Consistent Estimation of Multiple Breakpoints in Dependence Measures*

Marvin Borsch[1], Alexander Mayer[2], and Dominik Wied[3]

[1,3] *Institute of Econometrics and Statistics, University of Cologne*

[2] *Department of Economics, Università Ca' Foscari Venezia*

March 31, 2023

Abstract

This paper proposes different methods to consistently detect multiple breaks in copula-based dependence measures. Starting with the classical binary segmentation, also the more recent wild binary segmentation (WBS) is considered. For binary segmentation, consistency of the estimators for the location of the breakpoints as well as the number of breaks is proved, taking filtering effects from AR-GARCH models explicitly into account. Monte Carlo simulations based on a factor copula as well as on a Clayton copula model illustrate the strengths and limitations of the procedures. A real data application on recent Euro Stoxx 50 data reveals some interpretable breaks in the dependence structure.

Keywords: (Wild) Binary Segmentation, Clayton Copula, Factor Copula, AR-GARCH residuals

*Financial support by Deutsche Forschungsgemeinschaft (DFG grant 'Strukturbrüche und Zeitvariation in hochdimensionalen Abhängigkeitsstrukturen') is gratefully acknowledged. Moreover, we are grateful to helpful comments from the Editor Atsushi Inoue, an associate editor and two referees. The computations were implemented in Matlab, parallelized and performed using CHEOPS, the DFG-funded (Funding number: INST 216/512/1FUGG) High Performance Computing (HPC) system of the Regional Computing Center at the University of Cologne (RRZK).

1 Introduction

For asset allocation in financial markets and risk management purposes, dependence measures are of great interest. For example, it is necessary to estimate the variance-covariance matrix of asset returns to construct mean-variance efficient portfolios as introduced by Markowitz (1952). In recent years, also copula-based/non-linear dependencies Nelsen (2006, Chapter 5) became popular, e.g. using a conditional multivariate version of Spearman's ρ as in Schmid and Schmidt (2007) or Penzer et al. (2012). Therefore, estimators of these measures are crucial. However, in financial time series, breaks in dependence measures can occur, see e.g. Longin and Solnik (1995). These happen quite frequently in times of crisis, e.g. during the financial crisis in 2008 or more recently in the beginning of 2020 when the Corona pandemic began. This phenomenon is generally known as the diversification meltdown. Hence, consistent estimation of the location of those break points as well as their number need to be considered in order to correctly estimate the dependencies between assets (in between the breaks). These better estimators might be useful e.g. for Value at Risk forecasting, see Berens et al. (2015).

In this paper, we propose methods for dating multiple breaks in copula-based dependence measures, which are broadly applicable. They are based on the binary segmentation (BS) algorithm that goes back to Vostrikova (1981) and which since then has been adopted frequently in the literature; see, e.g. Bai (1997). In principle, we follow the algorithm used in Galeano and Tsay (2009) or in Galeano and Wied (2014, 2017), who consider breaks in correlations. The principle is to sequentially apply the 'moment-based' test from Manner et al. (2019) for detecting breakpoints in copula-based dependence measures for filtered data. Filtering means that residuals from univariate AR-GARCH models are considered. More specifically, the maximum of this CUSUM-statistic within the subintervals is compared to some critical value. Here, and similar to Bai and Perron (1998, Section 4.3), we derive the asymptotic distribution of the maximum statistic under the null hypothesis that there is no *additional* break to take multiple testing into account. In doing so, we use recent results by Nasri et al. (2022) to derive consistency of our procedure for the location as well as the number of breaks from primitive conditions. We also consider the wild binary segmentation (WBS) algorithm of Fryzlewicz (2014), which improves the detection of change points, especially for multiple breaks in close proximity.

The literature for structural changes in dependence measures is wide. Wied et al. (2012), for instance, test for breaks in the correlation using an extended functional delta method, while Aue et al. (2009) test for changes in the covariance matrix. Tests for nonlinear dependence measures like Spearman's ρ are proposed in Wied et al. (2014), Kojadinovic et al. (2016), Manner et al. (2019) or Stark and Otto (2022). In contrast to this, the question of estimating the number of breaks and dating breaks in copula-based dependence measures has found somewhat less interest in the literature up to now. Some of the mentioned tests provide ad-hoc possibilities for dating breakpoints, but there is a lack in results on the (asymptotic) validity of these procedures.

Our setup allows for considering several copula-based dependence measures including, for example, Spearman's ρ, Gini's γ, Spearman's footrule, (smoothed) quantile dependencies or measures of asymmetry. They may be calculated group-specific, taking previous knowledge of the variable structure into account. Moreover, we follow suggestions by Bücher et al. (2014) for calculating ranks sequentially to improve power. Our test statistic is based on filtered return data instead of observable time series; an idea also recently employed by Barassi et al. (2020) to analyze changes in the conditional correlation.

As our procedure is based on a CUSUM-type statistic, our paper also contributes to this literature. Starting with Shewhart (1931) or Roberts (1959), CUSUM tests were first used in engineering and optimality properties were derived (Shiryaev, 1963). Brown et al. (1975) and Ploberger and Krämer (1992) apply this principle in seminal papers to breaks in the linear regression model, followed by numerous modifications and refinements.

One prominent model, where our procedure can be applied, is the factor copula framework as introduced in Oh and Patton (2017). This is especially useful in high-dimensional applications due to the sparse amount of parameters in the factor structure. Abstaining from a multivariate Gaussian distribution assumption, it can capture tail risk as well as the leverage effect. However, our procedure can also be applied for other types of data, such as data generated from a conditional correlation model. Our procedure is nonparametric is the sense that we do not test specific model assumptions, but check whether there are breaks in the nonparametrically estimated dependence measures.

2 Finding Multiple Breaks in Dependence Measures

2.1 Modeling Breaks in Dependence Measures

Suppose we observe a sample $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_T$ of length T with a cross section $\boldsymbol{Y}_t \coloneqq (Y_{1,t}, \ldots, Y_{n,t})'$ of n financial assets so that each $n \times 1$ vector $\boldsymbol{Y}_t$, $t \in \{1, \ldots, T\}$, obeys a parametric location-scale specification

$$\boldsymbol{Y}_t = \boldsymbol{\mu}_t(\boldsymbol{\phi}^0) + \boldsymbol{\sigma}_t(\boldsymbol{\phi}^0)\boldsymbol{\eta}_t, \tag{1}$$

where $\boldsymbol{\eta}_t \coloneqq (\eta_{1,t}, \ldots, \eta_{n,t})'$ are innovations with $\mathsf{E}[\eta_{i,t}] = 0$ and $\mathsf{var}[\eta_{i,t}] = 1$ for any $i \in \{1, \ldots, n\}$. Similar to Chen and Fan (2006), Oh and Patton (2013), and Oh and Patton (2017), we assume that the conditional mean and variance functions—respectively given by $\boldsymbol{\mu}_t(\boldsymbol{\phi}^0) \coloneqq (\mu_{1,t}(\boldsymbol{\phi}^0), \ldots, \mu_{n,t}(\boldsymbol{\phi}^0))'$ and $\boldsymbol{\sigma}_t(\boldsymbol{\phi}^0) \coloneqq \mathsf{diag}\{\sigma_{1,t}(\boldsymbol{\phi}^0), \ldots, \sigma_{n,t}(\boldsymbol{\phi}^0)\}$—are (a) parametrically known up to a finite dimensional parameter $\boldsymbol{\phi}^0$ and (b) measurable with respect to $\mathcal{F}_{t-1} \coloneqq \sigma(\{\boldsymbol{Y}_j, j < t\})$. The innovations $\boldsymbol{\eta}_t$, on the other hand, are jointly independent of $\mathcal{F}_{t-1}$. This setting allows for many AR-GARCH specifications commonly encountered in financial econometrics.

Assuming continuous margins $\mathsf{F}_{i,t}(x) \coloneqq \mathsf{P}(\eta_{i,t} \leq x)$, $x \in \mathbb{R}$, we can, given the location-scale structure of Eq. (1) and the independence of $\boldsymbol{\eta}_t$ from $\mathcal{F}_{t-1}$, rephrase the conditional joint distribution as

$$\mathsf{P}(Y_{1,t} \leq y_1, \ldots, Y_{n,t} \leq y_n \mid \mathcal{F}_{t-1}) = \mathsf{C}_t(\mathsf{F}_{1,t}(y_t), \ldots, \mathsf{F}_{n,t}(y_n)), \quad y_i \in \mathbb{R}, \tag{2}$$

where the copula $\mathsf{C}_t(\boldsymbol{u}) = \mathsf{P}(\boldsymbol{U}_t \leq \boldsymbol{u})$, $\boldsymbol{u} \in [0,1]^n$, $\boldsymbol{U}_t := (U_{1,t}, \ldots, U_{n,t})'$ for $U_{i,t} := \mathsf{F}_{i,t}(\eta_{i,t})$, captures uniquely the dependence among the n variates in $\boldsymbol{\eta}_t$; see Patton (2006). In what follows, we aim at finding change points in certain dependence measures that can be solely expressed in terms of the copula. As argued, for example, in Bücher et al. (2014) or Kojadinovic et al. (2016), classical nonparametric tests based on sequential empirical processes have little power against alternatives that leave the margins unchanged and only involve a change in the copula.

More specifically, for a finite collection of disjoint sets $\{\mathcal{G}_1, \ldots, \mathcal{G}_G\}$ partitioning the cross-sectional index set $\{1, \ldots, n\}$ and a collection of suitable bivariate functions $h_l : [0,1]^2 \to \mathbb{R}$, $l \in \{1, \ldots, H\}$, we focus on averaged rank-based dependence measures of the type

$$m_{t,g,l} := \frac{1}{\binom{|\mathcal{G}_g|}{2}} \sum_{\substack{1 \leq i < j \leq n \\ i,j \in \mathcal{G}_g}} m_{i,j,t,l}, \; m_{i,j,t,l} := \mathsf{E}[h_l(U_{i,t}, U_{j,t})], \; g \in \{1, \ldots, G\}. \tag{3}$$

We aim at locating and estimating change points of the $p \times 1$ vector $\boldsymbol{m}_t := (\boldsymbol{m}'_{t,1}, \ldots, \boldsymbol{m}'_{t,G})'$, $\boldsymbol{m}_{t,g} := (m_{t,g,1}, \ldots, m_{t,g,H})'$, $p := GH$.

To illustrate, measures of asymmetry with $h(u,v) = |u+v-1|^k \mathsf{sign}(u+v-1)$, $k \geq 3$ (see Rosco and Joe (2013)), or Spearman's ρ, Spearman's footrule (ϱ), and Gini's γ fit into this setup with $h_\rho(u,v) = 12uv - 3$, $h_\varrho(u,v) = 1 - 3|u-v|$, and $h_\gamma(u,v) = 2(|u+v-1| - |u-v|)$ for $(u,v) \in [0,1]^2$, respectively; see, e.g., Nelsen (2006). The group structure is often used in empirical studies, where some degree of homogeneity within each group can be expected so that $m_{i,j,t,l} = m_{g,t,l}$ for all $i, j \in \mathcal{G}_g$. Group structures are commonly used in empirical work; see, e.g. Oh and Patton (2017), Opschoor et al. (2021), or Oh and Patton (2021). Averaged dependence measures are used frequently and have been found to perform well; see, e.g. Schmid and Schmidt (2007), Quessy (2009), Kojadinovic et al. (2016), or Manner et al. (2019).

Importantly, the above-mentioned dependence measures do not depend on the univariate margins, i.e.

$$m_{i,j,t,l} = \int_{[0,1]^2} h_l(u,v)\, \mathsf{d}\mathsf{C}_{i,j,t}(u,v), \tag{4}$$

where $\mathsf{C}_{i,j,t}(u_i,u_j) \coloneqq \mathsf{C}_t(\boldsymbol{u}^{(i,j)})$, with $\boldsymbol{u}^{(i,j)} = (1,\ldots,1,u_i,1,\ldots,1,u_j,1\ldots,1)'$ for some $n \times 1$ vector $\boldsymbol{u} \in [0,1]^n$, denotes the bivariate marginal copula of C_t. A change in $\boldsymbol{m}_t$ over time must thus be due to a change in C_t. We therefore consider the following break-point scenario as summarized by Assumption A.

Assumption A

The following holds true:

(A1) *Consider a finite partition* $0 =: z_0^0 < z_1^0 < \cdots < z_{\ell^0}^0 < z_{\ell^0+1}^0 \coloneqq 1$. *If* $\lfloor z_{i-1}^0 T\rfloor \leq \lfloor t/T \rfloor \leq \lfloor z_i^0 T\rfloor$, $i \in \{1,\ldots,\ell^0+1\}$, *then* $\mathsf{C}_t(\boldsymbol{u}) = \mathsf{C}_i(\boldsymbol{u})$, *where* $\mathsf{C}_1,\ldots,\mathsf{C}_{\ell^0+1}$ *denote copulae, which are pairwise different in at least one point* $\boldsymbol{u} \in [0,1]^n$ *such that* $\boldsymbol{m}_t = \boldsymbol{g}(t/T)$, *where* $\boldsymbol{g} : [0,1] \to \mathbb{R}^p$ *is a vector-valued step function defined as*

$$\boldsymbol{g}(z) = \sum_{i=0}^{\ell^0} \boldsymbol{\gamma}_i^0 1\{z \in [z_i^0, z_{i+1}^0)\},\ \boldsymbol{g}(1) \coloneqq \boldsymbol{\gamma}_{\ell^0}^0.$$

The $p \times 1$ *level vectors* $\boldsymbol{\gamma}_0^0,\ldots,\boldsymbol{\gamma}_{\ell^0}^0$, *as well as* $z_1^0,\ldots,z_{\ell^0}^0$ *and* ℓ^0 *are finite constants independent of* T.

(A2) *The partial derivatives* $\dot{\mathsf{C}}_{i,k+1} \coloneqq \partial \mathsf{C}_{k+1}/\partial u_i$ *exist and are continuous on* $\{\boldsymbol{u} \in [0,1]^n : 0 < u_i < 1\}$, *for any* $k \in \{0,1,\ldots,\ell^0\}$ *and* $i \in \{1,\ldots,n\}$.

(A3) $\boldsymbol{\eta}_1,\ldots,\boldsymbol{\eta}_T$ *is a sample from a process* $\{\boldsymbol{\eta}_t\}_{t\in\mathbb{Z}}$ *of independently distributed* $n \times 1$ *random vectors* $\boldsymbol{\eta}_t = (\eta_{1,t},\ldots,\eta_{n,t})'$ *with strictly stationary and continuous univariate margins* $\eta_{i,t} \sim \mathsf{F}_i$, $i \in \{1,\ldots,n\}$.

The function $\boldsymbol{g}(\cdot)$ defined in part (A1) of Assumption A specifies the timing and the size of the changes in the dependence measures. This specification is tailored towards abrupt changes and allows for a change in only a subset of $\boldsymbol{m}_t$; see Galeano and Wied (2014, 2017) for a detailed discussion. Part (A2) is due to Segers (2012) and imposes a smoothness condition on the partial derivatives of the copula. Similar to Bücher et al. (2014) or Kojadinovic et al. (2016), by imposing part (A3) of Assumption A, we maintain stationary marginal distributions, i.e. the joint distribution is only affected by a break in the copula. The following remark illustrates the setting within the context of factor copulas.

Remark 1. *Recently, factor copula models gained some popularity as a parsimonious yet flexible way to model the cross-sectional dependence among a possibly large number of financial assets; see, e.g.* Oh and Patton (2013) *or* Opschoor et al. (2021). *According to this literature, one may assume that the unknown copula C_t governing Eq.* (2) *can be generated from an auxiliary factor model of the form $X_{i,t} = \beta_{i,t} Z_t + \varepsilon_{i,t}$, where Z_t, $\beta_{i,t}$, and $\varepsilon_{i,t}$ denote factor, factor loading, and idiosyncratic component, respectively. For example, the group structure mentioned above arises naturally within the 'block-equidependence' model from* Oh and Patton (2017) *where assets are driven by some latent market factor with different intensity in each group/industry resulting in equal dependence between the assets in each group but different dependence across groups. Breaks in dependence measures could result from structural changes in loadings, i.e. if $\beta_{i,t} \neq \beta_{i,t+1}$ for some $t \in \{1, \ldots, T\}$; an issue investigated below in more detail.*

2.2 Segmentation Algorithms

Suppose $\hat{\boldsymbol{\phi}}$ is a $\sqrt{T}$-consistent estimator of $\boldsymbol{\phi}^0$ so that we can define the residuals $\hat{\eta}_{i,t} := \sigma_{i,t}^{-1}(\hat{\boldsymbol{\phi}})(Y_{i,t} - \mu_{i,t}(\hat{\boldsymbol{\phi}}))$. Moreover, for any $1 \leq k < m \leq T$, let $\hat{U}_{i,t}^{k:m}$, $i \in \{1, \ldots, n\}$, $t \in \{k, \ldots, m\}$, denote the rank of $\hat{\eta}_{i,t}$ among $\hat{\eta}_{i,k}, \ldots, \hat{\eta}_{i,m}$. We then define for any $[a, b] \subseteq [0, 1]$, a sample analogue of (3) by

$$\hat{m}_{g,l}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor} := \frac{1}{\binom{|\mathcal{G}_g|}{2}} \sum_{\substack{1 \leq i < j \leq n \\ i,j \in \mathcal{G}_g}} \hat{m}_{i,j,l}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor}, \; l \in \{1, \ldots, H\}, \; g \in \{1, \ldots, G\}, \tag{5}$$

where

$$\hat{m}_{i,j,l}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor} := \frac{1}{\lfloor bT \rfloor - \lfloor aT \rfloor} \sum_{t=\lfloor aT \rfloor+1}^{\lfloor bT \rfloor} h_l(\hat{U}_{i,t}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor}, \hat{U}_{j,t}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor}),$$

so that we obtain the $p \times 1$ vector $\hat{\boldsymbol{m}}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor} := (\hat{\boldsymbol{m}}_1^{\lfloor aT \rfloor+1:\lfloor bT \rfloor\prime}, \ldots, \hat{\boldsymbol{m}}_G^{\lfloor aT \rfloor+1:\lfloor bT \rfloor\prime})'$, where $\hat{\boldsymbol{m}}_g^{\lfloor aT \rfloor+1:\lfloor bT \rfloor} := (\hat{m}_{g,1}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor}, \ldots, \hat{m}_{g,H}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor})'$. Next, define for a given interval $[a,b] \subseteq [0,1]$ with $0 \leq \lfloor aT \rfloor < \lfloor bT \rfloor \leq T$, the detector

$$\hat{M}(a,b;z) = \frac{\lfloor zT \rfloor - \lfloor aT \rfloor}{\sqrt{\lfloor bT \rfloor - \lfloor aT \rfloor}} \left\| \hat{\boldsymbol{m}}^{\lfloor aT \rfloor+1:\lfloor zT \rfloor} - \hat{\boldsymbol{m}}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor} \right\|_2,$$

and the fluctuation test statistic

$$\hat{M}(a,b) := \sup_{z \in [a,b]} \hat{M}(a,b;z) = \max_{z \in \Pi_T(a,b)} \hat{M}(a,b;z), \tag{6}$$

where $\Pi_T(a,b) := \{\lfloor aT \rfloor/T, (\lfloor aT \rfloor + 1)/T, \ldots, (\lfloor bT \rfloor - 1)/T, \lfloor bT \rfloor/T\}$, $\|\cdot\|_2$ is the Euclidean norm, and the conventions $\hat{M}(a,b;z) = 0$ if $z = a$ and $\Pi_T := \Pi_T(0,1)$ are used.

Equation (6) can be related to the test statistic used in Wied et al. (2014) up to a scaling factor for $h = h_\rho$ and $H = 1$. However, they calculate the ranks for Spearman's ρ for the subset $[1, \ldots, k]$ relative to the complete sample in contrast to our approach basing it only on the subset. Bücher et al. (2014) and Kojadinovic et al. (2016) provide empirical evidence showing more powerful results using ranks computed relative to the subset; see also Manner et al. (2019) for an application. The aforementioned papers all use averaged dependence measures as well while we allow for multiple averaged dependencies in several groups.

2.2.1 Binary Segmentation

The following break-point detection algorithm is closely related to the segmentation algorithm proposed by Galeano and Tsay (2009), Galeano and Wied (2014, 2017), which, in turn, are based on the seminal work of Vostrikova (1981). For a given upper tail probability α, the maximum of the test statistic from Eq. (6) is evaluated over $\ell + 1$ subintervals and then compared to some critical value $c_\alpha(\ell)$, which depends on the amount of breaks ℓ already found and which is formally derived from Proposition 2 below. This way, the issue of multiple testing is taken into account similar to the procedure in Bai and Perron (1998, Section 4.3). More formally, the algorithm can be summarized by the following two steps:

1. Obtain from Eq. (6) the test statistics $\hat{M}(0,1)$.

   (a) If the test statistic is statistically significant, i.e. if $\hat{M}_{0,1} > c_\alpha(0)$, where $c_\alpha(0)$ is the asymptotic critical value $c_\alpha(\ell)$ from Proposition 2 below with $\ell = 0$ for a given upper tail probability α, then a change is announced. Let $z_1 = \operatorname*{arg\,max}_{z \in \Pi_T} \hat{M}(0,1;z)$ be the break point estimator and go to Step 2.

   (b) If the test statistic is not statistically significant, the algorithm stops, and no change points are detected.

2. Let $z_1, \ldots, z_\ell$ be the ℓ change points in increasing order already found in previous iterations and $z_0 := 0$, $z_{\ell+1} := 1$. If $\max_{1 \le k \le \ell+1} \hat{M}(z_{k-1}, z_k) > c_\alpha(\ell)$, then a new change point is detected at the point fraction at which the value $\hat{M}(z_{k_{\max}-1}, z_{k_{\max}})$ is attained, where $k_{\max} = \operatorname*{arg\,max}_{1 \le k \le \ell+1} \hat{M}(z_{k-1}, z_k)$. Repeat this step until no more change points are found.

In other words, the whole data set is scanned for a change point, once a significant change point is found, the data set is split at that point. On the resulting subsamples the test statistic is computed again and the maximum is then compared to the according critical value. If the associated change point is significant the sample is split again at that point. This procedure stops when no significant change points can be found.

2.2.2 Analytical Results

In order to derive the asymptotic properties of the algorithm from Section 2.2.1, we impose the following assumptions.

Assumption B

For any $[a,b] \subseteq [0,1]$, *the function* $M(a,b;\cdot) : [a,b] \to \mathbb{R}_+$, $z \mapsto M(a,b;z)$, *where* $M(a,b;z) := \|\boldsymbol{M}(a,b;z)\|_2$, *with*

$$\boldsymbol{M}(a,b;z) = \frac{1}{b-a}\left(\int_a^z \boldsymbol{g}(t)\,\mathsf{d}t - \frac{z-a}{b-a}\int_a^b \boldsymbol{g}(t)\,\mathsf{d}t\right),$$

is either constant or has a unique maximum.

Assumption C

For any $l \in \{1,\dots,H\}$, *the following holds true.*

(C1) $h_l : [0,1]^2 \to \mathbb{R}$ *is of bounded variation in the sense of Hardy-Krause.*

(C2) $h_l : [0,1]^2 \to \mathbb{R}$ *is Lipschitz; i.e., for any* $(u_1,v_1) \in [0,1]^2$ *and* $(u_2,v_2) \in [0,1]^2$, $|h_l(u_1,v_1) - h_l(u_2,v_2)| \leq c_0(|u_1-u_2| + |v_1-v_2|)$ *for some constant* $c_0 \in (0,\infty)$.

Assumption B is similar to Galeano and Wied (2014, Assumption 2) and Galeano and Wied (2017, Assumption 7), and restricts the nature of the breaks. Assumption B is fulfilled if there is a 'dominating' break in any subset of components at the same time point. For examples where the assumption is fulfilled or violated see Galeano and Wied (2014) for the case that $\boldsymbol{g}(t)$ is a scalar. To provide some intuition, note that, as shown in the appendix, the function $M(a,b;z)$, specified in Assumption B, coincides with the probability limit of the scaled detector, i.e. $|\hat{M}(a,b;z)/\sqrt{\lfloor bT \rfloor - \lfloor aT \rfloor} - M(a,b;z)| = o_p(1)$ uniformly on $[a,b]$. Assumption C is satisfied by the above-mentioned dependence measures. Part (C1) enables the use of an integration by parts formula for bivariate integrals; see, e.g. Fermanian et al. (2004) and Berghaus et al. (2017). Part (C2) is a convenient assumption, which covers the dependence measures described below Eq. (3), but rules out certain dependence measures like quantile dependence. However, smoothed versions of quantile dependencies or differences thereof would be allowed, i.e. for a given quantile $\tau \in [0,1]$ one could use some smooth link function $h_{\tau,a} : [0,1]^2 \mapsto [0,1]$ (e.g. the logistic function) with tuning parameter a such that $h_{\tau,a}(U,V) = 1\{U \leq \tau, V \leq \tau\}$ as $a \searrow 0$.

In order to discuss the next assumption, it is convenient to first introduce the *infeasible* sequential empirical copula process

$$\tilde{\mathbb{B}}_k(z,\boldsymbol{u}) \coloneqq \frac{1}{\sqrt{\lfloor z_{k+1}^0 T\rfloor - \lfloor z_k^0 T\rfloor}} \sum_{t=\lfloor z_k^0 T\rfloor}^{\lfloor zT\rfloor} (1\{\boldsymbol{U}_t \le \boldsymbol{u}\} - \mathsf{C}_{k+1}(\boldsymbol{u}))$$

for any $z \in [z_k^0, z_{k+1}^0]$. Note that, by part (A3) of Assumption A, Theorem 2.12.1 of van der Vaart and Wellner (1996) in conjunction with the continuous mapping theorem yields $\tilde{\mathbb{B}}_k \rightsquigarrow \mathbb{B}_k$ in $\ell^\infty([z_k^0, z_{k+1}^0] \times [0,1]^n)$, $k \in \{0,1,\ldots,\ell^0\}$, where $\mathbb{B}_k$ is a C_{k+1}-Kiefer process, i.e. $\mathbb{B}_k$ is a tight mean-zero Gaussian process with

$$\mathsf{cov}[\mathbb{B}_k(z,\boldsymbol{u}), \mathbb{B}_k(s,\boldsymbol{v})] = \varphi_k(z\wedge s)(\mathsf{C}_{k+1}(\boldsymbol{u}\wedge\boldsymbol{v}) - \mathsf{C}_{k+1}(\boldsymbol{u})\mathsf{C}_{k+1}(\boldsymbol{v}))$$

and $\varphi_k(z) \coloneqq (z - z_k^0)/(z_{k+1}^0 - z_k^0)$ for any $z, s \in [z_k^0, z_{k+1}^0]$, and $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^n$. The following Assumption D ensures that the sequential *residual* empirical copula process

$$\tilde{\mathbb{C}}_k(z,\boldsymbol{u}) \coloneqq \frac{1}{\sqrt{\lfloor z_{k+1}^0 T\rfloor - \lfloor z_k^0 T\rfloor}} \sum_{t=\lfloor z_k^0 T\rfloor}^{\lfloor zT\rfloor} (1\{\hat{\boldsymbol{U}}_t^{\lfloor z_k^0 T\rfloor+1:\lfloor z_{k+1}^0 T\rfloor} \le \boldsymbol{u}\} - \mathsf{C}_{k+1}(\boldsymbol{u}))$$

converges weakly in $\ell^\infty([z_k^0, z_{k+1}^0] \times [0,1]^n)$ to

$$\mathbb{C}_k(z,\boldsymbol{u}) \coloneqq \mathbb{B}_k(z,\boldsymbol{u}) - \sum_{i=1}^n \mathbb{B}_k(z,\boldsymbol{u}^{(i)})\dot{\mathsf{C}}_{i,k+1}(\boldsymbol{u}),\ \boldsymbol{u}^{(i)} = (1,\ldots,1,u_i,1,\ldots,1),$$

which is a crucial ingredient for the development of the asymptotic theory; see Nasri et al. (2022). Note that $\mathbb{C}_k$ is unaffected by the estimation error associated with $\hat{\boldsymbol{\phi}}$.

Assumption D

The following holds true:

(D1) $(\tilde{\mathbb{B}}_k, \sqrt{T}(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}^0)) \rightsquigarrow (\mathbb{B}_k, \Theta)$, $k \in \{0,1,\ldots,\ell^0\}$, *where* Θ *is a tight random vector;*

(D2) $\sup_{x\in\mathbb{R}} |x\mathsf{f}_i(x)| < \infty$, $i \in \{1,\ldots,n\}$, *where* $\mathsf{f}_i = \partial\mathsf{F}_i/\partial x_i$*;*

(D3) *Eq.* (1) *describes a stationary-ergodic finite-order AR-GARCH model;*

Assumption D is a slightly modified set of regularity conditions that can be found in Nasri et al. (2022, Appendix B). Part (D1) is satisfied by commonly used quasi maximum likelihood estimators of AR-GARCH models; see, e.g. Francq and Zakoïan (2004); part (D2) restricts the tails of the densities; part (D3) could be generalized at the extent of an increase in technicalities; see Nasri et al. (2022).

Akin to Bai (1997, Proposition 2) and Galeano and Wied (2017, Theorem 2), we first derive T-consistency of the change point estimator assuming knowledge of the number of breaks ℓ^0. The idea of the proof is to exploit that the fluctation test statistic needs different centering terms in the case of breakpoints. These centering terms lead to a non-constant limit function under appropriate scaling, to whose maximum the maximum of the test statistic converges. Superconsistency of argmax-type breakpoint estimator is common in the literature; the intuition is that the limit function is not continuous in the breakpoint.

Proposition 1

Suppose Assumptions A-D *hold and let* $\hat{z}_1 < \cdots < \hat{z}_{\ell^0}$ *denote the change-points found upon executing the algorithm in Section* 2.2.1 *without testing for significance. Then, for every* $k \in \{1, 2, \ldots, \ell^0\}$ *and* $\varepsilon > 0$*, there exists a finite constant* $M \in (0, \infty)$ *and an integer* $N \in \mathbb{N}_+$ *such that for all* $T > N$*,* $\mathsf{P}(|\hat{z}_k - z_k^0| > M/T) < \varepsilon$*.*

Once consistency is established, Proposition 1 can be used to derive asymptotic critical values. Similar to Bai and Perron (1998, Section 4.3), Proposition 2 summarizes the limiting distribution of the statistic under the null of ℓ^0 breaks. The proof relies on applying established results on the convergence of empirical copula processes to subsequent intervals.

Proposition 2

Suppose Assumptions A-D *hold. Then,*

$$\max_{k \in \{0,1,\ldots,\ell^0\}} \sup_{z \in [\hat{z}_k, \hat{z}_{k+1}]} \hat{M}(\hat{z}_k, \hat{z}_{k+1}; z) \xrightarrow{d} \max_{k \in \{0,1,\ldots,\ell^0\}} \sup_{z \in [z_k^0, z_{k+1}^0]} \mathbb{M}(z_k^0, z_{k+1}^0; z).$$

The limiting process $\mathbb{M}(z_k^0, z_{k+1}^0; z)$, $z \in [z_k^0, z_{k+1}^0]$, *is given by*

$$\mathbb{M}(z_k^0, z_{k+1}^0; z) := \sqrt{\sum_{g=1}^{G} \sum_{l=1}^{H} \left[\binom{|\mathcal{G}_g|}{2}^{-1} \sum_{\substack{1 \le i < j \le n \\ i,j \in \mathcal{G}_g}} \int_{[0,1]^2} \mathbb{D}_{k,i,j}(z, u, v) \, \mathsf{d}h_l(u, v) \right]^2}$$

for $\mathbb{D}_{k,i,j,l}(z, u_i, u_j) := \mathbb{D}_k(z, \boldsymbol{u}^{(i,j)})$, *with* $\mathbb{D}_k(z, \boldsymbol{u}) = \mathbb{C}_k(z, \boldsymbol{u}) - \varphi_k(z)\mathbb{C}_k(z_{k+1}^0, \boldsymbol{u})$.

Importantly, the limiting distribution is unaffected by the sampling uncertainty resulting from the first-step estimation of $\boldsymbol{\phi}^0$, which reflects a known result from multivariate dynamic copula models where, under certain conditions, semi-parametric estimators based on estimated residuals behave asymptotically as if the unknown innovations were used; see e.g. Chen and Fan (2006). Hence, as Nasri et al. (2022, Section 3) suggests, we can compute the critical values $c_\alpha(\ell)$ based on the traditional IID-bootstrap. The use of the critical values $c_\alpha(\ell)$ ensures a correctly sized test under the null of no further breaks.

An alternative approach for dealing with the multiple testing issue in a binary segmentation framework can be found, e.g. in Galeano and Tsay (2009) and Galeano and Wied (2014), using a Šidák correction on the alpha level to keep the significance level constant for multiple tests.

Finally, the number of breaks found with the help of the segmentation algorithm, $\hat{\ell}$, say, consistently estimates the true amount ℓ^0 if the significance level α tends to zero at a sufficiently slow rate.

Proposition 3

Suppose Assumptions A-D *hold. In addition,* $c_\alpha(\ell) := c_{\alpha,T}(\ell) \to \infty$ *yet* $c_\alpha(\ell) = o(\sqrt{T})$ *as* $T \to \infty$. *Then,* $\mathsf{P}(\hat{\ell} = \ell^0) \to 1$ *as* $T \to \infty$.

Proposition 1 shows the consistency in the location of the break while Proposition 3 shows consistency in the number of changes. The assumption $c_\alpha(\ell) = o(\sqrt{T})$ is common in the literature and arises from the necessity to have critical values such that the test statistic converges in the case of no breaks and diverges in the case of breaks; see Bai (1997), Galeano and Wied (2014), Galeano and Wied (2017). By construction the algorithm induces some overestimation bounded by the alpha level. In order for that to vanish asymptotically the assumption is needed. In finite samples the statistician can choose an appropriate significance level.

Remark 2. *Since bootstrapping the limiting distribution from Proposition* 2 *at each iteration of the algorithm is time consuming, one might resort to a computationally more efficient alternative information-criterion approach. To sketch the procedure, fix some prespecified (large) number* $L \in \mathbb{N}_+$ *that is independent of* T*. Let* $\hat{\boldsymbol{k}} := (\hat{k}_1, \ldots, \hat{k}_L)'$ *denote (in increasing order) the break points obtained from executing the algorithm* $L+1$ *iterations without checking for significance. Suppose* L *has been chosen large enough so that* $\ell^0 \leq L$*. Then, although the number of changes has been overestimated, we can still infer consistency of a subtuple of* $\hat{\boldsymbol{z}} = \hat{\boldsymbol{k}}/T$ *by borrowing principles from* Andrews (1999)*. More specifically, introduce the set of selection vectors* $\mathcal{C}_L := \left\{\boldsymbol{c} = (c_1, \ldots, c_L)' \in \mathbb{R}^L \mid c_j \in \{0,1\},\, 1 \leq j \leq L\right\}$*. Let* $\hat{\boldsymbol{k}}_c := (\hat{k}_{1,c}, \ldots, \hat{k}_{|\boldsymbol{c}|_1,c})' \in \mathbb{N}^{|\boldsymbol{c}|_1}$ *denote the* $|\boldsymbol{c}|_1 \times 1$ *vector of (in increasing order) break-point candidates selected by* $\boldsymbol{c} \in \mathcal{C}_L$*, with* $|\boldsymbol{c}|_1 := \sum_{j=1}^{L} c_j$*, where* $\boldsymbol{c}$ *is a* $L \times 1$ *break-point selection vector whose* j*th entry is one if a break point from* $\hat{\boldsymbol{k}}$ *is included in* $\hat{\boldsymbol{k}}_c$ *and zero otherwise. Hence, we aim at finding that selection vector* $\boldsymbol{c}^0$ *such that* $\hat{\boldsymbol{z}}_{c^0} \xrightarrow{p} \boldsymbol{z}^0$*, where* $\boldsymbol{z}^0 := (z_1^0, \ldots, z_{\ell^0}^0)$*. To do so, we define for any* $\boldsymbol{c} \in \mathcal{C}_L$ *a break-point 'information criterion'* $\mathsf{IC}_T(\boldsymbol{c}) := \hat{S}^2(\boldsymbol{c}) + h(|\boldsymbol{c}|_1)\kappa_T$*, where* $\hat{S}(\boldsymbol{c}) := \hat{S}(\boldsymbol{k}_c)$*, with*

$$\hat{S}(\boldsymbol{k}_c) := \max_{1 \leq j \leq |\boldsymbol{c}|_1 + 1} \max_{k_{c,j-1} < t \leq k_{c,j}} \hat{S}(k_{c,j-1}, k_{c,j}; t)$$

$$\hat{S}(k_{c,j-1}, k_{c,j}; t) := \frac{t - k_{c,j-1}}{\sqrt{k_{c,j} - k_{c,j-1}}} \|\hat{\boldsymbol{m}}^{k_{c,j-1}+1:t} - \hat{\boldsymbol{m}}^{k_{c,j-1}+1:k_{c,j}}\|_2, \quad \hat{k}_{0,c} := 1,\ \hat{k}_{|\boldsymbol{c}|_1+1,c} := T.$$

Imposing similar conditions on $h(\cdot)$ *and the penalty* $\kappa_T \to \infty$ *as in* Andrews (1999) *and* Andrews and Lu (2001)*, it can be shown that* $(\hat{\boldsymbol{k}}_{\hat{c}}/T, |\hat{\boldsymbol{c}}|_1) \xrightarrow{p} (\boldsymbol{z}^0, \ell^0)$*; see Appendix C in the working paper version of this article for details.*

2.2.3 Wild Binary Segmentation

A more recent approach for multiple change point detection has been proposed by Fryzlewicz (2014), the so called *wild* binary segmentation; see also Fryzlewicz (2020). Since test statistics, in the context of multiple break point detection, are often tailored against single change point alternatives, these tests might not perform well in terms of power when the process contains several change points. This is due to the fact that different change points can offset each other; see, e.g. Fryzlewicz (2014, section 2). To counteract this issue, m random intervals are considered by the wild binary segmentation algorithm. As a slight modification, one can add also the complete subsample that would have been used by the ordinary binary segmentation of Section 2.2.1. Within these intervals, the test statistic of Eq. (6) is calculated and the interval associated with the largest statistic announces a change point candidate that can be checked for significance. If a significant change point is found, the sample is split and the procedure is repeated on these subsamples similarly to the binary segmentation until no significant change point is left. In doing so, we hope that at least one 'favorable' random interval designated to find a specific break is generated; i.e. an interval that contains only one change point and is as large as possible. Therefore, this procedure adapts the wild binary segmentation from Fryzlewicz (2020) to the same type of algorithm introduced in Section 2.2.1.

3 Monte Carlo Simulations

To analyze the performance of the segmentation procedures, we conduct Monte Carlo simulations. The first setting follows closely Oh and Patton (2017), where a 'block-equidependence' framework is proposed. More specifically, we assume that the unknown copula can be generated by an auxiliary one-factor model (see also Remark 1):

$$X_{i,t} = b_{i,t} Z_t + \varepsilon_{i,t},\ i = 1, \dots, n,\ t = 1, \dots, T. \tag{7}$$

The loadings are group-specific, i.e. $b_{i,t} = b_{j,t} = \beta_{g,t}$ if $i, j \in \mathcal{G}_g$, $g = 1, \ldots, G$, and $\{\mathcal{G}_1, \ldots, \mathcal{G}_G\}$ partitions the cross-sectional index set. We consider $G = 4$ groups of equal size $|\mathcal{G}_g| = 4$, so that $n = 16$. The common factor is distributed according to Hansen's skew t-distribution $Z_t \overset{\text{IID}}{\sim} Skewt(\nu, \lambda)$ with $\nu = 4$ degrees of freedom and skewness parameter $\lambda = -0.5$, while for the idiosyncratic errors $\varepsilon_{i,t} \overset{\text{IID}}{\sim} t(\nu)$. A similar parametrization for the common factor and the idiosyncratic errors can be found in Oh and Patton (2013).We mostly focus on the case of $T \in \{1000, 1500\}$.

In the following several change point scenarios are analyzed, where changes in the dependence measures are induced by an abrupt change in the factor loadings that drive the dependence across marginals. Table 1 contains the loadings $\beta_{g,t}$ for the according intervals. Here, group 1 and 2 have opposite loadings, group 3 does not contain any change point and group 4 contains a larger change compared to the other groups. This should reflect different behaviours of different industries.

loadings	0 breaks	1 break		2 breaks			3 breaks			
	$[1, T]$	$[1, t_1]$	$(t_1, T]$	$[1, 0.4T]$	$(0.4T, 0.6T]$	$(0.6T, T]$	$[1, 0.3T]$	$(0.3T, 0.5T]$	$(0.5T, 0.7T]$	$(0.7T, T]$
$\beta_{1,t}$	0.8	0.8	0.5	0.8	0.5	0.8	0.6	0.8	0.4	0.6
$\beta_{2,t}$	0.5	0.5	0.8	0.5	0.8	0.5	0.6	0.4	0.8	0.6
$\beta_{3,t}$	0.5	0.5	0.5	0.5	0.5	0.5	0.6	0.6	0.6	0.6
$\beta_{4,t}$	1	0.5	1	0.5	1	0.5	0.6	1	0.2	0.6
ρ_1	0.3581	0.3578	0.2020	0.3577	0.2013	0.3574	0.2556	0.3559	0.1476	0.2558
ρ_2	0.2017	0.2012	0.3573	0.2006	0.3569	0.2017	0.2548	0.1469	0.3550	0.2565
ρ_3	0.2015	0.2018	0.2009	0.2018	0.2000	0.2009	0.2556	0.2563	0.2532	0.2559
ρ_4	0.4472	0.2008	0.4457	0.2006	0.4445	0.2016	0.2551	0.4454	0.0519	0.2541
$\bar{\rho}$	0.3021	0.2404	0.3015	0.2402	0.3007	0.2404	0.2553	0.3011	0.2019	0.2555

Table 1: Loadings from the DGP (7) with 4 groups in the corresponding interval. ρ_i contains the cross-sectional average of Spearman's ρ in group i averaged over 1000 simulations. $\bar{\rho}$ is the average over ρ_i.

In the second setting, we simulate data from a modified Clayton copula, defined by $C(u_1, \ldots, u_d) = \prod_{i=1}^{d} u_i^{1-\theta_i} \left(1 + \sum_{i=1}^{d} \left(u_i^{-\gamma\theta_i} - 1\right)\right)^{-1/\gamma}$. The modification was introduced by Liebscher (2008) and makes the usual Clayton copula potentially more asymmetric. This allows for analyzing whether our procedure can detect changes in the asymmetric dependence, a feature which might be relevant in risk management (stronger asymmetry might be an indicator for a period of financial turmoil), but has not been considered often in the literature as far as we know. The parameters are chosen as $\theta = 0.5$, $\theta = 0.5 \to 0.65$, $\theta = 0.5 \to 0.65 \to 0.5$, $\theta = 0.5 \to 0.65 \to 0.35 \to 0.5$ for the zero, one, two and three break scenario respectively, similarly to Liebscher (2008, Example 1) for $\theta = \theta_1 = \cdots = \theta_d$. Moreover, $\gamma = 2$.

In all cases, we consider residuals of the DGP in (1) by estimating AR-GARCH parameters using quasi maximum likelihood. More specifically, we follow Oh and Patton (2013) and consider an AR(1)-GARCH(1,1) structure with Gaussian innovations $\eta_{i,t}$; i.e., the dgp for each marginal time series is given by $Y_{i,t} = \phi_0 + \phi_1 Y_{i,t-1} + \sigma_{it}\eta_{i,t}$, where $\sigma_{i,t}^2 = \omega + \beta\sigma_{i,t-1}^2 + \alpha\sigma_{i,t-1}^2\eta_{i,t-1}^2$, with $\phi = (\phi_0, \phi_1, \omega, \beta, \alpha)' = (0.01, 0.05, 0.05, 0.85, 0.1)'$. The cross-sectional dependence structure of the innovations $\boldsymbol{\eta}_t := (\eta_{1,t}, \ldots, \eta_{n,t})'$, is governed by the factor copula of Eq. (7) or the modified Clayton copula.

Analyzed are change points in Spearman's ρ and in the vector of Spearman's ρ and the measure for asymmetry described in Section 2 with $k = 3$. The results are very similar in both cases, which indicates that the power for Spearman's ρ is considerably higher than for the asymmetry measure. We only present the results for both measures, believing that this case is more interesting for applications. In both cases, bivariate dependence measures are averaged among each group. This set-up leads to a vector-valued step function as in Assumption (A1). For the critical values an IID-bootstrap is used analogously to Manner et al. (2019, Section 2.4); see also the discussion below Proposition 2.

Table 2 shows the frequency of under- or overestimation of underlying breaks for the four scenarios in both copula models for $T = 1000$. Included are the binary segmentation (BS), the wild binary segmentation (WBS) with $m = 20$ random intervals and the wild binary segmentation where the complete (sub-)interval is taken as one of the 'random' intervals (WBS_{BS}).

We first describe the results for the factor copula in the left panel. In the $1break$-scenario, the binary segmentation has the highest detection rate, in particular if the breaks are close

	Factor copula						Clayton copula					
br	0	1			2	3	0	1			2	3
						BS						
-3						0.037						0.039
-2					0.385	0.016					0.652	0.013
-1		0.058	0.000	0.028	0.008	0.289		0.237	0.001	0.287	0.024	0.325
0	0.942	0.810	0.960	0.802	0.559	0.629	0.943	0.645	0.958	0.597	0.292	0.603
1	0.057	0.131	0.035	0.167	0.045	0.029	0.051	0.115	0.039	0.115	0.032	0.018
2	0.001	0.001	0.005	0.003	0.003	0.000	0.005	0.003	0.002	0.001	0.000	0.002
3	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000
						WBS						
-3						0.005						0.003
-2					0.059	0.425					0.212	0.257
-1		0.470	0.002	0.375	0.246	0.443		0.546	0.045	0.584	0.318	0.567
0	0.945	0.498	0.944	0.574	0.655	0.124	0.931	0.418	0.913	0.373	0.438	0.164
1	0.049	0.031	0.053	0.048	0.035	0.003	0.061	0.036	0.039	0.038	0.031	0.009
2	0.006	0.001	0.001	0.003	0.005	0.000	0.008	0.000	0.002	0.005	0.001	0.000
3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000
						WBS_{BS}						
-3						0.007						0.003
-2					0.071	0.251					0.222	0.124
-1		0.386	0.000	0.250	0.083	0.516		0.457	0.005	0.504	0.168	0.587
0	0.954	0.570	0.951	0.681	0.801	0.218	0.938	0.492	0.955	0.434	0.581	0.282
1	0.042	0.042	0.048	0.066	0.044	0.007	0.058	0.048	0.038	0.057	0.028	0.004
2	0.004	0.001	0.001	0.003	0.001	0.001	0.003	0.003	0.002	0.005	0.001	0.000
3	0.000	0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000

Table 2: Over- and underestimation frequencies of the amount of true change points $\hat{\ell} - \ell^0$ with $\alpha = 0.05$ and $T = 1000$, break fractions for one break: $z_1 = 0.15, 0.5, 0.85$, for two breaks: $z_1 = 0.4, z_2 = 0.6$, for three breaks: $z_1 = 0.3, z_2 = 0.5, z_3 = 0.7$.

to the beginning or end of the sample. In this case, the wild binary segmentation often does not detect the break. Sometimes, BS detects more than one break. Overall, WBS_{BS} performs slightly better than WBS.

Turning to the $2breaks$-scenario, WBS usually detects at least one change point, while BS in some cases does not detect any breaks. Again, WBS_{BS} improves the performance of WBS. The improvement arises since WBS_{BS} is able to detect the remaining breakpoint after the first one is found. This is due to the fact that, after the first break is found, there is only a single break left, for which one the complete subinterval should be considered. Looking at the BS, we see that, if breaks are found, then both of them. Here, the same argumentation as for the WBS_{BS} case applies. However, quite often no break is found, showing that several breaks in close proximity can offset each other leading to no detection.

	Factor copula						Clayton copula					
br	0	1			2	3	0	1			2	3
						BS						
-3						0.004						0.005
-2					0.195	0.002					0.510	0.001
-1		0.007	0.000	0.002	0.000	0.076		0.113	0.000	0.096	0.005	0.131
0	0.971	0.869	0.967	0.862	0.767	0.895	0.954	0.733	0.968	0.735	0.458	0.846
1	0.029	0.121	0.033	0.133	0.038	0.022	0.045	0.145	0.031	0.164	0.027	0.017
2	0.000	0.003	0.000	0.003	0.000	0.001	0.001	0.009	0.001	0.005	0.000	0.000
3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
						WBS						
-3						0.001						0.000
-2					0.004	0.155					0.069	0.073
-1		0.234	0.000	0.158	0.073	0.505		0.387	0.003	0.363	0.211	0.522
0	0.963	0.735	0.962	0.794	0.890	0.329	0.967	0.577	0.957	0.572	0.691	0.392
1	0.036	0.030	0.037	0.049	0.032	0.010	0.030	0.036	0.040	0.065	0.029	0.013
2	0.001	0.001	0.001	0.001	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
						WBS_{BS}						
-3						0.000						0.001
-2					0.001	0.052					0.079	0.016
-1		0.099	0.000	0.049	0.007	0.401		0.289	0.000	0.243	0.082	0.405
0	0.965	0.855	0.968	0.880	0.946	0.530	0.960	0.645	0.965	0.658	0.815	0.561
1	0.034	0.045	0.032	0.069	0.043	0.016	0.039	0.064	0.035	0.094	0.023	0.016
2	0.000	0.001	0.000	0.002	0.003	0.001	0.001	0.002	0.000	0.005	0.001	0.001
3	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 3: Over- and underestimation frequencies of the amount of true change points $\hat{\ell} - \ell^0$ with $\alpha = 0.05$ and $T = 1500$, break fractions for one break: $z_1 = 0.15, 0.5, 0.85$, for two breaks: $z_1 = 0.4, z_2 = 0.6$, for three breaks: $z_1 = 0.3, z_2 = 0.5, z_3 = 0.7$.

This issue is further explained and illustrated in Fryzlewicz (2014), see also the discussion in Bai and Perron (2003), and is a clear disadvantage of the widely used binary segmentation in the case of two breaks. At the same time, this shows the strength of the wild binary segmentation which in this scenario also has a larger detection rate of the correct amount of changes.

Turning to the $3breaks$-scenario, BS correctly detects the amount of breaks more often than the wild binary segmentation algorithms. Under closer inspection, BS usually first finds (if the test statistic is rejected) the middle break z_2; see the change in Spearman's ρ in Table 1. The resulting subintervals contain only one breakpoint each, which the algorithm can easily detect, explaining the results.

Similar comments with respect to the relative performance of the three procedures apply

BS	1 break z_1=0.15	z_1=0.5	z_1=0.85	2 breaks z_1=0.4 z_2=0.6	3 breaks z_1=0.3 z_2=0.5 z_3=0.7
$T = 500$	13.38	3.32	9.21	8.34	16.88
$T = 1000$	6.72	1.80	6.16	3.67	8.55
$T = 1500$	3.99	1.40	3.91	2.28	3.59
WBS					
$T = 500$	11.27	4.50	7.61	16.55	21.55
$T = 1000$	6.27	2.29	3.97	7.78	18.28
$T = 1500$	3.75	1.65	2.77	3.37	14.20
WBS_{BS}					
$T = 500$	8.79	3.88	8.13	14.12	21.52
$T = 1000$	5.21	2.09	4.89	4.44	16.66
$T = 1500$	3.41	1.39	3.43	2.21	10.58

Table 4: Hausdorff distance d_H for the factor copula model using Spearman's ρ and the asymmetry measure averaged over 1000 iterations multiplied by 100

to the results for the Clayton copula in the right panel. Overall, the power is lower in most cases for the Clayton copula. Note that in our setting, the maximal breaks of Spearman's ρ found in a certain group are larger in the factor copula model, while the breaks in the Clayton copula model are larger on average. So, the results indicate that rather the large breaks drive the breakpoint detection. Of course, the difference between the two models might change if other parameter combinations were chosen.

Table 3 shows the results for $T = 1500$ and illustrates that all procedures are consistent for the number of change points.

Following Okui and Wang (2021), among others, Table 4 shows the accuracy of the location of the estimated breaks measured by the average Hausdorff distance d_H of the change point fractions (multiplied by 100), exemplary for the factor copula model:

$$d_H(\boldsymbol{z}^0, \hat{\boldsymbol{z}}) = \max\left\{ \max_{1\le j\le \ell^0} \min_{1\le k\le \hat{\ell}} |z_j^0 - \hat{z}_k|, \max_{1\le k\le \hat{\ell}} \min_{1\le j\le \ell^0} |z_j^0 - \hat{z}_k| \right\}$$

with $\boldsymbol{z}^0 := \{z_1^0, \ldots, z_{\ell^0}^0\}$ and $\hat{\boldsymbol{z}} := \{\hat{z}_1, \ldots, \hat{z}_{\hat{\ell}}\}$. The accuracy measured by the Hausdorff distance confirms our previous findings. Moreover, Table 4 provides finite sample evidence for the consistency of the change point locations for the segmentation algorithms as the

Finance	Allianz, Axa, Banco Santander, BBVA, Deutsche Bank, Deutsche Boerse, BNP Paribas, Generali, ING Groep, Intesa Sanpaolo, Muenchner Rueck., Societe Generale, Unicredit
Energy	E.ON, Enel, Eni, Suez, Iberdrola, Repsol, RWE, TotalEnergies
Telecom and Media	Deutsche Telekom, Orange, Telecom Italia, Telefonica, Vivendi
Consumer Retail	Anheuser-Busch, Carrefour, Danone, L'Oreal, LVMH

Table 5: Included stocks by industry

Hausdorff distance converges to zero with sample size.

4 Empirical Application

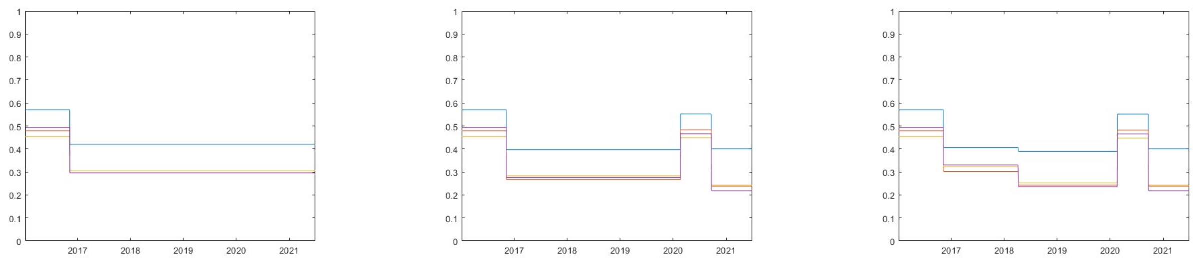
For our empirical application, data from Thomson Reuters Eikon is used that covers the recent COVID-19 pandemic where financial markets plummeted and quickly recovered, giving us reason to believe in at least one change point of the dependence structure between assets due to a diversification meltdown. More specifically, return data of companies from the four largest industry sectors of the EURO STOXX 50 is used from 01.01.2016 until 30.06.2021 resulting in $T = 1{,}433$ trading days and $n = 31$ assets; see Table 5. Serial dependence and GARCH effects of the marginals are filtered out using an AR(1)-GARCH(1,1) specification with t-innovation. Averaging over each of the 31 estimated parameters results in a process of the form $Y_t = 0.0002 - 0.0377 Y_{t-1} + \varepsilon_t$, with $\sigma_t^2 = 0.0000 + 0.8830\sigma_{t-1}^2 + 0.0891\varepsilon_{t-1}^2$. Similarly to the Monte Carlo simulations, we consider Spearman's ρ both alone and in combination with the asymmetry measure. It turns out that the detected dates are the same in both cases. This means that our procedure does not detect any more time-varying behavior in the asymmetry structure compared to the case, where only Spearman's ρ is considered. The p-values are the same in both cases, therefore, for ease of exposition, we focus on Spearman's ρ in the following.

A significance level of $\alpha = 0.05$ is chosen with $B = 100$ bootstrap repetitions. For the

Detector	Dependence measures			
BS	7.11.16			
WBS	7.11.16	20.02.20	21.09.20	
WBS_{BS}	7.11.16	6.04.18	19.02.20	21.09.20

Table 6: Breakpoints for Spearman's ρ and both Spearman's ρ and the asymmetry measure. The detected change points are the same in both cases.

WBS algorithm, 50 random intervals are used. Table 6 shows the change points found for all three procedures. All procedures find a change point at 07.11.2016 which closely coincides

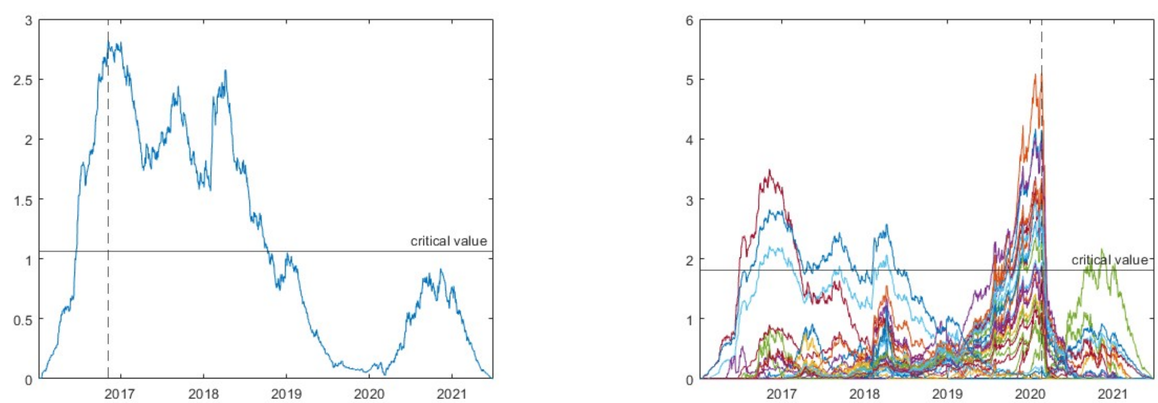


(a) BS (b) WBS (c) WBS_{BS}

Figure 1: Rank correlation averaged among industries
blue: Finance red: Energy
purple: Consumer Retail yellow: Telecom and Media

with the presidential election 2016 and the election of Donald Trump as 45th US President on the 8th of November 2016. The stock market not only in the US but also in Europe reacted bullish after the election. The WBS algorithms additionally find breaks at the beginning of the coronavirus pandemic in accordance with the lockdown of Wuhan in January 2020, while the break date 20.02.2020 coincides with a massive price drop of the EURO STOXX 50. Note, that the second change point candidate of the BS is similar to the one from WBS, namely 23.01.2020, with a p-value of 0.07. Both WBS algorithms also detect a break at 21.09.2020 which coincides with the second wave of the COVID-19 pandemic and the resulting implementation of new restrictions; e.g. 20.09.2020 and 14.10.2020 new restrictions were announced in the UK and France, respectively. The WBS_{BS} algorithm additionally finds a break at 06.04.2018, which corresponds with a sell-off starting on Wall Street and spreading to Europe resulting in a price drop in assets at the beginning of February 2018.

The WBS algorithms find quite similar change points, while the BS algorithm only finds a single change point. To provide some intuition: After BS finds the first change point at the beginning of the data set, a situation with two remaining changes (in the second subsample) emerge for which the binary segmentation either finds both changes or misses both, which is in line with the results of the Monte Carlo simulations for two underlying change points. Figure 1b, 1c show the estimated values of Spearman's ρ between the changes revealing a dependence measure process similar to the data generating process of Section 3. Finally, the p-values of WBS and WBS_{BS} after the change points have been found are 0.39 and 0.22, respectively, suggesting we do not miss any further change point.



(a) BS (b) WBS_{BS}

Figure 2: Test statistic for all time points using (a) binary segmentation and (b) wild binary segmentation with 50 random intervals. The dashed line marks the maximal value of the test statistic and therefore the first change point candidate.

For Spearman's ρ, Figure 2 illustrates the test statistic of the binary (BS) and wild binary segmentation (WBS_{BS}) in the first step to find the first change point. BS finds the first break at the end of 2016. Although the test statistic shows signs of a change at the end of 2020, it does not beomce very large. The test statistic for WBS_{BS} announces a first change point at the beginning of 2020. However, also for the other change points the test statistic exceeds the critical value. Notice that WBS_{BS} includes the test statistic of the binary segmentation.

5 Conclusion

Procedures for the detection of multiple change points for copula-based dependence measures have been analyzed. Focusing on the binary segmentation, consistency results for the amount as well as the location of change points are provided. Binary and wild binary segmentation are used to study change points in Spearman's ρ and an asymmetry measure covering the recent COVID-19 pandemic.

One interesting avenue for future research could be a more detailed investigation of the information criterion of Remark 2 to find a suitable penalty term leading to a more time efficient procedure for finding multiple change points. Moreover, checking for breaks in the marginals in conjunction with a segmentation algorithm to estimate the marginals between the detected changes, could be another interesting question.

A Proofs

A.1 Proof of Proposition 1

Proof. Proposition 1 follows from Lemma A.1 and A.2 below, in conjunction with the proof of Galeano and Wied (2017, Theorem 1).

Lemma A.1

For any given interval $[a, b] \subseteq [0, 1]$,

$$\sup_{z \in [a,b]} \left| \frac{\hat{M}(a, b; z)}{\sqrt{\lfloor bT \rfloor - \lfloor aT \rfloor}} - M(a, b; z) \right| = O_p(1/\sqrt{T}).$$

Proof of Lemma A.1. Consider

$$\frac{\hat{M}(a, b; z)}{\sqrt{\lfloor bT \rfloor - \lfloor aT \rfloor}} = \sqrt{\sum_{g=1}^{G} \sum_{l=1}^{H} \left[\frac{1}{\binom{|\mathcal{G}_g|}{2}} \sum_{\substack{1 \le i < j \le n \\ i,j \in \mathcal{G}_g}} \frac{\lfloor zT \rfloor - \lfloor aT \rfloor}{\lfloor bT \rfloor - \lfloor aT \rfloor} (\hat{m}_{i,j,l}^{\lfloor zT \rfloor + 1 : \lfloor bT \rfloor} - \hat{m}_{i,j,l}^{\lfloor aT \rfloor + 1 : \lfloor bT \rfloor}) \right]^2},$$

fix an arbitrary pair $i, j \in \mathcal{G}_g, 1 \le i < j \le n$, for some $g \in \{1, \ldots, G\}$. Next, choose some $l \in \{1, \ldots, H\}$ and define for any given interval $[a, b] \subseteq [0, 1]$ the *infeasible* dependence measure

$$\tilde{m}_{i,j,l}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor} = \frac{1}{\lfloor bT \rfloor - \lfloor aT \rfloor} \sum_{t = \lfloor aT \rfloor+1}^{\lfloor bT \rfloor} h_l(U_{i,t}, U_{j,t}). \tag{A.1}$$

Note that, by the triangle inequality in conjunction with Assumptions C, one gets

$$\begin{aligned} |\hat{m}_{i,j,l}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor} - \tilde{m}_{i,j,l}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor}| &\le \frac{c_0}{\lfloor bT \rfloor - \lfloor aT \rfloor} \sum_{t = \lfloor aT \rfloor+1}^{\lfloor bT \rfloor} \sum_{k \in \{i,j\}} |\hat{U}_{k,t}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor} - U_{k,t}| \\ &= \frac{c_0}{\lfloor bT \rfloor - \lfloor aT \rfloor} \sum_{t = \lfloor aT \rfloor+1}^{\lfloor bT \rfloor} \frac{1}{\sqrt{T}} \sum_{k \in \{i,j\}} |\hat{\mathbb{F}}_k(a, b, \eta_{k,t})| \end{aligned} \tag{A.2}$$

where

$$\hat{\mathbb{F}}_i(a, b, x_i) = \hat{\mathbb{F}}(a, b, \boldsymbol{x}^i), \quad \boldsymbol{x}^i = (\infty, \ldots, \infty, x_i, \infty, \ldots, \infty)$$

for $i \in \{1, \ldots, n\}$, with $\hat{\mathbb{F}}(a, b, \boldsymbol{x}) = \sqrt{T}(\hat{F}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor} - \mathsf{F})(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathbb{R}^n$ and $(a, b) \in \Delta$, with $\Delta = \Delta_{0,1}$, where $\Delta_{a,b} = \{(\alpha, \beta) \in [a, b]^2 : 0 \le a \le b \le 1\}$ and

$$\hat{F}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor}(\boldsymbol{x}) = \frac{1}{\lfloor bT \rfloor - \lfloor aT \rfloor} \sum_{t=\lfloor aT \rfloor+1}^{\lfloor bT \rfloor} 1\{\hat{\boldsymbol{\eta}}_t \le \boldsymbol{x}\}.$$

Now, by part (A3) of Ass. A and Nasri and Rémillard (2019, Corollary 1), $\mathsf{max}_{k \in \{i,j\}} \mathsf{sup}_{(a,b,x) \in \Delta \times \mathbb{R}} |\hat{\mathbb{F}}_k(a, b, x)| = O_p(1)$. Hence, by Eq. (A.2), $\mathsf{sup}_{(a,b) \in \Delta} |\hat{m}_{i,j}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor} - \tilde{m}_{i,j}^{\lfloor aT \rfloor+1:\lfloor bT \rfloor}| = O_p(1/\sqrt{T})$. Next, assume, for brevity[1], two $\ell^0 = 2$ break points $0 =: z_0^0 < z_1^0 < z_2^0 < z_3^0 := 1$, $n = 2$, and $H = 1$; i.e., $p = 1$. Then either **1.** $a \le z_1^0 < z_2^0 \le b$, **2.** $z_2^0 \le a$,

[1] The following can be readily extended to $\ell^0 > 2$: Similar to the case with $\ell^0 = 2$ break points, one has to distinguish between $\ell^0 + 1 + \ell^0(\ell^0 + 1)/2$ cases; i.e., $\ell^0 + 1$ cases with no breaks located within $[a, z]$ and $\ell^0(\ell^0 + 1)/2$ cases with at least one break located within $[a, z]$.

3. $z_1^0 \le a < z_2^0 \le b$, **4.** $z_1^0 \le a < b \le z_2^0$, **5.** $a \le z_1^0 < b \le z_2^0$, or **6.** $z_1^0 \ge b$. Hence,

$$\begin{aligned}\int_a^z g(t)\,\mathsf{d}t =& 1\{a \le z_1^0 < z_2^0 \le z\}[(z_1^0 - a)\gamma_0^0 + (z_2^0 - z_1^0)\gamma_1^0 + (z - z_2^0)\gamma_2^0] \\ &+ 1\{a \le z_1^0 < z \le z_2^0\}[(z_1^0 - a)\gamma_0^0 + (z - z_1^0)\gamma_1^0] \\ &+ 1\{z_1^0 \le a < z_2^0 \le z\}[(z_2^0 - a)\gamma_1^0 + (z - z_2^0)\gamma_2^0] \\ &+ (b-a)[1\{z_1^0 \ge z\}\gamma_0^0 + 1\{z_1^0 \le a < z \le z_2^0\}\gamma_1^0 + 1\{z_2^0 \le a\}\gamma_2^0].\end{aligned} \tag{A.3}$$

Now, consider

$$\begin{aligned}(\lfloor zT\rfloor - \lfloor aT\rfloor)&\tilde{m}_{i,j}^{\lfloor aT\rfloor+1:\lfloor zT\rfloor} \\ =& 1\{a \le z_1^0 < z_2^0 \le z\}\Big[(\lfloor z_1^0T\rfloor - \lfloor aT\rfloor)\tilde{m}_{i,j}^{\lfloor aT\rfloor+1:\lfloor z_1^0T\rfloor} \\ &+ (\lfloor z_2^0T\rfloor - \lfloor z_1^0T\rfloor)\tilde{m}_{i,j}^{\lfloor z_1^0T\rfloor+1:\lfloor z_2^0T\rfloor} \\ &+ (\lfloor zT\rfloor - \lfloor z_2^0T\rfloor)\tilde{m}_{i,j}^{\lfloor z_2^0T\rfloor+1:\lfloor zT\rfloor}\Big] \\ &+ 1\{a \le z_1^0 < z \le z_2^0\}\Big[(\lfloor z_1^0T\rfloor - \lfloor aT\rfloor)\tilde{m}_{i,j}^{\lfloor aT\rfloor+1:\lfloor z_1^0T\rfloor} \\ &+ (\lfloor zT\rfloor - \lfloor z_1^0T\rfloor)\tilde{m}_{i,j}^{\lfloor z_1^0T\rfloor+1:\lfloor zT\rfloor}\Big] \\ &+ 1\{z_1^0 \le a < z_2^0 \le z\}\Big[(\lfloor z_2^0T\rfloor - \lfloor aT\rfloor)\tilde{m}_{i,j}^{\lfloor aT\rfloor+1:\lfloor z_2^0T\rfloor} \\ &+ (\lfloor zT\rfloor - \lfloor z_2^0T\rfloor)\tilde{m}_{i,j}^{\lfloor z_2^0T\rfloor+1:\lfloor zT\rfloor}\Big] \\ &+ (1\{z_1^0 \ge z\} + 1\{z_1^0 \le a < z \le z_2^0\} + 1\{z_2^0 \le a\})\tilde{m}_{i,j}^{\lfloor aT\rfloor+1:\lfloor zT\rfloor}.\end{aligned} \tag{A.4}$$

Next, introduce the *infeasible* sequential empirical copula process

$$\tilde{\mathbb{B}}_{i,j,t}(a,b,u,v) = \frac{1}{\sqrt{\lfloor bT\rfloor - \lfloor aT\rfloor}} \sum_{t=\lfloor aT\rfloor}^{\lfloor bT\rfloor} (1\{U_{i,t} \le u, U_{j,t} \le v\} - \mathsf{C}_{i,j,t}(u,v)).$$

By part (A1) of Ass. A, one has $\mathsf{C}_{i,j,t} = \mathsf{C}_{i,j,k}$ for all $\lfloor aT\rfloor \le t \le \lfloor bT\rfloor$ and thus $\tilde{\mathbb{B}}_{i,j,t}(\alpha,\beta,u,v) = \tilde{\mathbb{B}}_{i,j,k}(\alpha,\beta,u,v)$ for any $(\alpha,\beta) \in \Delta_{a,b}$, $[a,b] \subseteq [z_k^0, z_{k+1}^0]$, $k \in \{0,1,2\}$, and $m_{i,j,t} = \gamma_k^0 =$

$\int_{[0,1]^2} h(u,v)\,\mathsf{d}\mathsf{C}_{i,j,k}(u,v)$. Moreover, by Fermanian et al. (2004, Theorem 6), one gets

$$\sqrt{T}(\tilde{m}_{i,j}^{\lfloor \alpha T\rfloor+1:\lfloor \beta T\rfloor} - m_{i,j,k}) = \int_{[0,1]^2} \tilde{\mathbb{B}}_{i,j,k}(\alpha,\beta,u,v)\,\mathsf{d}h(u,v) + O(1/\sqrt{T}).$$

Since, by Ass. D, $\tilde{\mathbb{C}}_{i,j,k}$ converges weakly for any $k \in \{1,\dots,\ell\}$ on $\Delta_{a,b} \times [0,1]^2$ to a tight Gaussian process [see also Bücher and Kojadinovic (2016)], the claim follows. □

Lemma A.2

Define $\hat{z}_{a,b} = \operatorname*{arg\,max}_{z\in[a,b]} \hat{M}(a,b;z)$*. For any* $[a,b] \subseteq [0,1]$*, the change point estimator* $\hat{z}_{a,b}$ *is consistent for the dominating change point in* $[a,b]$ *defined as* $z^\star_{a,b} := \operatorname*{arg\,max}_{z\in[a,b]} M(a,b;z)$*.*

A.2 Proof of Proposition 2

Proof. The proof follows from part Lemma B.1 and part (1) of Lemma B.2 below.

Lemma B.1

If there is no break in $[z_k^0, z_{k+1}^0]$*,* $k \in \{0,1,\dots,\ell^0\}$*, then*

$$\max_{k\in\{0,1,\dots,\ell^0\}} \sup_{z\in[z_k^0,z_{k+1}^0]} \hat{M}_T(z_k^0, z_{k+1}^0; z) \xrightarrow{d} \max_{k\in\{0,1,\dots,\ell^0\}} \sup_{z\in[z_k^0,z_{k+1}^0]} \mathbb{M}(z_k^0, z_{k+1}^0; z).$$

Proof of Lemma B.1. Set $\hat{M}_k(z) := \hat{M}(z_k^0, z_{k+1}^0; z)$. Now, for some $k \in \{0,1,\dots,\ell^0\}$, $l \in \{1,\dots,H\}$, $g \in \{1,\dots,G\}$, and any pair $i,j \in \mathcal{G}_g$, $i \neq j$, one has

$$\begin{aligned}\hat{m}_{i,j,l}^{\lfloor z_k^0 T\rfloor+1:\lfloor zT\rfloor} &- \hat{m}_{i,j,l}^{\lfloor z_k^0 T\rfloor+1:\lfloor z_{k+1}^0 T\rfloor} \\ &= \int_{[0,1]^2} (\hat{C}_{i,j}^{\lfloor z_k^0 T\rfloor+1:\lfloor zT\rfloor} - \hat{C}_{i,j}^{\lfloor z_k^0 T\rfloor+1:\lfloor z_{k+1}^0 T\rfloor})(u,v)\,\mathsf{d}h_l(u,v) + O_p(1/T),\end{aligned} \tag{A.5}$$

using Fermanian et al. (2004); see also Berghaus et al. (2017). Moreover, consider

$$\begin{aligned}\hat{\mathbb{D}}_k(z,\boldsymbol{u}) &:= \frac{\lfloor zT\rfloor - \lfloor z_k^0 T\rfloor}{\sqrt{\lfloor z_{k+1}^0 T\rfloor - \lfloor z_k^0 T\rfloor}} (\hat{C}^{\lfloor z_k^0 T\rfloor+1:\lfloor zT\rfloor} - \hat{C}^{\lfloor z_k^0 T\rfloor+1:\lfloor z_{k+1}^0 T\rfloor})(\boldsymbol{u}) \\ &= \hat{\mathbb{C}}_k(z,\boldsymbol{u}) - \varphi_k(z)\hat{\mathbb{C}}_k(z_{k+1}^0,\boldsymbol{u}),\end{aligned} \tag{A.6}$$

where

$$\hat{\mathbb{C}}_k(z,\boldsymbol{u}) := \frac{1}{\sqrt{\lfloor z_{k+1}^0 T\rfloor - \lfloor z_k^0 T\rfloor}} \sum_{t=\lfloor z_k^0 T\rfloor+1}^{\lfloor zT\rfloor} (1\{\hat{\boldsymbol{U}}_t^{\lfloor z_k^0 T\rfloor+1:\lfloor zT\rfloor} \le \boldsymbol{u}\} - \mathsf{C}_{k+1}(\boldsymbol{u})) \tag{A.7}$$

for $z \in [z_k^0, z_{k+1}^0]$. We can deduce from Bücher and Kojadinovic (2016) and Nasri et al. (2022) in conjunction with the continuous mapping theorem, that $\hat{\mathbb{C}}_k \rightsquigarrow \mathbb{C}_k$ in $\ell^\infty([z_k^0, z_{k+1}^0]\times[0,1]^n)$. Hence, $\hat{\mathbb{D}}_k \rightsquigarrow \mathbb{D}_k$ in $\ell^\infty([z_k^0, z_{k+1}^0]\times[0,1]^n)$, where $\mathbb{D}_k(z,\boldsymbol{u}) = \mathbb{C}_k(z,\boldsymbol{u}) - \varphi_k(z)\mathbb{C}_k(z_{k+1},\boldsymbol{u})$. Therefore,

$$\frac{\lfloor zT\rfloor - \lfloor z_k^0 T\rfloor}{\sqrt{\lfloor z_{k+1}^0 T\rfloor - \lfloor z_k^0 T\rfloor}} (\hat{m}_{i,j,l}^{\lfloor z_k^0 T\rfloor+1:\lfloor zT\rfloor} - \hat{m}_{i,j,l}^{\lfloor z_k^0 T\rfloor+1:\lfloor z_{k+1}^0 T\rfloor}) \rightsquigarrow \int_{[0,1]^2} \mathbb{D}_{k,i,j}(z,u,v)\,\mathsf{d}h_l(u,v), \tag{A.8}$$

so that $\sup_{z\in[z_k^0,z_{k+1}^0]} \hat{M}_k(z) \xrightarrow{d} \sup_{z\in[z_k^0,z_{k+1}^0]} \mathbb{M}_k(z)$. Since $\{\mathbb{M}_k(z)\}_{k=0}^{\ell^0}$ are independent and Gaussian, the claim follows by the continuous mapping theorem.

Lemma B.2

(1) *If there is no break in* $[z_k^0, z_{k+1}^0]$ *for all* $k \in \{0,1,\ldots,\ell^0\}$, *then*

$$\max_{0\le k\le \ell} \left| \sup_{z\in[z_k^0,z_{k+1}^0]} \hat{M}(z_k^0, z_{k+1}^0; z) - \sup_{z\in[\hat{z}_k,\hat{z}_{k+1}]} \hat{M}(\hat{z}_k,\hat{z}_{k+1}; z)\right| = O_p(1/\sqrt{T}).$$

(2) *If there is at least one break in* $[z_k^0, z_{k+1}^0]$ *for at least one* $k \in \{0,1,\ldots,\ell\}$, *then*

$$\max_{0\le k\le \ell} \left| \sup_{z\in[z_k,z_{k+1}]} \hat{M}(z_k^0, z_{k+1}^0; z) - \sup_{z\in[\hat{z}_k,\hat{z}_{k+1}]} \hat{M}(\hat{z}_k,\hat{z}_{k+1}; z)\right| = O_p(1).$$

Proof of Lemma B.2. We assume, *w.l.o.g.*, that $\ell^0 = 1$; i.e., $0 =: z_0^0 < z_1^0 < z_2^0 := 1$. Set

$\hat{M}_1(z) \coloneqq \hat{M}(0, z_1^0; z)$, $\hat{M}_{\hat{z}_1}(z) \coloneqq \hat{M}(0, \hat{z}_1; z)$ and define

$$Z_1(\hat{z}_1) \coloneqq \sup_{z \in [0, z_1^0]} \hat{M}_1(z) - \sup_{z \in [0, \hat{z}_1]} \hat{M}_{\hat{z}_1}(z).$$

Next, define $D_T \coloneqq \{z : z \in [0,1], |z - z_1| \leq c_0/T\}$ for some $c_0 \in (0, \infty)$ and note that for any $\delta > 0$ and $\epsilon > 0$

$$\begin{aligned} \mathsf{P}(|Z_1(\hat{z}_1)| > \delta) &\leq \mathsf{P}(\hat{z}_1 \notin D_T) + \mathsf{P}(\{|Z_1(\hat{z}_1)| > \delta\} \cap \{\hat{z}_1 \in D_T\}) \\ &\leq \epsilon + \mathsf{P}\Big(\sup_{\zeta \in D_T} |Z_1(\zeta)| > \delta \Big). \end{aligned} \tag{A.9}$$

Hence, it suffices to show that $|Z_1(\zeta)| = o_p(1)$ uniformly in $\zeta \in D_T$. To see this, note that

$$\begin{aligned} \left| \sup_{z \in [0, z_1]} \hat{M}_1(z) - \sup_{z \in [0, \zeta]} \hat{M}_\zeta(z) \right| &\leq \sup_{z \in [0, z_1]} \left| \hat{M}_1(z) - \hat{M}_\zeta(z) \right| \\ &\quad + \left| \sup_{z \in [0, z_1]} \hat{M}_\zeta(z) - \sup_{z \in [0, \zeta]} \hat{M}_\zeta(z) \right| \coloneqq A(\zeta) + B(\zeta), \end{aligned} \tag{A.10}$$

say. Begin with $A(\zeta)$. **Suppose, *w.l.o.g.*, that $\zeta > z_1$.** Consider

$$\begin{aligned} \left| \hat{M}_1(z) - \hat{M}_\zeta(z) \right| &= \frac{\lfloor zT \rfloor}{\sqrt{\lfloor \zeta T \rfloor}} \left| \sqrt{\frac{\lfloor \zeta T \rfloor}{\lfloor z_1 T \rfloor}} \| \hat{\boldsymbol{m}}^{1:\lfloor zT \rfloor} - \hat{\boldsymbol{m}}^{1:\lfloor z_1 T \rfloor} \|_2 - \| \hat{\boldsymbol{m}}^{1:\lfloor zT \rfloor} - \hat{\boldsymbol{m}}^{1:\lfloor \zeta T \rfloor} \|_2 \right| \\ &\leq \sqrt{1 - \frac{\lfloor z_1 T \rfloor}{\lfloor \zeta T \rfloor}} \hat{M}_1(z) + \frac{\lfloor zT \rfloor}{\sqrt{\lfloor \zeta T \rfloor}} A_1(\zeta), \end{aligned} \tag{A.11}$$

where $A_1(\zeta) \coloneqq \| \hat{\boldsymbol{m}}^{1:\lfloor \zeta T \rfloor} - \hat{\boldsymbol{m}}^{1:\lfloor z_1 T \rfloor} \|_2$; for the inequality, we expanded the term in absolute values on the right-hand side of the equality with $\| \hat{\boldsymbol{m}}^{1:\lfloor zT \rfloor} - \hat{\boldsymbol{m}}^{1:\lfloor z_1 T \rfloor} \|_2$, we used the (reverse) triangle inequality and $\sqrt{u} - \sqrt{v} \leq \sqrt{u - v}$ for $u \geq v \geq 0$. It follows that $\sup_{z \in [0, z_1]} \hat{M}_1(z) = O_p(1)$ if there is no break in $[0, z_1]$ or $\sup_{z \in [0, z_1]} \hat{M}_1(z)/\sqrt{T} = O_p(1)$ if there is at least one break in $[0, z_1]$. Turning to $A_1(\zeta)$, the triangle inequality and $\|\boldsymbol{v}\|_2 \leq n \max_{1 \leq i \leq n} |v_i|$ for any $\boldsymbol{v} \in \mathbb{R}^n$

yields

$$A_1(\zeta) \leq p \max_{1\leq l \leq H} \max_{1 \leq g \leq G} \max_{i,j \in \mathcal{G}_g} \Bigg[\frac{1}{\lfloor \zeta T \rfloor} \sum_{t=\lfloor z_1 T \rfloor}^{\lfloor \zeta T \rfloor} \left| h_l(\hat{U}_{i,t}^{1:\lfloor \zeta T \rfloor}, \hat{U}_{j,t}^{1:\lfloor \zeta T \rfloor}) \right| + \left(\frac{\lfloor \zeta T \rfloor}{\lfloor z_1 T \rfloor} - 1 \right) \frac{1}{\lfloor z_1 T \rfloor} \sum_{t=1}^{\lfloor z_1 T \rfloor} \left| h_l(\hat{U}_{i,t}^{1:\lfloor \zeta T \rfloor}, \hat{U}_{j,t}^{1:\lfloor \zeta T \rfloor}) \right| + \frac{1}{\lfloor z_1 T \rfloor} \sum_{t=1}^{\lfloor z_1 T \rfloor} \left| h_l(\hat{U}_{i,t}^{1:\lfloor \zeta T \rfloor}, \hat{U}_{j,t}^{1:\lfloor \zeta T \rfloor}) - h_l(\hat{U}_{i,t}^{1:\lfloor z_1 T \rfloor}, \hat{U}_{j,t}^{1:\lfloor z_1 T \rfloor}) \right| \Bigg]$$

say. It follows from Ass. C that $\max_{1\leq l \leq H} \sup_{(u,v)\in[0,1]^2} |h_l(u,v)| \leq c_0$ for some $c_0 \in (0,\infty)$. Hence the first to summands on the preceding display are $O(1/T)$ uniformly in $\zeta \in D_T$. Turning to the third summand, note that, by Ass. (C2), one gets for some $c_0 \in (0,\infty)$

$$\frac{1}{\lfloor z_1 T \rfloor} \sum_{t=1}^{\lfloor z_1 T \rfloor} \left| h_l(\hat{U}_{i,t}^{1:\lfloor \zeta T \rfloor}, \hat{U}_{j,t}^{1:\lfloor \zeta T \rfloor}) - h_l(\hat{U}_{i,t}^{1:\lfloor z_1 T \rfloor}, \hat{U}_{j,t}^{1:\lfloor z_1 T \rfloor}) \right| \leq 2c_0 \max_{k\in\{i,j\}} \sup_{x\in\mathbb{R}} \left| \hat{F}_k^{1:\lfloor \zeta T \rfloor}(x) - \hat{F}_k^{1:\lfloor z_1 T \rfloor}(x) \right| = O_p(1/T). \tag{A.12}$$

Putting the above together yields, uniformly in $\zeta \in D_T$, $A(\zeta) = O_p(1/\sqrt{T})$ if there is no break and $A(\zeta) = O_p(1)$ if there is at least one break in $[0,1]$. Turning to $B(\zeta)$, a simple calculation reveals that

$$B(\zeta) \leq \sup_{z\in[z_1,\zeta]} \left| \hat{M}_\zeta(z) - \hat{M}_\zeta(z_1) \right| \leq \sup_{z\in[z_1,\zeta]} \frac{\lfloor zT \rfloor}{\sqrt{\zeta T}} \| \hat{\boldsymbol{m}}^{1:\lfloor zT \rfloor} - \hat{\boldsymbol{m}}^{1:\lfloor z_1 T \rfloor} \|_2 = O_p(1/\sqrt{T}),$$

noting that there is no nontrivial break in $[z_1, \zeta]$.

A.3 Proof of Proposition 3

The proof follows from Lemma B.2 using the same arguments used in the proof of Galeano and Wied (2017, Theorem 3).

References

Andrews, D. W. and Lu, B. (2001). Consistent model and moment selection procedures for gmm estimation with application to dynamic panel data models. *Journal of Econometrics*, 101:123–164.

Andrews, D. W. K. (1999). Consistent moment selection procedures for generalized method of moments estimation. *Econometrica*, 67:543–563.

Aue, A., Hörmann, S., Horváth, L., and Reimherr, M. (2009). Break detection in the covariance structure of multivariate time series models. *The Annals of Statistics*, 37:4046–4087.

Bai, J. (1997). Estimating multiple breaks one at a time. *Econometric Theory*, 13:315–352.

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66:47–78.

Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1):1–22.

Barassi, M., Horváth, L., and Zhao, Y. (2020). Change-point detection in the conditional correlation structure of multivariate volatility models. *Journal of Business & Economic Statistics*, 38:340–349.

Berens, T., Weiß, G., and Wied, D. (2015). Testing for structural breaks in correlations: Does it improve value-at-risk forecasting? *Journal of Empirical Finance*, 32:135–152.

Berghaus, B., Bücher, A., and Volgushev, S. (2017). Weak convergence of the empirical copula process with respect to weighted metrics. *Bernoulli*, 23:743–772.

Brown, R., Durbin, J., and Evans, J. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society Series B*, 37:149–192.

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

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

Chen, X. and Fan, Y. (2006). Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics*, 135:125–154.

Fermanian, J.-D., Radulovic, D., and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli*, 10:847–860.

Francq, C. and Zakoïan, J.-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10:605–637.

Fryzlewicz, P. (2014). Wild binary segmentation for multiple change-point detection. *The Annals of Statistics*, 42:2243–2281.

Fryzlewicz, P. (2020). Detecting possibly frequent change-points: Wild binary segmentation 2 and steepest-drop model selection. *Journal of the Korean Statistical Society*, 49:1027–1070.

Galeano, P. and Tsay, R. S. (2009). Shifts in individual parameters of a GARCH model. *Journal of Financial Econometrics*, 8:122–153.

Galeano, P. and Wied, D. (2014). Multiple break detection in the correlation structure of random variables. *Computational Statistics & Data Analysis*, 76:262–282.

Galeano, P. and Wied, D. (2017). Dating multiple change points in the correlation matrix. *Test*, 26:331–352.

Kojadinovic, I., Quessy, J.-F., and Rohmer, T. (2016). Testing the constancy of spearman's rho in multivariate time series. *Annals of the Institute of Statistical Mathematics*, 68:929–954.

Liebscher, E. (2008). Construction of asymmetric multivariate copulas. *Journal of Multivariate Analysis*, 99:2234–2250.

Longin, F. and Solnik, B. (1995). Is the correlation in international equity returns constant: 1960–1990? *International Money and Finance*, 14:3–26.

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

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7:77–91.

Nasri, B. R. and Rémillard, B. N. (2019). Copula-based dynamic models for multivariate time series. *Journal of Multivariate Analysis*, 172:107–121.

Nasri, B. R., Rémillard, B. N., and Bahraoui, T. (2022). Change-point problems for multivariate time series using pseudo-observations. *Journal of Multivariate Analysis*, 187:104857.

Nelsen, R. B. (2006). *An introduction to copulas.* Springer, New York.

Oh, D. H. and Patton, A. J. (2013). Simulated method of moments estimation for copula-based multivariate models. *Journal of the American Statistical Association*, 108:689–700.

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

Oh, D. H. and Patton, A. J. (2021). Dynamic factor copula models with estimated cluster assignments. *Journal of Econometrics, forthcoming.*

Okui, R. and Wang, W. (2021). Heterogeneous structural breaks in panel data models. *Journal of Econometrics*, 220:447–473.

Opschoor, A., Lucas, A., Barra, I., and van Dijk, D. (2021). Closed-form multi-factor copula models with observation-driven dynamic factor loadings. *Journal of Business & Economic Statistics*, 39:1066–1079.

Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, 47:527–556.

Penzer, J., Schmid, F., and Schmidt, R. (2012). Measuring large comovements in financial markets. *Quantitative Finance*, 12:1037–1049.

Ploberger, W. and Krämer, W. (1992). The cusum test with ols residuals. *Econometrica*, 60(2):271–285.

Quessy, J.-F. (2009). Theoretical efficiency comparisons of independence tests based on multivariate versions of spearman's rho. *Metrika*, 70:315–338.

Roberts, S. (1959). Control chart tests based on geometric moving averages. *Technometrics*, 1(3):239–250.

Rosco, J. and Joe, H. (2013). Measures of tail asymmetry for bivariate copulas. *Statistical Papers*, 54:709–726.

Schmid, F. and Schmidt, R. (2007). Multivariate conditional versions of spearman's rho and related measures of tail dependence. *Journal of Multivariate Analysis*, 98:1123–1140.

Segers, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli*, 18:764–782.

Shewhart, W. (1931). *Economic Control of Quality of Manufactured Products*. D. Van Nostrand Company, New York.

Shiryaev, A. (1963). On optimum methods in quickest detection problems. *Theory of Probability and Its Applications*, 8:22–46.

Stark, F. and Otto, S. (2022). Testing and dating structural changes in copula-based dependence measures. *Journal of Applied Statistics*, 49:1121–1139.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak convergence and empirical processes.* Springer, New York.

Vostrikova, L. (1981). Detecting 'disorder' in multidimensional random processes. *Sov. Dokl. Math.*, 24:55–59.

Wied, D., Dehling, H., van Kampen, M., and Vogel, D. (2014). A fluctuation test for constant spearman's rho with nuisance-free limit distribution. *Computational Statistics & Data Analysis*, 76:723–736.

Wied, D., Krämer, W., and Dehling, H. (2012). Testing for a change in correlation at an unknown point in time using an extended functional delta method. *Econometric Theory*, 28:570–589.