Testing for Structural Breaks in Factor Copula Models

Hans Manner, Florian Stark and Dominik Wied

Abstract

We propose new fluctuation tests for detecting structural breaks in factor copula models and analyze the behavior under the null hypothesis of no change. In the model, the joint copula is given by the copula of random variables which arise from a factor model. This is particularly useful for analyzing data with high dimensions. Parameters are estimated with the simulated method of moments (SMM). Due to the discontinuity of the SMM objective function, it is not trivial to derive a functional limit theorem for the parameters. We analyze the behavior of the tests in Monte Carlo simulations and a real data application. In particular, it turns out that our test is more powerful than nonparametric tests for copula constancy in high dimensions.

Keywords: Factor Copula Model, Fluctuation Test, Simulated Method of Moments

This Version: August 23, 2017., Universität zu Köln, Institut für Ökonometrie und Statistik, E-Mail: `manner@statistik.uni-koeln.de (Hans Manner)`, `fstark3@uni-koeln.de (Florian Stark)`, `dwied@uni-koeln.de (Dominik Wied)`. Research is supported by Deutsche Forschungsgemeinschaft (DFG grant "Strukturbrüche und Zeitvariation in hochdimensionalen Abhängigkeitsstrukturen").

1. INTRODUCTION

Analyzing time-variant parameters in models for financial data e.g. returns, is a current research topic. In particular, several tests for constant dependencies have recently been developed, see e.g. Bücher and Ruppert (2013) for the case of copulas Dehling, Vogel, Wendler, and Wied (2016) for the case of Kendall's tau. The main motivation for such tests is that dependencies usually increase in times of crises. So, they can be applied to detect and quantify contagion between different financial markets or to construct optimal portfolios in portfolio management.

In this paper, we consider factor copula models which have been recently proposed by Oh and Patton (2013) and Krupskii and Joe (2013), whereas we focus on the first approach. In such models, the joint copula between random variables is given by the copula of random variables which arise from a factor model. The variable parameters are factor loadings and possible distribution parameters of factor and error term distributions.

The advantage of these models is that they can be used in relatively high dimensional applications and nevertheless capture the dependence structure by a low number of parameters. In the suitability for high-dimensional data, factor copula models are similar to hierarchical Archimedean copulas (see Savu and Trede, 2010) and vine copulas (see Bedford and Cooke, 2002). We focus on factor copula models to have both considerable model flexibility and the possibility to perform statistical inference.

For the estimation of the model parameters, we use the simulated method of moments (SMM), which is different to standard method of moments applications, since the theoretical moment-counterparts are simulated and not as usual analytically derived. This makes asymptotic theory more difficult, as deriving consistency and asymptotic distribution results of the estimators. The reason is, that the objective function is not continuous and furthermore not differentiable in the parameters and standard asymptotic approaches can not be used here.

We propose a new fluctuation test, where successively parameter estimators are compared to the parameter estimates of the full sample and we then analyze the behavior of the test under the null hypothesis of no change. In contrast to formerly proposed nonparametric tests for constant copulas by e.g. (Bücher, Kojadinovic, Rohmer, and Segers, 2014), our test is of parametric nature.

It is not trivial to derive the asymptotic distribution of the test statistic. Due to the non-smoothness of the objective function, we can not make use of a Taylor expansion approach to derive the distribution under the null. To tackle this issue we propose a new construction principle inspired by (Newey and McFadden, 1994). These new functional limit theorems hold in general for SMM estimation and are therefore of broader interest.

We analyze size and power properties of our test in Monte Carlo simulation in various situations. Moreover, we compare the fluctuation test based on parameter estimators with a test based on the moment functions and with the (Bücher et al., 2014)-test. While (Bücher et al., 2014) has better properties for low dimensions, our test performs better in high dimensions, if the number of parameters is kept fixed. This reflects the fact that the drawback of having to estimate the model with simulated methods is more and more compensated with increasing dimensions. If the number of dimensions is kept fixed, one simply has more data for estimating the model, while, on the other hand, in a nonparametric copula constancy test, the complexity of the estimated objects increase. Finally, we use the test in a real-data application on daily returns of international banks.

2. TESTING FOR CONSTANCY OF FACTOR COPULA MODELS

We first describe the factor copula model and its estimation, before we turn to the problem of testing for structural changes.

2.1. Factor copula models and SMM estimation

In this article we consider the same class of data-generating process as in Oh and Patton (2013). In this class the dynamic of the marginal distributions is covered by a parameter vector ϕ_0 and each variable can have time varying conditional mean $\mu_t(\phi_0)$ and variance $\sigma_t(\phi_0)$. The dynamic of the joint distribution of the residuals η_t, namely the copula $C(., \theta_t)$, is covered by the unknown parameters θ_t for $t = 1, \dots, T$. The data-generating process is given by

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

with varying conditional mean $\boldsymbol{\mu}_t(\phi_0) := [\mu_{1t}(\phi_0), \dots, \mu_{Nt}(\phi_0)]'$, varying conditional variance $\boldsymbol{\sigma}_t(\phi_0) := \mathsf{diag}\{\sigma_{1t}(\phi_0), \dots, \sigma_{Nt}(\phi_0)\}$ and $[\eta_{1t}, \dots, \eta_{Nt}] =: \boldsymbol{\eta}_t \overset{\text{iid}}{\sim} \boldsymbol{F}_\eta = C(F_1(\eta_1), \dots, F_N(\eta_N); \theta_t)$, with marginal distributions F_i, where $\boldsymbol{\mu}_t$ and $\boldsymbol{\sigma}_t$ are $\mathcal{F}_{t-1}$-measurable and independent of η_t. $\mathcal{F}_{t-1}$ is the sigma field containing information from the past $\{\boldsymbol{Y}_{t-1}, \boldsymbol{Y}_{t-2}, \dots\}$. Note that the $r \times 1$ vector ϕ_0 is $\sqrt{T}$ consistently estimable, which is fulfilled by many multivariate time series models, e.g. ARCH and GARCH models and the estimator is denoted as $\hat{\phi}$.

We are interested in estimating the $p \times 1$ vectors $\theta_t \in \Theta$ of the copula, by the residual information $\{\hat{\boldsymbol{\eta}}_t := \boldsymbol{\sigma}_t^{-1}(\hat{\boldsymbol{\phi}})[\boldsymbol{Y}_t - \boldsymbol{\mu}_t(\hat{\phi})]\}_{t=1}^T$ from the data and information genereated by simulations from the factor copula model $C(., \theta_t)$ for all t, implied by the following factor structure

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

with $X_{it} = \sum_{k=1}^{K} \beta_{ik}^t Z_{kt} + q_{it}$, where $q_{it} \overset{iid}{\sim} F_{\boldsymbol{q}_t}(\alpha_{\boldsymbol{q}_t})$ and $Z_{kt} \overset{init}{\sim} F_{\boldsymbol{Z}_{kt}}(\gamma_{kt})$ for $i = 1, \dots, N$, $t = 1, \dots, T$ and $k = 1, \dots, K$. Note that Z_{kt} and q_{it} are independent $\forall i, k, t$ and the Copula for $\boldsymbol{X}_t$ is given by

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

with marginal distributions $G_{it}(.,\theta_t)$ and $\theta_t = \left[\{\{\beta_{ik}^t\}_{i=1}^N\}_{k=1}^K, \alpha_{\boldsymbol{q}_t}', \gamma_{1t}', \ldots, \gamma_{Kt}'\right]'$. In principle we allow θ_t to be time-varying, having a piecewise constant model in mind. We assume that this implied copula governs the dependence of $\boldsymbol{Y}_t$, noting that we are not interested in the marginals G_{it} itself. For the estimation, we use the simulated method of moments (SMM) to obtain estimators $\theta_{sT,S}$ of $\theta_{\lfloor sT \rfloor} = \theta_t$. The estimators are defined as

$$\theta_{sT,S} := \underset{\theta\in\Theta}{\mathsf{arg\,min}}\; Q_{sT,S}(\theta),$$

where $Q_{sT,S}(\theta) := g_{sT,S}(\theta)'\hat{W}_{sT} g_{sT,S}(\theta)$, $g_{sT,S}(\theta) := \hat{m}_{sT} - \tilde{m}_S(\theta)$ and $\hat{W}_{sT}$ a positive definite weight matrix. $\hat{m}_{sT}$ are $k\times 1$ vectors of averaged pairwise dependence measures $\hat{m}_{sT}^{ij}$, computed with the residuals $\{\hat{\eta}_t\}_{t=1}^{\lfloor sT \rfloor}$ and $\tilde{m}_S(\theta)$ is the corresponding vector of dependence measures, computed with $\{\tilde{\eta}_t\}_{t=1}^S$, using S simulations from $\boldsymbol{F}_{\boldsymbol{X}_t}$, where $\hat{m}_{sT}^{ij}$ consisting of the pairwise dependence measures defined below. For the dependence measures of the pair (η_i, η_j), we use Spearman's rank ρ^{ij} and quantile dependence λ_q^{ij}, these are defined as

$$\begin{aligned}
\rho^{ij} :=& 12\int_0^1\int_0^1 C_{ij}(u_i,v_j)du_i dv_j - 3\\
\lambda_q^{ij} :=& \begin{cases} P[F_i(\eta_i)\le q|F_j(\eta_j)\le q] = \frac{C_{ij}(q,q)}{q}, & q\in(0,0.5]\\ P[F_i(\eta_i)> q|F_j(\eta_j)> q] = \frac{1-2q+C_{ij}(q,q)}{1-q}, & q\in(0.5,1)\end{cases}.
\end{aligned}$$

The sample counterparts are defined as

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

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

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

2.2. Null hypothesis and testing setup

We are interested in testing

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

with the test statistic $S_{T,S}$, defined as

$$S_{T,S} := \sup_{s \in [\varepsilon,1]} P_{sT,S} := \sup_{s \in [\varepsilon,1]} s^2 T(\theta_{sT,S} - \theta_{T,S})'(\theta_{sT,S} - \theta_{T,S}) \tag{2.1}$$

$$\simeq \max_{\lfloor \varepsilon T \rfloor \leq t \leq T} \left(\frac{t}{T}\right)^2 T(\theta_{t,S} - \theta_{T,S})'(\theta_{t,S} - \theta_{T,S}), \tag{2.2}$$

where $\theta_{sT,S}$ is the SMM estimator up to the information at time point $t = \lfloor sT \rfloor$, T the sample size of the data, S the number of simulations in the SMM and $\varepsilon > 0$. Note that analytically ε has to be choosen strictly greater than zero and thus $s \in [\varepsilon, 1]$ to apply the required limit theorems for our proof of the asymptotic distribution. In the finite sample case ε should be choosen in a way, that we have enough information up to time point $\lfloor \varepsilon T \rfloor$, to receive reasonable estimators.

The null hypothesis of no parameter change is rejected if the test statistic (2.1) is "too large", i.e., if the successively estimated parameters fluctuate too much over time. In the following subsection, we will derive the analytical limit of the test statistic under the null hypothesis. For applications, we propose to use a bootstrap approximations, which is described in the after next subsection.

2.3. Asymptotic analysis

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

Assumption 1.

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

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

The assumption is similar to Assumption 1 in (Oh and Patton, 2013), but the assumption on the copula is relaxed.

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

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

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

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

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

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

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

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

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

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

This assumption is similar to Assumption 2 in (Oh and Patton, 2013), only part (i) is more restrictive. We need this because we consider successively estimated parameters.

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

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

ii) The space Θ of all θ is compact.

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

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

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

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

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

iii) $\forall s \in [\varepsilon, 1], \varepsilon > 0$: $g_{sT,S}(\theta_{sT,S})'\hat{W}_{sT}g_{sT,S}(\theta_{sT,S}) \leq \inf_{\theta\in\Theta} g_{sT,S}(\theta)'\hat{W}_{sT}g_{sT,S}(\theta) + o_p^*((s^2T)^{-1})$, where $o_p^*((s^2T)^{-1})$ converges on the right hand side to zero and is therefore strictly positive.

With these assumptions, we can formulate our main theorem:

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

$$s\sqrt{T}\,(\theta_{sT,S} - \theta_0) \overset{d}{\Longrightarrow} A^*(s)$$

as $T, S \to \infty$ in the space of Càdlàg functions on the interval $[\epsilon, 1]$ and $\frac{S}{T} \to k \in (0, \infty)$ or $\frac{S}{T} \to \infty$. Here, $A^*(s) = (G'WG)^{-1}\,G'W(A(s) - \frac{s}{\sqrt{k}}A(1))$, $A(s)$ is a Gaussian process defined in the proof of Lemma 7 in the appendix and θ_0 the value of all θ_t under the null.

With Theorem 1 we obtain the asymptotic distribution under the null of our test statistic.

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

$$S_{T,S} = \sup_{s\in[\varepsilon,1]} s^2T(\theta_{sT,S} - \theta_{T,S})'(\theta_{sT,S} - \theta_{T,S}) \overset{d}{\longrightarrow} \sup_{s\in[\varepsilon,1]} (A^*(s) - sA^*(1))'(A^*(s) - sA^*(1))$$

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

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

$$\tilde{s} = \operatorname*{argmax}_{s\in[\varepsilon,1]} s^2T(\theta_{sT,S} - \theta_{T,S})'(\theta_{sT,S} - \theta_{T,S}).$$

2.4. Bootstrap distribution

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

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

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

iii) Calculate the distribution bootstrap version of our test statistic

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

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

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

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

iv) Compute B versions of $K^{(b)}$ and determine the critical value K by e.g.

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

Note that by construction, the bootstrap distribution of the test statistic is mainly obtained by calculating B versions of the moment process $\frac{t}{T} \sqrt{T} \left(\hat{m}_t^{(b)} - \hat{m}_T^{(b)} \right)$, which can be calculated fast and direct from the data and it is therefore not necessary to solve B minimization problems, like on the left hand side, which requires a high computing effort.

3. SIMULATION EVIDENCE

We now want to investigate the size and power of our constructed test, therefore we consider the simple factor copula model

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

where $Z_t \overset{init}{\sim} \mathsf{Skew\ t}\,(\sigma^2, \nu^{-1}, \lambda)$ and $q_t \overset{iid}{\sim} \mathsf{t}\,(\nu^{-1})$ for $t = 1, \ldots, T$ and we fix $\sigma^2 = 1$, $\nu^{-1} = 0.25$ and $\lambda = -0.5$, such that our model is parametrized by the factor loading parameter $\theta_t = \beta_t$. For the estimation of the sequential parameters θ_t for $t = \varepsilon T, \ldots, T$ in the test statistic, we use the SMM approach, with $S = 25 \cdot T$ simulations to match the simulated dependence measures values with the dependence measures values of the data, with sample size T. For this we use five dependence measures, namely Spearman's rank correlation and the 0.05, 0.10, 0.90, 0.95 quantile dependence measures, averaged across all pairs. Note that the burn in period $\lfloor \varepsilon T \rfloor$ has to be choosen sufficiently large, to derive reasonable estimators for $\theta_{\lfloor \varepsilon T \rfloor}$ in our test statistic. Table 1 and Table 2 show the strong fluctuations in the $\theta_{T,S}$ for small sample sizes, we therefore recommend to choose $\varepsilon = 0.2$, as this leads to reasonable size and power properties, as Table 3 and Table 4 show.

T	50	100	150	200	300	350	400	1000
Min	-1.431	-1.282	-1.037	- 0.793	0.712	0.697	0.702	0.817
Max	1.724	1.641	1.522	1.615	1.415	1.333	1.392	1.200
Median	0.859	0.977	0.958	0.984	0.986	0.974	0.987	0.995
Bias	0.220	0.043	0.042	0.010	0.006	0.018	0.007	0.004
relative Bias	0.220	0.043	0.042	0.010	0.006	0.018	0.007	0.004

Table 1: Statistical analysis of $\theta_{T,S}$ for different sample size T, $N = 5$ and $\theta_0 = 1$ for 1000 simulation runs.

T	50	100	150	200	300	350	400	1000
Min	-0.656	-0.546	-0.565	-0.521	-0.450	-0.446	0.424	0.405
Max	1.084	0.930	0.872	0.946	0.789	0.749	0.735	0.639
Median	0.414	0.479	0.477	0.489	0.493	0.485	0.492	0.496
Bias	0.168	0.060	0.049	0.021	0.006	0.010	0.004	0.002
relative Bias	0.338	0.121	0.098	0.041	0.013	0.022	0.008	0.004

Table 2: Statistical analysis of $\theta_{T,S}$ for different sample size T, $N = 5$ and $\theta_0 = 0.5$ for 1000 simulation runs.

For our purpose, we chose the number of bootstrap replications $B = 1000$, to capture the asymptotic behavior of the distribution and we test the behavior under the null for $\theta_0 = 1$ and $\theta_0 = 0.5$. In the course of this, we vary the sample size T and the dimension N.

To demonstrate the size and the power of the test, we choose the 0.95 quantile of the bootstrap distribution and repeat the test 301 times. The results of the rejection rate under the null are presented in Table 3 for $\theta_0 = 1$ and $\theta = 0.5$, for various combinations of the sample size T and dimension N.

	$\theta_0 = 1$	$N = 5$	$N = 10$	$N = 20$	$\theta_0 = 0.5$	$N = 5$	$N = 10$	$N = 20$
$T = 500$								
	θ_t	0.066	0.056	0.053	θ_t	0.102	0.079	0.056
	m_t	0.030	0.039	0.056	m_t	0.029	0.036	0.046
	B	0.049	0.053	0.049	B	0.046	0.036	0.033
$T = 1000$								
	θ_t	0.056	0.046	0.069	θ_t	0.089	0.049	0.049
	m_t	0.049	0.043	0.076	m_t	0.046	0.033	0.056
	B	0.066	0.056	0.076	B	0.043	0.046	0.056
$T = 1500$								
	θ_t	0.056	0.069	0.066	θ_t	0.073	0.059	0.043
	m_t	0.049	0.063	0.066	m_t	0.056	0.056	0.049
	B	0.053	0.069	0.066	B	0.046	0.056	0.069

Table 3: Rejection rate for $\theta_0 = 1.0$ and $\theta_0 = 0.5$ for the parameter Test (θ_t) with $\varepsilon = 0.2$, the moment function test (m_t) and the nonparametric test of Bücher et al. (B)

Table 3 reveals that for increasing sample size and dimension the rejection rate tends more and more to 0.05. Note that a higher burn in period leads to a slightly better asymptotic behaviour, especially for smaller T and N, due to fewer variation in the numerical minimization procedure and distribution estimation.

The increasing rejection rate for some combinations of T and N can be reasoned by the non optimal choice of the step size $\varepsilon_{T,S}$ in the estimation of the non continuous derivative matrix $g_{T,S}$, see Oh and Patton (2015). In the computation of the numerical derivative we fix $\varepsilon_{T,S}$ to 0.1 for all combinations, which is much higher than standard step sizes used in computing numerical derivatives, this can lead to a smaller valued scale matrix $(\hat{G}'\hat{W}_T\hat{G})^{-1}\hat{G}'\hat{W}_T$ and hence smaller quantile values yielding a higher rejection rate, for some combinations of T and N. Nevertheless it is necessary in order to handle the non smoothness of $g_{T,S}$. For a optimal choice of the step size see Hong, Mahajan, and Nekipelov (2015). Furthermore, note that the non smoothness of $g_{T,S}$ leads to a non smooth objective function $Q_{T,S}$, see Figure 3.1, which is difficult to optimize with standard minimization procedures, leading to problems in the parameter estimation.

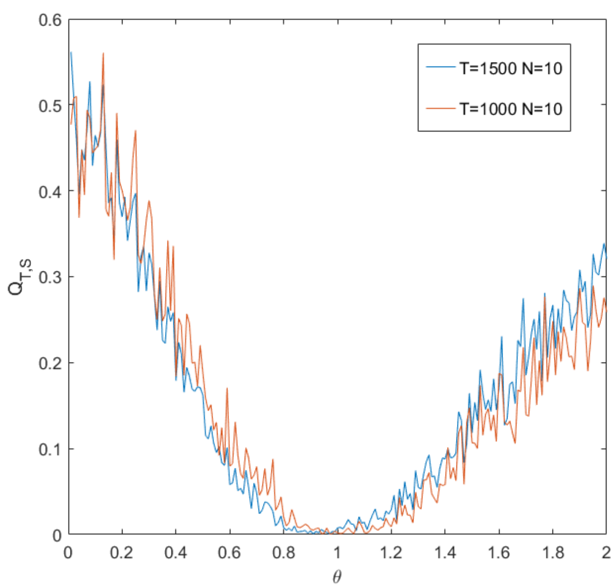

Figure 3.1: Objective function $Q_{T,S}(\theta)$ for T=1000/1500 and N=10.

To emphasize the power of the test, we generate data with a break point at $\frac{T}{2}$ for different sample sizes, where the data is simulated with $\theta_t = 1$ for $t \in \{\varepsilon T, \ldots, \frac{T}{2}\}$, denoted as θ_0 and after that, we increase the parameter to $\theta_t = \{1.2, 1.4, 1.6, 1.8, 2.0\}$ for $t \in \{\frac{T}{2} + 1, \ldots, T\}$, denoted as θ_1. Table 4 and Figure 3.2 reveal, that the power of the test increases clearly with a larger sample size and at the same time the size is improved, nevertheless the dimension N is small.

	N=5	$\theta_0 = 1$	$\theta_1 = 1.2$	$\theta_1 = 1.4$	$\theta_1 = 1.6$	$\theta_1 = 1.8$	$\theta_1 = 2.0$
$T = 500$							
	θ_t	0.066	0.272	0.551	0.833	0.963	0.993
	m_t	0.030	0.173	0.452	0.771	0.940	0.987
	B	0.049	0.272	0.727	0.946	0.996	1.000
$T = 1000$							
	θ_t	0.056	0.352	0.781	0.980	1.000	1.000
	m_t	0.049	0.285	0.717	0.966	1.000	1.000
	B	0.066	0.481	0.946	1.000	1.000	1.000
$T = 1500$							
	θ_t	0.056	0.488	0.950	1.000	1.000	1.000
	m_t	0.049	0.382	0.923	0.996	1.000	1.000
	B	0.053	0.667	0.996	1.000	1.000	1.000

Table 4: Rejection rate for $\theta_0 = 1.0$ and $N = 5$ for the parameter test (θ_t) with $\varepsilon = 0.2$, the moment function test (m_t) and the nonparametric test of Bücher et al. (B)

	N=40	$\theta_0 = 1$	$\theta_1 = 1.2$	$\theta_1 = 1.4$	$\theta_1 = 1.6$	$\theta_1 = 1.8$	$\theta_1 = 2.0$
$T = 500$							
	θ_t	0.043	0.302	0.691	0.910	0.996	1.000
	B	0.059	0.225	0.588	0.903	0.996	1.000

Table 5: Rejection rate for $\theta_0 = 1.0$ and $N = 40$ for the parameter test (θ_t) with $\varepsilon = 0.2$ and the nonparametric test of Bücher et al. (B)

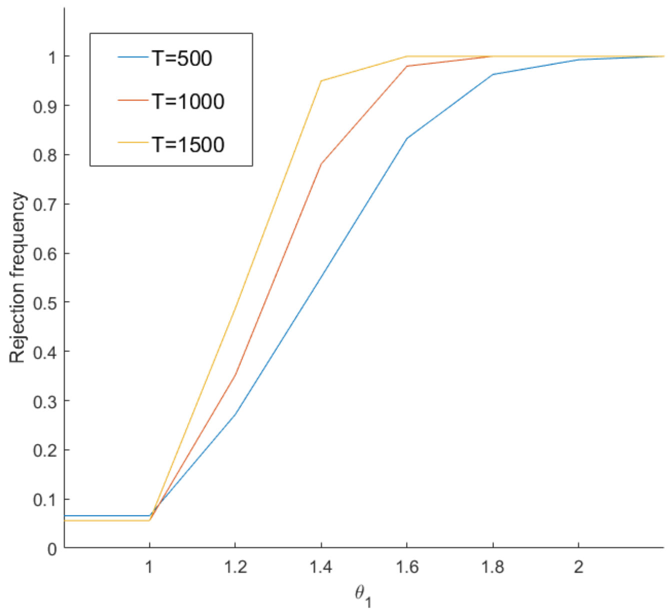

Figure 3.2: Power function for $\theta_0 = 1, T = 500/1000/1500, N = 5$ and $\varepsilon = 0.2$.

Another possibility is to directly test on the moment functions, consisting of the dependence measures, since the distribution is already known. We therefore use

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

$$\xrightarrow{d} \sup_{s\in[\varepsilon,1]} (A(s) - sA(1))'(A(s) - sA(1)) \quad T, S \to \infty,$$

where $A(s)$ is the Gaussian process defined in the proof of Lemma 7 in the appendix. To demonstrate the size and the power of the test, we again chose the 0.95 quantile of the bootstrap distribution and repeat the test 301 times. The results of the rejection rate under the null are presented in Table 3 for $\theta_0 = 1$ and $\theta = 0.5$. Table 3 reveals that for increasing sample size and dimension the rejection rate tends to 0.05. Note that the results for the size are mostly better then the results of the parameter test, reasoned by the existing fluctuation of the parameter and derivation estimation in the parameter testing.

To emphasize the power of the test, we again generate data with a break point at $\frac{T}{2}$, where the data is simulated with $\theta_t = 1$ for $t \in \{\varepsilon T, \dots, \frac{T}{2}\}$, denoted as θ_0 and after that, we increase the parameter to $\theta_t = \{1.2, 1.4, 1.6, 1.8, 2.0\}$ for $t \in \{\frac{T}{2}+1, \dots, T\}$, denoted as θ_1. Table 4 reveal, that the power of the test increases with a larger sample size and at the same time the size is improved. Noticeable is the weaker power of the test, compared to the parameter test.

We now want to compare our constructed parameter test to a non parametric test of (Bücher et al., 2014), where the change point detection in this test is sensitive to changes in the copula of the multivariate continuous observations. We repeat the test 301 times by using the already implemented and online available R package of the copula test. The results of the rejection rate under the null are presented in Table 3 for $\theta_0 = 1$ and $\theta = 0.5$. Table 3 reveals that for a smaller sample size and dimension the rejection rate is closer to 0.05, than our constructed parametric test, however for larger T and N it is the other way around. Due to the SMM estimation, our test needs more information to compensate the upcoming fluctuations in the parameter estimation, as mentioned above. To emphasize the power of the test, we again generate data with a break point at $\frac{T}{2}$ and we use the same settings for θ_0 and θ_1 as used before. Table 4 reveal that the belonging power function increases more than the power function of our test, especially for smaller T in the case of small $N = 5$.
On the other hand for larger dimensions, as Table 5 reveals, the rejection rate of our constructed test increases, where the power of the copula based non parametric test decreases. To conclude, the non parametric test has better properties for low dimensions while our test performs better in high dimensions, if the number of parameters is kept fixed, this can be reasoned due to the more available data information in the SMM estimation and on the other hand, in a nonparametric copula constancy test, the complexity of the estimated objects increase.

4. REAL DATA EXAMPLE

4.1. 1D-Model

In this section we apply our test to a real data application. We use daily stock log return data over a time span ranging from May 2002 to July 2013 of ten large banks namely Citigroup, HSBC Holdings ($), UBS-R, Barclays, BNP Paribas, HSBC Holdings (ORD), Mitsubishi, Royal Bank, Credit Agricole, Bank of America and JP Morgan implying $T = 2900$ and $N = 10$.

To determine the residuals of the series, we first estimate an AR(1)-GARCH(1,1) model

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

for $t = 1, \ldots, 2900$, $i = 1, \ldots, 10$ and assume the same one factor copula model as described in Section 3 equation (3.1).

For a primary parameter analysis we estimate the common factor parameter in a rolling window of size 650 and the result can be seen in Figure 4.3, where the date is referring to the end of the considered window. As we can see the parameter increases strongly between July 2006 and June 2009, in the time span of the financial crisis, from approximately 0.65 up to 0.92, indicating a strong change of the residual dependencies. We now use our constructed test to test for a break in the common factor loading and choose the burn in period as 500, implying $\varepsilon = 0.17$.

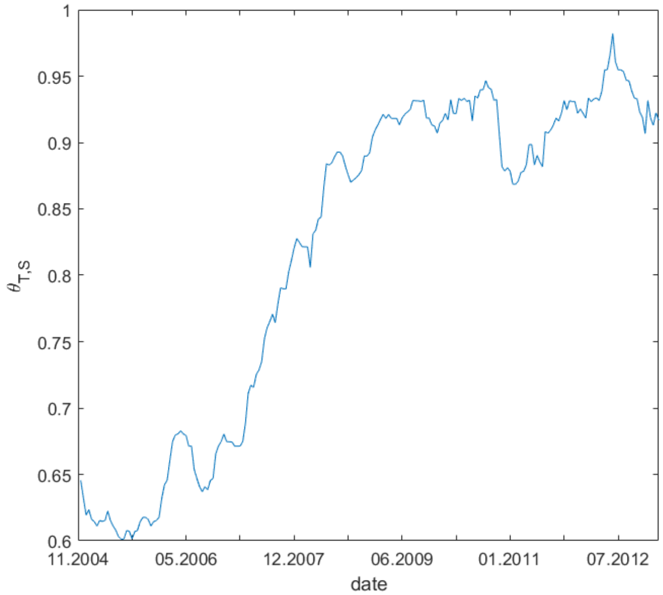

Figure 4.3: Rolling window parameter estimation for a window of size 650 in a data set of size T=2900 and dimension $N = 10$.

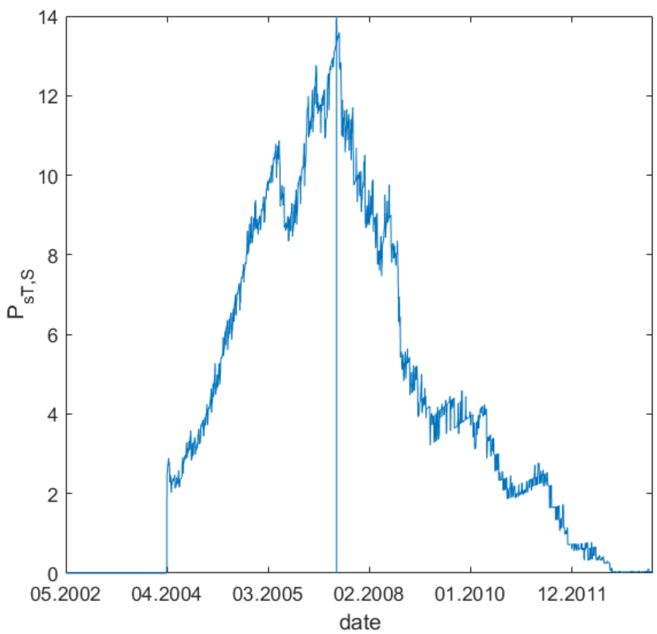

Figure 4.4: $P_{sT,S}$ for data with $T = 2900, N = 10$ and $\varepsilon = 0.17$, detected breakpoint in July 2007.

Figure 4.4 is a plot of $P_{sT,S}$ for every time point $\lfloor sT \rfloor$ between our considered time period starting at April 2004, due to consideration of the burn in period. The test statistic value is around 13.94 indicating a break point in July 2007 ($\lfloor \tilde{s}T \rfloor = 1438$) and the belonging 0.95-quantile value is 1.53, hence the null hypothesis of no change in the parameter is clearly rejected. One possible reason for the detected break point at around July 2007 can be the well known start of the financial crisis in the Summer of 2007, also compare the strong change of the parameter estimation in Figure 4.3 around this time point.

4.2. 2D-Model

In this section we apply our test to another real data application. We use daily stock log return data over a time span ranging from January 2002 to June 2012 from the EURO STOXX 50 of ten large financial firms namely Allianz, Generali, AXA, BBVA, Banco Santander, BNP Paribas, Deutschebank, Deutsche Börse, ING Groep and Intesa Sanpaolo as well as ten large industrial firms namely Airliquide, Arcelor, Bayer, E.ON, Enel, Eni, Iberdrola, Repsol, RWE and Total, implying $T = 2724$ and $N = 20$.

To determine the residuals of the series, we again estimate an AR(1)-GARCH(1,1) model and assume the following two factor copula model:

$$[X_{1t}, \ldots, X_{Nt}]' =: X_t = \begin{pmatrix} \beta_{1t} & 0 \\ 0 & \beta_{2t} \end{pmatrix} \begin{pmatrix} \boldsymbol{Z}_t \\ \boldsymbol{Z}_t \end{pmatrix} + \boldsymbol{q}_t,$$

where $Z_t \overset{init}{\sim} \mathsf{Skew\ t}\,(\sigma^2, \nu^{-1}, \lambda)$ and $q_t \overset{iid}{\sim} \mathsf{t}\,(\nu^{-1})$ for $t = 1, \ldots, T$ and we fix $\sigma^2 = 1$, $\nu^{-1} = 0.25$ and $\lambda = -0.5$, such that our model is parametrized by the factor loading parameters $\theta_t = (\beta_{1t}\ \beta_{2t})'$.

For the estimation of the sequential parameters θ_t for $t = \varepsilon T, \ldots, T$ in the test statistic, we again use the SMM approach, with $S = 25 \cdot T$ simulations with sample size T. We again use five dependence measures, namely Spearman's rank correlation and the 0.05, 0.10, 0.90, 0.95 quantile dependence measures, in a block eqidependence model as in Oh and Patton (2015), leading to a total number of ten dependence measures, reasond by the existence of two groups (financial and industrial sector each with ten companies). We check the size and power of our test for the two dimensional model for $T = 1000$, $N = 20$ and $\varepsilon = 0.2$ and we receive similar results as in the one factor model, compare Table 6.[1]

	$\theta_0 = [1.0\ 0.5]$	$\theta_1 = [1.2\ 0.5]$	$\theta_1 = [1.4\ 0.5]$	$\theta_1 = [1.6\ 0.5]$	$\theta_1 = [1.8\ 0.5]$
$T = 1000$	0.066	0.461	0.887	0.996	1.000

Table 6: Empirical power for $\theta_0 = [1\ 0.5]$, $N = 20$, $\varepsilon = 0.2$ and group size 10.

For a primary parameter analysis we estimate the two factor parameters for the finacial and industrial sector in a rolling window of size 500 and the result can be seen in Figure 4.5, where the date is referring to the end of the considered window. The two parameters increase strongly between July 2006 and July 2010, in the time span of the financial crisis, from approximately 0.8 up to 1.2 for the industrial sector and from approximately 0.9 up to 1.4 for the financial sector, indicating a strong change of the residual dependencies.

We now apply our test to the above described application, with the two factor model and choose the burn in period as 500, implying $\varepsilon = 0.18$.

Figure 4.6 is a plot of $P_{sT,S}$ for every time point $\lfloor sT \rfloor$ between our considered time period starting at December 2003 due to consideration of the burn in period. The test statistic value is around 41.31 indicating a break point in August 2008 ($\lfloor \tilde{s}T \rfloor = 1718$) and the belonging 0.95-quantile value is 5.13, hence the null hypothesis of no change in the parameters is clearly

[1] In rare cases, the SMM estimation yields very extreme unnatural estimates in the case of two or more parameters. To exclude them from the test statistic $S_{T,S}$, we eliminate all values $P_{sT,S}$ which are larger than the median plus 5 times the quarter-quantile difference of all $P_{sT,S}$.

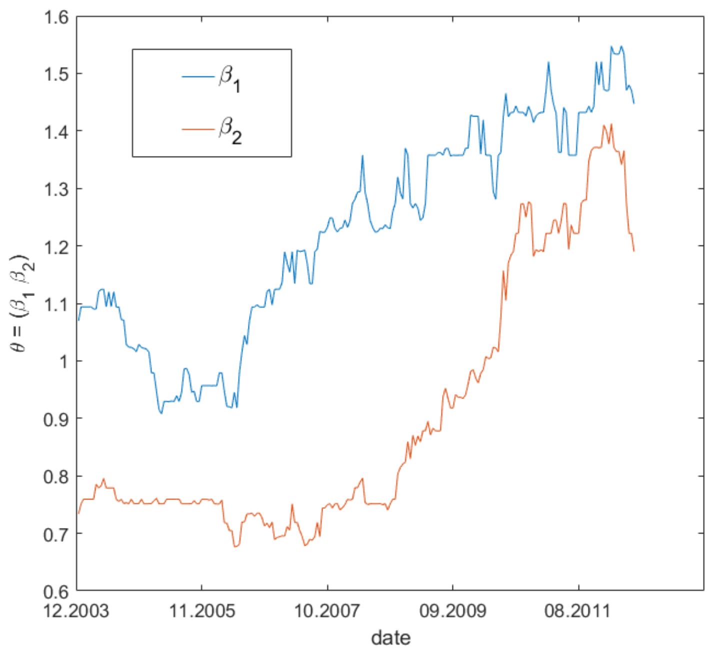

Figure 4.5: Rolling window parameter estimation for a window of size 500 in a data set of size $T = 2724$ and dimension $N = 20$.

rejected. One possible reason for the detected break point at around August 2008 can be the well known high of the financial crisis in the Summer of 2008, also compare the strong change of the parameters estimation in Figure 4.5 around this time point.

5. SUMMARY

We constructed new fluctuation tests for detecting structural breaks in factor copula models and analyzed the behavior under the null hypothesis of no change. Due to the discontinuity of the SMM objective function, it is not trivial to derive a functional limit theorem for the model parameters. Further we analyzed the behavior of the tests in Monte Carlo simulations and applied the test to real data applications.

In future reserach, our work could be extended in several interesting directions. First, one could derive a monitoring procedure for detecting parameter changes in an online-setup. Second, it would be interesting to explore the usefulness of our test in risk management

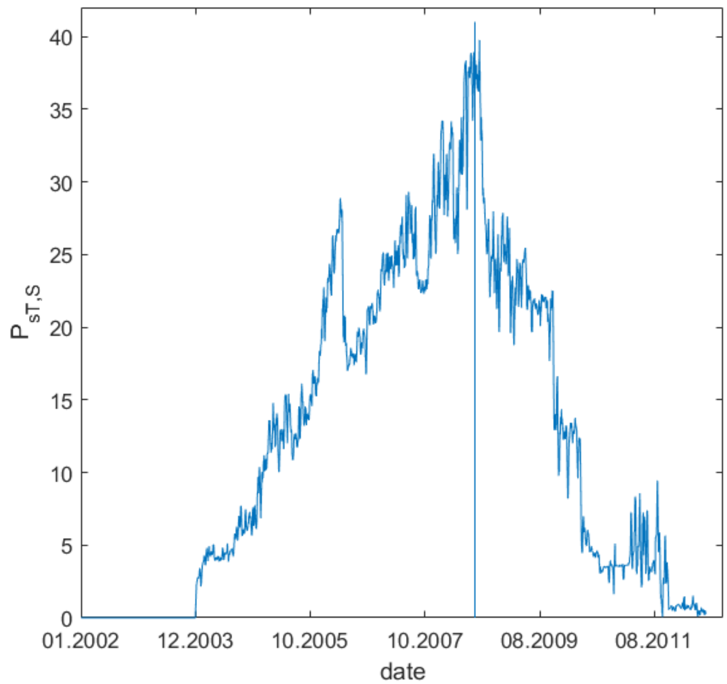

Figure 4.6: $P_{sT,S}$ for data with $T = 2724, N = 20$ and $\varepsilon = 0.18$, detected breakpoint in August 2008.

applications like the forecast of value at risk (VaR) and expected shortfall (ES). Finally, it would be worthwhile, but also mathematically demanding to derive appropriate tests in the case of time-varying marginal distributions.

A. ADDITIONAL RESULTS AND PROOFS

Theorem 1 is proved in different steps. First, we provide a consistency result in Lemma 2. Then, Theorem 4, which is based on Theorem 3, yields a general convergence result for SMM estimators. Lemma 6, which is based on Lemma 5 provides stochastic equicontinuity for the objective function in a general SMM setting. Finally, Lemma 7 yields distribution results for the empirical moments in our specific problem. All these results are then used for proving Theorem 1.

Lemma 2. If $\hat{\theta}_{T,S} \xrightarrow{a.s.} \theta_0, \qquad T, S \to \infty$, then

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

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

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

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

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

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

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

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

Theorem 3. Under the null hypothesis $H_0 : \theta_1 = \theta_2 = \cdots = \theta_T$, suppose that $\forall s \in [\varepsilon, 1], \varepsilon > 0 \quad Q_{sT,S}(\theta_{sT,S}) \geq \sup_{\theta\in\Theta} Q_{sT,S}(\theta) - o_p^*((s^2T)^{-1})$, $\sup_{s\in[\varepsilon,1)} \|\hat{\theta}_{sT,S} - \theta_0\| \xrightarrow{p} 0$, $T, S \to \infty$ and:

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

ii) $\theta_0(= \theta_1 = \cdots = \theta_T)$ are interior points of Θ

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

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

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

$\Rightarrow s\sqrt{T}(\theta_{sT,S} - \theta_0) \xrightarrow{d} A^*(s) \qquad \forall s \in [\varepsilon, 1], \varepsilon > 0$ and $A^*(s) = H^{-1}A(s)$,

where $A(s)$ is a continuous Gaussian process.

Proof. For simplification set $Q := Q_0$ and $\hat{Q} := Q_{sT,S}$. We first show the limitation $s\sqrt{T}\|\theta_{sT,S}-\theta_0\| = O_p(1)$. With a Taylor-expansion of $Q(\theta)$ around θ_0 and knowing $\nabla_\theta Q(\theta_0) =$

0, due to condition i), we receive $Q(\theta) = Q(\theta_0) + \frac{1}{2}(\theta - \theta_0)' H(\theta - \theta_0) + o(\|\theta - \theta_0\|^3)$. We also know from condition i) and iii), that $\exists C > 0 : (\theta - \theta_0)' H(\theta - \theta_0) + o(\|\theta - \theta_0\|^3) \leq -C\|\theta - \theta_0\|^2$ $\Rightarrow Q(\theta_{sT,S}) \leq Q(\theta_0) - C\|\theta_{sT,S} - \theta_0\|^2$ and we obtain

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

Consider

$$
\begin{aligned}
\Big(\|\theta_{sT,S} - \theta_0\| + O_p(s^{-1}T^{-\frac{1}{2}})\Big)^2 &= \|\theta_{sT,S} - \theta_0\|^2 + \|\theta_{sT,S} - \theta_0\|O_p(s^{-1}T^{-\frac{1}{2}}) + O_p(s^{-2}T^{-1}) \\
&\overset{(\star)}{\leq} \|\theta_{sT,S} - \theta_0\|O_p(s^{-1}T^{-\frac{1}{2}}) + o_p^*((s^2T)^{-1}) + O_p(s^{-2}T^{-1}) \\
&\leq O_p(s^{-2}T^{-1})
\end{aligned}
$$

$$
\Rightarrow \Big|\|\theta_{sT,S} - \theta_0\| + O_p(s^{-1}T^{-\frac{1}{2}})\Big| \leq O_p(s^{-1}T^{-\frac{1}{2}}), \quad \forall s \in [\varepsilon, 1] \qquad (\star\star)
$$

and we get

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

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

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

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

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

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

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

From the previous result (++), we know the convergence ordering of the remainder term of the approximation in (+++). So we receive

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

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

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

Because $\theta_{sT,S}, \tilde{\theta}_{sT,S} \in \Theta$, the convergence ordering of the remainder terms are known and

$H = H(\theta_0)$ is negatively definite and non singular

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

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

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

$$\Rightarrow s\sqrt{T}(\theta_{sT,S} - \theta_0) \xrightarrow{p} -H^{-1}s\sqrt{T}\hat{D}_{sT} \xrightarrow[iv)]{d} -H^{-1}A(s) = A^*(s).$$

□

Theorem 4. Under the null hypothesis $H_0 : \theta_1 = \theta_2 = \cdots = \theta_T$, suppose that

$\forall s \in [\varepsilon, 1], \varepsilon > 0 : \quad g_{sT,S}(\theta_{sT,S})'\hat{W}_{sT}g_{sT,S}(\theta_{sT,S}) \leq \inf_{\theta\in\Theta} g_{sT,S}(\theta)'\hat{W}_{sT}g_{sT,S}(\theta) + o_p^*((s^2T)^{-1})$,

$\sup_{s\in[\varepsilon,1]} \|\hat{\theta}_{sT,S} - \theta_0\| \xrightarrow{p} 0$, $sup_{s\in[\varepsilon,1]} \|\hat{W}_{sT} - W\| \xrightarrow{p} 0$, $\quad T, S \to \infty$ and:

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

ii) $\theta_0(= \theta_1 = \cdots = \theta_T)$ are interior points of Θ

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

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

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

$\Rightarrow s\sqrt{T}(\theta_{sT,S} - \theta_0) \xrightarrow{d} A^*(s) \qquad \forall s \in [\varepsilon, 1], \varepsilon > 0$

and $A^*(s) = (G'WG)^{-1}G'WA(s)$,

where $A(s)$ is a continuous Gaussian process.

Proof. The Theorem follows by verifying the conditions of Theorem 3. Set $\hat{Q}(\theta) := Q_{sT}(\theta) := -\frac{1}{2}\hat{g}(\theta)'\hat{W}_{sT}\hat{g}(\theta) + \hat{\Delta}_{sT}(\theta)$ with $\hat{g}(\theta) := g_{sT,S}(\theta)$ and $Q(\theta) := Q_0(\theta) := -\frac{1}{2}g(\theta)'Wg(\theta)$ with $g(\theta) := g_0(\theta)$. With a Taylor-expansion of $g(\theta)$ around θ_0

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

we obtain

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

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

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

noting that $Q(\theta)$ is maximized at θ_0, it follows $H(\theta_0) = -G'WG$, where H is a non singular negative definite matrix. Because H is by construction a nonsingular negative definite matrix, $\exists$ neighborhood of θ_0, where $Q(\theta)$ has a unique maximum at θ_0 with $Q(\theta_0) = 0$.

$\Rightarrow$ Conditions i), ii) and iii) of Theorem 3 are satisfied. By choosing $\hat{D}_{sT} = -G'\hat{W}_{sT}g_{sT,S}(\theta_0)$ it follows, $\forall s \in [\varepsilon, 1]$,

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

thus condition iv) of Theorem 3 is fulfilled. Now we define

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

and we get

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

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

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

The remainder term is just chosen in this way, that $\hat{Q}(\theta)$ is consistent with $-\frac{1}{2}\hat{g}(\theta)'\hat{W}_{sT}\hat{g}(\theta)$, which is shown in the next window and that we get the right convergence ordering, when checking condition v) of Theorem 3. First notice that by condition v) $\forall\delta>0 \quad \sup_{s\in[\varepsilon,1]}\sup_{\|\theta-\theta_0\|<\delta}\|\hat{\varepsilon}(\theta)\| = o_p(s^{-1}T^{-\frac{1}{2}})$, furthermore

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

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

With the consistency of $\hat{Q}(\theta)$ we can show the initial condition of Theorem 3

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

Finally we have to check condition v) of Theorem 3, for that we calculate

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

inserting (+) and $Q(\theta) = -\frac{1}{2}g(\theta)'Wg(\theta)$, sorting, triangle inequality,

choosing $\hat{D}_{sT} = -G'\hat{W}_{sT}\hat{g}(\theta_0)$ and size up the resulting terms, leads to

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

Now we have

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

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

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

For r_2, we obtain

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

Considering r_3 yields

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

For r_4, it holds

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

Finally, for r_5,

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

□

Lemma 5. Under Assumption 1, 2, 3.ii) and 3.iii)

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

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

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

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

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

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

$$\exists K : |m_0(\theta_1) - m_0(\theta_2)| \leq K\|\theta_1 - \theta_2\|. \qquad (\star\star)$$

Now consider

$$\begin{aligned}
|g_{sT,S}(\theta_1) - g_{sT,S}(\theta_2)| =& |\hat{m}_{sT} - \tilde{m}_S(\theta_1) - \hat{m}_{sT} + \tilde{m}_S(\theta_2)| \\
=& |\tilde{m}_S(\theta_2) - \tilde{m}_S(\theta_1)| = |\tilde{m}_S(\theta_1) - \tilde{m}_S(\theta_2)| \\
\underset{c.s.}{\overset{(\star)}{\leq}}& |m_0(\theta_1) - m_0(\theta_2)| + |o_p(1)| \\
\overset{(\star\star)}{\leq}& K\|\theta_1 - \theta_2\| + |o_p(1)| \\
=& \left(K + \frac{|o_p(1)|}{\|\theta_1 - \theta_2\|}\right)\|\theta_1 - \theta_2\| \\
=:& B\|\theta_1 - \theta_2\|.
\end{aligned}$$

ii) For some $\delta > 0$

$$\Rightarrow \limsup_{T,S\to\infty} E\left(B^{2+\delta}\right) = \limsup_{T,S\to\infty} E\left(\left(K + \frac{|o_p(1)|}{\|\theta_1 - \theta_2\|}\right)^{2+\delta}\right) < \infty.$$

$\square$

Lemma 6. Under Assumption 1, 2, 3.ii) and 3.iii), for $\frac{S}{T} \to \infty$ or $\frac{S}{T} \to k \in (0, \infty)$,

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

Proof. By Lemma 5)i) $\{g_{sT,S}(\theta) : \theta \in \Theta\}$ is Lipschitz-continuous $\forall s \in [\varepsilon, 1], \varepsilon > 0$ and so a Type II class of functions in Andrews (1994). By Theorem 2 of Andrews $\{g_{sT,S}(\theta) : \theta \in \Theta\}$ satisfies Pollard's entropy condition with envelope

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

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

Furthermore $g_{sT,S}(\theta)$ is bounded and by Lemma 5)ii) it holds

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

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

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

□

Lemma 7. We consider the dependence measures Spearman's rho and quantile dependence measures, which are functions only depending on bivariate copulas.

Under the null and Assumption 1 and 2,

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

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

Proof. The proof follows the idea of Bücher et al. (2014) and we only consider the limit process for $T \to \infty$.

By Proposition 3.3 in (Bücher et al., 2014) (+) the sequential empirical copula of the N-dimensional random vectors fulfills

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

where $\mathbf{u} \in [0,1]^N$, $\mathbf{u}^{(j)} \in [0,1]^N$ defined by $\mathbf{u}_i^{(j)} = \mathbf{u}_j$, if $i = j$ and 1 otherwise and $\hat{\mathbf{F}}^s(\hat{\eta}_t) = (\hat{F}_1^s(\hat{\eta}_{1t}), \ldots, \hat{F}_N^s(\hat{\eta}_{Nt}))$. Here, $\hat{F}_j^s$ denotes the marginal empirical distribution function of the j-th component calculated from data up to time point $[sT]$. Moreover $\mathbb{B}(s, \mathbf{u})$ is a tight centered continuous Gaussian process with $\mathbb{B}(0, \mathbf{u}) = 0$ and

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

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

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

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

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

is continuous and we directly obtain

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

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

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

□

Proof of Theorem 1

The proof follows by checking the conditions of Theorem 4. The initial conditions of Theorem 4 follow by Assumption 4.iii) and Lemma 2.

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

ii) $\theta_0(= \theta_1 = \cdots = \theta_T)$ are interior points of Θ given by Assumption 4.i).

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

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

$$\begin{aligned} s\sqrt{T}g_{sT,S}(\theta_0) =& s\sqrt{T}(\hat{m}_{sT} - \tilde{m}_S(\theta_0)) \\ =& s\sqrt{T}(\hat{m}_{sT} - m_0(\theta_0)) + s\sqrt{T}(m_0(\theta_0) - \tilde{m}_S(\theta_0)) \\ =& s\sqrt{T}(\hat{m}_{sT} - m_0(\theta_0)) - \frac{\sqrt{T}}{\sqrt{S}}s\sqrt{S}(\tilde{m}_S(\theta_0) - m_0(\theta_0)) \\ \xrightarrow[\text{Lemma 7}]{d}& A(s) \end{aligned}$$

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

$$\begin{aligned} s\sqrt{T} g_{sT,S}(\theta_0) =& s\sqrt{T}(\hat{m}_{sT} - \tilde{m}_S(\theta_0)) \\ =& s\sqrt{T}(\hat{m}_{sT} - m_0(\theta_0)) + s\sqrt{T}(m_0(\theta_0) - \tilde{m}_S(\theta_0)) \\ =& s\sqrt{T}(\hat{m}_{sT} - m_0(\theta_0)) - \frac{\sqrt{T}}{\sqrt{S}} s\sqrt{S}(\tilde{m}_S(\theta_0) - m_0(\theta_0)) \\ \xrightarrow[Lemma\ 7]{d}& A(s) - \frac{s}{\sqrt{k}} A(1), \end{aligned}$$

combined we get

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

v) We know by Lemma 6, that for $\frac{S}{T} \to \infty$ or $\frac{S}{T} \to k \in (0,\infty)$ $v_{sT,S}(\theta) = \sqrt{sT}[g_{sT,S}(\theta) - g_0(\theta)]$ is stochastically equicontinuous $\forall s \in [\varepsilon, 1], \varepsilon > 0$.

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

Furthermore the inequality

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

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

Finally we obtain

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

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

This completes the proof. □

REFERENCES

Bedford, T. and R. Cooke (2002): "Vines - A New Graphical Model for Dependent Random Variables," *Annals of Statistics*, 30, 1031–1068.

Boudebda, S. and T. Zari (2013): "Strong Approximation of empirical Copula Processes by Gaussian Processes," *Université Pierre et Marie Curie*.

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

Bücher, A. and M. Ruppert (2013): "Consistent Testing for a Constant Copula Under Strong Mixing Based on the Tapered Block Multiplier Techniqu," *Journal of Multivariate Analysis*, 116, 208–229.

Bücher, A. and S. Volgushev (2011): "Empirical and sequential empirical copula processes under serial dependence," 44.

Dehling, H., D. Vogel, M. Wendler, and D. Wied (2016): "Testing for Changes in Kendall's tau," *Econometric Theory, forthcoming.*

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

Hong, H., A. Mahajan, and D. Nekipelov (2015): "Extremum estimation and numerical derivatives," *Journal of Econometrics*, 188, 250–263.

Inoue, A. (2001): "Testing for Distributional Change in Time Series," *Econometric Theory*, 17, 156–187.

Joe, H. (1997): "Multivariate Models and Multivariate Dependence Concepts," *Monographs on Statistics and Applied Probability*, 73.

Kiefer, J. (1955): "Distance tests with good power for the nonparametric k-sample problem," *Ann. Math. Stat.*, 26.

Krupskii, P. and H. Joe (2013): "Factor Copula Models for Multivariate Data," *Journal of Multivariate Analysis*, 120, 85–101.

Nelsen, R. B. (1998): "An introduction to copulas," *Lecture notes in statistics*, 139.

Newey, W. and D. McFadden (1994): "Large Sample Estimation and Hypothesis Testing," *Handbook of Econometrics*, 4.

Oh, D. and A. Patton (2013): "Simulated Method of Moments Estimation for Copula-Based Multivariate Models," *Journal of the American Statistical Association*, 108.

——— (2015): "Modelling Dependence in High Dimensions with Factor Copulas," *Finance and Economics Discussion Series 2015-051*, 051.

Rémillard, B. (2010): "Goodness-of-Fit Tests for Copulas of Multivariate Time Series," *Mathematics Subject Classification*.

Savu, C. and M. Trede (2010): "Hierarchies of Archimedean Copulas," *Quantitative Finance*, 10, 295–304.

van der Vaart, A. (1998): "Asymptotic Statistics," *Cambridge University Press*.

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