Noname manuscript No.
(will be inserted by the editor)

Testing for Relevant Dependence Change in Financial Data: A CUSUM Copula Approach

Tim Kutzker · Florian Stark · Dominik Wied

Received: date / Accepted: date

Abstract We propose a new non-parametric test for detecting relevant breaks in copula functions. We assume that the data is driven by two non-equal copulas C_1 and C_2. Under the null hypothesis, the copula difference within an appropriate norm is smaller than a certain positive adjustable threshold Δ. Within the alternative hypothesis, the copula difference exceeds the fixed value Δ. The test is based on a cumulative sum approach of the empirical copula with sequentially estimated marginals. We propose a bootstrap procedure to compute critical values. The Monte Carlo simulation indicates that the test results in a reasonable sized and powered testing procedure. A real data application of the DAX30 up to cross sectional dimension $N = 30$ shows the test's ability to detect relevant break points.

Keywords: Relevant change, Copula, Break testing, Bootstrap, CUSUM
JEL codes: C12, C13, C32

Research is supported by Deutsche Forschungsgemeinschaft MA 7225/1-1, AOBJ 628937 (DFG grant "Strukturbrüche und Zeitvariation in hochdimensionalen Abhängigkeitsstrukturen")

Tim Kutzker
Universitätsstr. 24, 50937 Cologne, Germany
E-mail: tim.kutzker@uni-koeln.de

Florian Stark, *corresponding author*
Universitätsstr. 24, 50937 Cologne, Germany
E-mail: fstark3@uni-koeln.de

Dominik Wied
Universitätsstr. 24, 50937 Cologne, Germany
E-mail: dwied@uni-koeln.de

1 Introduction

It is well known that dependencies within a portfolio increase in times of financial crisis (cf. Aloui et al. (2011)). From a portfolio manager point of view the increase of the dependencies is disadvantageous, which is known as the diversification effect. In fact, investors are interested in decreasing the dependencies by rescheduling the portfolio to lower the risk of losses. One of many approaches to detect those changes in the dependence structure is to test for changes in the copula function. For instance, Busetti and Harvey (2011), Brodsky et al. (2009) and Krämer and van Kampen (2011) have designed nonparametric tests for breaks in the copula in a fixed point considering N-dimensional random vectors. Bücher and Ruppert (2013) extended their approaches by testing for overall constancy of the copula in the case of known marginal distributions, while the test of constancy suggested in Bücher et al. (2014) considers sequentially estimated marginals. Wied et al. (2013) propose a test for changes in Spearman's rho, Dehling et al. (2017) consider a test for changes in Kendall's tau. Manner et al. (2019) construct a parametric test for detecting breaks in the parameters of factor copula models. The above mentioned tests can be applied to detect and quantify contagions between different financial markets or to construct optimal portfolios.

All the proposed methods test for the "classical" hypothesis, meaning that they test for stationarity in a sequence of random vectors $\{\mathbf{X}_j\}_{j=1}^T$ with $\mathbf{X}_j \in \mathbb{R}^N$, i.e.

$$H_0 : \mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_T \sim F.$$

with the alternative in the simplest case of one structural breakpoint in time (cf. Dette and Wied (2016))

$$H_1 : \mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_j \sim F_1 \text{ and } \mathbf{X}_{j+1}, \ldots, \mathbf{X}_T \sim F_2,$$

where the distribution function changes from F_1 to F_2 with $F_1 \overset{d}{\neq} F_2$ at time $j \in \{1, \ldots T\}$, i.e. F_1 and F_2 are not equal in distribution.
A general issue of such hypothesis testing is the consistency problem, i.e. any consistent test will detect any arbitrary small change in the parameters if the sample size is sufficiently large. This discrepancy was mentioned for the first time in 1938 by Berkson (1938).

Beyond that, in the case of small changes the rejection of the null might result in an unnecessary break point estimation and an expensive adjustment of the considered model. In practice, small changes in the data might not be crucial, since they do not necessarily add up to significant changes. Thus, the gain derived by the detected break point could be negatively overcompensated by the costs of adjusting the model (e.g. in case of portfolio theory these can be interpreted as transaction costs) or to be short and to the point: Significance does not necessarily imply relevance.

Therefore, we impose the more realistic assumption that our sequence of random vectors $\{\mathbf{X}_j\}_{j=1}^T$ with $\mathbf{X}_j \in \mathbb{R}^N$ with $j \in \{1, \ldots T\}$ is driven by the distribution function F_1 and F_2, i.e. $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_j \sim F_1$ and $\mathbf{X}_{j+1}, \ldots, \mathbf{X}_T \sim F_2$ for some $j \in \{1, \ldots T\}$ such that

$$H_0 : \|F_1 - F_2\| \leq \Delta \quad \text{versus} \quad H_1 : \|F_1 - F_2\| > \Delta, \tag{1}$$

where $\|.\|$ is an appropriate norm and $\Delta > 0$ a fixed adjustable size. The framework in (1) allows for a break in the data (classical break point tests do not) and the adjustable size Δ could serve as a measure to control for the extent of the change.

Dette and Wied (2016) proposed a general approach to this problem. Later on, Dette et al. (2018) and Dette and Gösmann (2017) extended this to the detection of changes in second-order characteristics and to high-dimensional models, respectively. Motivated by their analysis, we are interested in augmenting the literature of testing for relevant breaks in the copula of random vectors by a nonparametric testing procedure that detects relevant changes in the copula function with sequentially estimated marginal distributions. Thus, the testing problem is given by

$$H_0 : \|C_1 - C_2\| \leq \Delta \quad \text{versus} \quad H_1 : \|C_1 - C_2\| > \Delta,$$

where $C_1 \neq C_2$ are copulas and $\Delta > 0$ fixed. As the copula measures the dependence between random variables, we therefore test whether the dependence structure changes more than some given threshold Δ.

Coming back to portfolio management, a small increase in the dependence structure of a portfolio does not necessarily indicate the need to reschedule the portfolio, since transaction costs could overcompensate the benefits of the new, more risk diversified portfolio. Only a relevant change in the dependence structure, i.e. the copula difference within a certain norm is larger than Δ, should result in rescheduling the portfolio.

In our empirical application we analyzed the German DAX30 data of cross sectional dimension $N = 30$ between January 2003 and July 2015. Here, Δ could be interpreted as the largest admissible copula difference such that the relevant change hypothesis is not rejected. Every other choice of Δ that is smaller leads to a rejection of the null hypothesis. As a result, Δ can also be considered as a measure that quantifies the extent of a crisis (given that dependencies of financial returns are usually larger in times of crises).

The rest of the paper is structured as follows. Section 2 introduces the considered null hypothesis and test statistic, where Section 3 presents the bootstrap procedure to determine critical values to perform the test. Results from the Monte Carlo simulations can be found in Section 4. Section 5 presents our empirical application and Section 6 concludes. A Supplemental Appendix provides theoretical background.

2 Relevant change and test statistic

In this section we introduce the null hypothesis, the assumptions and the the relevant change characteristic of our testing procedure in a fully non-parametric setting.
Let $\mathbf{X}_1, \ldots, \mathbf{X}_T$ denote N-dimensional random vectors and $\mathbf{U}_1, \ldots, \mathbf{U}_T$ the vector of the marginal distributions, i.e. $\mathbf{U}_t := (F_1(X_{t1}), \ldots, F_N(X_{tN}))$ for $t = 1, ..., T$ where $F_i(\cdot)$ is the i-th marginal such that

$$\begin{aligned} \mathbf{U}_1, \ldots, \mathbf{U}_{\lfloor sT \rfloor} &\sim C_1(\mathbf{u}) \\ \mathbf{U}_{\lfloor sT \rfloor+1}, \ldots, \mathbf{U}_T &\sim C_2(\mathbf{u}), \end{aligned} \tag{2}$$

where $\mathbf{u} \in [0,1]^N$ and $C_1, C_2 : [0,1]^N \to [0,1]$ are copulas which capture the dependencies between the components of $\mathbf{X}_1, ..., \mathbf{X}_{\lfloor sT \rfloor}$ and $\mathbf{X}_{\lfloor sT \rfloor+1}, ..., \mathbf{X}_T$, respectively. Here, $\lfloor sT \rfloor$ denotes the change point in time, where T is the size of the sample and $s \in (0,1)$. Note, that the model set-up (2) is valid under both the null and the alternative hypothesis. In order to achieve reliable results, classical concepts of dependencies (e.g. $\mathbf{U}_1, ..., \mathbf{U}_T$ is stationary and strong mixing with coefficients converging sufficiently fast to 0) are not applicable any more in the setting of detecting relevant changepoints, because the general model set-up (2) of relevant changepoint analysis allows the sequence $\mathbf{U}_1, ..., \mathbf{U}_T$ to be non-stationary. That is why we have to impose the assumption of a triangular array, that is α-mixing[1].
To aggregate over $\mathbf{u}$, we consider the L^2-norm $\|\cdot\|_{L^2}$. Thus, the null hypothesis of no relevant change in the copula function is given by

$$H_0 : \|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2} \leq \Delta$$

versus the alternative

$$H_1 : \|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2} > \Delta,$$

where $\|.\|_{L^2}$ is the L^2-norm and $\Delta > 0$ fixed. For every $\mathbf{u} \in [0,1]^N$ and $t \in (0,1)$ the cumulative sum (CUSUM) type process for detecting changes in the copula is then

$$\hat{\mathbb{U}}^*_T(t, \mathbf{u}) := t(1-t)\left(\frac{1}{\lfloor tT \rfloor}\sum_{i=1}^{\lfloor tT \rfloor} Z_i^{1:\lfloor tT \rfloor}(\mathbf{u}) - \frac{1}{T - \lfloor tT \rfloor}\sum_{i=\lfloor tT \rfloor+1}^{T} Z_i^{\lfloor tT \rfloor+1:T}(\mathbf{u})\right), \tag{3}$$

where $Z_i^{t_1:t_2}(\mathbf{u}) := \mathbb{1}\{\hat{F}_1^{t_1:t_2}(X_{i1}) \leq u_1, \ldots, \hat{F}_N^{t_1:t_2}(X_{iN}) \leq u_N\}$, $t_1 < t_2 \in \{1, \ldots, T\}$ for $i = 1, \ldots, T$ and $\mathbb{1}\{\cdot\}$ the indicator function. Here $\hat{F}_j^{t_1:t_2}(\cdot)$ is the empirical distribution function, using data information between t_1 and t_2 and is defined as

$$\hat{F}_j^{t_1:t_2}(x) := \frac{1}{t_2 - t_1 + 1}\sum_{i=t_1}^{t_2} \mathbb{1}\{X_{ij} \leq x\}, \ j = 1, ..., N.$$

[1] Due to the fact that this discussion is very technical, we shifted the details to the Supplemental Appendix.

For the derivation of our testing procedure we now consider $\hat{\mathbb{U}}_T(t, \mathbf{u})$ defined as

$$\hat{\mathbb{U}}_T(t, \mathbf{u}) := t(1-t) \left(\frac{1}{\lfloor tT \rfloor} \sum_{i=1}^{\lfloor tT \rfloor} Z_i(\mathbf{u}) - \frac{1}{T - \lfloor tT \rfloor} \sum_{i=\lfloor tT \rfloor + 1}^{T} Z_i(\mathbf{u}) \right), \tag{4}$$

where $Z_i(\mathbf{u}) := \mathbb{1}\{F_1(X_{i1}) \leq u_1, \ldots, F_N(X_{iN}) \leq u_N\}$, $t_1 < t_2 \in \{1, \ldots, T\}$ with F_i as known marginals, $i = 1, \ldots, T$. For fixed $s \in (0,1)$, a straightforward calculation yields[2]

$$\lim_{T\to\infty} \mathbb{E}[\hat{\mathbb{U}}_T(t, \mathbf{u})] = \begin{cases} s(1-t)\left(C_1(\mathbf{u}) - C_2(\mathbf{u})\right), & s \leq t \\ t(1-s)\left(C_1(\mathbf{u}) - C_2(\mathbf{u})\right), & s > t, \end{cases} \tag{5}$$

where we have to distinguish between data before and after the breakpoint $\lfloor sT \rfloor$. In the next step, we want to eliminate the quantile and time dimension $\mathbf{u}$ and t, respectively. For this purpose, we consider the L^2-norm and obtain

$$L(t) := \lim_{T\to\infty} \mathbb{E}[\|\hat{\mathbb{U}}_T(t, \mathbf{u})\|_{L^2}^2] = \begin{cases} s^2(1-t)^2 \|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}^2, & t > s \\ (1-s)^2 t^2 \|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}^2, & t \leq s, \end{cases}$$

for every norm of the type $\|f(\cdot, \mathbf{u})\|_{L^2}^2 := \int_{[0,1]^N} f(\cdot, \mathbf{u})^2 d\mathbf{u}$. Integrating out t yields

$$\int_0^1 L(t)dt = \frac{s^2(1-s)^2}{3} \|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}^2. \tag{6}$$

Thus, integrating out t from the empirical counterpart $\hat{L}_T(t) := \|\hat{\mathbb{U}}_T(t, \mathbf{u})\|_{L^2}^2$ yields the test statistic $\hat{\kappa}_T$ for the initial problem of detecting the relevant change

$$\hat{\kappa}_T := \int_0^1 \hat{L}_T(t)dt. \tag{7}$$

Due to the fact that our test statistic mainly consists of an integral over t, we disregard the possibility to trim the sample in some way, as it is sometimes done in the breakpoint literature.

We use $\hat{s} := \underset{s\in(0,1)}{\operatorname{argmax}} \|\hat{\mathbb{U}}_T(s, \mathbf{u})\|_{L^2}$ as the natural argmax estimator for the changepoint location fraction s[3]. We reject the null hypothesis of no relevant change if the test statistic (7) less the adjusted centering $\frac{s^2(1-s)^2}{3} \|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}^2$ deviates too far from zero. If the marginal distributions are known the limiting distribution of the process

$$\sqrt{T} \left(\int_0^1 \hat{L}_T(t)dt - \frac{s^2(1-s)^2}{3} \|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}^2 \right), \tag{8}$$

is normal which is shown in the Supplemental Appendix.

Due to the high computational effort in high dimensions using the L^2-norm it could be reasonable to only test for specific points $\mathbf{q}$ in the copula, e.g. $\mathbf{q}$ could be chosen as the value that maximizes the copula difference, i.e. $\mathbf{q} := \sup_{\mathbf{u}\in[0,1]^N} |C_1(\mathbf{u}) - C_2(\mathbf{u})|$. For this purpose we fix $\mathbf{q} = (q_1, \ldots, q_N)'$. What we call quantile counterpart of the process (8) is then given by

$$\sqrt{T} \left(\int_0^1 \hat{L}_T^{\mathbf{q}}(t)dt - \frac{1}{3} s^2(1-s)^2 (C_1(\mathbf{q}) - C_2(\mathbf{q}))^2 \right), \tag{9}$$

[2] For the very detailed derivation of the testing procedure we refer to the Supplemental Appendix.

[3] Note, $\hat{s}$ is a superconsistent estimator of the changepoint fraction s with convergence rate T (cf. Dette and Wied (2016)).

where $\hat{L}_T(t)$ from (8) is replaced by its quantile version $\hat{L}_T^{\mathbf{q}}(t) := (\hat{\mathbb{U}}_T(s, \mathbf{q}))^2$ for $\mathbf{q} \in [0,1]^N$ fixed. Accordingly, the test statistic $\hat{\kappa}_T^{\mathbf{q}}$ is then defined as

$$\hat{\kappa}_T^{\mathbf{q}} := \int_0^1 \hat{L}_T^{\mathbf{q}}(t)dt. \tag{10}$$

Since the limit distributions of the processes (8) and (9) are not known in case of unknown marginals, we suggest a bootstrap procedure. The null hypothesis will be rejected if the expression in (8) or (9) is greater than the value of the corresponding quantile, which can be obtained by applying the bootstrap procedure presented in Section 3. The test holds the size level if the fixed adjustable threshold Δ is chosen as $\|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}$ or for the quantile case $|C_1(\mathbf{u}) - C_2(\mathbf{u})|$. For Δ smaller than this threshold the test is oversized while a larger Δ results in a lower rejection rate. In the application later on, we set $\mathbf{q} = 0.6 \cdot (1, \dots, 1)$, which is in line with our Monte Carlo simulations. An Δ chosen in this way can be used, for example, to assess the extent of a crisis.

Our Monte Carlo simulations below confirm that the bootstrap results in a reasonably sized and powered testing procedure. For the bootstrap we consider the L^2-norm, but this can be easily adjusted to the quantile version simply by interchanging the L^2-norm with the absolute value $|\cdot|$ for fixed $\mathbf{q} \in [0,1]^N$.

3 Bootstrap and Testing procedure

The bootstrap is based on the natural estimators of the respective terms of the process (8) or (9). We assume that our sample $\{\mathbf{X}_i\}_{i=1}^T$ is serially independently distributed or residual data from pre-estimated time series models e.g. GARCH adjusted data. Further, $\{\mathbf{X}_i\}_{i=1}^T$ is compounded of $\{\mathbf{X}_i\}_{i=1}^{\lfloor sT \rfloor}$ and $\{\mathbf{X}_i\}_{i=\lfloor sT \rfloor+1}^T$, such that there is only one breakpoint location in $\lfloor sT \rfloor$, $s \in (0,1)$ and $\{\mathbf{X}_i\}_{i=1}^{\lfloor sT \rfloor} \sim C_1(F(\mathbf{X}))$ and $\{\mathbf{X}_i\}_{i=\lfloor sT \rfloor+1}^T \sim C_2(F(\mathbf{X}))$. Then, the bootstrap procedure suggests the following course of action:

i) Estimate the breakpoint location $\lfloor sT \rfloor$ by $\lfloor \hat{s}T \rfloor$, where $\hat{s}$ is determined by

$$\hat{s} := \underset{s \in (0,1)}{\operatorname{argmax}} \|\hat{\mathbb{U}}_T^*(s, \mathbf{u})\|_{L^2}. \tag{11}$$

Sample separately with replacement from $\{\mathbf{X}_i\}_{i=1}^{\lfloor \hat{s}T \rfloor}$ and $\{\mathbf{X}_i\}_{i=\lfloor \hat{s}T \rfloor+1}^T$ to obtain B bootstrap samples $\{\mathbf{X}_i^{(b)}\}_{i=1}^T$, for $b = 1, \dots, B$.

ii) Estimate the break point location $\lfloor \hat{s}_b T \rfloor$ for each bootstrap sample $\{\mathbf{X}_i^{(b)}\}_{i=1}^T$, for $b = 1, \dots, B$, using adjusted (11).

iii) Estimate the copula difference $\Delta_C^b = \|\hat{C}_b^{1:\lfloor \hat{s}_b T \rfloor}(\mathbf{u}) - \hat{C}_b^{\lfloor \hat{s}_b T \rfloor+1:T}(\mathbf{u})\|_{L^2}$ for each bootstrap sample $\{\mathbf{X}_i^{(b)}\}_{i=1}^T$, for $b = 1, \dots, B$, where $\hat{C}_b^{t_1:t_2}$ is the empirical copula estimate with sequentially estimated marginals, using the data from t_1 to t_2.

iii) Calculate the bootstrap versions of the centred expressions (8) or (9)

$$K^{(b)} := \sqrt{T} \left(\int_0^1 \hat{L}_T^{*b}(t)dt - \frac{1}{3}\hat{s}_b^2(1-\hat{s}_b)^2 \Delta_C^b \right),$$

with $\hat{L}_T^{*b}(t) := \|\hat{\mathbb{U}}_T^{*b}(s, \mathbf{u})\|_{L^2}^2$, where $\hat{\mathbb{U}}_T^{*b}(s, \mathbf{u})$ is the bootstrap analogue of (3), using $\{\mathbf{X}_i^{(b)}\}_{i=1}^T$.

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

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

where $q \in (0,1)$.

With the above described bootstrap procedure we can calculate critical values for (8) and (9). The testing procedure is as follows: We reject the null of no relevant change $\|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2} \leq \Delta$ if

$$\hat{\kappa}_T > \frac{\hat{s}^2(1-\hat{s})^2}{3}\Delta^2 + \frac{b_{1-\alpha}}{\sqrt{T}}, \tag{12}$$

where $b_{1-\alpha}$ is the $1-\alpha$ quantile of the bootstrap distribution. Note that the critical values obtained by the bootstrap remain stochastically bounded both under the null and the alternative hypothesis, as the test statistic is always correctly centered.
The bootstrap and testing procedure can be easily adapted for the quantile case by adapting step i) - iii). The test given in equation (12) is an exact level α test if Δ is chosen as the copula difference $||C_1(\mathbf{u}) - C_2(\mathbf{u})||_{L^2}$ or $|C_1(\mathbf{q}) - C_2(\mathbf{q})|$. Otherwise, the size is smaller than α. In particular, $\hat{\kappa}_T$ converges weakly to a degenerated random variable if the copula difference is equal to zero and the Davies problem is present, i.e. the break point is unidentified under the null hypothesis. Consequently, the level of the proposed tests have practically size zero, whereas classical stationarity tests hold the asymptotic α-level. Thus, the power of the classical tests is usually larger than the power of the relevant change tests considered here. For practitioners we suggest to run a classical test first, e.g. Bücher and Ruppert (2013) for the case of known marginals and Bücher et al. (2014) in the case of sequentially estimated marginals. If the test rejects the null of stationarity, i.e. the copula difference is significantly larger than zero, estimate the break fraction and apply the proposed relevant change test. This two-step procedure has the drawback, however, that the statistical properties are not clear.

4 Quantile- and L^2-Simulations

In this section we want to analyze finite sample properties of our proposed relevant testing procedure, where we simulate multivariate data up to dimension $N = 30$ using a factor copula model following Oh and Patton (2017). We consider both serially independently distributed and residual data.

4.1 Serially Independently Distributed Data

In this subsection we conduct two major Monte Carlo simulations. First, we consider the following simple DGP

$$\mathbf{X}_t = [X_{1t}, X_{2t}]' = N_2(\mathbf{0}, \Sigma_t(\rho)), \tag{13}$$

where $N_2(\mathbf{0}, \Sigma_t(\rho))$ with $t = 1, ..., T$ describes the bivariate normal distribution with expectation vector zero and covariance matrix $\Sigma_t(\rho) = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ and $\rho \in [-1, 1]$. We set ρ equal to -0.3 for $t = 1, ..., \frac{T}{2}$ and $\rho = 0.8$ for $t = \frac{T}{2} + 1, ..., T$. Thus, the breakpoint $\lfloor sT \rfloor$ is chosen at $\frac{T}{2}$. We restrict the size analysis in this subsection to the two dimensional case $N = 2$. The following size study presents both L^2-norm based results and an analysis where we consider the specific point $\mathbf{q} = (0.6, 0.6)$. Note, the closer the quantile is to its boundaries, i.e. 0 or 1, the more observations are needed. Critical values of our tests are computed using the bootstrap algorithms from Sections 3 with $B = 300$ bootstrap replications. The tests are performed at the $\alpha = 0.05, 0.1$ significance level using 301 Monte Carlo replications. 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).
Table 1 presents the results of the relevant change tests under the null with Δ chosen as the estimated copula difference $|C_1(\mathbf{q}) - C_2(\mathbf{q})|$, where C_1 and C_2 are estimated by the consistent copula estimator

$$\hat{C}(\mathbf{u}) = \frac{1}{t_2 - t_1} \sum_{i=t_1}^{t_2} \mathbb{1}\{\hat{F}_1^{t_1:t_2}(X_{i1}) \leq u_1, \ldots, \hat{F}_N^{t_1:t_2}(X_{iN}) \leq u_N\}, \tag{14}$$

using realizations $\{\mathbf{X}_1, \ldots \mathbf{X}_{\lfloor \hat{s}T \rfloor}\}$ and $\{\mathbf{X}_{\lfloor \hat{s}T \rfloor + 1}, \ldots \mathbf{X}_T\}$. The breakpoint $\lfloor \hat{s}T \rfloor$ is estimated by

$$\hat{s} := \underset{s \in (0,1)}{\operatorname{argmax}} |\hat{\mathbb{U}}_T(s, \mathbf{q})|. \tag{15}$$

Table 2 reports the results of the relevant change tests under the null, where the functional difference between the copulas is determined by the L^2-norm. Similar to the quantile case, we consider for the size analysis $\mathit{\Delta} := \|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}$ and accordingly $\hat{s} := \underset{s\in(0,1)}{\operatorname{argmax}} \|\hat{\mathbb{U}}_T(s, \mathbf{u})\|_{L^2}$. Overall, the tests show good size properties and converges to the predetermined rejection level α if T gets larger. For smaller T, the differences to α are slightly, but not dramatically larger and it does not appear necessary to consider size corrections. For the power analysis, we use the quantile based test and consider two different scenarios.

Table 1: Size using quantile version

	$T = 300$	$T = 500$	$T = 750$	$T = 1000$
q_{95}	0.06	0.06	0.04	0.05
q_{90}	0.12	0.11	0.10	0.10

Table 1 reports the rejection rate of the relevant change test for data generated with the DGP described in (13) using $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{q} = (0.6, 0.6)$. In total, we conducted 301 Monte Carlo replications.

Table 2: Size using the L^2-norm

	$T = 300$	$T = 500$	$T = 750$	$T = 1000$
q_{95}	0.06	0.06	0.04	0.06
q_{90}	0.11	0.11	0.10	0.12

Table 2 reports the rejection rate of the relevant change test for data generated with the DGP described in (13) using $B = 300$ bootstrap replications. The copula difference is determined using the L^2-norm. In total, we conducted 301 Monte Carlo replications.

In the first scenario we keep $\mathit{\Delta}$ fix and vary ρ in the DGP (13). In the second scenario we vary $\mathit{\Delta}$ and keep the DGP (13) fixed.

The upper panel of Table 3 depicts the first scenario. In this case we determine $\mathit{\Delta}_0$ as the copula difference at the point $\mathbf{q} = (0.6, 0.6)$ generated by the DGP (13) with $\rho = -0.3$ before and $\rho = 0.8$ after the break point at $\frac{T}{2}$. We now vary $\rho \in \{-0.4, -0.5, -0.6, -0.7\}$ before the break point and the results of the rejection rate can be seen in the upper panel of table 3 for different sample sizes.

The lower panel of table 3 depicts the second scenario. After determining the quantile value under the null, we decrease the tolerance $\mathit{\Delta}$ in the test (12) by $\mathit{\Delta} = d \cdot \mathit{\Delta}_0$ with $d \in \{0.95, 0.9, 0.85, 0.8\}$.

Note, that in both cases the rejection rate of the relevant change test holds the size level α and the rejection rate tends to 1 for increasing sample size T and decreasing d or ρ. This is the expected behavior as the null and alternative hypothesis differ the more the smaller d or ρ are. In the second major MC simulation, we consider our data to be jointly distributed with a one factor copula model following Oh and Patton (2017), where the marginal distributions are in general unknown and the copula is implied by the following factor structure

$$\mathbf{X}_t = [X_{1t}, \dots, X_{Nt}]' = \boldsymbol{\beta}_t Z + \boldsymbol{q}, \tag{16}$$

with $\boldsymbol{\beta}_t = \beta_t \cdot (1 \dots, 1)'$ is a parameter vector of size N, $Z \overset{i.i.d.}{\sim} \mathsf{Skew\ t}\left(\nu^{-1}, \lambda\right)^4$ and $\boldsymbol{q} = [q_{1t}, \dots, q_{Nt}]'$ with $q_{it} \overset{i.i.d.}{\sim} \mathsf{t}\left(\nu^{-1}\right)$ for $i = 1, \dots, N$ and $t = 1, \dots, T$. 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 $t = 1, \dots, T$. The DGP in (16) provides left skewed and fat tailed data, which is a common property in financial data applications and also in

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

Table 3: Power

Power Analysis varying ρ					
	$\rho = -0.3$	$\rho = -0.4$	$\rho = -0.5$	$\rho = -0.6$	$\rho = -0.7$
$T = 300$	0.06	0.48	0.83	0.98	1.00
$T = 500$	0.06	0.62	0.94	1.00	1.00
$T = 750$	0.04	0.69	0.98	1.00	1.00
$T = 1000$	0.05	0.84	0.99	1.00	1.00
Power Analysis varying Δ					
	$\Delta = \Delta_0$	$\Delta = 0.95 \cdot \Delta_0$	$\Delta = 0.9 \cdot \Delta_0$	$\Delta = 0.85 \cdot \Delta_0$	$\Delta = 0.8 \cdot \Delta_0$
$T = 300$	0.06	0.31	0.63	0.88	0.97
$T = 500$	0.06	0.33	0.75	0.94	1.00
$T = 750$	0.04	0.36	0.83	0.96	1.00
$T = 1000$	0.05	0.42	0.86	1.00	1.00

Table 3 reports the rejection rate of the quantile relevant change test for data generated with the DGP described in (13) using $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{q} = (0.6, 0.6)$. Varying $\Delta = d \cdot \Delta_0$, where $d = \{0.95, 0.9, 0.85, 0.8\}$ (lower panel) and $\rho = \{-0.4, -0.5, -0.6, -0.7\}$ in $\Sigma_t(\rho)$ (upper panel) for $t = \lfloor \frac{T}{2} \rfloor + 1, ..., T$. In total, we conducted 301 Monte Carlo repetitions.

line with our application in Section 5. We construct a break at $\frac{T}{2}$, where θ_0 denotes the parameter value of the model before and θ_1 the parameter value after the break. For our simulation study we choose $\theta_0 = 1$ and $\theta_1 = 2$. Note again, the test is an exact level α test if and only if Δ is chosen as the copula difference. Table 4 reports the results of the relevant change test under the null, where the functional difference is computed with the help of the L^2-norm. Table 4 shows, that the test using the proposed bootstrap procedure holds the size level using the DGP (16). As expected, the size converges to the corresponding rejection level $\alpha \in \{0.05, 0.1\}$ as T gets larger. This characteristic also holds for Table 5.

Table 4: Size using the L^2-norm

		$T = 300$	$T = 500$	$T = 750$	$T = 1000$
$N = 2$	q_{95}	0.03	0.04	0.04	0.04
	q_{90}	0.07	0.08	0.11	0.11
$N = 3$	q_{95}	0.03	0.05	0.04	0.05
	q_{90}	0.06	0.10	0.12	0.10
$N = 5$	q_{95}	0.02	0.04	0.07	0.05
	q_{90}	0.04	0.09	0.13	0.09

Table 4 shows the rejection rate of the relevant change test for the DGP (16) using $B = 300$ bootstrap replication. In total, we conducted 301 Monte Carlo repetitions.

In this case, we set Δ equal to the copula difference evaluated at the specific point $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)$, where $(1, \ldots, 1)'$ is a N-dimensional vector. The experiment is repeated in Table 6 for $\mathbf{q} = 0.1 \cdot (1, \ldots, 1)$. Considering such particular quantiles provides the advantage to conduct high dimensional data analysis with comparatively moderate computational efforts. Thus, the relevant change test is especially suitable for high dimensional data applications. In practice, for instance, the specific quantile $\mathbf{q}$ can be chosen as the quantile that maximizes the copula difference. The simulations show that a higher T is necessary to avoid size distortions if q is close to 0. Table 7 presents size results for the setting of Table 5 with the modification that the break already occurs at $\frac{T}{4}$. Here, the empirical size is slightly further away from the nominal size, but the differences are minor.

The size analysis in the factor copula setting is completed by analyzing the two-step procedure mentioned at the end of Section 3. This means we first perform the non-parametric copula constancy test proposed in Bücher et al. (2014) and state the rejection frequency. Then, for the rejected runs, we apply the relevant change test, where Δ is chosen as the estimated copula difference, and again state the frequency of rejections, cf. Table 8. The frequency in the first step gives the empirical power of the Bücher et al. (2014) test, which tends to 1 for increasing T. The frequency in the second step gives the empirical size of the relevant change test, which is close to the nominal size.

Table 5: Size using quantile version

		$T = 300$	$T = 500$	$T = 750$	$T = 1000$
$N = 2$	q_{95}	0.06	0.04	0.05	0.06
	q_{90}	0.14	0.10	0.13	0.12
$N = 3$	q_{95}	0.08	0.07	0.06	0.06
	q_{90}	0.15	0.13	0.12	0.12
$N = 5$	q_{95}	0.04	0.04	0.05	0.05
	q_{90}	0.10	0.10	0.13	0.12
$N = 30$	q_{95}	0.03	0.04	0.05	0.06
	q_{90}	0.08	0.10	0.09	0.11

Table 5 reports the rejection rate of the relevant change test for data generated with the DGP described in (16) using $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$. In total, we conducted 301 Monte Carlo repetitions.

Table 6: Size using quantile version

		$T = 1000$	$T = 2000$	$T = 4000$
$N = 2$	q_{95}	0.08	0.09	0.08
	q_{90}	0.18	0.17	0.16
$N = 3$	q_{95}	0.09	0.10	0.07
	q_{90}	0.17	0.18	0.15
$N = 5$	q_{95}	0.06	0.09	0.05
	q_{90}	0.12	0.19	0.13
$N = 30$	q_{95}	0.04	0.05	0.05
	q_{90}	0.11	0.12	0.12

Table 6 reports the rejection rate of the relevant change test for data generated with the DGP described in (16) using $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{q} = 0.1 \cdot (1, \ldots, 1)'$. In total, we conducted 301 Monte Carlo repetitions.

For the power analysis of the quantile based test we consider two different scenarios. First, we set a fixed Δ while we increase the copula difference by increasing the parameter θ_1 after the break. Second, we keep the parameter values $\theta_0 = 1$ and $\theta_1 = 2$ fixed and decrease Δ, while the starting point for Δ is equal to the implied copula difference at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$.

Table 9 reports the rejection rate of the test (12) using the 95%$-$quantile of the proposed bootstrap distribution in Section 3. The first column depicts the rejection rate under the null hypothesis. The values of the other columns are obtained by increasing the corresponding copula parameter $\theta_1 \in \{2.2, 2.4, 2.6, 2.8\}$, while Δ remains fixed to the initial copula difference, i.e. $\theta_0 = 1$ and $\theta_1 = 2$.

Table 9 illustrates that the power of the test (12) generally increases not only if T but also if the cross sectional dimension N increases. For example, the scenario $N = 30$, $T = 750$ and $\theta_1 = 2.6$ always rejects the null hypothesis, i.e. the rejection rate is equal to 1. This is also expected, since we increase the parameter in the factor copula model (16) for each component. Consequently, the error is effectively

Table 7: Size using quantile version

		$T = 300$	$T = 500$	$T = 750$	$T = 1000$
$N = 2$	q_{95}	0.06	0.08	0.04	0.06
	q_{90}	0.14	0.15	0.11	0.12
$N = 3$	q_{95}	0.07	0.06	0.05	0.06
	q_{90}	0.15	0.13	0.10	0.11
$N = 5$	q_{95}	0.06	0.06	0.05	0.05
	q_{90}	0.14	0.13	0.12	0.11
$N = 30$	q_{95}	0.06	0.08	0.07	0.07
	q_{90}	0.09	0.13	0.13	0.12

Table 7 reports the rejection rate of the relevant change test for data generated with the DGP described in (16) using $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{q} = 0.6 \cdot (1, \dots, 1)'$. The break is constructed at $\frac{T}{4}$. In total, we conducted 301 Monte Carlo repetitions.

Table 8: Pretest

	$T = 300$		$T = 500$	
	Pre	Delta	Pre	Delta
$N = 2$	0.43	0.09	0.65	0.08
$N = 3$	0.63	0.05	0.89	0.04
$N = 5$	0.76	0.03	0.95	0.04

Table 8 shows the rejection rate of the relevant change test for the DGP (16) with a break at $\frac{T}{2}$ with $\theta_0 = 1$ and $\theta_1 = 1.6$ using $B = 300$ bootstrap replication on a significance level of 0.05. In total, we conducted 301 Monte Carlo repetitions. First we performed the Bücher et al. (2014) test (Pre) and for the rejected runs we applied the relevant change test (Delta). The Δ is chosen as the estimated copula difference.

added up which leads to the gain in power.
Table 10 provides the power analysis for the setting of Table 9 for a break at $\frac{T}{4}$. According to the expectations, the table shows that the power increases for an increasing θ or T. However, the empirical power is lower compared to the setting with a break at $\frac{T}{2}$.
Finally, Table 11 analyzes the rejection rate if Δ decreases while the copula difference remains fixed. The value Δ_0 is equal to the copula difference computed at the point $\mathbf{q} = 0.6 \cdot (1, \dots, 1)'$. Now, we decrease Δ stepwise, i.e. $\Delta = d \cdot \Delta_0$ with $d \in \{0.95, 0.9, 0.85, 0.8, 0.75\}$. Table 11 shows, that the rejection rate tends to 1 if T increases. Moreover, the power is generally higher for larger N.

4.2 Residual Data

In this subsection we consider residual data X_t from pre-estimated time series models for $t = 1, \dots, T$. For our simulation we consider a GARCH(1,1) model, i.e.

$$\begin{aligned} r_{it} &= \sigma_{it} X_{it} \\ \sigma_{it}^2 &= \alpha_0 + \alpha_1 r_{i,t-1}^2 + \beta_1 \sigma_{i,t-1}^2 \end{aligned} \tag{17}$$

for $i = 1, \dots, N$ and $t = 1, \dots, T$. To get serial correlated data we first simulate residual data using the factor copula model (16) with a break constructed at $\frac{T}{3}$ and $\theta_0 = 1$ and $\theta_1 = 2$. Then, we transform the residual data in serial correlated data r_{it} using the GARCH(1,1) model with fixed parameter values $\alpha_0 = \frac{1}{10}, \alpha_1 = \frac{1}{15}$ and $\beta_1 = \frac{1}{3}$ for $i = 1, \dots, N$ and $t = 1, \dots, T$.
With the simulated serial correlated data r_{it} we estimate the time series models using a GARCH(1,1) model and determine the residual data X_{it} for $i = 1, \dots, N$, which is used to perform the test. We vary the sample size $T = 1000, 2000, 4000$ and cross sectional dimension $N = 3, 5, 10$. The results can be seen

Table 9: Power

		$\theta_1 = 2.0$	$\theta_1 = 2.2$	$\theta_1 = 2.4$	$\theta_1 = 2.6$	$\theta_1 = 2.8$
$N = 2$	$T = 300$	0.06	0.46	0.70	0.80	0.89
	$T = 500$	0.04	0.48	0.75	0.87	0.93
	$T = 750$	0.05	0.56	0.81	0.93	0.97
	$T = 1000$	0.06	0.57	0.87	0.97	1.00
$N = 3$	$T = 300$	0.08	0.49	0.70	0.86	0.95
	$T = 500$	0.07	0.44	0.75	0.89	0.96
	$T = 750$	0.06	0.56	0.81	0.97	0.99
	$T = 1000$	0.06	0.65	0.94	0.99	1.00
$N = 5$	$T = 300$	0.04	0.42	0.70	0.86	0.95
	$T = 500$	0.04	0.50	0.82	0.94	0.99
	$T = 750$	0.05	0.60	0.95	1.00	1.00
	$T = 1000$	0.05	0.67	0.94	1.00	1.00
$N = 30$	$T = 300$	0.03	0.56	0.80	0.93	0.97
	$T = 500$	0.04	0.55	0.92	0.99	1.00
	$T = 750$	0.05	0.68	0.97	1.00	1.00
	$T = 1000$	0.06	0.78	0.99	1.00	1.00

Table 9 reports the rejection rate of the quantile relevant change test for data generated with the DGP described in (16) using $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$ varying $\theta_1 \in \{2.0, 2.2, 2.4, 2.6, 2.8\}$ with $\theta_0 = 1$ in the DGP (16). In total, we conducted 301 Monte Carlo repetitions.

Table 10: Power break at $\frac{T}{4}$

		$\theta_1 = 2.0$	$\theta_1 = 2.2$	$\theta_1 = 2.4$	$\theta_1 = 2.6$	$\theta_1 = 2.8$
$N = 2$	$T = 1000$	0.08	0.27	0.53	0.72	0.86
	$T = 2000$	0.08	0.41	0.76	0.94	0.98
$N = 3$	$T = 1000$	0.07	0.30	0.71	0.82	0.95
	$T = 2000$	0.08	0.45	0.84	0.96	0.99
$N = 5$	$T = 1000$	0.06	0.33	0.67	0.90	1.00
	$T = 2000$	0.07	0.51	0.89	0.99	1.00

Table 10 reports the rejection rate of the quantile relevant change test for data generated with the DGP described in (16) using $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$ varying $\theta_1 \in \{2.0, 2.2, 2.4, 2.6, 2.8\}$ with $\theta_0 = 1$ in the DGP (16) where the break point ist constructed at $\frac{T}{4}$. In total, we conducted 301 Monte Carlo repetitions.

in Table 12, which indicates that the test using residual data holds the size level.
Table 13 presents size results for the case of breaks in the GARCH parameters, where the GARCH residuals are calculated by means of the known GARCH parameters. Here, the coefficient β from (17) increases from 0.4 to 0.7 at $\frac{T}{2}$. Also in this case, the empirical size is close to the nominal size.
Finally, the power of our test in the case of GARCH residuals (with constant parameters) is examined in Table 14. The power is slightly lower than in the case without GARCH effects (Table 9), but also converges to 1 for increasing θ_1.

5 Application

In this section, we apply the quantile based test to a multivariate data set of cross-sectional dimension $N = 30$. First, we apply a $GARCH(1,1)$ filter to the daily aggregated stock log-returns over a time span

Table 11: Power

		$\Delta = \Delta_0$	$\Delta = 0.95{\cdot}\Delta_0$	$\Delta = 0.9 \cdot \Delta_0$	$\Delta = 0.85{\cdot}\Delta_0$	$\Delta = 0.80{\cdot}\Delta_0$	$\Delta = 0.75{\cdot}\Delta_0$
$N = 2$	$T = 300$	0.06	0.18	0.36	0.57	0.72	0.95
	$T = 500$	0.04	0.19	0.37	0.55	0.72	0.94
	$T = 750$	0.05	0.15	0.35	0.56	0.74	0.95
	$T = 1000$	0.06	0.22	0.43	0.64	0.80	0.98
$N = 3$	$T = 300$	0.08	0.17	0.37	0.57	0.77	0.95
	$T = 500$	0.07	0.15	0.32	0.54	0.73	0.96
	$T = 750$	0.06	0.20	0.41	0.64	0.84	0.97
	$T = 1000$	0.06	0.17	0.45	0.60	0.87	0.98
$N = 5$	$T = 300$	0.04	0.15	0.29	0.51	0.69	0.90
	$T = 500$	0.04	0.16	0.35	0.56	0.75	0.94
	$T = 750$	0.05	0.19	0.43	0.63	0.83	0.99
	$T = 1000$	0.05	0.17	0.43	0.73	0.92	1.00
$N = 30$	$T = 300$	0.03	0.18	0.32	0.47	0.65	0.88
	$T = 500$	0.04	0.15	0.34	0.55	0.76	0.95
	$T = 750$	0.05	0.15	0.42	0.66	0.84	0.98
	$T = 1000$	0.06	0.19	0.50	0.75	0.91	1.00

Table 11 reports the rejection rate of the quantile relevant change test for data generated with the DGP described in (16) using $B = 300$ bootstrap replications and 301 Monte Carlo repetitions. The copula difference is evaluated at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$, while $\Delta = d \cdot \Delta_0$ with $d \in \{0.95, 0.9, 0.85, 0.8, 0.75\}$.

Table 12: Size using quantile version for GARCH-data

		$T = 1000$	$T = 2000$	$T = 4000$
$N = 2$	q_{95}	0.06	0.07	0.06
	q_{90}	0.15	0.14	0.12
$N = 3$	q_{95}	0.05	0.07	0.06
	q_{90}	0.12	0.13	0.13
$N = 5$	q_{95}	0.05	0.06	0.06
	q_{90}	0.11	0.13	0.11
$N = 10$	q_{95}	0.06	0.05	0.05
	q_{90}	0.11	0.10	0.10

Table 12 reports the rejection rate of the relevant change test where residual data from pre-estimated GARCH(1,1) models is considered. The copula difference is evaluated at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$. In total, we conducted $B = 300$ bootstrap replications and 701 Monte Carlo repetitions.

ranging from January 2003 to July 2015 from the German DAX30, implying $T = 3200$ and $N = 30$[5]. Second, we estimate a possible break point location in our $GARCH(1,1)$ adjusted data set, using (15), with the quantile $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$. This gives $\lfloor \hat{s}T \rfloor = 1884$ (15.02.2011), cf. the black solid line in Figure 1. The first estimated break point in February 2011 now divides the data set into two parts (Dec. 04 - Feb. 11 and Feb. 11 - Jul. 15). As the test indeed indicates a break (see below), we repeat the breakpoint estimation in each part and obtain $\lfloor \hat{s}T \rfloor = 676$ (17.05.2006) for the first part and $\lfloor \hat{s}T \rfloor = 2653$ (26.02.2014) for the second part, respectively, which are both represented by the black dotted line in Figure 1. We do not search further for any breaks, as it seems to be unlikely that there are more than three change points, see Manner et al. (2019).

In the next step, we calculate (for each interval with one estimated breakpoint inside) $\Delta^{smallest}$ for each estimated break point as the smallest Δ for which the null hypothesis of no relevant change cannot be rejected, i.e., $|C_1(\mathbf{0.6}) - C_2(\mathbf{0.6})| \leq \Delta^{smallest}$. The number of bootstrap replications is 300. For each estimated break point, we also calculate the difference of the two resulting empirical copulas δ for

[5] We adjusted the estimate for 5% of their outliers by setting these values equal to the expected value.

Table 13: Size using quantile version for GARCH-data with breaks

		$T = 1000$	$T = 2000$	$T = 4000$
$N = 2$	q_{95}	0.04	0.06	0.05
	q_{90}	0.09	0.09	0.11
$N = 3$	q_{95}	0.03	0.05	0.05
	q_{90}	0.08	0.11	0.09
$N = 5$	q_{95}	0.03	0.06	0.05
	q_{90}	0.08	0.12	0.11
$N = 10$	q_{95}	0.04	0.06	0.07
	q_{90}	0.13	0.14	0.13

Table 13 reports the rejection rate of the relevant change test where residual data from pre-estimated GARCH(1,1) models is considered. The coefficient of β_1 increases from 0.4 to 0.7 for each time series at $\frac{T}{2}$. The copula difference is evaluated at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$. In total, we conducted $B = 300$ bootstrap replications and 701 Monte Carlo repetitions.

Table 14: Power GARCH

Power Analysis varying θ of the DGP (16) using GARCH-residuals						
		$\theta_1 = 2.0$	$\theta_1 = 2.2$	$\theta_1 = 2.4$	$\theta_1 = 2.6$	$\theta_1 = 2.8$
$N = 2$	$T = 1000$	0.05	0.46	0.73	0.90	0.99
	$T = 2000$	0.07	0.47	0.82	0.98	1.00
$N = 3$	$T = 1000$	0.05	0.46	0.76	0.94	0.99
	$T = 2000$	0.06	0.50	0.93	1.00	1.00
$N = 5$	$T = 1000$	0.04	0.47	0.82	0.96	1.00
	$T = 2000$	0.05	0.64	0.97	1.00	1.00

Table 14 reports the rejection rate of the quantile relevant change test for data generated with the DGP described in (16) using $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$ varying $\theta_1 \in \{2.0, 2.2, 2.4, 2.6, 2.8\}$ with $\theta_0 = 1$ in the DGP (16) using GARCH-Residuals. In total, we conducted 301 Monte Carlo repetitions.

$\mathbf{u} = 0.6 \cdot (1, ..., 1)'$, i.e. $\delta := |\hat{C}_1(\mathbf{0.6}) - \hat{C}_2(\mathbf{0.6})|$ and the change of the pairwise averaged Spearman's rhos before and after the estimated break point. The results of $\Delta^{smallest}$, δ and the change of the pairwise averaged Spearman's rhos can be found in Table 16. Table 15 provides the estimated Spearman's rho using the initial dataset for the breakpoint in Feb. 11 and the resulting subdatasets for the change in Spearman's rho in May 06 and Feb. 14. Given $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$, one possible reason for the first

Table 15: Empirical Values

	Dec. 04 - May 06	May 06 - Feb. 11	Feb.11 - Feb. 14	Feb. 14 - Jul. 15	Dec. 04 - Feb. 11	Feb.11 - Jul. 15
$\rho_{Spearman}$	0.3469	0.4216	0.4677	0.5237	0.3998	0.4841

Table 15 reports the mean of Spearman's ρ in the corresponding intervals of the $GARCH(1, 1)$ adjusted log returns.

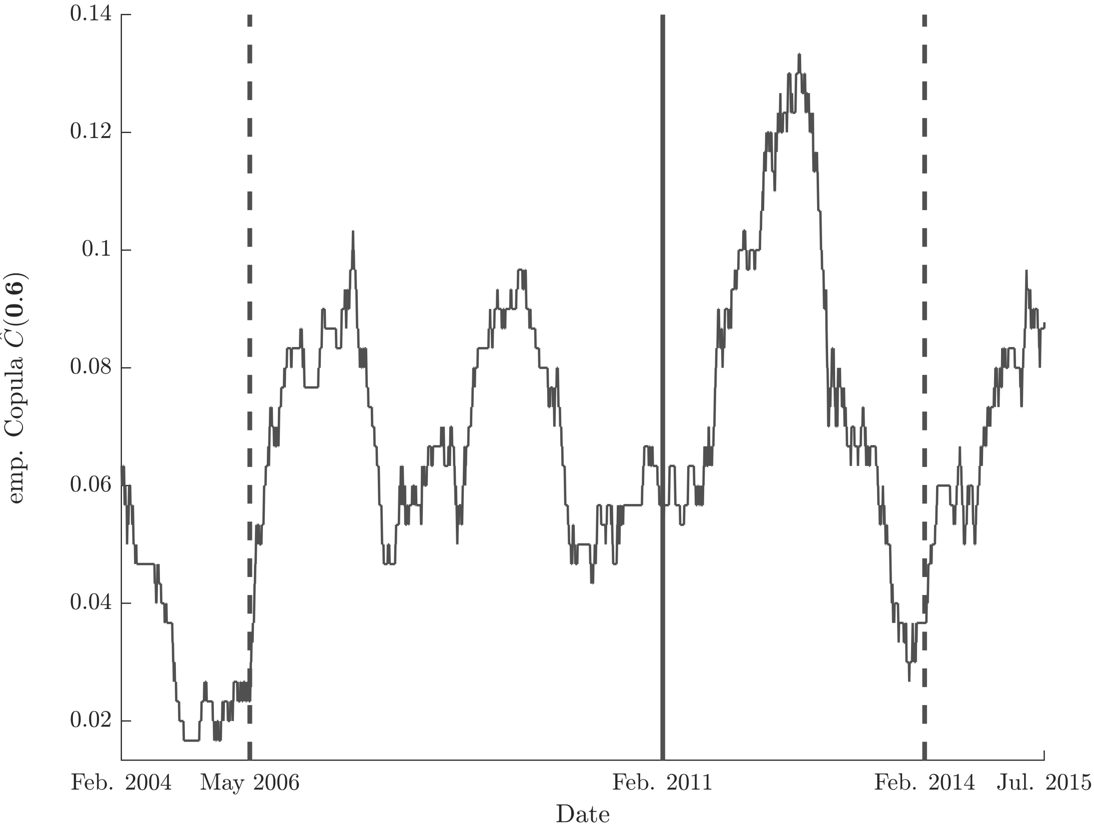
estimated break point in Feb. 2011 can be the beginning of the Euro crisis, i.e. the period in which considerable peaks of several Euro government bond yields were observed. Both, Delta and Spearman's rho are at their highest values here, which suggests that this crisis is having the strongest impact. It is well known that dependencies increase in times of crisis. The fact that Spearman's rho is rising strongly and positively after the break point in Feb. 11 is therefore an indication. For the break point in May 06,

Table 16: Empirical Break Points

	May 06	Feb. 11	Feb. 14
$\Delta^{smallest}$	0.0098	0.0293	0.0273
δ	0.0292	0.0316	0.0387
$\rho^{diff}_{Spearman}$	0.0747	0.0843	0.0560

Table 16 reports the smallest Delta $\Delta^{smallest}$ such that the null hypothesis of no relevant break cannot be rejected. It also provides the empirical copula difference before and after the break point. The last row depicts the change in Spearman's rho before and after the estimated break points.

Fig. 1: Value of the empirical Copula at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)$



Value of the empirical copula defined in (14) evaluated at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)$, computed in a rolling window of size 300. The estimated breakpoint, using (15), is displayed with the vertical black solid line (15.02.2011). The two dotted lines represent estimated breakpoints of the resulting subdatasets on 17. May 2006 and 26. February 2014, respectively. Observed data between January 2003 and July 2015, implying $T = 3200$ and $N = 30$.

$\Delta^{smallest}$ is equal to 0.0098, while the difference in Spearman's rho from the period Dec. 04 - May 06 and May 06 - Feb. 11 is equal to 0.0747. Analogously, for the estimated break point in Feb. 14, $\Delta^{smallest}$ is equal to 0.0273, while the change in Spearman's rho is equal to 0.0560. To sum up, if Δ is chosen to be the smallest value for which the null hypothesis of no relevant change cannot be rejected, the testing procedure provides a formula to determine Δ biuniquely. In addition, we observe that large values of

$\Delta^{smallest}$ are related to large values of Spearman's rho. This means, the test can not only be used to test for relevant changes in the copula, but also as a selection tool to assess the effects of breaks.

6 Conclusion

In summary, the classical break point testing framework has two severe issues: On the one hand it considers a null which is theoretically never fulfilled and on the other hand any consistent test detects any arbitrary small change if the sample size is sufficiently large. Relevant change point analysis offers a way out.

We propose a new non-parametric test for detecting relevant breaks in copula functions, where the hypothesis is of the form $H_0 : \|C_1(\mathbf{u}) - C_2(\mathbf{u})\| \leq \Delta$ versus $H_1 : \|C_1(\mathbf{u}) - C_2(\mathbf{u})\| > \Delta$ with Δ a positive adjustable size to allow for difference in the copulas C_1 and C_2. Here, the norm in the hypothesis represents two different approaches: Either it measures the distance of the copulas given a certain value $\mathbf{q}$ or it equals the L^2-norm.

As a starting point, we consider a natural CUSUM-type test statistic fitting to the underlined testing problem. For the estimation of the limiting distribution, we construct a new non-parametric bootstrap based on natural estimates of the constructed testing process, which is applicable in the case of unknown sequentially estimated marginal distributions.

In the case where the copula distance is measured at a given value $\mathbf{q}$, we consider simulated data up to cross sectional dimension $N = 30$. For the L^2-norm, we investigate the behavior of our test up to $N = 5$. The Monte Carlo simulations show considerable size and power properties for both serially independent and residual data.

In our empirical application we analyze German DAX30 data of cross sectional dimension $N = 30$ between January 2003 and July 2015. Here, Δ is interpreted as the smallest admissible copula difference for which the relevant change hypothesis cannot be rejected. Every other choice of Δ that is smaller leads to a rejection of the null hypothesis. Cutting the empirical data into three parts leads to a detection of the very start of the financial crisis in 2006, the start of the Euro crisis in 2011 and to a break in 2014 given that the quantile $\mathbf{q}$ is chosen to be equal to $0.6 \cdot (1, \ldots, 1)'$.

In conclusion, Δ can be regarded not only as the upper bound of an admissible copula distance, but also as a measure of the extent of a crisis.

References

Aloui R, Aïssa MSB, Nguyen DK (2011) Global financial crisis, extreme interdependences, and contagion effects: The role of economic structure? Journal of Banking & Finance 35(1):130–141

Berkson J (1938) Some difficulties of interpretation encountered in the application of the chi-square test. J Am Statist Ass 33:526–536

Brodsky B, Penikas H, Safaryan I, et al. (2009) Detection of structural breaks in copula models. Applied Econometrics 16(4):3–15

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

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

Busetti F, Harvey A (2011) When is a copula constant? a test for changing relationships. Journal of Financial Econometrics 9:106–131

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

Dette H, Gösmann J (2017) Relevant change points in high dimensional time series. working paper, arXiv: 170404614

Dette H, Wied D (2016) Detecting changes in time series models. Journal of the Royal Statistical Society, Series B 78:371–394

Dette H, Wu W, Zhou Z (2018) Change point analysis of second order characteristics in non-stationary time series. Statistica Sinica, forthcoming

Hansen EB (1994) Autoregressive density estimation. International Economic Review 35:705–730

Krämer W, van Kampen M (2011) A simple nonparametric test for structural change in joint tail probabilities. Economics Letters 110:245–247

Manner H, Stark F, Wied D (2019) Testing for structural breaks in factor copula models. Journal of Econometrics 208:324–345

Oh D, Patton A (2017) Modelling dependence in high dimensions with factor copulas. Journal of Business and Economic Statistics 35:139–154

Wied D, Dehling H, van Kampen M, Vogel D (2013) A fluctuation test for constant spearmans's rho with nuisance-free limit distribution. Computational Statistics and Data Analysis 76:723–736