}_{Nt}))$, the sequential empirical copula process

$$\frac{1}{\sqrt{T}} \left[\sum_{t=1+\lfloor (s-m)T \rfloor}^{\lfloor sT \rfloor} \mathbb{1}\{\hat{\mathbf{F}}^{1+(s-m)T:st}(\hat{\eta}_t) \leq \mathbf{u}\} - C(\mathbf{u}) \right]$$

converges in distribution to some limit process $A^*(s, \mathbf{u})$ on $[0,1]^N \times [m, 1]$

Parts i) to vi) of this assumption are similar to Assumption 2 in Oh and Patton (2013), only part (i) is more restrictive. We need this because we consider successively estimated parameters. Part vii) ensures that the empirical copula process of the residuals has some well defined limit. Note that Assumption vii) is plausible and follows from a combination of the results in Bücher et al. (2014) and Remillard (2017). The next assumption is needed for consistency of the successively estimated parameters. It is the same as Assumption 3 in Oh and Patton (2013) with the difference that part (iv) is adapted to our situation. Note that part i) ensures the identifiability of the factor model.

Assumption 8. i) $g_0(\theta)$ is the probability limit function of $g_{.,S}(\theta)$ for $T, S \to \infty$, e.g. $g_{.,S}(\theta) = g_{1:mT,S}(\theta)$ and it holds $g_0(\theta) = 0$ only for $\theta = \theta_0$ (the value of all θ_t under the null).

ii) The space Θ of all θ is compact.

iii) Every bivariate marginal copula $C_{ij}(u_i, u_j; \theta)$ of $\mathbf{C}(u;\theta)$ is Lipschitz-continuous for $(u_i, u_j) \in (0,1) \times (0,1)$ on Θ.

iv) The sequential weighting matrix $\hat{W}_{(s-m)T:sT}$ is $O_p(1)$ and $\sup_{s\in[m,1]} \|\hat{W}_{(s-m)T:sT} - W\| \xrightarrow{p} 0$ for $m \geq \varepsilon > 0$.

v) It holds for the moment simulating function $\tilde{m}_S(\theta)$ that, for $\theta_1, \theta_2 \in \Theta$,

$$|\tilde{m}_S(\theta_1) - \tilde{m}_S(\theta_2)| \leq C_S \|\theta_1 - \theta_2\|$$

with a random variable C_S that is independent of $\theta_1 - \theta_2$ and that fulfills $E(C_S^{2+\delta}) < \infty$ for some $\delta > 0$.

Finally, we need an assumption for distributional results, which is the same as Assumption 4 in Oh and Patton (2013) with a difference in part iii).

Assumption 9. i) θ_0 is an interior point of Θ.

ii) $g_0(\theta)$ is differentiable at θ_0 with derivative G such that $G'WG$ is non singular.

iii) $\forall s \in [m,1], \varepsilon > 0$: $\quad g_{.,S}(\theta_{(s-m)T:sT,S})'\hat{W}g_{.,S}(\theta_{(s-m)T:sT,S}) = \inf_{\theta\in\Theta} g_{.,S}(\theta)'\hat{W}g_{.,S}(\theta) + d_T$, where $d_T = o_p^*((m^2T)^{-1})$ (instead of $o_p((m^2T)^{-1})$) indicates that the remainder term on the right hand side tends to zero *and* is non-negative.

Proof of Theorem 5

We consider the dependence measures Spearman's rho and quantile dependence measures, which are functions only depending on bivariate copulas.

Under the null and all mentioned assumptions, we first want to show

$$m\sqrt{T}(\hat{m}_{(s-m)T:sT} - m_0(\theta_0)) \stackrel{d}{\Longrightarrow} A(s), \qquad T \to \infty, \quad \forall s \in [m,1], m \geq \varepsilon > 0$$

where $A(s)$ is a Gaussian process and θ_0 the value of all θ_t under the null.

By Assumption 7 vii) (1) the sequential empirical copula of the N-dimensional random vectors fulfills

$$\begin{aligned}\mathbb{C}_T :=& m\sqrt{T}\left[\hat{C}_{1+(s-m)T:sT}(\mathbf{u}) - C(\mathbf{u})\right]\\ =& \frac{1}{\sqrt{T}}\left[\sum_{t=1+\lfloor (s-m)T\rfloor}^{\lfloor sT\rfloor} \mathbb{1}\{\hat{\mathbf{F}}^{1+(s-m)T:sT}(\hat{\eta}_t) \leq \mathbf{u}\} - C(\mathbf{u})\right]\\ \xrightarrow[(1)]{d}& = A^*(s,\mathbf{u}), \quad T\to\infty, \quad \forall s\in[m,1], m\geq\varepsilon>0,\end{aligned}$$

where $\mathbf{u}\in[0,1]^N$ and $\hat{\mathbf{F}}^{1+(s-m)T:sT}(\hat{\eta}_t) := (\hat{F}_1^{1+(s-m)T:sT}(\hat{\eta}_{1t}),\ldots,\hat{F}_N^{1+(s-m)T:sT}(\hat{\eta}_{Nt}))$. Here, $\hat{F}_j^{1+(s-m)T:sT}$ denotes the marginal empirical distribution function of the j-th component and $\hat{C} := \hat{C}_{1+(s-m)T:sT}(\mathbf{u})$ the empirical copula both calculated from the data between the time point $1+\lfloor (s-m)T\rfloor$ and time point $\lfloor sT\rfloor$. Note that Spearman's rho between the i-th and j-th component is given by

$$12\int_0^1\int_0^1 C(1,\ldots,1,u_i,1,\ldots,1,u_j,1,\ldots,1)du_idu_j - 3$$

and that the quantile dependencies are projections of the N-dimensional copula onto one specific point divided by some prespecified constant. Define the function $m^{ij}(C)$ as the function which generates a vector of all considered dependence measures (Spearman's rho and/or quantile dependencies for different levels) between the i-th and j-th component out of the copula C. Without loss of generality consider the equicontinuity case (averaging over all possible pairs, for details see Oh and Patton (2017)), then the function

$$\begin{aligned} m(C): D[0,1]^N &\rightarrow \mathbb{R}^k\\ C &\rightarrow m(C) = \frac{2}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} m^{ij*}(C)\end{aligned}$$

is continuous and we directly obtain

$$
\begin{aligned}
m\sqrt{T}(\hat{m}_{1+(s-m)T:sT} - m_0(\theta)) =& m\sqrt{T}\left[m(\hat{C}) - m(C)\right] \\
&\overset{d}{\Longrightarrow} \frac{2}{N(N-1)}\left(\sum_{i,j} m^{ij}(A^*(s,\mathbf{u}))\right) =: A(s)
\end{aligned}
$$

as $T \to \infty$ with $s \in [m,1], m \geq \varepsilon > 0$. Here, $m^{ij}(\cdot)$ is the same function as $m^{ij*}(\cdot)$ with the only difference that the formula for Spearman's rho between the i-th and j-th component is replaced by

$$
12\int_0^1\int_0^1 C(1,\ldots,1,u_i,1,\ldots,1,u_j,1,\ldots,1)du_i du_j.
$$

Then we receive for $\frac{S}{T} \to k \in (0,\infty]$ and $T, S \to \infty$

$$
\begin{aligned}
m\sqrt{T}g_{1+(s-m)T:sT,S}(\theta) &= m\sqrt{T}\left(\hat{m}_{1+(s-m)T:sT} - \tilde{m}_S(\theta)\right) \\
=& m\sqrt{T}\left(\hat{m}_{1+(s-m)T:sT} - m_0(\theta)\right) - m\sqrt{T}\left(\tilde{m}_S - m_0(\theta)\right) \\
=& m\sqrt{T}\left(\hat{m}_{1+(s-m)T:sT} - m_0(\theta)\right) - \sqrt{\frac{T}{S}}m\sqrt{S}\left(\tilde{m}_S - m_0(\theta)\right) \\
\overset{d}{\Longrightarrow}& A(s) - \frac{m}{\sqrt{k}}B,
\end{aligned}
$$

where $B = N(0,\Sigma_0)$ is a centered Gaussian distribution with covariance matrix Σ_0, for details see Oh and Patton (2013). The limit result then follows with the same steps as in the proof of Theorem 4 from Section 3, using the given limit result for $m\sqrt{T}g_{1+(s-m)T:sT,S}(\theta)$ and replacing the scale factor $s\sqrt{T}$ by $m\sqrt{T}$.

This completes the proof. □

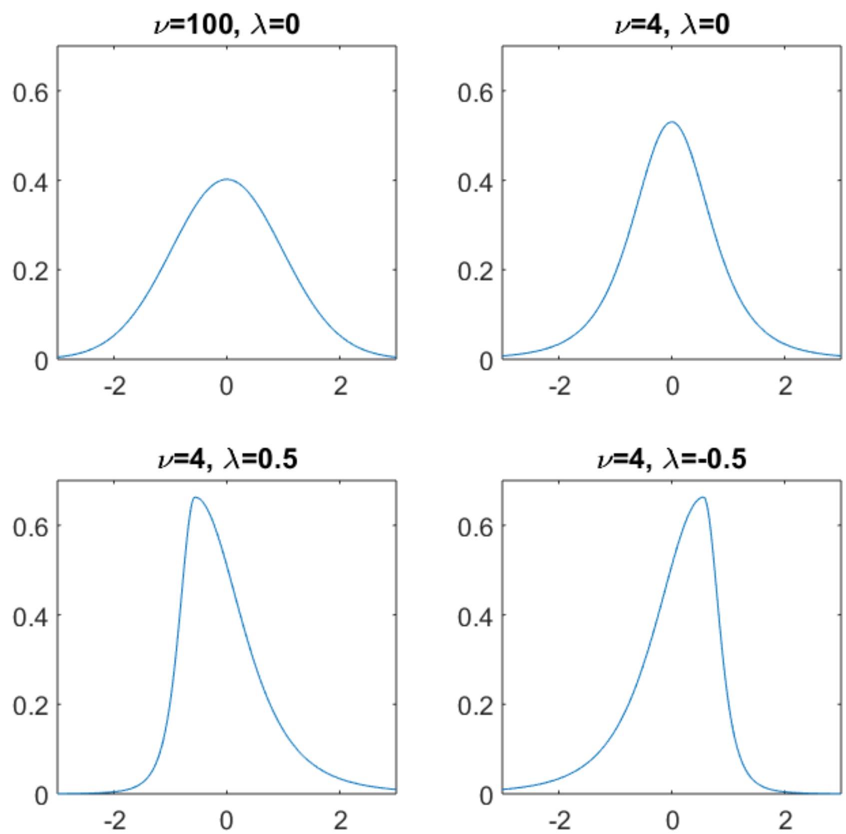
7.2. Skewed t Distribution

The Skewed t distribution, introduced by (Hansen, 1994), has the following density function:

$$f_{\nu,\lambda}(x) = \begin{cases} bc\left(1+\frac{1}{\nu-2}\left(\frac{bx+a}{1-\lambda}\right)^2\right)^{-\frac{v+1}{2}}, & \text{for } x < -\frac{a}{b} \\ bc\left(1+\frac{1}{\nu-2}\left(\frac{bx+a}{1+\lambda}\right)^2\right)^{-\frac{v+1}{2}}, & \text{for } x \geq -\frac{a}{b}, \end{cases}$$

where

Figure 7.17: Skewed $t(\nu, \lambda)$ Density Plots



Note: Skewed $t(\nu, \lambda)$ density plots for different values of λ and ν.

$$a = 4\lambda c \frac{\nu - 2}{\nu - 1}$$

$$b = 1 + 3\lambda^2 - a^2$$

$$c = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi(\nu - 2)}\Gamma\left(\frac{\nu}{2}\right)}$$

and $\Gamma(\cdot)$ is the Gamma function given by

$$\Gamma(t) = \int_0^\infty y^{t-1} e^{-y} dy.$$

The degrees of freedom parameter ν takes values between 2 and ∞ and controls the tail behavior of the distribution. The skeweness parameter λ takes values between -1 and 1 and controles the skewness of the distribution, while $\lambda > 0$ implies right-skewness, $\lambda < 0$ implies left-skewness and $\lambda = 0$ implies symmetry of the distribution. The Skewed Student's t distribution is able to capture asymmetry and heavy-tailedness which is beneficial for analyzing financial data.

Note that the normal and Student's t distribution can be received easely as special cases of the Skewed t distribution by setting $\lambda = 0$ and $\nu = \infty$ for the normal case and $\lambda = 0$ and $\nu \neq \infty$ for the Student's t case. Density plots of the Skewed-t distribution for different combinations of ν and λ can be seen in Figure 7.17.

8. LIST OF TABLES

10. LIST OF SYMBOLS

General notations

$\bar{\mathbb{R}}$	The extended real line $[-\infty, \infty]$
$X \times Y$	The Cartesian product of a space X and Y
$\bar{\mathbb{R}}^N$	The N dimensional space constructed using the Cartesian product $\bar{\mathbb{R}} \times \cdots \times \bar{\mathbb{R}}$
$\mathbf{I}$	The unit interval $[0, 1]$
$\mathbf{I}^N$	The N dimensional space constructed using the Cartesian product $\mathbf{I} \times \cdots \times \mathbf{I}$
A^{-1}	Inverse of a matrix A
A'	The transposed of a matrix A
$\lfloor a \rfloor$	Lower Gaussian brackets of a skalar a
$X \xrightarrow{d} Y$	Convergence in distribution of X to Y
$X \xrightarrow{p} Y$	Convergence in probability of X to Y
$X \xrightarrow{a.s.} Y$	Almost sure convergence of X to Y
$X \overset{d}{\Longrightarrow} Y$	Process convergence in distribution of X to Y
$\mathsf{det}(A)$	Determinant of a matrix A
$\mathsf{exp}(A)$	Exponent of a matrix A
$\mathsf{vec}(A)$	Columns of the matrix A stacked on top of each other in one vector
e_k	The k-th unit vector for the underlying dimension
$O_p(\cdot)$	Denotes a sequence which is bounded in probability
$o_p(\cdot)$	Denotes a sequence which converges to zero in probability
$o_p^*(\cdot)$	Denotes a sequence which converges to zero in probability and is nonnegative
$\boldsymbol{a}$	Denotes an N-dimensional vector $\boldsymbol{a} = (a_1, \ldots, a_N)$

$\boldsymbol{a} \leq \boldsymbol{b}$	If $a_i \leq b_i$ for all entries of $\boldsymbol{a}$ and $\boldsymbol{b}$
$\mathsf{Cov(X,Y)}$	Covariance between the random vectors X and Y
$\mathsf{Var(X)}$	Variance of the random vector X
$\mathbb{E}(X)$	The expectation of a random vector X
$f = \mathcal{O}(a)$	$\frac{f}{a}$ is bounded (Landau-Symbol)
$\mathbb{1}\{\cdot\}$	Equals one if the expression in brackets is true
$\frac{\partial f(x)}{\partial x}_{\vert x=x_0}$	The partial derivative of a function $f(x)$ in x at $x = x_0$
$\sup_{x \in D} f(x)$	Supremum of a function $f(x)$ over all elements $x \in D$
$\inf_{x \in D} f(x)$	Infimum of a function $f(x)$ over all elements $x \in D$
$\max_{x \in D} f(x)$	Maxmimum of a function $f(x)$ over all elements $x \in D$
$\min\{A\}$	Minimum of an amount A
$\arg\min_{x \in D} f(x)$	Determines the point where $f(x)$ reaches its minimum
$\arg\max_{x \in D} f(x)$	Determines the point where $f(x)$ reaches its maximum

Copulas and cdf

$\tilde{C}(u_1, \ldots, u_N)$	The N-dimensional subcopula
$C(u_1, \ldots, u_N)$	The N-dimensional copula
$F(x_1, \ldots, x_N)$	The N-dimensional joint distribution function
$F_i(x_i)$	The i-th marginal distribution function
$\hat{C}_{k:l}(u_1, \ldots u_N)$	The N-dimensional empirical copula using data information from k to l
$\hat{F}_{ti}^{k:l}(X_{ti})$	The i-th empirical distribution function using data information from k to l evaluated at time point t
$C_{ij}(u, v)$	Bivariate copula for inputs u and v
$\hat{C}_{ij}^{(\cdot)}(u, v)$	Empirical bivariate copula, using data for a specific time span for inputs u and v

Dependence measures

$\rho(X_i, X_j)$	Pearson's linear correlation coefficient between X_i and X_j
$\rho_S^{ij}(X_i, X_j)$	Spearman's rank correlation coefficient between X_i and X_j
$\hat{\rho}_S^{ij}(X_i, X_j)/\hat{\rho}^{ij}(X_i, X_j)$	Empirical Spearman's rank correlation coefficient between X_i and X_j
$\tau(X_i, X_j)$	Kendall's tau between X_i and X_j
$\hat{\tau}(X_i, X_j)$	Empirical Kendall's tau between between X_i and X_j
$\lambda_q^{ij}(X_i, X_j)$	Quantile dependence between X_i and X_j for a quantile value $q \in (0, 1)$
$\hat{\lambda}_q^{ij}(X_i, X_j)$	Empirical quantile dependence between X_i and X_j for a quantile value $q \in (0, 1)$

Model and parameters

ϕ_0	Marginal data model parameter vector
$\hat{\phi}$	Estimated marginal data model parameter vector
θ_t	Time varying factor copula model parameter for the time point t
θ_0	True factor copula model parameter under the null
$\hat{\theta}_{(\cdot),S}$	Factor copula model parameter estimate using data for a specific time span and S simulations from the copula model,
α_q	Parameter vector of the idiosyncratic factor q within the factor copula model
γ_k	Parameter vector of the k-th common factor Z_k within the factor copula model
$\beta_{(\cdot)}$	Factor loading paramaters within the factor copula model
$S(i)$	Sorts an asset i of the portfolio to its determinated group
r_{it}	log-return value of asset i at a time point t
$\mu_{it}(\phi_i)$	i-th time varying conditional mean term with data parameter vector ϕ_i at time point t
$\sigma_{it}(\phi_i)$	i-th time varying conditional standard deviation term with

	data parameter vector ϕ_i at time point t
$\eta_{it}(\phi_i)$	i-th residual of the time series model at time point t
$\hat{\eta}_{it}$	i-th sample residual from the inversed pre estimated time series model at time point $t = \{1, \ldots, T\}$
$\tilde{\eta}_{ij}$	i-th simulated residual from the copula model for simulation $j = \{1, \ldots, S\}$
Numbers and scaling	
T	Sample size
S	Number of simulations from the copula model
N	Number of cross sectional dimensions
s	A scalar between $\varepsilon > 0$ and 1, for locating $t = sT$
$\lambda_T(n, m)$	Scaling factor for the empirical processes
$\varepsilon_{T,S}$	Step size used within the numerical derivative
k_g	Number of elements in group g for the block equidependence model
m^*	Number of moments used in the dependence vectors
p	Number of parameters within the copula model
ε	Lower bound of the location scalar s
B	Number of bootstrap replications
$\tilde{s}$	Retro-perspective break fraction estimate
τ_T	Stopping time of the monitoring procedure for a monitored period of size T
m	A scalar, between $\varepsilon > 0$ and 1, determining the size of the initial sample
k	Ratio between S and T
G^*	Number of groups of the block equidependence model
k^*	True unknown break point within the period spanned by

	mT and τ_T (monitoring)
$\hat{k}$	Estimated break point within the period spanned by mT and τ_T(monitoring)
$\hat{s}_a$	Estimated break fraction using the dependence setting a
$\hat{K}_a$	Pivot confidence interval using the dependence setting a
SMM and asymptotics	
$g_{(\cdot),S}(\theta)$	Difference of the data and simulated dependence vectors using data for a specific time span and S simulations from the copula model, with parameter θ
$\hat{m}_{(\cdot)}$	Data dependence vector using data for a specific time span
$\tilde{m}_S(\theta)$	Simulated dependence vector, using S simulations from the copula model and parameter θ
$Q_{(\cdot),S}(\theta)$	Objective function of the SMM procedure using data for a specific time span and S simulations from the copula model with copula parameter θ
W	The positive definite weighting matrix used in the objective function within the SMM procedure
$\hat{W}_{(\cdot)}$	The positive definite weighting matrix estimate used in the objective function within the SMM procedure
G	The derivative matrix of $g_0(\theta)$ in θ_0
$\hat{G}$	The estimated derivative matrix of $g_{(\cdot),S}(\theta)$ in $\theta_{(\cdot),S}$
$A^*(s)$	Limiting distribution of the parameter process for a value s between $\varepsilon > 0$ and 1
$A(s)$	Limiting distribution of the dependence vector process for a value s between $\varepsilon > 0$ and 1

Aas, K., C. Czado, A. Frigessi, and H. Bakken (2009): "Pair-copula constructions of multiple dependence," *Insurance: Mathematics and Economics*, 44, 182–198.

Almeida, C., C. Czado, and H. Manner (2016): "Modeling high-dimensional time-varying dependence using dynamic D-vine models," *Applied Stochastic Models in Business and Industry*, 32, 621–638.

Andrews, D. W. K. (1993): "Tests for parameter instability and structural change with unknown change point," *Econometrica*, 61, 821–856.

Bauwens, L., S. Laurent, and J. V. K. Rombouts (2006): "Multivariate GARCH models: a survey," *Journal of Applied Econometrics*, 21, 79–109.

Bedford, T. and R. Cooke (2002): "Vines - A new graphical model for dependent random variables," *Annals of Statistics*, 30, 1031–1068.

Berens, T., D. Wied, G. Weiss, and D. Ziggel (2014): "A new set of improved Value-at-Risk backtests," *Journal of Banking and Finance*, 48, 29–41.

Brechmann, E. and U. Schlepsmeier (2013): "Modeling dependence with C- and D-Vine copulas: The R Package CDVine," *Journal of Statistical Software*, 52.

Brechmann, E. C. and C. Czado (2013): "Risk management with high-dimensional vine copulas: An analysis of the Euro Stoxx 50," *Statistics & Risk Modeling*, 30, 307–342.

B.Schweizer and A.Sklar (1983): "Probabilistic metric spaces," *Elsevier Science Publishing Co.*

Bücher, A. and I. Kojadinovic (2016): "A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing," *Bernoulli*, 22(2), 927–968.

Bücher, A., I. Kojadinovic, T. Rohmer, and J. Segers (2014): "Detecting changes in cross-sectional dependence in multivariate time series," *Journal of Multivariate Analysis*, 132, 111–128.

Bücher, A. and M. Ruppert (2013): "Consistent testing for a constant copula under strong mixing based on the tapered block multiplier techniqu," *Journal of Multivariate Analysis*, 116, 208–229.

Chen, X. and Y. Fan (2006): "Estimation of copula-based semiparametric time series models," *Journal of Econometrics*, 130(2).

Cherubini, U., E. Luciano, and W. Vecchiato (2004): "Copula methods in finance," *John Wiley and Sons.*

Chollete, L., A. Heinen, and A. Valdesogo (2009): "Modeling international financial returns with a multivariate regime switching copula," *Journal of Financial Econometrics*, 7, 437–480.

Chu, C., M. Stinchcombe, and H. White (1996): "Monitoring structural change," *Econometrica*, 64 (5), 1045-1065.

Creal, D., S. Koopman, and A. Lucas (2013): "Generalized autoregressive score models with applications," *Journal of Applied Econometrics*, 28(5), 777–795.

Creal, D. and R. Tsay (2015): "High-dimensional dynamic stochastic copula models," *Journal of Econometrics*, 189, 335–345.

Czado, C. (2010): "Pair-copula constructions of multivariate copulas," *Workshop on Copula Theory and its Applications, Springer.*

Czado, C., U. Schepsmeier, and A. Min (2012): "Maximum likelihood estimation of mixed C-vine pair copula with application to exchange rates," *Statistical Modelling*, 12(3).

Dehling, H., D. Vogel, M. Wendler, and D. Wied (2017): "Testing for changes in Kendall's tau," *Econometric Theory*, 33, 1352–1386.

Dias, A. and P. Embrechts (2004): "Change-point analysis for dependence structures in finance and insurance," in *Risk Measures for the 21st Century*, ed. by G. Szegoe, Wiley Finance Series, 321–335.

Dowd, K. and D. Blake (2006): "After VaR: The theory, estimation and insurance applications of quantile-based risk measures," *The journal of risk and insurance*, 73, 193–229.

Embrechts, P., F. Lindskog, and A. McNeil (2001): "Modelling dependence with copulas and applications to risk management," *Department of Mathematics, ETHZ.*

Engle, R. F. (2002): "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models," *Journal of Business and Economic Statistics*, 20, 339–350.

Fan, Y. and A. J. Patton (2014): "Copulas in econometrics," *Annual Review of Economics*, 6, 179–200.

Galeano, P. and D. Wied (2013): "Monitoring correlation change in a sequence of random variables," *Journal of Statistical Planning and Inference*, 143, 186–196.

Garcia, R. and G. Tsafack (2011): "Dependence structure and extreme comovements in international equity and bond markets," *Journal of Banking and Finance*, 35, 1954–1970.

Gray, S. (1996): "Modeling the conditional distribution of interest rates as a regime-switching process," *Journal of Financial Economics*, 42, 27–62.

Hafner, C. M. and H. Manner (2012): "Dynamic stochastic copula models: Estimation, inference and applications," *Journal of Applied Econometrics*, 7, 269–295.

Hafner, C. M. and O. Reznikova (2010): "Efficient estimation of a semiparametric dynamic copula model," *Computational Statistics and Data Analysis*, 54, 2609–2627.

Hansen, E. B. (1994): "Autoregressive density estimation," *International Economic Review*, 35, 705–730.

Hoga, Y. and D. Wied (2017): "Sequential monitoring of the tail behavior of dependent data," *Journal of Statistical Planning and Inference*, 182, 29–49.

Hušková, M. and C. Kirch (2008): "Bootstrapping confidence intervals for the change-point of time series," *Journal of time series analysis*, 29, 947–972.

Joe, H. (1996): "Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters," in *Distributions with Fixed Marginals and Related Topics*, ed. by L. Rüschendorf and B. Schweizer and M. D. Taylor.

——— (1997): "Multivariate models and dependence concepts," *Monographs on Statistics and applied probability*, 73.

Kimberling, C. (1974): "A probabilistic interpretation of complete monotonicity," *Aequationes Math.*, 10.

Klaassen, F. (2002): "Improving GARCH volatility forecasts with regime- switching GARCH," *Empirical Economics*, 27, 363–394.

Krupskii, P. and H. Joe (2013): "Factor copula models for multivariate data," *Journal of Multivariate Analysis*, 120, 85–101.

Kurowicka, D. and R. Cooke (2006): *Uncertainty analysis with high dimensional dependence modelling*, Chichester: Wiley.

Kurowicka, D. and H. Joe (2011): *Dependence modeling - Handbook on Vine Copulae*, Singapore: World Scientific Publishing Co.

Kurozumi, E. (2017): "Monitoring parameter constancy with endogenous regressors," *Journal of Time Series Analysis*, 38, 791–805.

Kutzker, T., F. Stark, and D. Wied (2018): "A non-parametric CUSUM-type test for testing relevant change in copulas," .

Manganelli, S. and R. Engle (2004): "CAViaR: Conditional autoregressive value at risk by regression quantiles," *Journal of Business and Economic Statistics*, 22, 4.

Manner, H., F. Stark, and D. Wied (2018): "A monitoring procedure for detecting structural breaks in factor copula models," .

——— (2019): "Testing for structural breaks in factor copula models," *Journal of Econometrics*, 208(2), 324–345.

McNeil, A. and R. Frey (2000): "Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach," *Journal of Empirical Finance*, 7, 270–300.

McNeil, A. and J. Nešlehová (2009): "Multivariate archimedian copulas, d-monotone functions and l_1- Norm symmetric distributions," *The Annals of Statistics*, 3059–3097.

McNeil, A., R.Frey, and P.Embrechts (2010): "Quantitative risk management: concepts techniques and tools," *Princeton university press*.

Murry, J. S., D. B. Dunson, L. Carin, and J. E. Lucas (2013): "Bayesian gaussian copula factor models for mixed data," *Journal of the American Statistical Association*, 108, 656–665.

NELSON, R. B. (2006): "An introduction to copulas," *Springer*, Second Edition.

NEWEY, W. AND D. MCFADDEN (1994): "Large sample estimation and hypothesis testing," *Handbook of Econometrics*, 4.

OH, D. AND A. PATTON (2013): "Simulated method of moments estimation for copula-based multivariate models," *Journal of the American Statistical Association*, 108, 689–700.

——— (2017): "Modelling dependence in high dimensions with factor copulas," *Journal of Business and Economic Statistics*, 35, 139–154.

OH, D. H. AND A. J. PATTON (2018): "Time-varying systemic risk: Evidence from a dynamic copula model of CDS spreads," *Journal of Business and Economic Statistics*, 36, 181–195.

OKHRIN, O., Y. OKHRIN, AND W. SCHMID (2013): "On the structure and estimation of hierarchical Archimedean copulas," *Journal of Econometrics*, 173, 189–204.

PAPE, K., D. WIED, AND P. GALEANO (2017): "Monitoring multivariate variance changes," *Journal of Empirical Finance*, 39(A), 54–68.

PATTON, A. J. (2006): "Modelling asymmetric exchange rate dependence," *International Economic Review*, 47, 527–556.

——— (2012): "A Review of copula models for economic time series," *Journal of Multivariate Analysis*, 110, 4–18.

QU, Z. AND P. PERRON (2007): "Estimating and testing structural changes in multivariate regressions," *Econometrica*, 75, 459–502.

REMILLARD, B. (2017): "Goodness-of-Fit tests for copulas of multivariate time series," *Econometrics*, 35, 139–154.

Savu, C. and M. Trede (2010): "Hierarchies of Archimedean copulas," *Quantitative Finance*, 10, 295–304.

Stark, F. (2018): "On the applicability of a nonparametric test for constant copula-based dependence measures: Dating break points and analyzing different dependence measure sets," .

Stöber, J. and C. Czado (2011): "Pair copula constructions," in *Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications*, ed. by M. Scherer and J.-F. Mai, Singapore: World Scientific Publishing, 187–232.

Stöber, J. and C. Czado (2014): "Regime switches in the dependence structure of multidimensional financial data," *Computational Statistics & Data Analysis*, 76, 672–686.

Stöber, J., H. Joe, and C. Czado (2013): "Simplified pair copula constructions - Limitations and extensions," *Journal of Multivariate Analysis*, 119, 101–118.

Wied, D. (2013): "CUSUM-type testing for changing parameters in an autoregressive model for stock returns," *Journal of Time Series Analysis*, 34, 221–229.

Wied, D., H. Dehling, M. van Kampen, and D. McFadden (2013): "A fluctuation test for constant Spearmans's rho with nuisance-free limit distribution," *Computational Statistics and Data Analysis*, 76, 723–736.

Yu, J. and R. Meyer (2006): "Multivariate stochastic volatility models: Bayesian estimation and model comparison," *Econometric Reviews*, 25, 361–384.

12. EIDESSTATTLICHE VERSICHERUNG

Hiermit versichere ich an Eides Statt, dass ich die vorgelegte Dissertation mit dem Titel

Detecting Structural Breaks in Factor Copula Models and in Vectors of Dependence Measures

selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Aussagen, Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise entgeltlich/unentgeltlich geholfen:
Prof. Dr. Hans Manner und Prof. Dr. Dominik Wied (gemeinsame Projektarbeit für die Inhalte in Kapitel 3 und Kapitel 4, finanziert durch die Deutsche Forschungsgemeinschaft (DFG)).
Weitere Personen waren an der inhaltlichmateriellen Erstellung der vorliegenden Dissertation nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die Dissertation wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt. Ich versichere, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe

Datum, Ort | Unterschrift