Testing Constant Cross-Sectional Dependence with Time-Varying Marginal Distributions in Parametric Models

Matthias Kaldorf *
University of Cologne

Dominik Wied †
University of Cologne

October 31, 2018

Abstract

The paper proposes parametric two-step procedures for assessing the stability of cross-sectional dependency measures in the presence of potential breaks in the marginal distributions. The procedures are based on formerly proposed sup-LR tests in which restricted and unrestricted likelihood functions are compared with each other. We derive suitable test statistics in different settings, i.e., in the case of bivariate normal and t distributions as well as in the case of copulae. The properties of the test statistics (size, power and the relevance of the residual effect) are analyzed and compared with existing methods in various Monte Carlo simulations.

Key words: Cumulated Sums; Empirical Copula; sup-LR Test; Structural Break; Two-Step Procedure

JEL classification: C58 (Financial Econometrics)

1 Introduction

Testing stability of cross-sectional dependence in multivariate time series models has received considerable attention over recent years, both in terms of methodological advance and in applications. In financial econometrics, those methods find application to asset price data subject to financial crisis or policy shocks. In the context of financial crisis, this phenomenon is usually called *shift contagion* and has been formally analyzed first by King and Wadhwani [1990] who use recursively calculated sample correlations to assess stability of the correlation over the considered sample. In an important contribution, Forbes and Rigobon [2002] stress that in equity markets increases in volatility of some equity market often precede an increase in correlation (or some other dependency measure). Therefore, before testing for constant correlation, potential changes in variances have to be taken into account. These procedures test constancy of the marginal distributions in a first step, eliminate potential structural

*Center for Macroeconomic Research, University of Cologne, Germany, email: matthias.kaldorf@.uni-koeln.de

†Institute for Econometrics and Statistics, University of Cologne, Germany, email: dwied@uni-koeln.de.

changes in the margins by suitable transformations and then test constancy of the cross-sectional dependence in step two. We will call such procedures *two-step testing procedures* in the following.

In general, there exist two fundamentally different approaches in a structural change context: likelihood-ratio-type tests that rely on some parametric model and tests imposing hypothesis on moments or quantile exceedance-probabilities, that use cumulated sums of empirical counterparts.

A seminal contribution for the first approach is Andrews [1993] who derived asymptotic tests for (partial) structural changes in a generalized method of moments framework. One class of these tests are supremum likelihood ratio type (sup-LR-type) tests. In a multivariate model, parameters are partitioned into those that change under the null hypothesis of constancy and the alternative and nuisance-parameters that are invariant under null and alternative hypothesis. For any change-point candidate, the sample is divided into two sub-samples and parameter stability is rejected, if the difference between two GMM objective functions becomes too large. This method has first been applied in the context of constant correlation by Dias and Embrechts [2004].

Within the latter class one can distinguish tests imposing constancy on cross-moments of the multivariate system and tests imposing constancy of the copula. Stability is rejected if the fluctuations in the cumulated sums of their empirical counterparts exceed certain critical values. Important contributions in an econometric context include Aue et al. [2009] for covariances, Wied et al. [2012b] for correlations, Rémillard [2010] and Bücher et al. [2014] for copulae.

In both frameworks, two-step procedures have been proposed. While the latter framework was tackled in Demetrescu and Wied [2018+], Blatt et al. [2015] worked in the first framework by analyzing shift contagion in VAR models using the multivariate normal distribution. Our aim is to continue the work in Blatt et al. [2015] by deriving appropriate supLR-type statistics in different parametric models, which are typically used for financial time series data. The motivation for using such tests is that, if the assumed model is correct and different from a normal distribution model, a parametric test might have higher power than a nonparametric test. To the best of our knowledge, test statistics for different multivariate distributions have not been explored in the literature before.

Critical values for the tests are obtained by bootstrap approximations. In particular, these bootstrap approximations are used because it cannot be expected that the usage of transformed/standardized data (using piecewise constant variance estimators, GARCH-residuals or empirical cumulative distribution functions) in step 2 leaves the asymptotics unaltered. While Demetrescu and Wied [2018+] derived analytic results for the residual effects in nonparametric models (see also Duan and Wied, 2018), the first contribution of this paper is to quantify the impact of transforming the original data by Monte Carlo simulations.

Secondly, after showing that the residual effect matters quantitatively in commonly used models, we move to compare power of the different approaches after correcting for the invalidity of standard asymptotics. Moreover, the Monte Carlo study compares the ability of parametric and non-parametric procedures to detect and date structural changes in the sequential setup we are interested in. Such a simulation study extends simulation results in Galeano and Peña [2007] who compare Gaussian sup-LR and fluctuation tests in the case of variance/covariance changes. Apart from Demetrescu and Wied [2018+], we also include the recently proposed tests for constant copulas by Bücher et al. [2014] (modified for a two-step-procedure) in our comparison.

If test and estimation procedures are constructed with applications to finance or macroeconomics in mind, it is natural to study their behaviour in settings that feature typical characteristics of financial data. In a first simulation study parametric and non-parametric methods building on the joint distribution function are examined using data generated from Gaussian and t-distributions, resembling financial return data with low observation frequency. Performance of copula-based methods is compared under non-linear dependence by generating data from t-copulae in a second simulation study. Due to its practical relevance, dimensionality effects are taken into account. Some attention is devoted to the situations where the test procedures are applied under misspecification, such as choosing the wrong joint distribution function or copula within the sup-LR framework.

Many results however carry over from models using the joint distribution function or empirical moments to models using copulae: more elaborate parametric models should only be used if the sample size permits reliable estimation and one is reasonably confident about the goodness-of-fit of the model. If one is confident in applying parametric methods, copula-based sup-LR tests can be preferred to distribution-function based sup-LR tests. Should one operate in smaller samples, sup-LR tests under the simplest parametric assumption, namely that of a multivariate Gaussian distribution or copula, provide a suitable alternative to non-parametric methods. In many situations, simulation evidence does not show uniform dominance of either test, so one might use both tests and use a simple error correction scheme like the Boole-Bonferroni method to correct for multiple testing of the same hypothesis. Results are more obvious when it comes to estimating change-points: in almost every considered case, the parametric sup-LR method yields better estimators in terms of bias and variance, irrespectively of correct or incorrect model specification. If precise knowledge on the timing of regime-shifts is of central importance, for example in a portfolio management situation, one should rather use a parametric method. Even if one does not want to rely on a certain specification, one could still use the sup-LR framework under a Gaussian copula assumption to achieve useful results.

The structure of the paper is as follows: First, section 2 introduces the hypotheses pairs used in the two-step procedure. In section 3, we introduce the sup-LR test framework used for our applications and give some analytical high-level background. Next, we discuss certain parametric specifications, e.g. the bivariate t-distribution. Section 4 presents the simulation studies, while illustrations of the discussed methods are given in section 5, using commodity and stock return data from the 1990s. Section 6 gives a conclusion, while the supplementary material gives an overview about the different non-parametric test frameworks (moment-based fluctuation and empirical copula tests) used in the simulation study. With regard to the moment-based fluctuation test, there are some new analytical derivations. Moreover, the supplementary material provides some results on using the sup-LR framework under volatility clustering and, based on this results, a second application on daily stock returns of EURO STOXX 50 stock returns.

2 Setup

Let $X_t \sim (\delta_{(t)}, \theta_{(t)})$ be a multivariate random variable with dimension m, $\delta_{(t)}$ denote a parameter vector shaping its dependence structure and $\theta_{(t)} = [\theta'_{1,(t)}, ..., \theta'_{m,(t)}]'$ denote a vector consisting of the parameters shaping the marginal distributions, indexed by i. Let the change-point corresponding to dimension i be denoted by l_i and l_D the change-point of the dependency-structure and further assume $l_1 \leq l_2 \leq ... \leq l_m \leq l_D$. The particular order of change-points merely eases notation and does not lead to loss of generality, since the asymptotics in sequential procedures are not affected from switching the change-point order. Denoting the time index by t, one formally has

$$
\begin{aligned}
&X_t \overset{i.i.d.}{\sim} (\theta_{1,1}, \theta_{2,1}, \cdots, \theta_{m,1}, \delta_{D,1}) \quad \text{for } t = 1, ..., l_1 \\
&X_t \overset{i.i.d.}{\sim} (\theta_{1,2}, \theta_{2,1}, \cdots, \theta_{m,1}, \delta_{D,1}) \quad \text{for } t = l_1 + 1, ..., l_2 \\
&\cdots\cdots \\
&X_t \overset{i.i.d.}{\sim} (\theta_{1,2}, \theta_{2,2}, \cdots, \theta_{m,2}, \delta_{D,1}) \quad \text{for } t = l_m + 1, ..., l_D \\
&X_t \overset{i.i.d.}{\sim} (\theta_{1,2}, \theta_{2,2}, \cdots, \theta_{m,2}, \delta_{D,2}) \quad \text{for } t = l_D, ..., n
\end{aligned}
$$

with at most $m + 1$ asymptotically distinct break points. Note that, defining $\lambda_i := \frac{l_i}{n}$, two change-points are asymptotically distinct if $\lambda_1 \neq \lambda_2$ as $n \to \infty$. The ordering of the break dates reflects a situation where *shift contagion* is present: at first, there is a change in mean and variance of one variable (e.g. a stock market index of country A), followed by a change in the second (country B), third (country C) variable and so forth. The correlation between both markets changes at some later point in time, l_D. The following hypothesis pairs are relevant for the sequential procedures under consideration and have appeared in this or slightly different forms throughout the existing literature:

Hypothesis Pair 1 (Marginal Distributions)**.** For every margin i, we test:

$$
\begin{aligned}
&H_0 : \theta_{i,(1)} = ... = \theta_{i,(n)} \quad \text{against} \\
&H_1 : \theta_{i,1} = \theta_{i,(1)} = ... = \theta_{i,(l_i)} \neq \theta_{i,(l_i+1)} = ... = \theta_{i,(n)} = \theta_{i,2}
\end{aligned}
$$

The CUSUM of squares test from Wied et al. [2012a] obtains if $\theta_i = \sigma_i^2$.

Hypothesis Pair 2 (Dependency, Constant Margins)**.**

$$
\begin{aligned}
&H_0 : \delta_{(1)} = ... = \delta_{(n)} \quad \text{against} \\
&H_1 : \delta_1 = \delta_{(1)} = ... = \delta_{(l_D)} \neq \delta_{(l_D+1)} = ... = \delta_{(n)} = \delta_2 \\
&\text{with } \theta_{i,(1)} = ... = \theta_{i,(n)} \quad \forall\, i = 1, ..., m
\end{aligned}
$$

If Pearson's correlation matrix is used to measure cross-sectional dependency, Hypothesis Pair 2 is in line with the test proposed in Wied et al. [2012b] who extend the covariance test from Aue et al. [2009] to moment hypothesis on correlations. In a shift contagion situation it comes natural to extend the first two hypothesis pairs into a joint framework:

Hypothesis Pair 3 (Two-Step Testing Procedure)**.**

$$
H_0 : \delta_{(1)} = ... = \delta_{(n)}
$$

$$\theta_{1,(1)} = ... = \theta_{1,(l_1)} \neq \theta_{1,(l_1+1)} = ... = \theta_{1,(n)}$$
$$\cdots \quad \cdots \quad \cdots$$
$$\theta_{m,(1)} = ... = \theta_{m,(l_m)} \neq \theta_{m,(l_m+1)} = ... = \theta_{m,(n)} \text{ against}$$
$$H_1 : \delta_1 = \delta_{(1)} = ... = \delta_{(l_D)} \neq \delta_{(l_D+1)} = ... = \delta_{(n)} = \delta_2$$
$$\theta_{1,(1)} = ... = \theta_{1,(l_1)} \neq \theta_{1,(l_1+1)} = ... = \theta_{1,(n)}$$
$$\cdots \quad \cdots \quad \cdots$$
$$\theta_{m,(1)} = ... = \theta_{m,(l_m)} \neq \theta_{m,(l_m+1)} = ... = \theta_{m,(n)}$$

Hypothesis Pair 3 allows for changes in the marginal distributions under both the null and alternative hypothesis. In particular, there is no stationarity under the null hypothesis. Under the null, we have constant dependence and under the alternative, we have a two-regime model in the dependence structure.

3 Framework

In this section, we propose a framework which uses fully specified parametric models in order to evaluate parameter stability. The framework goes back to the seminal contribution of Andrews [1993] who suggests a method which is essentially applicable for all GMM-type estimators, such as maximum-likelihood and pseudo-maximum-likelihood. The framework is the following: The sample is divided into two sub-samples for any $j = \pi \cdot n,\ \pi \in [\underline{\pi}, \overline{\pi}]$, where parameters are divided into those that change under the null and alternative hypothesis and nuisance parameters that are invariant under null and alternative, denoted by η. Parameter constancy is tested by forming a sequence of likelihood-ratio test statistics for all change point candidates. The testing function is given by the log-likelihood function and the test statistic for a fixed j is given by the difference of the log-likelihood under the restricted and the unrestricted ML- or pseudo-ML-estimator. No restriction here means that the parameter, which is tested for constancy, is calculated based on $X_1, \ldots, X_j$ and $X_{j+1}, \ldots, X_n$ separately. We note that, while the framework is based on Andrews [1993], that paper does not look directly at these test statistics, but on "supLR-type" statistics which are based on the differences of GMM-objective functions evaluated at the restricted and unrestricted estimator. We use the likelihood functions themselves in order to avoid calculating the scores for each of our parametric models.
In the following, we shortly present the parametric frameworks for the first two hypothesis pairs, which are known from Andrews [1993]. Hypothesis Pair 3 is discussed afterwards.
Testing constancy of marginal distributions, i.e. Hypothesis Pair 1, is performed with the test statistic

$$A_j := A_j(\hat{\theta}_0, \hat{\theta}_1, \hat{\theta}_2, \hat{\eta}) := A_j(X; \hat{\theta}_0, \hat{\theta}_1, \hat{\theta}_2, \hat{\eta}) := 2\big(L(X; \hat{\theta}_1, \hat{\theta}_2, \hat{\eta}) - L(X; \hat{\theta}_0, \hat{\eta})\big) \tag{3.1}$$

with

$$L(X; \hat{\theta}_1, \hat{\theta}_2, \hat{\eta}) \quad = \quad \sum_{t=1}^{j} l_t(\hat{\theta}_1, \hat{\eta}) + \sum_{t=j+1}^{n} l_t(\hat{\theta}_2, \hat{\eta}) \text{ and}$$

$$L(X;\hat{\theta}_0,\hat{\eta}) = \sum_{t=1}^{n} l_t(\hat{\theta}_0,\hat{\eta}).$$

Here, $l_i(\cdot)$ denotes the contribution to the log-likelihood from observation i. Moreover, $\hat{\theta}_1 = \hat{\theta}_{1,j}$ is the ML-estimator for θ based on $X_1,\dots,X_j$, where $j := [\pi n]$ (the floor function is omitted in the following for brevity) and $\hat{\theta}_2 = \hat{\theta}_{2,j}$ the one based on $X_{j+1},\dots,X_n$, $\hat{\theta}_0$ the one based on $X_1,\dots,X_n$ and $\hat{\eta}$ the ML estimator based on $X_1,\dots,X_n$ for some constant nuisance parameter η.

Hypothesis Pair 2 is tested with the similar test statistic

$$A_j(\hat{\delta}_0,\hat{\delta}_1,\hat{\delta}_2,\hat{\eta}) = 2\big(L(X;\hat{\delta_1},\hat{\delta_2},\hat{\eta}) - L(X;\hat{\delta_0},\hat{\eta})\big) \tag{3.2}$$

We now state a theorem concerning the asymptotic distribution of the sequence of LR-statistics. The result can be indirectly inferred from Andrews [1993] and Andrews and Ploberger [1995], as Andrews and Ploberger [1995] state that the asymptotic distribution is the same as that of the sup-Wald and the sup-LM statistics which were introduced in Andrews [1993] and shown to converge to the same limit in our Theorem 1. On the other hand, we provide a direct proof. Before, some standard assumptions are imposed.

Assumption 1. For Hypothesis Pair 1, it holds under the null hypothesis:

1. The true parameter θ_0 lies inside a set $\Theta \subset \mathbb{R}^k$.
2. The estimators $\hat{\theta}_1$, $\hat{\theta}_2$ and $\hat{\theta}_0$ fulfill a central limit theorem, i.e.,

   $$\sqrt{n}\begin{pmatrix}\hat{\theta}_1-\theta_0\\ \hat{\theta}_2-\theta_0\\ \hat{\theta}_0-\theta_0\end{pmatrix}$$

   converges to

   $$\begin{pmatrix}\frac{1}{\pi}H^{-1/2}\Gamma_k(\pi)\\ \frac{1}{1-\pi}H^{-1/2}\Gamma_k(1-\pi)\\ H^{-1/2}\Gamma_k\end{pmatrix}$$

   with $H = -\lim_{n\to\infty}\frac{1}{n}\sum_{i=t}^{n}\frac{\partial^2}{\partial\theta\partial\theta'}l_t(\theta_0)$ and Γ_k denoting a k-dimensional vector of independent Brownian motions.
3. The third derivatives of $l_t(\theta)$ with respect to θ exist and are uniformly bounded for $\theta \in \Theta$ and $= 1,\dots,n$.

A similar assumption holds for Hypothesis Pair 2.

Note that the expression of the limits appears to be natural if it can be assumed that $\hat{\theta}_{1,j}-\theta_0$ can be linearized as $\left(\sum_{t=1}^{j}\frac{\partial^2}{\partial\theta\partial\theta'}l_t(\theta_0)\right)^{-1}\sum_{t=1}^{j}\frac{\partial}{\partial\theta}l_t(\theta_0)+o_p(\sqrt{n})$ and $\hat{\theta}_{2,j}-\theta_0$ can be linearized as $\left(\sum_{t=j+1}^{n}\frac{\partial^2}{\partial\theta\partial\theta'}l_t(\theta_0)\right)^{-1}\sum_{t=j+1}^{n}\frac{\partial}{\partial\theta}l_t(\theta_0)+o_p(\sqrt{n})$. This is the case in the models we consider in 3.

Theorem 1. *Under Assumption 1, it holds for the sequence of LR-statistics that*

$$A_{\pi n} \Rightarrow_d \frac{(\Gamma_k(\pi)-\pi\Gamma_k(1))'(\Gamma_k(\pi)-\pi\Gamma_k(1))}{\pi(1-\pi)}, \tag{3.3}$$

in $D[0,1]$, the space of càdlàg-functions over the unit interval. Moreover, $k = \dim(\theta_0)$ for Hypothesis Pair 1 and $k = \dim(\delta_0)$ for Hypothesis Pair 2. The limit process is called a standardized tied-down Bessel process of order k, denoted by $\mathcal{B}_k(\pi)$

Proof: It holds

$$A_{\pi n} = \left(\hat{\theta} - \hat{\theta}_1\right)' \sum_{t=1}^{\pi n} \frac{\partial^2}{\partial\theta\partial\theta'} l_t(\hat{\theta}_1) \left(\hat{\theta} - \hat{\theta}_1\right) + \left(\hat{\theta} - \hat{\theta}_2\right) \sum_{t=\pi n+1}^{n} \frac{\partial^2}{\partial\theta\partial\theta'} l_t(\hat{\theta}_2) \left(\hat{\theta} - \hat{\theta}_2\right) + o_p(1).$$

This term converges in distribution to

$$\begin{aligned} & \left(\sqrt{\pi}\Gamma_k(1) - \frac{1}{\sqrt{\pi}}\Gamma_k(\pi)\right)' \left(\sqrt{\pi}\Gamma_k(1) - \frac{1}{\sqrt{\pi}}\Gamma_k(\pi)\right) \\ + \quad & \left(\sqrt{1-\pi}\Gamma_k(1) - \frac{1}{\sqrt{1-\pi}}\Gamma_k(1-\pi)\right)' \left(\sqrt{1-\pi}\Gamma_k(1) - \frac{1}{\sqrt{1-\pi}}\Gamma_k(1-\pi)\right), \end{aligned}$$

which has the same distribution as

$$\frac{(\Gamma_k(\pi) - \pi\Gamma_k(1))'(\Gamma_k(\pi) - \pi\Gamma_k(1))}{\pi(1-\pi)}.$$

□

As we have the factor $\pi(1-\pi)$ in the denominator, it is clear that $\Pi = [\underline{\pi}, \overline{\pi}]$ has to be a strict subset of the unit interval. To test the null hypothesis of parameter constancy against a single unknown change point, the sup-functional is applied to the test sequence of LR-test statistics and, in the framework of Andrews [1993],

$$\sup_{\underline{\pi}\cdot n \le j \le \overline{\pi}\cdot n} A_j \to_d \sup_{\Pi} \mathcal{B}_k(\pi). \tag{3.4}$$

So the null hypothesis is rejected when the $(1-\alpha)$-quantile associated with the limiting process (3.3), defined by $c_\alpha = P(\sup_{\pi\in\Pi} \mathcal{B}_k > c_\alpha) = \alpha$ is exceeded. Critical values are tabulated in Andrews [1993] and depend on the degrees of freedom of the limiting process and the considered interval Π of candidate change points. In every situation considered in the following, the supremum of $\{A_j\}$ is also used to estimate the change-point by

$$\hat{l} = \arg\sup_{\underline{\pi} n \le j \le \overline{\pi} n} A_j \tag{3.5}$$

In practical applications sup- and argsup-functional are replaced by the max- and argmax-functional, respectively. Following the suggestion of Andrews [1993] the set of potential change points is chosen to be $\Pi = [0.15, 0.85]$, the general case will however be maintained in the notation.

Our testing idea for analyzing Hypothesis Pair 3 is similar to that of Hypothesis Pair 2. The test statistic is given by

$$\sup_{\underline{\pi}\cdot n \le j \le \overline{\pi}\cdot n} A_j \tag{3.6}$$

with A_j given in (3.2), but with the original data replaced by appropriate residuals. Here, "residuals" implies that we transform marginal time series such that they do not exhibit breaks any more and also the dependence structure does not change. This means that we are interested in the dependence structure of $Z_t = f(Y_t, t/n, \theta)$, but can only observe $X_t = f(Y_t, t/n, \hat{\theta})$. The type of transformations as well as the particular test statistics depend on the respective parametric specification, which are derived in the following subsections. For example, the setting allows for time-varying marginal variances if $\theta = \theta_1$ for $t \leq j = \pi n$ and $\theta = \theta_2$ for $t > j = \pi n$. Demetrescu and Wied [2018+] discuss such models in detail and also argue analytically and with numerical evidence why it is not possible to test Hypothesis Pair 3 with the standard method of Andrews [1993] who assumes stationarity under the null hypothesis.

Allowing for unknown marginal parameters, which have to be estimated, introduces complications concerning the limit distribution. As pointed out in many studies on that matter, using estimated parameters and change-point locations in the first step potentially affects estimation of parameters and change-point locations in the second step, see Qu and Perron [2007], Chan et al. [2009] and Demetrescu and Wied [2018+].

In our setting, the reason for getting a residual effect is the following: Define with $\tilde{\delta}(\pi) = (\hat{\delta}_1, \hat{\delta}_2)$ the vector of unrestricted ML-estimators, such that $l_t(\tilde{\delta}(\pi), \hat{\theta}) := l_t(\hat{\delta}_1, \hat{\theta}, \hat{\eta})$ for $t \leq j = \pi n$ and $l_t(\tilde{\delta}(\pi), \hat{\theta}) := l_t(\hat{\delta}_2, \hat{\theta}, \hat{\eta})$ for $t > j = \pi n$. To ease notation, we write $\hat{\delta} := \hat{\delta}_0$ and omit the dependency of the likelihood contributions on the nuisance parameter $\hat{\eta}$. We impose an additional assumption, which appears to be natural given that both $\tilde{\delta}(\pi)$ and $\hat{\delta}$ are consistent for δ under the null hypothesis. The most crucial part of this assumption is part 3, which we illustrate in section 3.1 below.

Assumption 2. For Hypothesis Pair 3, it holds under the null hypothesis:

1. $\hat{\theta}_1, \hat{\theta}_2$ and θ_0 satisfy a central limit theorem similar to Assumption 1.2, with H replaced by

$$H^* = -\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2}{\partial\theta\partial\theta'} l_t(\tilde{\delta}(\pi), \theta_0).$$

2. The third derivatives of $l_t(\cdot, \theta)$ with respect to θ exist and are uniformly bounded for $\theta \in \Theta$ and $i = 1, \ldots, n$.

3. The process

$$B_{\pi n} := \sqrt{n}(\hat{\theta} - \theta) \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial}{\partial\theta} \left(l_t(\tilde{\delta}(\pi), \theta) - l_t(\hat{\delta}, \theta) \right)$$

converges to some limit process $R(\pi)$ in $D[0, 1]$.

4. The process $C_{\pi n} := \frac{\partial^2}{\partial\theta\partial\theta'} \frac{1}{n} \sum_{t=1}^{n} \left(l_t(\tilde{\delta}(\pi), \theta) - l_t(\hat{\delta}, \theta) \right)$ converges to zero in probability.

Then, we have the following theorem:

Theorem 2. *Under Assumption 2, it holds for $A_j(\tilde{\delta}(\pi), \hat{\delta}, \hat{\theta}) := A_j(\tilde{\delta}(\pi), \hat{\delta}, \hat{\theta}, \hat{\eta})$ that*

$$A_{\pi n}(\tilde{\delta}(\pi), \hat{\delta}, \hat{\theta}, \hat{\eta}) \Rightarrow_d \mathcal{B}_k(\pi) + R(\pi).$$

with $k = \dim(\delta_0)$ and the residual effect is given by $R(\pi)$.

Proof: A Taylor approximation of $A_{\pi n}(\tilde{\delta}(\pi), \hat{\delta}, \hat{\theta}, \hat{\eta})$ in the third component around θ yields

$$\begin{aligned}
A_{\pi n}(\tilde{\delta}(\pi), \hat{\delta}, \hat{\theta}) &= \sum_{t=1}^{n} \Big(l_t(\tilde{\delta}(\pi), \hat{\theta}) - l_t(\hat{\theta}, \hat{\theta}) \Big) \\
&= \sum_{t=1}^{n} \Big(l_t(\tilde{\delta}(\pi), \theta) - l_t(\hat{\theta}, \theta) \Big) \\
&+ \sum_{t=1}^{n} \left(\frac{\partial}{\partial \theta} \Big(l_t(\tilde{\delta}(\pi), \theta) - l_t(\hat{\delta}, \theta) \Big) (\hat{\theta} - \theta) \right) \\
&+ \frac{1}{2} \sum_{t=1}^{n} (\hat{\theta} - \theta)' \frac{\partial^2}{\partial \theta \partial \theta'} \Big(l_t(\tilde{\delta}(\pi), \theta) - l_t(\hat{\delta}, \theta) \Big) (\hat{\theta} - \theta) \\
&+ o_p(1) \\
&= A_{\pi n}(\tilde{\delta}(\pi), \hat{\delta}, \theta) + B_{\pi n} + \frac{1}{2} C_{\pi n} + o_p(1).
\end{aligned}$$

It holds that

$$C_n(\pi) = \sqrt{n}(\hat{\theta} - \theta)' \frac{\partial^2}{\partial \theta \partial \theta'} \frac{1}{n} \sum_{t=1}^{n} \Big(l_t(\tilde{\delta}(\pi), \theta) - l_t(\hat{\delta}, \theta) \Big) \sqrt{n}(\hat{\theta} - \theta)$$

Then, $C_n(\pi) \Rightarrow_p 0$ and $B_n(\pi) \Rightarrow_d R_(\pi)$. So, $A_{\pi n} \Rightarrow \mathcal{B}_k(\pi) + R(\pi)$. □

Simulation evidence supports using a residual bootstrap scheme, which leads to correctly sized tests. Under the assumption of proper transformation prior to step two, we can use a simple residual bootstrap scheme, which is now briefly lined out: Let a sample from $\hat{Z}_1, ..., \hat{Z}_n$, drawn with replacement, be denoted by $\hat{Z}_1^*, ..., \hat{Z}_n^*$. For any bootstrap repetition b, let the sup-LR test statistic from (3.1) or (3.2) be denoted by $\sup A_j^b$, such that the p-value follows as

$$\hat{p} = \frac{1}{B} \sum_{b=1}^{B} 1_{\{\sup A_j^b > \sup A_j\}} \tag{3.7}$$

If the estimation error in the first step could be ignored, it would be reasonable to use the same critical values as in Hypothesis Pair 2, because the difference of estimated parameters under alternative and null remains exactly the same.

Testing parameter constancy is straightforward in financial return data, if it can be reasonably assumed that these data represent draws from a weakly stationary distribution. Note that the i.i.d. assumption is not crucial here as long as weak stationarity is satisfied, since correctly-sized can be obtained by using an appropriate covariance-matrix, as pointed out for example by Blatt et al. [2015]. With this preliminary remarks out of the way, the sup-LR framework obviously requires assumption of a particular parametric model, while the particular moments which are subject to structural changes need to be specified in the fluctuation test framework

3.1 Gaussian Distribution

The easiest choice is a multivariate Gaussian, parametrized in terms of means, variances and correlations. Although most likely not the best choice in many cases, it provides a good starting point and offers some useful insight into more complicated models. We impose the regularity condition, that variances are bounded away from zero. Decomposing the entire covariance matrix into

$$\Sigma = S'PS = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_m \end{pmatrix} \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{12} & 1 & \cdots & \rho_{2n} \\ \vdots & \ddots & 1 & \vdots \\ \rho_{1n} & \rho_{2n} & \cdots & 1 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_m \end{pmatrix}$$

enables us to easily separating inference on marginal parameters and correlation matrix by writing

$$\begin{aligned}
&X_t \overset{i.i.d.}{\sim} N(\theta_{1,1}, \theta_{1,1}, \cdots, \theta_{m,1}, P_1) \quad \text{for } t = 1, ..., l_1 \\
&X_t \overset{i.i.d.}{\sim} N(\theta_{1,2}, \theta_{2,1}, \cdots, \theta_{m,1}, P_1) \quad \text{for } t = l_1 + 1, ..., l_2 \\
&\cdots\cdots \\
&X_t \overset{i.i.d.}{\sim} N(\theta_{1,2}, \theta_{2,2}, \cdots, \theta_{m,2}, P_1) \quad \text{for } t = l_m + 1, ..., l_D \\
&X_t \overset{i.i.d.}{\sim} N(\theta_{1,2}, \theta_{2,2}, \cdots, \theta_{m,2}, P_2) \quad \text{for } t = l_D, ..., n
\end{aligned}$$

and testing Hypothesis Pair 3 with $\theta_i = (\sigma_i^2, \mu_i)$ and $\delta_i = P_i$. We start with testing constant margins: from the probability density of a Gaussian random variable

$$f(X_{i,t}; \mu_i, \sigma_i^2) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\Big(- \frac{(X_{i,t} - \mu_i)^2}{2\sigma_i^2} \Big)$$

the log-Likelihood for full-sample estimation is given by

$$L(X_i, \mu_0, \sigma_0^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma_0^2) - \sum_{t=1}^{n} \frac{(X_{t,i} - \mu_0)^2}{2\sigma_0^2}$$

Dividing the sample at any j yields

$$\begin{aligned}
L(X_i, \mu_{i,1}, \mu_{i,2}, \sigma_{i,1}^2, \sigma_{i,2}^2) = &-\frac{j}{2}\log(2\pi) - \frac{j}{2}\log(\sigma_{i,1}^2) - \sum_{t=1}^{j} \frac{(X_{i,t} - \mu_{i,1})^2}{2\sigma_{i,1}^2} \\
&- \frac{n-j}{2}\log(2\pi) - \frac{n-j}{2}\log(\sigma_{i,2}^2) - \sum_{t=j+1}^{n} \frac{(X_{i,t} - \mu_{i,2})^2}{2\sigma_{i,2}^2}
\end{aligned}$$

for the log-likelihood. It should be noted, that under serial independence the log-likelihood is completely separated in terms of $(\mu_{i,1}, \sigma_{i_1}^2)$ and $(\mu_{i,2}, \sigma_{i,2}^2)$, so maximum-likelihood estimators are derived for each sub-sample the usual way. Evaluating the difference of the log-likelihood under full-sample and partial-sample estimators gives after some simplifications the test statistic for a fixed j:

$$A_j(X_i; \hat{\mu}_{i,0}, \hat{\mu}_{i,1}, \hat{\mu}_{i,2}, \hat{\sigma}_{i,0}^2, \hat{\sigma}_{i,1}^2, \hat{\sigma}_{i,2}^2) = n\log(\hat{\sigma}_{i,0}^2) - j\log(\hat{\sigma}_{i,1}^2) - (n-j)\log(\hat{\sigma}_{i,2}^2) \quad (3.8)$$

The limiting process of $\{A_j\}$ is of the form of equation (3.3) and has $k = 2$ degrees of freedom. It can be easily checked, that Assumption 1 holds in this case. Conditional on the test decision, the data are standardized by:

$$\hat{Z}_{i,t} = \frac{X_{i,t} - \hat{\mu}_{i,1}\mathbb{1}_{t\leq\hat{l}_1} - \hat{\mu}_{i,2}\mathbb{1}_{t>\hat{l}_1}}{\sqrt{\hat{\sigma}^2_{i,1}\mathbb{1}_{t\leq\hat{l}_1} + \hat{\sigma}^2_{i,2}\mathbb{1}_{t>\hat{l}_1}}} \text{ if a break is detected or } \hat{Z}_t = \frac{X_{i,t} - \hat{\mu}_i}{\hat{\sigma}_i} \text{ else} \tag{3.9}$$

For the piecewise standardized data, full-sample and partial-sample ML-estimators follow from the simplified log-likelihood, now given by

$$L(\hat{Z}; P_0) = -\frac{n}{2}\log|P_0| - \frac{1}{2}\sum_{t=1}^{n} \hat{Z}_t' P_0^{-1}\hat{Z}_t$$

$$L(\hat{Z}; P_1, P_2) = -\frac{j}{2}\log|P_1| - \frac{1}{2}\sum_{t=1}^{j} \hat{Z}_t' P_1^{-1}\hat{Z}_t - \frac{n-j}{2}\log|P_2| - \frac{1}{2}\sum_{t=j+1}^{n} \hat{Z}_t' P_2^{-1}\hat{Z}_t$$

yielding

$$\hat{P}_0 = \frac{1}{n}\sum_{t=1}^{n}\begin{pmatrix} \hat{Z}_{1,t}\hat{Z}_{2,t} & \cdots \\ \cdots & \hat{Z}_{m-1,t}\hat{Z}_{m,t}\end{pmatrix}, \text{ and}$$

$$\hat{P}_1 = \frac{1}{j}\sum_{t=1}^{j}\begin{pmatrix} \hat{Z}_{1,t}\hat{Z}_{2,t} & \cdots \\ \cdots & \hat{Z}_{m-1,t}\hat{Z}_{m,t}\end{pmatrix}, \quad \hat{P}_2 = \frac{1}{n-j}\sum_{t=j+1}^{n}\begin{pmatrix} \hat{Z}_{1,t}\hat{Z}_{2,t} & \cdots \\ \cdots & \hat{Z}_{m-1,t}\hat{Z}_{m,t}\end{pmatrix}$$

where it was used that $\sum_{t=1}^{j}\hat{Z}^2_{i,t} = 1$ and $\sum_{t=j+1}^{n}\hat{Z}^2_{i,t} = 1$ for every dimension i by construction of $\hat{Z}$. Given j, the likelihood-ratio test statistic for centered and standardized Gaussian data is obtained as

$$A_j = n \cdot \log(|\hat{P}_0|) - j \cdot \log(|\hat{P}_1|) - (n-j)\cdot\log(|\hat{P}_2|) \tag{3.10}$$

Had one based the test statistic on the unobserved Z_t, the critical value associated with the sup-functional $\sup_{\underline{\pi}\cdot n\leq j\leq\overline{\pi}\cdot n} A_j$ would be given by $\sup_{\pi\in\Pi}\mathcal{B}_{(m-1)m/2}(\pi)$.

Remark 1. For the two-dimensional case we illustrate why Assumption 2.3 is a plausible assumption. Without loss of generality, assume that $t < \pi n$. Note that dropping this assumption would lead to gradients and information matrices with block structure. The results remain unchanged, expressions become substantially more cumbersome and do not add much to the argument. For details, see Demetrescu and Wied [2018+].
The log-likelihood contribution of observation t can be written as

$$l_t(\rho, \hat{\theta}) = \frac{1}{2}\frac{1}{1-\rho^2}\left(\frac{(X_{1,t}-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(X_{1,t}-\mu_1)(X_{2,t}-\mu_2)}{\sigma_1\sigma_2} + \frac{(X_{2,t}-\mu_2)^2}{\sigma_2^2}\right)$$

We will consider the case of a break in dimension $i = 1$ at $\hat{l} = \hat{\pi}_1 n$ with $1 < \hat{l}_1 < n$ and constant parameters for $i = 2$, so $\theta = (\mu_{1,1}, \mu_{1,2}, \sigma^2_{1,1}, \sigma^2_{1,2}, \mu_2, \sigma^2_2)$. In order to verify

Assumption 2.3, write

$$
\frac{\partial}{\partial\theta} l_t(\rho,\theta) = \frac{1}{1-\rho^2}
\begin{pmatrix}
\mathbb{1}_{t\le\hat{l}_1}\left(\rho\frac{X_{2,t}-\mu_2}{\sigma_{1,1}\sigma_2} - \frac{X_{1,t}-\mu_{1,1}}{\sigma_{1,1}^2}\right) \\
\mathbb{1}_{t>\hat{l}_1}\left(\rho\frac{X_{2,t}-\mu_2}{\sigma_{1,2}\sigma_2} - \frac{X_{1,t}-\mu_{1,2}}{\sigma_{1,2}^2}\right) \\
\mathbb{1}_{t\le\hat{l}_1}\left(\rho\frac{(X_{1,t}-\mu_{1,1})(X_{2,t}-\mu_2)}{\sigma_{1,1}^3\sigma_2} - \frac{(X_{1,t}-\mu_{1,1})^2}{\sigma_{1,1}^4}\right) \\
\mathbb{1}_{t>\hat{l}_1}\left(\rho\frac{(X_{1,t}-\mu_{1,2})(X_{2,t}-\mu_2)}{\sigma_{1,2}^3\sigma_2} - \frac{(X_{1,t}-\mu_{1,2})^2}{\sigma_{1,2}^4}\right) \\
\mathbb{1}_{t\le\hat{l}_1}\left(\rho\frac{X_{1,t}-\mu_{1,1}}{\sigma_{1,1}\sigma_2} - \frac{X_{2,t}-\mu_2}{\sigma_2^2}\right) + \mathbb{1}_{t>\hat{l}_1}\left(\rho\frac{X_{1,t}-\mu_{1,2}}{\sigma_{1,2}\sigma_2} - \frac{X_{2,t}-\mu_2}{\sigma_2^2}\right) \\
\mathbb{1}_{t\le\hat{l}_1}\left(\rho\frac{(X_{1,t}-\mu_{1,1})(X_{2,t}-\mu_2)}{\sigma_{1,1}\sigma_2^3} - \frac{(X_{2,t}-\mu_2)^2}{\sigma_2^4}\right) + \\
\mathbb{1}_{t>\hat{l}_1}\left(\rho\frac{(X_{1,t}-\mu_{1,2})(X_{2,t}-\mu_2)}{\sigma_{1,2}\sigma_2^3} - \frac{(X_{2,t}-\mu_2)^2}{\sigma_2^4}\right)
\end{pmatrix}
$$

$$
:= \frac{1}{1-\rho^2} G_\rho
$$

The first gradient of $l_t(\cdot)$ w.r.t. $\mu_{1,1}$ and $\sigma_{1,1}^2$ is zero for $t > \hat{l}_1$ and vice versa. This is not the case for μ_2 and σ_2^2, which affect all likelihood-contributions but to varying degree, depending on the sub-sample. To make this explicit, the difference in gradients of the log-likelihoods with respect to ρ can be written as

$$
\begin{aligned}
\sum_{t=1}^{n} \frac{\partial}{\partial\theta}\Big(l_t(\tilde{\rho},\theta) - l_t(\hat{\rho},\theta)\Big) &= \sum_{t=1}^{\hat{l}_1}\Big(\frac{1}{1-\tilde{\rho}^2}G_{\tilde{\rho}} - \frac{1}{1-\hat{\rho}^2}G_{\hat{\rho}}\Big) + \sum_{t=\hat{l}_1+1}^{n}\Big(\frac{1}{1-\tilde{\rho}^2}G_{\tilde{\rho}} - \frac{1}{1-\hat{\rho}^2}G_{\hat{\rho}}\Big) \\
&= \sum_{t=1}^{\hat{l}_1}\frac{1}{1-\hat{\rho}^2}\Big(G_{\hat{\rho}} - G_{\tilde{\rho}}\Big) + G_{\tilde{\rho}}\Big(\frac{1}{1-\hat{\rho}^2} - \frac{1}{1-\tilde{\rho}^2}\Big) \\
&+ \sum_{t=\hat{l}_1+1}^{n}\frac{1}{1-\hat{\rho}^2}\Big(G_{\hat{\rho}} - G_{\tilde{\rho}}\Big) + G_{\tilde{\rho}}\Big(\frac{1}{1-\hat{\rho}^2} - \frac{1}{1-\tilde{\rho}^2}\Big)
\end{aligned}
$$

Since both $\hat{\rho}$ and $\tilde{\rho}$ are consistent estimators and $\lim_{n\to\infty} G_{\tilde{\rho}} = \lim_{n\to\infty} G_{\hat{\rho}}$ for both sub-samples, the likelihood functions also converge to the same limit, i.e. their differences converge to zero. This is a necessary condition for the fact that

$$
\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial}{\partial\theta}\Big(l_t(\tilde{\rho},\theta) - l_t(\hat{\rho},\theta)\Big)
$$

converges to a limit process in $D[0,1]$. Since this limit process is non-standard, appropriately correcting $A_j(\tilde{\delta},\hat{\delta},\hat{\theta})$ for the residual effect is difficult. This feature gives rise to the bootstrap schemes laid out previously.

3.2 Gaussian Copula

Alternatively step 2 can also be based on the copula associated with the Gaussian distribution assumption. Step 1 remains unchanged, however the data are now (piecewise) transformed onto the copula scale by

$$
\begin{aligned}
\hat{U}_{i,t} &= F(X_i, \hat{\mu}_{i,1}, \hat{\sigma}_{i,1}) \quad \text{for } i = 1, ..., \hat{l}_i \\
\hat{U}_{i,t} &= F(X_i, \hat{\mu}_{i,2}, \hat{\sigma}_{i,2}) \quad \text{for } i = \hat{l}_i + 1, ..., n \quad \text{if the test rejects} \\
\hat{U}_{i,t} &= F(X_i, \hat{\mu}_{i,0}, \hat{\sigma}_{i,0}) \quad \text{for } i = 1, ..., n \quad \text{if not}
\end{aligned} \tag{3.11}
$$

The pseudo-observations are then used to estimate the dependency parameter (i.e. the correlation matrix) of the Gaussian copula under the null and alternative hypothesis. Consider next the density of the Gaussian copula

$$f(\hat{U}_t; P) = |R|^{-\frac{1}{2}} \exp\Bigg(-\frac{1}{2}\hat{U}_t'(R^{-1} - I)\hat{U}_t\Bigg)$$

from where the full-sample log-likelihood

$$L(\hat{U}; P_0) = -\frac{n}{2}|R| - \frac{1}{2}\sum_{t=1}^{n}\hat{U}_t'(R_0^{-1} - I)\hat{U}_t$$

and the partial-sample log-likelihood

$$L(\hat{U}; P_1, P_2) = -\frac{j}{2}|R_1| - \frac{n-j}{2}|R_2| - \frac{1}{2}\sum_{t=1}^{j}\hat{U}_t'(R_1^{-1} - I)\hat{U}_t - \frac{1}{2}\sum_{t=j+1}^{n}\hat{U}_t'(R_2^{-1} - I)\hat{U}_t$$

are obtained. Let $\hat{R}_0$, $\hat{R}_1$ and $\hat{R}_2$ denote the ML-estimators for the correlation matrix of the full sample and each sub-sample. Evaluating the log-likelihood at the respective parameter estimates gives the test statistic for a fixed j as

$$A_j = 2\big(L(\hat{U}; \hat{R}_1, \hat{R}_2) - L(\hat{U}; \hat{R}_0)\big). \tag{3.12}$$

Had one based the test statistic on the unobserved Z_t, the critical value associated with the sup-functional $\sup_{\underline{\pi}\cdot n \leq j \leq \overline{\pi}\cdot n} A_j$ would be given by $\sup_{\pi\in\Pi} \mathcal{B}_{(m-1)m/2}(\pi)$.

3.3 Bivariate t-Distribution

In many financial applications with moderate observation frequencies (e.g. monthly or weekly), the heavy-tailed t-distribution yields a better fit than the Gaussian distribution, see Cont [2001] who collects empirical facts on asset returns. Therefore we now turn to the problem of testing parameter stability under the assumption to observe data from

$$\begin{aligned}
(X_{1,t}, X_{2,t}) &\overset{i.i.d.}{\sim} t(\mu_{1,1}, \mu_{2,1}, \xi_{11,1}, \xi_{22,1}, \rho_1, \nu) \quad \text{for } t = 1, ..., l_1 \\
(X_{1,t}, X_{2,t}) &\overset{i.i.d.}{\sim} t(\mu_{1,2}, \mu_{2,1}, \xi_{11,2}, \xi_{22,2}, \rho_1, \nu) \quad \text{for } t = l_1 + 1, ..., l_2 \\
(X_{1,t}, X_{2,t}) &\overset{i.i.d.}{\sim} t(\mu_{1,2}, \mu_{2,2}, \xi_{11,2}, \xi_{22,2}, \rho_1, \nu) \quad \text{for } t = l_2 + 1, ..., l_D \\
(X_{1,t}, X_{2,t}) &\overset{i.i.d.}{\sim} t(\mu_{1,2}, \mu_{2,1}, \xi_{11,2}, \xi_{22,2}, \rho_2, \nu) \quad \text{for } t = l_D, ..., n
\end{aligned}$$

where ν denotes the degrees of freedom and Ξ denotes the dispersion matrix, such that the covariance matrix follows as $\Sigma = \frac{\nu}{\nu-2}\Xi$. Similar to the Gaussian case, we impose that ν and ξ are bounded away from zero. The correlation coefficient satisfies

$$\rho_{12} = \frac{\frac{\nu}{\nu-2}\xi_{12}}{\sqrt{\frac{\nu}{\nu-2}\xi_{11}} \cdot \sqrt{\frac{\nu}{\nu-2}\xi_{22}}} = \frac{\xi_{12}}{\sqrt{\xi_{11}}\sqrt{\xi_{22}}}$$

and the t-distribution is equivalently parametrized - similar to the covariance decomposition in the Gaussian case - in terms of the cross-dispersion and the correlation:

$$\Xi = \begin{pmatrix} \xi_{11} & \sqrt{\xi_{11}}\sqrt{\xi_{22}}\rho \\ \sqrt{\xi_{11}}\sqrt{\xi_{22}}\rho & \xi_{22} \end{pmatrix} = \begin{pmatrix} \xi_{11} & \xi_{12} \\ \xi_{12} & \xi_{22} \end{pmatrix}$$

By the properties of the multivariate t-distribution, each marginal distribution i satisfies

$$\begin{aligned} X_{i,t} &\overset{i.i.d.}{\sim} t(\mu_{i,1}, \xi_{i,1}, \nu) \quad \text{for } t = 1, ..., l_i \\ X_{i,t} &\overset{i.i.d.}{\sim} t(\mu_{i,2}, \xi_{i,2}, \nu) \quad \text{for } t = l_i + 1, ..., n \end{aligned}$$

and can test 1 by setting $\theta_i = (\mu_i, \xi_i)$. From the distributional assumption, the probability density is given by

$$f(X_t; \mu, \Xi, \nu) = \frac{\Gamma(\frac{\nu+m}{2})}{\Gamma(\frac{\nu}{2})(\pi\nu)^{m/2}|\Xi|^{0.5}} \Big(1 + \frac{1}{\nu}(X_t - \mu)'\Xi^{-1}(X_t - \mu)\Big)^{(-\frac{\nu+m}{2})}$$

from where the marginal density of dimension i follows as

$$f(X_{i,t}; \mu_i, \xi_i, \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu\xi_i^2}} \Big(1 + \frac{(X_{i,t} - \mu_i)^2}{\nu\xi_i^2}\Big)^{(-\frac{\nu+1}{2})}$$

Although degrees of freedom are assumed to be constant in, they nevertheless have to be estimated in finite samples. This is done before testing marginal distributions by maximizing the log-likelihood associated with the joint distribution. No closed-form solution exists for maximizing the log-likelihood, so one has to use numerical methods to find the ML-estimator for μ, Ξ and ν. We refer to Liu and Rubin [1995] for a detailed description of the EMCE-algorithm typically used in this context. Let $\bar{\nu}$ denote the ML-estimator of the degrees of freedom for the multivariate distribution, which is now fixed when testing Hypothesis Pair 1 for each margin. Using the same separation of the log-likelihood in terms of (μ_1, ξ_1) and (μ_2, ξ_2) as in the Gaussian case, the EMCE-algorithm delivers the corresponding ML-estimator. Full-sample estimators $(\hat{\mu}_0, \hat{\xi}_0)$ are obtained accordingly, which are plugged back into the LR-statistic together with $(\hat{\mu}_1, \hat{\mu}_2, \hat{\xi}_1, \hat{\xi}_2)$. After omitting constants one obtains for a fixed j

$$\begin{aligned} A_{i,j} &= 2\big((X_i; \hat{\mu}_{i,1}, \hat{\mu}_{i,2}, \hat{\xi}_{i,1}, \hat{\xi}_{i,2}, \bar{\nu}) - L(X_i; \hat{\mu}_{i,0}, \hat{\xi}_{i,0}, \bar{\nu})\big) \\ &= n \cdot \log(\hat{\xi}_{i,0}^2) - j \cdot \log(\hat{\xi}_{i,1}^2) - (n-j) \cdot \log(\hat{\xi}_{i,2}^2) \\ &- (\bar{\nu}+1) \sum_{t=1}^{j} \log\Big(1 + \frac{1}{\bar{\nu}}(\frac{X_{i,t} - \hat{\mu}_{i,1}}{\hat{\xi}_{i,1}})^2\Big) - (\bar{\nu}+1) \sum_{t=j+1}^{n} \log\Big(1 + \frac{1}{\bar{\nu}}(\frac{X_{i,t} - \hat{\mu}_{i,2}}{\hat{\xi}_{i,2}})^2\Big) \\ &+ (\bar{\nu}+1) \sum_{t=1}^{n} \log\Big(1 + \frac{1}{\bar{\nu}}(\frac{X_{i,t} - \hat{\mu}_{i,0}}{\hat{\xi}_{i,0}})^2\Big) \end{aligned} \tag{3.13}$$

and Hypothesis Pair 1 is tested using

$$\sup_{\underline{\pi}\cdot n \le j \le \overline{\pi}\cdot n} A_j \to_d \sup_{\Pi} \mathcal{B}_2(\pi)$$

Testing constant dependency is specified by recognizing that $\delta_D = \rho_{12}$ under the assumption of constant degrees of freedom. Since the multivariate t-distribution is a location-scale family in (μ, ξ), a standardization similar to the Gaussian case

$$\hat{Z}_{i,t} = \frac{X_{i,t} - \hat{\mu}_{i,1}\mathbb{1}_{t\leq\hat{l}_1} - \hat{\mu}_{i,2}\mathbb{1}_{t>\hat{l}_1}}{\sqrt{\hat{\xi}_{i,1}\mathbb{1}_{t\leq\hat{l}_1} + \hat{\xi}_{i,2}\mathbb{1}_{t>\hat{l}_1}}} \text{ if a break is detected or } \hat{Z}_t = \frac{X_{i,t} - \hat{\mu}_i}{\sqrt{\hat{\xi}_i}} \text{ else} \quad (3.14)$$

leaves us with

$$\hat{Z}_t \overset{i.i.d.}{\sim} t(0, 0, 1, 1, \rho, \overline{\nu})$$

since for standardized data $\rho = \xi_{12}$. The log-likelihood simplifies considerably, so a simple line search on the first-order condition now suffices to obtain full-sample and partial-sample estimators $\hat{\xi}_{12}$. The LR-statistic for a constant j is given by

$$\begin{aligned} A_j(\hat{Z}, \hat{\rho}_0, \hat{\rho}_1, \hat{\rho}_2) = {} & n \log(1 - \hat{\rho}_0^2) - j \log(1 - \hat{\rho}_1^2) - (n-j)\log(1 - \hat{\rho}_1^2) \\ & + (\overline{\nu} + 2) \sum_{t=1}^{n} \log\left(1 + \frac{1}{\overline{\nu}} \frac{\hat{Z}_{1,t}^2 - 2\hat{\rho}_0 \hat{Z}_{1,t}\hat{Z}_{2,t} + \hat{Z}_{2,t}^2}{1 - \hat{\rho}_0^2}\right) \\ & - (\overline{\nu} + 2) \sum_{t=1}^{j} \log\left(1 + \frac{1}{\overline{\nu}} \frac{\hat{Z}_{1,t}^2 - 2\hat{\rho}_1 \hat{Z}_{1,t}\hat{Z}_{2,t} + \hat{Z}_{2,t}^2}{1 - \hat{\rho}_1^2}\right) \\ & - (\overline{\nu} + 2) \sum_{t=j+1}^{n} \log\left(1 + \frac{1}{\overline{\nu}} \frac{\hat{Z}_{1,t}^2 - 2\hat{\rho}_2 \hat{Z}_{1,t}\hat{Z}_{2,t} + \hat{Z}_{2,t}^2}{1 - \hat{\rho}_2^2}\right) \end{aligned} \quad (3.15)$$

and the test statistic against a single break follows as

$$\sup_{\underline{\pi}\cdot n \leq j \leq \overline{\pi}\cdot n} A_j$$

Again, $\sup_{\Pi} \mathcal{B}_1(\pi)$ would emerge as the asymptotic distribution, if (3.15) would be based directly on observed data. Extensions to the multivariate case are obtained analogously to the Gaussian case. Because of the high computational effort, the lack of additional insight and the more flexible way to handle t-distributed random variables presented in the next section, this is not pursued further.

3.4 t-Copula

Instead, we use a consequent extension to model marginal and joint distribution in separate steps using the concept of copulae and allows to relax the assumption of constant degrees of freedom. More specifically it assumed that the observations are sampled from univariate t-distributions, while the underlying DGP is a t-copula, see Demarta and McNeil [2005]. Maintaining the (piecewise) i.i.d. assumption we now have

$$\begin{aligned} (X_{1,t}, X_{2,t}) &\overset{i.i.d.}{\sim} t(\mu_{1,1}, \xi_{1,1}, ..., \mu_{m,1}, \xi_{m,1}, P_1, \vec{\nu}) \quad \text{for } t = 1, ..., l_1 \\ (X_{1,t}, X_{2,t}) &\overset{i.i.d.}{\sim} t(\mu_{1,2}, \xi_{1,2}, ..., \mu_{m,1}, \xi_{m,1}, P_1, \vec{\nu}) \quad \text{for } t = l_1 + 1, ..., l_2 \\ & \quad \cdots\cdots \\ (X_{1,t}, X_{2,t}) &\overset{i.i.d.}{\sim} t(\mu_{1,2}, \xi_{1,2}, ..., \mu_{m,2}, \xi_{m,2}, P_1, \vec{\nu}) \quad \text{for } t = l_m + 1, ..., l_D \\ (X_{1,t}, X_{2,t}) &\overset{i.i.d.}{\sim} t(\mu_{1,2}, \xi_{1,2}, ..., \mu_{m,2}, \xi_{m,2}, P_2, \vec{\nu}) \quad \text{for } t = l_D, ..., n \end{aligned}$$

where $\vec{\nu} = (\nu_1, ..., \nu_m, \nu_D)$ is an $m+1$-vector of the degrees of freedom, which are assumed to be constant over time but now are free to vary in the cross-section. By separating the degrees-of-freedom of marginal and joint distribution we implicitly introduce a two-stage model with t-distributed marginals and a t-copula. The correlation matrix P directly parametrizes the t-copula and can no longer be obtained from the covariance matrix. Step 1 of the Sequential Procedure now requires separate ML-estimation at the margins instead of estimating ν over the joint distribution and keeping it constant for every margin. ν is not necessarily fixed, but can be part of the ML-estimation in each sub-sample (again using an ECME-algorithm). Testing constant marginal distributions is almost unchanged, (3.15) is now computed using the full-sample ML-estimator $\hat{\nu}_i$ for each margin rather than $\overline{\nu}$. Conceptually identical to the Gaussian copula case, the observed data are now transformed by the cumulative distribution function of the t-distribution, denoted $F(X_{i,t}, \theta_{i,t})$ evaluated at the ML-estimator $\hat{\theta}_i$:

$$\begin{aligned}
&\hat{U}_{i,t} = F(X_i, \hat{\mu}_{i,1}, \hat{\sigma}_{i,1}, \hat{\nu}_i) \quad \text{for } i = 1, ..., \hat{l}_i \\
&\hat{U}_{i,t} = F(X_i, \hat{\mu}_{i,2}, \hat{\sigma}_{i,2}, \hat{\nu}_i) \quad \text{for } i = \hat{l}_i + 1, ..., n \quad \text{if the test rejects} \\
&\hat{U}_{i,t} = F(X_i, \hat{\mu}_{i,0}, \hat{\sigma}_{i,0}, \hat{\nu}_i) \quad \text{for } i = 1, ..., \hat{n} \quad \text{if not}
\end{aligned} \tag{3.16}$$

The test for a constant t-copula is now based on $\hat{U}$. From the probability density of the t-copula

$$c(\hat{U}_t; P, \nu) = \frac{\Gamma(\frac{\nu+m}{2})\big(\Gamma(\frac{\nu}{2})\big)^{m-1}}{\big(\Gamma(\frac{\nu+1}{2})\big)^m |P|^{0.5}} \left(\prod_{i=1}^{m} \big(1 + \frac{\hat{Y}_{i,t}^2}{\nu}\big)^{\frac{\nu+1}{2}} \right) \left(1 + \frac{1}{\nu} \hat{Y}_t' P^{-1} \hat{Y}_t \right)$$

with $\hat{Y}_{i,t} = F^{-1}(\hat{U}_{i,t}, \nu)$ denoting the quantile function of a standardized t-distribution and $l\Gamma$ the log $-\Gamma$-function, the log-likelihood follows as

$$\begin{aligned}
L(\hat{U}; P_0, \nu) =& n \cdot \left(l\Gamma(\frac{\nu+m}{2}) + (m-1) l\Gamma(\frac{\nu}{2}) - m \cdot l\Gamma(\frac{\nu+1}{2}) - 0.5 \log |P_0| \right) \\
&+ \sum_{t=1}^{j} \left(\frac{\nu+1}{2} \sum_{i=1}^{m} \log \left(1 + \frac{\hat{Y}_{i,t}^2}{\nu} \right) - \frac{\nu+m}{2} \log \left(1 + \frac{\hat{Y}_t' P_0^{-1} \hat{Y}_t}{\nu_0} \right) \right)
\end{aligned} \tag{3.17}$$

for the full-sample and

$$\begin{aligned}
L(\hat{U}; P_1, P_2, \nu) =& j \cdot \left(l\Gamma(\frac{\nu+m}{2}) + (m-1) l\Gamma(\frac{\nu}{2}) - m \cdot l\Gamma(\frac{\nu+1}{2}) - 0.5 \log |P_1| \right) \\
&+ \sum_{t=1}^{j} \left(\frac{\nu_0+1}{2} \sum_{i=1}^{m} \log \left(1 + \frac{\hat{Y}_{i,t}^2}{\nu_0} \right) - \frac{\nu_1+m}{2} \log \left(1 + \frac{\hat{Y}_t' P_1^{-1} \hat{Y}_t}{\nu_0} \right) \right) \\
&(n-j) \cdot \left(l\Gamma(\frac{\nu+m}{2}) + (m-1) l\Gamma(\frac{\nu}{2}) - m \cdot l\Gamma(\frac{\nu+1}{2}) - 0.5 \log |P_2| \right) \\
&+ \sum_{t=j+1}^{n} \left(\frac{\nu+1}{2} \sum_{i=1}^{m} \log \left(1 + \frac{\hat{Y}_{i,t}^2}{\nu} \right) - \frac{\nu+m}{2} \log \left(1 + \frac{\hat{Y}_t' P_2^{-1} \hat{Y}_t}{\nu} \right) \right)
\end{aligned} \tag{3.18}$$

for the partial samples. Similar to the multivariate t-distribution, ML-estimation of (3.17) and (3.18) requires numerical methods, such as EM-algorithms. Let $(\hat{P}_0, \hat{\nu})$ and $(\hat{P}_1, \hat{P}_2)$

denote the full-sample and partial-sample ML-estimator, we have for a fixed j:

$$A_j = 2\left(L(\hat{U};\hat{P}_1,\hat{P}_2,\hat{\nu}) - L(\hat{U};\hat{P}_0,\hat{\nu}) \right) \tag{3.19}$$

The corresponding sup-LR test statistic would follow $\sup_{\pi\in\Pi} \mathcal{B}_{(m-1)m/2}(\pi)$ under the null hypothesis, if the residual effect could be ignored. Full ML-estimation of the t-copula is extremely time-consuming, particularly in higher dimensions, Demarta and McNeil [2005] therefore suggest a semi-parametric pseudo-ML procedure sharing the asymptotic properties of full ML-estimation. In a first step, the empirical Kendall's tau matrix $\hat{P}^\tau$ of the data transformed as in (3.17) is calculated as

$$\hat{P}^\tau = \begin{pmatrix} \hat{\rho}_\tau(\hat{Z}_1,\hat{Z}_1) & \cdots & \hat{\rho}_\tau(\hat{Z}_1,\hat{Z}_n) \\ \vdots & \ddots & \vdots \\ \hat{\rho}_\tau(\hat{Z}_n,\hat{Z}_1) & \cdots & \hat{\rho}_\tau(\hat{Z}_n,\hat{Z}_n) \end{pmatrix}$$

where each element is given as the empirical pairwise Kendall's tau coefficient.

$$\hat{\rho}_\tau(\hat{Z}_n,\hat{Z}_n) = \binom{n}{2}^{-1} \sum_{1\le t_1<t_2\le n} \operatorname{sign}\left((\hat{Z}_{t_1,i} - \hat{Z}_{t_2,i})(\hat{Z}_{t_1,j} - \hat{Z}_{t_2,j}) \right)$$

The empirical Kendall's tau matrix serves to construct a method-of-moments estimator for P by $P^* = \sin(\frac{\pi}{2}\hat{P}^\tau)$ and subsequently estimate ν_C, holding P^* fixed. As in the case of the multivariate t-distribution, a one-dimensional line search is required to compute $\hat{\nu}_C$. Following Mashal and Zeevi [2002] one can perform a simple bisection algorithm over the first-order condition of the log-likelihood with respect to ν_C. As pointed out by Mashal and Zeevi [2002] using Pseudo-ML-estimators affects the limit distribution. This is unproblematic in our case, because a bootstrap scheme, that could also be used to approximate the appropriate limit distribution in small samples, is already at hand.
Similar to the multivariate t-distribution with constant degrees of freedom we also consider the effects of misspecification of the t-copula. More precisely, we assume that the marginal distributions are correctly specified and tested but that the underlying copula is mistakenly assumed to be Gaussian, which (as a by-product) reduces the computational effort in higher dimensions. One could of course also discuss effects of misspecification at the margins. As this section is concerned with the performance of non-parametric and parametric copula tests, the issue is omitted here.

4 Power and Size Results

Finite sample properties of the parametric and nonparametric tests with the asymptotic and bootstrapped critical values are examined in a Monte Carlo study with 1000 repetitions for each parameter constellation. Every parametric model is covered in one subsection, the setup of the Monte Carlo study is identical for each model: at first empirical rejection rates under H_0 are reported, i.e. the correlation coefficient is kept constant while there are changes in marginal parameters (**scenario 1**). The extent to which marginal parameters change is controlled by a tuning parameter s, such that μ and σ^2 change simultaneously. In a second

study we compute empirical power under changing marginal parameters (**scenario 2**). The change points are chosen to be distinct for each margin and the joint distribution.
The nominal significance level α is set to 5 %, the corresponding are either taken from Kiefer [1959] for the fluctuation tests and Andrews [1993] for the sup-LR test or simulated using 1000 Monte Carlo repetitions on a discrete grid with 10.000 elements. Additionally the Monte Carlo average of each break point estimator and their respective Monte Carlo standard deviations are shown for scenario 2.

4.1 Gaussian Distribution

For the first simulation study we generate data from a m-dimensional Gaussian distribution with distinct change-points in marginal parameters and correlation. The sample sizes are set to 100, which seems reasonable for quarterly data, 500 which should be reached either in long time series or monthly data and 1500 to approximate asymptotic behaviour. Depending on the sample size the following timing of the regime shifts is chosen, mimicking a situation of financial contagion:

n	l_1	l_2	l_D
100	50	60	70
500	250	300	350
1500	750	900	1050

Using vector notation and the covariance decomposition of the multivariate Gaussian distribution $\Sigma = S'PS$, data is generated according to

$$X_t \overset{i.i.d.}{\sim} N\left(\begin{pmatrix} 0.05 \\ 0.05 \\ 0.05 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0.4 & 0.4 \\ 0.4 & 1 & 0.4 \\ 0.4 & 0.4 & 1 \end{pmatrix} \right) \quad \text{for } t = 1, ..., l_1$$

$$X_t \overset{i.i.d.}{\sim} N\left(\begin{pmatrix} 0.06 - 0.01s \\ 0.05 \\ 0.05 \end{pmatrix}, \begin{pmatrix} s & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0.4 & 0.4 \\ 0.4 & 1 & 0.4 \\ 0.4 & 0.4 & 1 \end{pmatrix} \right) \quad \text{for } t = l_1, ..., l_2$$

$$X_t \overset{i.i.d.}{\sim} N\left(\begin{pmatrix} 0.06 - 0.01s \\ 0.06 - 0.01s \\ 0.06 - 0.01s \end{pmatrix}, \begin{pmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & s \end{pmatrix}, \begin{pmatrix} 1 & 0.4 & 0.4 \\ 0.4 & 1 & 0.4 \\ 0.4 & 0.4 & 1 \end{pmatrix} \right) \quad \text{for } t = l_2, ..., l_D$$

$$X_t \overset{i.i.d.}{\sim} N\left(\begin{pmatrix} 0.06 - 0.01s \\ 0.06 - 0.01s \\ 0.06 - 0.01s \end{pmatrix}, \begin{pmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & s \end{pmatrix}, \begin{pmatrix} 1 & \rho_2 & \rho_2 \\ \rho_2 & 1 & \rho_2 \\ \rho_2 & \rho_2 & 1 \end{pmatrix} \right) \quad \text{for } t = l_D, ..., n$$

This is simplified to the bivariate and extended to the five-dimensional case accordingly. For scenario 1 the correlation is kept constant by setting $\rho_2 = 0.4$. In order to focus on the important aspects, the magnitude of parameter changes in each marginal distribution is identical and ranges over $s \in [0.2, 0.25, 0.5, 1, 2, 5]$. The case of $s = 1$ corresponds to testing Hypothesis Pair 2 while all other cases $s_1 = s_2 \neq 1$ test Hypothesis Pair 3 where H_0 is true. Results are shown in figure 4.1: the test for constant marginal distributions has higher power for X_1 than for X_2. This is consistent with both theory and previous studies, as we set $\lambda_1 = 0.5$ and $\lambda_2 = 0.6$.

Figure 4.1: Gaussian Distribution, Scenario 1: Rejection Rates under H_0

s	$n = 100$				$n = 500$				$n = 1500$			
	Fluctuation test		sup-LR test		Fluctuation test		sup-LR test		Fluctuation test		sup-LR test	
Margins	X_1	X_2	X_1	X_2	X_1	X_2	X_1	X_2	X_1	X_2	X_1	X_2
0.2	0.957	0.869	0.994	0.993	1	1	1	1	1	1	1	1
1/3	0.711	0.561	0.834	0.786	1	1	1	1	1	1	1	1
0.5	0.312	0.229	0.384	0.344	0.997	0.991	0.994	0.994	1	1	1	1
0.75	0.069	0.051	0.093	0.073	0.369	0.319	0.373	0.330	0.910	0.869	0.882	0.857
1	0.019	0.025	0.043	0.037	0.036	0.041	0.044	0.049	0.049	0.050	0.049	0.056
4/3	0.058	0.061	0.077	0.074	0.407	0.376	0.343	0.343	0.899	0.892	0.876	0.857
2	0.309	0.325	0.362	0.378	0.996	0.992	0.992	0.990	1	1	1	1
3	0.706	0.729	0.828	0.824	1	1	1	1	1	1	1	1
5	0.943	0.960	0.993	0.991	1	1	1	1	1	1	1	1

	Fluctuation test		sup-LR test		Fluctuation test		sup-LR test		Fluctuation test		sup-LR test	
$m = 2$	asym.	boot.	asym.	boot.	asym.	boot.	asym.	boot.	asym.	boot.	asym.	boot.
0.2	0.021	0.048	0.076	0.041	0.018	0.047	0.066	0.060	0.024	0.054	0.071	0.065
1/3	0.034	0.060	0.103	0.056	0.013	0.051	0.065	0.059	0.022	0.051	0.071	0.062
0.5	0.043	0.082	0.012	0.068	0.013	0.046	0.067	0.068	0.019	0.052	0.071	0.065
0.75	0.043	0.069	0.084	0.047	0.028	0.041	0.103	0.095	0.019	0.048	0.107	0.091
1	0.039	0.062	0.066	0.038	0.030	0.039	0.064	0.058	0.040	0.050	0.068	0.064
4/3	0.042	0.065	0.076	0.048	0.033	0.050	0.115	0.095	0.015	0.047	0.102	0.086
2	0.042	0.075	0.115	0.048	0.013	0.044	0.068	0.060	0.021	0.055	0.072	0.063
3	0.034	0.051	0.110	0.068	0.016	0.050	0.063	0.058	0.024	0.053	0.068	0.060
5	0.030	0.045	0.082	0.042	0.022	0.054	0.060	0.059	0.024	0.053	0.067	0.058

	Fluctuation test		sup-LR test		Fluctuation test		sup-LR test		Fluctuation test		sup-LR test	
$m = 3$	asym.	boot.	asym.	boot.	asym.	boot.	asym.	boot.	asym.	boot.	asym.	boot.
0.2	0.029	0.048	0.124	0.051	0.026	0.056	0.073	0.061	0.024	0.046	0.063	0.048
1/3	0.036	0.055	0.169	0.078	0.021	0.055	0.074	0.062	0.021	0.049	0.066	0.051
0.5	0.041	0.067	0.162	0.067	0.016	0.046	0.082	0.067	0.019	0.055	0.061	0.051
0.75	0.024	0.050	0.116	0.057	0.033	0.047	0.131	0.103	0.022	0.045	0.122	0.096
1	0.031	0.050	0.120	0.057	0.044	0.048	0.070	0.057	0.045	0.052	0.062	0.049
4/3	0.041	0.068	0.137	0.057	0.032	0.051	0.127	0.102	0.019	0.039	0.117	0.093
2	0.039	0.064	0.179	0.083	0.019	0.047	0.078	0.063	0.023	0.047	0.069	0.055
3	0.023	0.042	0.164	0.077	0.021	0.057	0.071	0.049	0.024	0.050	0.065	0.047
5	0.026	0.033	0.121	0.045	0.022	0.053	0.072	0.052	0.026	0.053	0.065	0.053

	Fluctuation test		sup-LR test		Fluctuation test		sup-LR test		Fluctuation test		sup-LR test	
$m = 5$	asym.	boot.	asym.	boot.	asym.	boot.	asym.	boot.	asym.	boot.	asym.	boot.
0.2	0.021	0.023	0.183	0.020	0.039	0.050	0.097	0.044	0.039	0.044	0.085	0.049
1/3	0.025	0.027	0.263	0.032	0.031	0.042	0.094	0.051	0.041	0.048	0.084	0.050
0.5	0.038	0.046	0.278	0.030	0.031	0.045	0.102	0.049	0.037	0.045	0.084	0.052
0.75	0.037	0.052	0.215	0.014	0.046	0.049	0.197	0.116	0.037	0.042	0.150	0.097
1	0.034	0.046	0.184	0.015	0.061	0.055	0.095	0.053	0.064	0.041	0.086	0.049
4/3	0.046	0.057	0.200	0.022	0.055	0.049	0.210	0.113	0.039	0.053	0.148	0.102
2	0.045	0.047	0.275	0.043	0.036	0.047	0.108	0.058	0.039	0.049	0.089	0.047
3	0.032	0.023	0.251	0.030	0.035	0.055	0.094	0.047	0.042	0.051	0.088	0.047
5	0.029	0.011	0.182	0.022	0.039	0.065	0.092	0.048	0.038	0.053	0.092	0.050

The simulations reveal presence of the residual effect for every sample size, which appears as soon as $s \neq 1$. In this case the bootstrap corrections increase power of the fluctuation test by 5 to 10 %, given a specific $\rho_2 \neq 0.4$, while the sup-LR test is corrected for the increased rejection rates under H_0. As n increases, we observe correctly sized test decisions at each margin in both test frameworks. To obtain results on empirical power, the correlation of the first regime is set to $\rho_1 = 0.4$ and ρ_2 varies symmetrically around ρ_1 from -0.1 to 0.9 in steps of 0.1. The dashed lines in figure 4.2 - 4.4 represent empirical power, if the incorrect asymptotic critical values are used, in this way one can quantify the residual effect on empirical power.

Figure 4.2: Gaussian Distribution, n=100, Scenario 2: Empirical Power

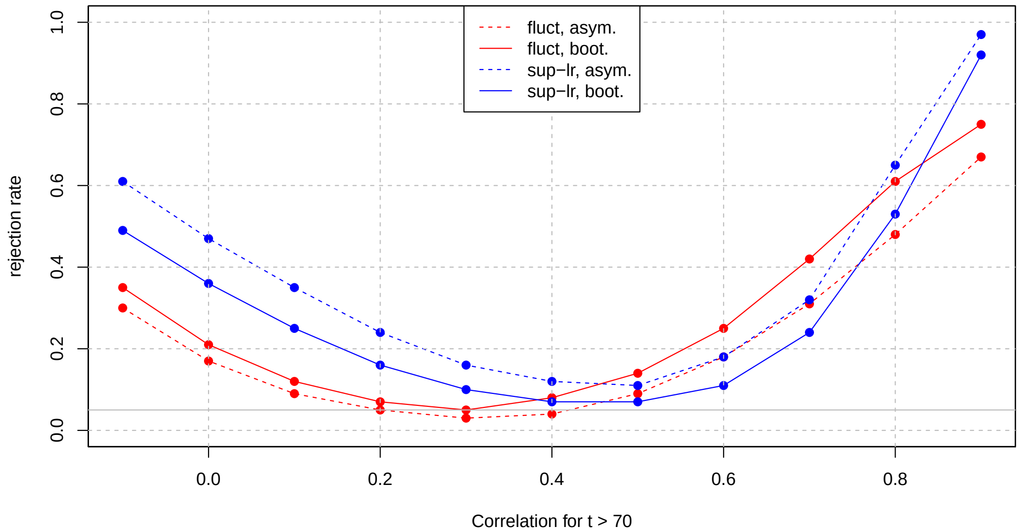


Figure 4.3: Gaussian Distribution, n=500, Scenario 2: Empirical Power

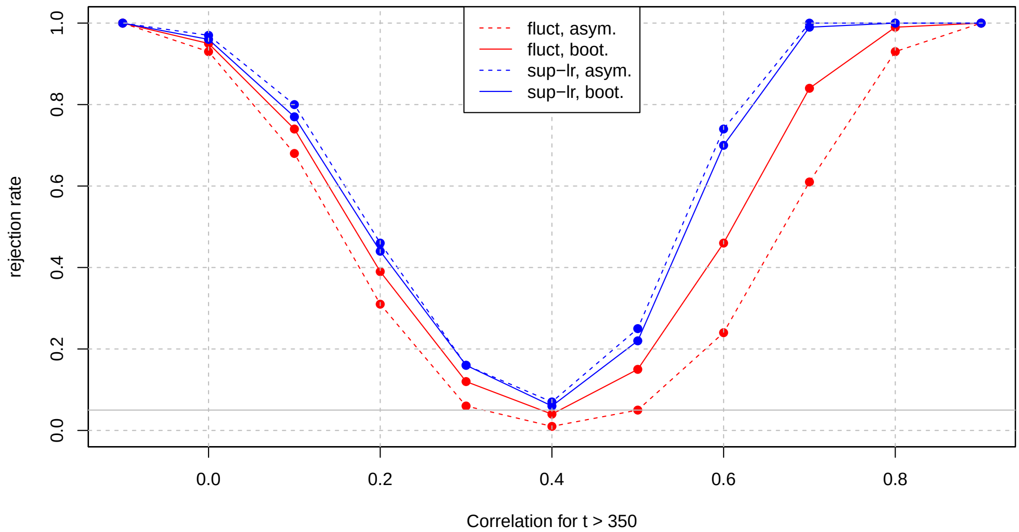
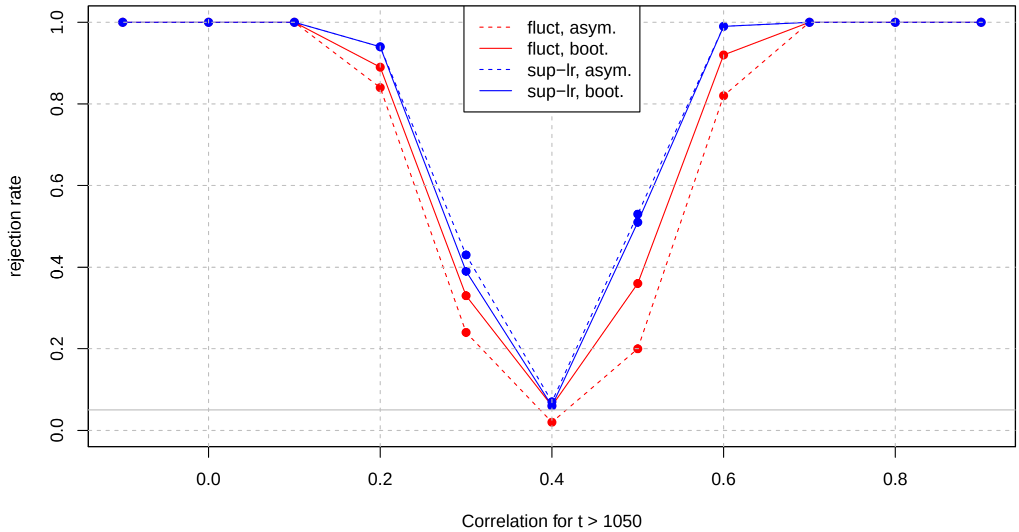


Figure 4.4: Gaussian Distribution, n=1500, Scenario 2: Empirical Power



If one compares the bootstrap-corrected versions indicated by solid lines, the results for testing constant correlation are inconclusive, at least in the bivariate case. Although the sup-LR test for constant marginal distributions outperforms the fluctuation test in small samples (see the upper panel of figure 4.1), this result does not carry over to the second step of the procedure. At larger samples, the picture is clearer: both procedures deliver similar results for testing at the marginal distributions and the parametric framework delivers significantly higher power when testing constant correlation. For example at $\rho_2 = 0.6$ the 95 %-confidence intervals are $[0.681, 0.728]$ for the sup-LR test and $[0.429, 0.491]$ for the fluctuation test at $n = 500$ and $[0.984, 0.996]$ for the sup-LR test and $[0.892, 0.928]$ for the fluctuation test at $n = 1500$.

Next we consider dimensionality effects for different sample sizes. As can be seen from the lower two panels of figure 4.1, the residual effect slowly declines with dimension m in the fluctuation test framework and even increases with m in the sup-LR-test framework, hence empirical power in figure 4.5 - 4.7 is compared only using the respective bootstrap schemes.

Figure 4.5: Multivariate Gaussian Distribution, n=100, Scenario 2: Empirical Power

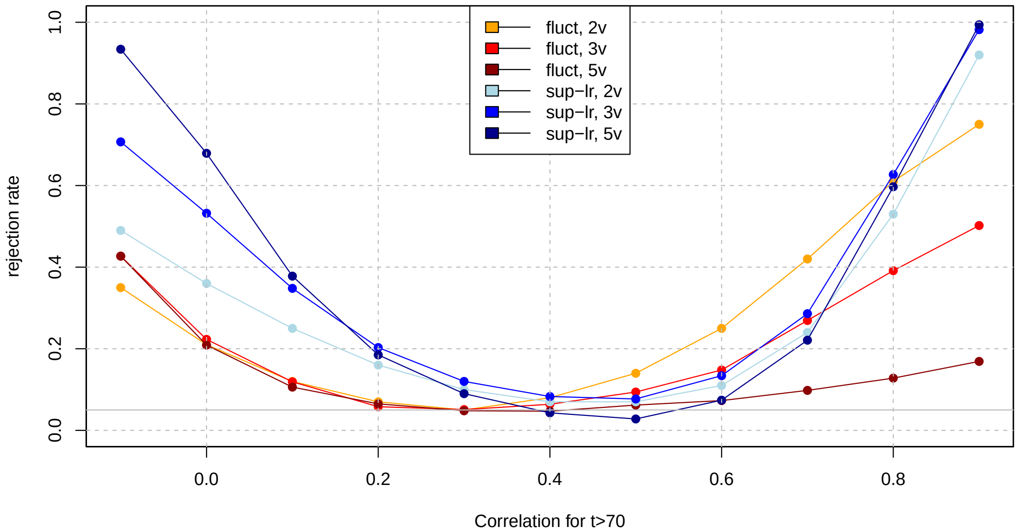


Figure 4.6: Multivariate Gaussian Distribution, n=500, Scenario 2: Empirical Power

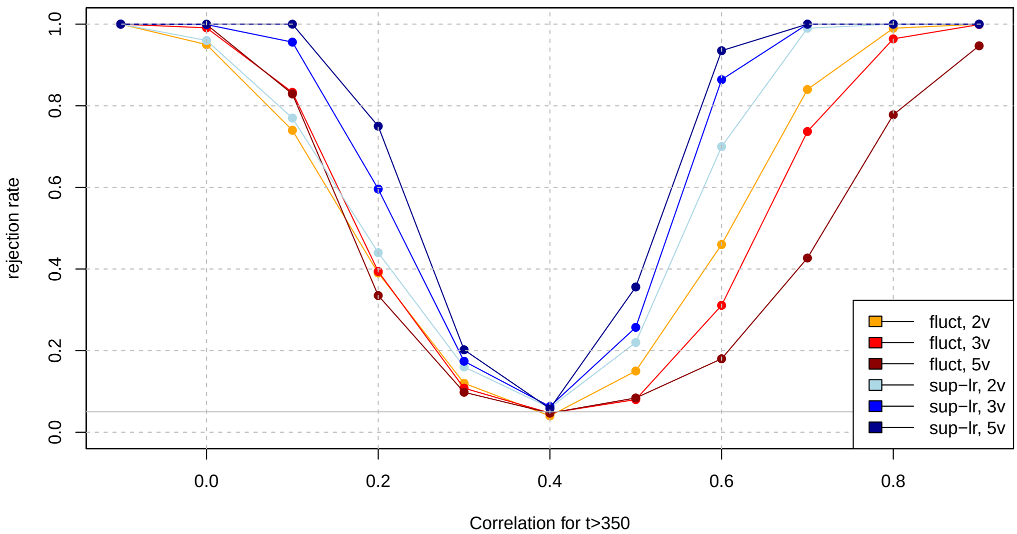
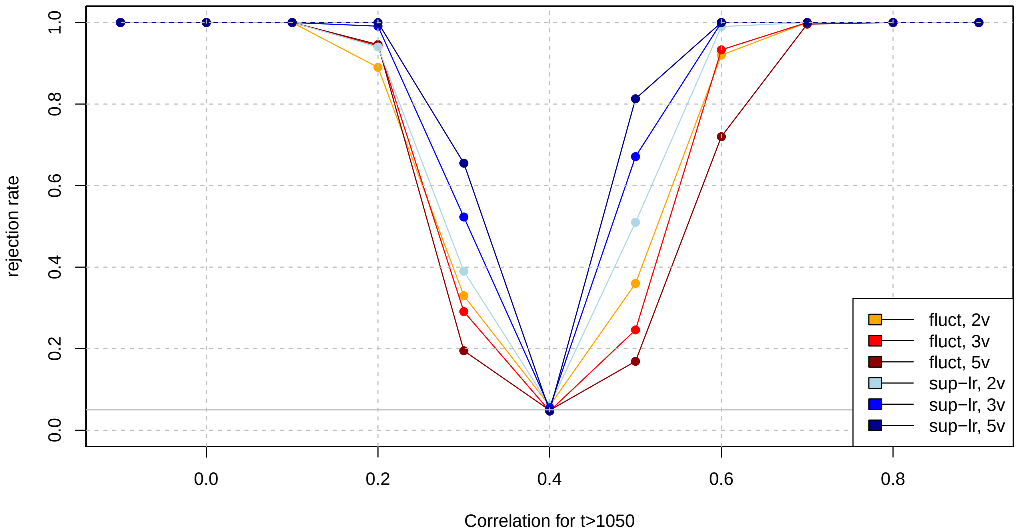


Figure 4.7: Multivariate Gaussian Distribution, n=1500, Scenario 2: Empirical Power



While both tests keep their size in case of no change in the dependency ($\rho_2 = 0.4$), the performance of the sup-LR test increases with the dimension m: in the 5-dimensional case, even moderate changes such as from $\rho_1 = 0.4$ to $\rho_2 = 0.6$ are detected almost every time, while the rejection rate is around 65 % in the bivariate case. The dimensionality effect is largely absent in the fluctuation test framework, where in the case of increasing correlation empirical power even declines with the dimension m.

Finally, in figure 4.8, we consider Monte Carlo bias and root mean-squared error of the break point estimator:

Figure 4.8: Gaussian Distribution, Scenario 2: Break Point Estimation

ρ_2	$n = 100$				$n = 500$				$n = 1500$			
	Fluctuation test		sup-LR test		Fluctuation test		sup-LR test		Fluctuation test		sup-LR test	
l_1	$bias(l_1)$	$rmse(l_1)$	$bias(l_1)$	$rmse(l_1)$	$bias(l_1)$	$rmse(l_1)$	$bias(l_1)$	$rmse(l_1)$	$bias(l_1)$	$rmse(l_1)$	$bias(l_1)$	$rmse(l_1)$
-0.1	6.85	14.11	2.36	16.27	11.17	22.28	4.81	25.50	11.54	25.68	3.70	21.64
0.1	6.78	13.96	2.27	16.41	11.14	21.97	4.23	25.58	11.42	25.54	3.83	21.07
0.3	6.75	14.16	2.35	16.57	10.92	21.50	4.17	25.35	11.51	25.88	3.82	21.08
0.5	6.87	14.18	2.48	16.72	11.00	21.70	4.22	25.44	11.34	25.09	3.91	20.90
0.7	6.80	14.14	2.37	16.78	10.92	21.47	4.00	26.08	11.46	25.30	4.07	20.61
0.9	6.76	14.25	2.59	16.90	11.00	21.69	4.39	26.99	11.67	25.56	4.04	20.42

	Fluctuation test		sup-LR test		Fluctuation test		sup-LR test		Fluctuation test		sup-LR test	
l_2	$bias(l_2)$	$rmse(l_2)$	$bias(l_2)$	$rmse(l_2)$	$bias(l_2)$	$rmse(l_2)$	$bias(l_2)$	$rmse(l_2)$	$bias(l_2)$	$rmse(l_2)$	$bias(l_2)$	$rmse(l_2)$
-0.1	2.27	12.20	-1.65	17.38	3.56	19.28	1.84	28.47	5.65	21.59	4.21	24.44
0.1	2.22	12.20	-1.74	17.55	3.50	18.94	1.94	28.10	5.91	21.66	4.17	24.67
0.3	2.19	12.57	-1.53	17.52	3.84	18.52	1.85	29.18	5.90	21.51	3.86	24.29
0.5	2.11	12.43	-1.56	17.46	4.16	18.71	2.32	29.78	6.05	22.53	4.10	24.48
0.7	2.19	12.49	-1.41	17.58	3.93	18.48	2.39	29.02	6.08	21.38	4.20	24.04
0.9	2.36	12.34	-1.73	17.73	4.12	18.69	1.95	28.98	5.88	22.14	4.41	25.24

	Fluctuation test		sup-LR test		Fluctuation test		sup-LR test		Fluctuation test		sup-LR test	
$m = 2$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$
-0.1	-3.86	14.26	-5.38	16.87	-13.46	31.19	-0.21	26.15	-16.76	37.13	2.04	18.83
0.1	-7.79	20.31	-11.11	23.07	-24.82	56.18	-11.31	60.42	-47.26	102.04	-2.70	66.38
0.3	-9.56	22.86	-15.53	27.36	-77.65	131.78	-73.80	135.84	-184.36	323.67	-122.75	301.75
0.5	-9.67	20.97	-16.40	29.00	-87.23	138.21	-64.60	122.59	-180.64	329.68	-88.24	252.58
0.7	-6.90	15.22	-8.81	22.72	-33.83	68.04	-4.98	31.31	-45.16	106.13	-2.81	22.59
0.9	-4.72	10.69	0.81	5.21	-16.24	37.83	-0.23	4.02	-16.21	43.73	0.09	3.45

					Fluctuation test		sup-LR test		Fluctuation test		sup-LR test	
$m = 3$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$
-0.1	-2.55	9.78	-3.57	12.62	-10.31	19.48	0.92	9.65	-13.47	25.87	1.54	7.93
0.1	-5.27	16.69	-9.94	21.59	-22.96	45.92	-3.34	37.48	-32.87	65.28	3.13	24.49
0.3	-8.58	19.48	-16.41	27.98	-82.95	125.35	-71.20	132.78	-167.12	287.54	-87.16	256.22
0.5	-9.42	17.62	-17.55	29.68	-96.40	134.51	-57.88	118.46	-205.71	320.80	-68.59	212.62
0.7	-8.28	14.64	-7.89	22.36	-37.18	70.05	-0.41	11.78	-33.52	84.18	-4.36	9.92
0.9	-6.60	11.68	1.27	3.03	-12.02	31.38	0.71	1.87	-11.35	37.17	0.55	1.91

					Fluctuation test		sup-LR test		Fluctuation test		sup-LR test	
$m = 5$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$
-0.1	-1.75	7.26	-0.83	6.45	-8.61	16.45	0.84	3.53	-9.89	18.76	1.01	3.31
0.1	-3.44	11.64	-6.38	17.47	-22.75	42.87	-1.27	18.98	-25.56	1.70	1.54	11.07
0.3	-7.27	15.40	-14.67	27.69	-76.89	110.93	-66.06	125.12	-191.27	290.56	-67.42	223.02
0.5	-10.88	16.68	-16.89	30.48	-103.63	128.80	-45.16	105.73	-265.41	344.97	-40.78	169.30
0.7	-11.89	16.60	-8.67	24.40	-64.10	90.31	0.52	7.07	-55.45	115.50	0.14	5.63
0.9	-12.65	16.93	1.46	2.70	-32.23	55.28	0.89	1.31	-12.60	36.70	0.88	1.16

Similar to the findings on empirical power, the fluctuation test outperforms the sup-LR test in estimating break-point locations at each margin for small sample sizes (upper two panels of figure 4.8), when it comes to the correlation change point, results however are switched: except for $n = 100$ and small ρ_2 both bias and variance are considerably smaller for the sup-LR test. The fluctuation test underestimates l_d even for a change from $\rho_1 = 0.4$ to $\rho_2 = 0.9$ and $n = 1500$ as compared to the sup-LR test, which has a negligible bias even for $\rho_2 = 0.7$ (second panel). Results in the higher-dimensional case (the bottom two panels) also favor the parametric framework: while using the sup-LR framework the regime shift is estimated very accurately in the 5-dimensional case compared to the (also precise) estimates in the two and three-dimensional set-up, break-point estimation is not considerably improved with m in the fluctuation test framework. This is especially true for situations of shift contagion, namely for increases in P. Since scenarios with potential shift contagion are usually associated with increasing correlation, the preceding findings suggest to use the sup-LR test, in particular when a precise estimation of the change-point is required.

Although there is some inconclusiveness for small samples we summarize from the simulation results laid out in this section, that in moderate to large samples the sup-LR test with the residual-bootstrap scheme has acceptable size properties under H_0 and outperforms the fluctuation test with a wild bootstrap scheme both in terms of detecting and estimating regime-shifts.

4.2 Bivariate t-Distribution

Similar to the bivariate Gaussian distribution examined in section 4.1, scenario 1 and 2 are adapted to a t-distribution with degrees of freedom fixed at $\nu = 5$ over the entire sample:

$$
\begin{aligned}
&(X_{1,t}, X_{2,t}) \overset{i.i.d.}{\sim} t_5(0.05, 0.05, 1, 1, 0.4) \quad \text{for } t = 1, ..., l_1 \\
&(X_{1,t}, X_{2,t}) \overset{i.i.d.}{\sim} t_5(0.06 - 0.01s, 0.05, s_1, 1, 0.4) \quad \text{for } t = l_1 + 1, ..., l_2 \\
&(X_{1,t}, X_{2,t}) \overset{i.i.d.}{\sim} t_5(0.06 - 0.01s; 0.06 - 0.01s, s, s, 0.4) \quad \text{for } t = l_2 + 1, ..., l_D \\
&(X_{1,t}, X_{2,t}) \overset{i.i.d.}{\sim} t_5(0.06 - 0.01s, 0.06 - 0.01s, s, s, \rho_2) \quad \text{for } t = l_D, ..., n
\end{aligned}
$$

For the timing of the regime-shift, the same values as in the Gaussian case are used.

Figure 4.9: t_5-Distribution, Scenario 1: Rejection Rates under H_0

s	Fluctuation test			sup-LR test, t, joint			sup-LR test, Gaussian		
$n = 100$	asym.	boot.	X_1	asym.	boot.	X_1	asym.	boot.	X_1
0.2	0.017	0.056	0.604	0.490	0.015	0.949	0.237	0.075	0.978
1/3	0.025	0.053	0.321	0.703	0.035	0.670	0.258	0.092	0.819
0.5	0.022	0.052	0.118	0.812	0.039	0.305	0.299	0.094	0.500
0.75	0.019	0.037	0.031	0.736	0.018	0.100	0.272	0.073	0.233
1	0.018	0.034	0.018	0.708	0.011	0.075	0.244	0.071	0.178
4/3	0.030	0.060	0.032	0.737	0.008	0.110	0.239	0.074	0.275
2	0.053	0.087	0.129	0.821	0.028	0.328	0.273	0.079	0.542
3	0.054	0.090	0.347	0.680	0.031	0.719	0.259	0.079	0.841
5	0.050	0.074	0.623	0.508	0.017	0.966	0.244	0.075	0.985

	Fluctuation test			sup-LR test, t, joint			sup-LR test, Gaussian		
$n = 500$	asym.	boot.	X_1	asym.	boot.	X_1	asym.	boot.	X_1
0.2	0.017	0.051	0.980	0.430	0.009	1	0.332	0.067	1
1/3	0.014	0.047	0.955	0.436	0.007	1	0.331	0.071	1
0.5	0.020	0.050	0.729	0.481	0.019	0.953	0.337	0.079	0.975
0.75	0.032	0.059	0.158	0.811	0.039	0.274	0.367	0.089	0.602
1	0.026	0.047	0.027	0.672	0.019	0.058	0.357	0.066	0.330
4/3	0.039	0.062	0.138	0.785	0.033	0.294	0.365	0.084	0.587
2	0.022	0.046	0.732	0.493	0.023	0.947	0.362	0.078	0.968
3	0.012	0.038	0.955	0.442	0.006	1	0.340	0.072	1
5	0.017	0.046	0.984	0.437	0.007	1	0.332	0.067	1

	Fluctuation test			sup-LR test, t, joint			sup-LR test, Gaussian		
$n = 1500$	asym.	boot.	X_1	asym.	boot.	X_1	asym.	boot.	X_1
0.2	0.018	0.059	0.999	0.404	0.010	1	0.399	0.062	1
1/3	0.017	0.052	0.997	0.402	0.014	1	0.402	0.064	1
0.5	0.020	0.052	0.975	0.442	0.054	0.982	0.398	0.069	1
0.75	0.038	0.065	0.420	0.700	0.646	0.706	0.442	0.075	0.867
1	0.040	0.052	0.035	0.638	0.022	0.056	0.435	0.069	0.406
4/3	0.037	0.062	0.425	0.722	0.066	0.654	0.435	0.072	0.847
2	0.019	0.058	0.981	0.436	0.056	0.974	0.404	0.085	1
3	0.021	0.060	0.994	0.402	0.008	1	0.393	0.062	1
5	0.022	0.052	0.998	0.406	0.010	1	0.394	0.063	1

The fluctuation test behaves similar to the Gaussian case when testing for constant cross-moments: the nominal level of 5 % is not reached under H_0 when asymptotic critical values are used. As before, the test shows good size properties under the wild bootstrap scheme. Attention has to be paid in the correctly specified sup-LR test. Although it possesses good power and size properties at the margins in step 1, using asymptotic critical values leads to rejection rates up to 80 % under H_0. Using the appropriate wild bootstrap scheme puts the empirical rejection rates into acceptable regions, but now constantly falling short of 5 % and decreasing towards zero if the margins vary strongly (see the bottom panel of figure 4.9). Similar to using the correct distributional assumption, testing in the Gaussian framework leads to severe size distortions of the sup-LR test using asymptotic critical values. The test keeps its size under H_0 if corrected by the wild bootstrap scheme and looks preferable in terms of size to the (computationally more intensive) sup-LR test under the correct distributional specification.

Figure 4.10: t_5-Distribution, n=100, Scenario 2: Empirical Power

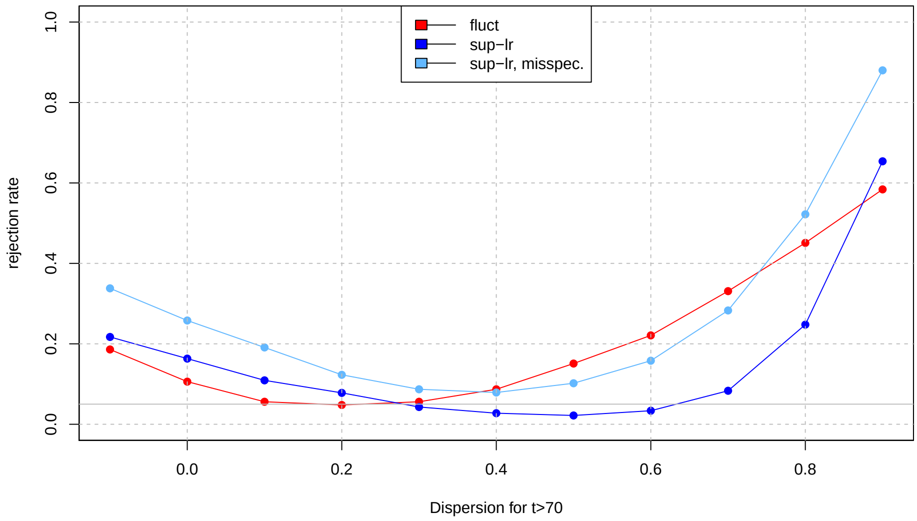


Figure 4.11: t_5-Distribution, n=500, Scenario 2: Empirical Power

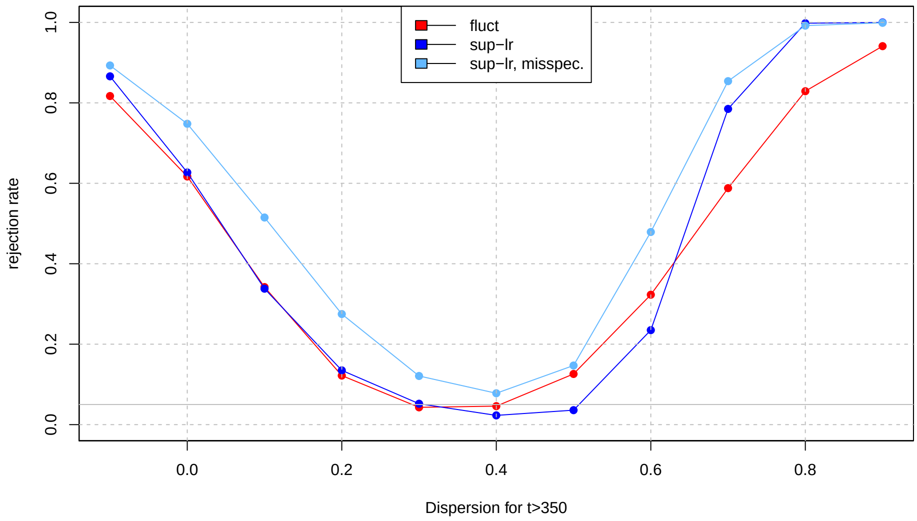
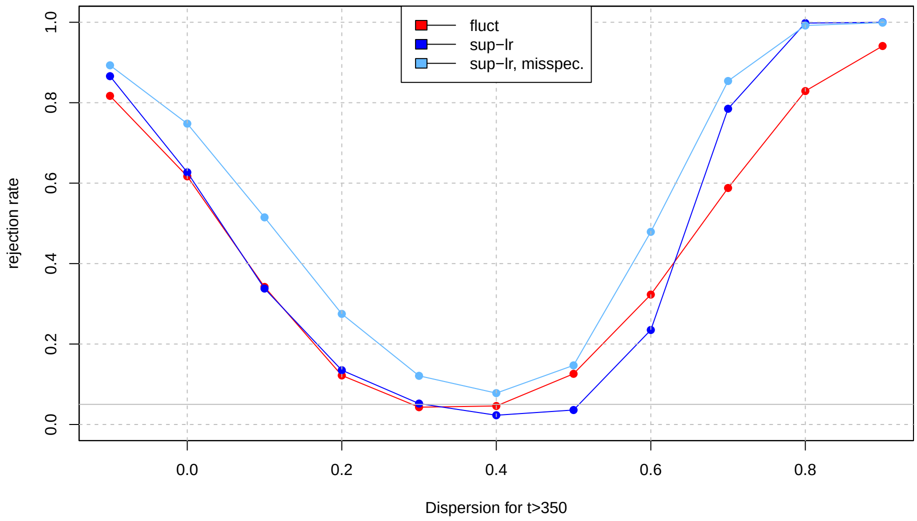


Figure 4.12: t_5-Distribution, n=1500, Scenario 2: Empirical Power



Findings on power draw a picture similar to the Gaussian case: there is some inconclusiveness in small samples, but figure 4.10 - 4.12 show the sup-LR test gaining power faster than the fluctuation test, even though it is found to be conservative. We defer the discussion of dimensionality effects in this parametric class to the next section on copulae and directly move on to the accuracy of break-point estimators. Figure 4.13 again shows superiority of the sup-LR tests over non-parametric methods in terms of bias and root mean-squared error of the break-point estimator for l_D and l_1 in larger samples. We further conclude that the more elaborate methods relying on the t-distribution should only be used, if the sample size is sufficiently large.

Figure 4.13: t_5-Distribution, Scenario 2: Break Point Estimation

ρ_2	Fluctuation test				sup-LR test				sup-LR test, Gaussian Distribution			
$n = 100$	$bias(l_D)$	$rmse(l_D)$	$bias(l_1)$	$rmse(l_1)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_1)$	$rmse(l_1)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_1)$	$rmse(l_1)$
-0.1	-4.10	16.34	7.13	16.73	-8.90	20.79	1.76	18.70	-8.77	21.74	1.59	18.88
0.1	-8.11	21.75	7.21	16.81	-13.32	24.69	1.82	18.87	-12.76	25.99	1.87	18.85
0.3	-9.53	22.65	7.28	16.99	-16.53	27.37	2.21	18.77	-15.62	28.37	1.64	18.74
0.5	-8.85	20.78	7.39	16.93	-15.70	27.03	1.95	18.60	-16.82	29.10	1.47	18.85
0.7	-7.58	17.72	7.40	16.68	-11.79	23.62	1.79	18.57	-9.18	22.68	1.31	18.85
0.9	-5.82	14.17	7.57	16.40	-2.81	10.36	1.67	18.72	-0.55	9.08	1.52	18.76

	Fluctuation test				sup-LR test, t-Distribution				sup-LR test, Gaussian Distribution			
$n = 500$	$bias(l_D)$	$rmse(l_D)$	$bias(l_1)$	$rmse(l_1)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_1)$	$rmse(l_1)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_1)$	$rmse(l_1)$
-0.1	-16.31	45.44	18.69	46.66	-1.93	41.62	4.74	41.11	-6.29	47.90	3.08	57.57
0.1	-36.85	87.20	17.60	45.19	-29.91	95.99	5.07	41.23	-32.02	92.43	3.45	57.78
0.3	-77.52	139.12	18.11	46.06	-85.34	154.96	5.01	41.40	-87.22	146.01	3.18	57.37
0.5	-78.12	130.81	19.05	46.58	-78.84	148.35	5.04	42.70	-73.01	132.82	2.89	57.44
0.7	-40.53	78.75	18.53	45.65	-3.93	42.24	5.25	42.78	-12.89	56.28	3.24	57.18
0.9	-23.50	49.67	18.56	46.33	0.18	4.72	5.06	42.47	0.77	8.88	3.69	56.37

	Fluctuation test				sup-LR test, t-Distribution				sup-LR test, Gaussian Distribution			
$n = 1500$ $bias(l_D)$	$rmse(l_D)$	$bias(l_1)$	$rmse(l_1)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_1)$	$rmse(l_1)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_1)$	$rmse(l_1)$	
-0.1	-32.67	77.77	27.20	67.95	-2.57	43.87	0.65	60.74	-1.62	54.17	9.01	67.65
0.1	-62.32	150.28	26.05	63.23	-13.95	110.93	2.04	50.37	-24.37	145.76	8.12	69.32
0.3	-222.37	382.55	26.11	62.42	-200.66	419.88	2.35	50.72	-193.88	382.50	5.80	68.18
0.5	-254.18	406.66	25.84	63.19	-187.67	403.02	1.93	50.54	-177.61	349.06	4.69	66.99
0.7	-76.73	170.33	27.72	63.86	-2.13	39.81	2.99	49.73	-7.04	59.83	4.16	66.12
0.9	-34.53	89.23	27.94	68.02	0.44	4.80	3.50	50.35	-0.12	6.22	4.31	65.85

Since sample sizes of 500 are hardly reached for monthly or quarterly data, we suggest using the misspecified Gaussian sup-LR test in a shift contagion scenario and employ the respective bootstrap method. As in the bivariate Gaussian case it could prove useful to additionally apply the fluctuation test, if a reduction in correlation is suspected. There also may be situations with more than two dimensions and a sample size too small to obtain reliable parameter estimates under the t-distribution specification, for example $m = 3$ and $n = 200$. Such cases are not formally considered here and are left for future research, based on the findings on the multivariate Gaussian one can suspect that - using an appropriate bootstrap scheme - the sup-LR test is preferable here. Should the sample be large enough to permit reliable estimation, the preceding findings favour the sup-LR test using the parametric approach.

4.3 t-Copula With t-Marginal Distributions

The third simulation study compares non-parametric and parametric tests for a constant copula. Specifically data are generated from a t_4-copula and subsequently transformed using the quantile function of a t_8-distribution $F_t^{-1}(\mu, \xi, \nu = 8)$

$$
\begin{aligned}
&U_t \overset{i.i.d.}{\sim} C_t(R_1, 4) && \text{for } t = 1, ..., l_D \\
&U_t \overset{i.i.d.}{\sim} C_t(R_2, 4) && \text{for } t = l_D, ..., n \\
&X_{1,t} = F_t^{-1}(U_{1,t}, 0.05, 1, 8) && \text{for } t = 1, ..., l_1 \\
&X_{1,t} = F_t^{-1}(U_{1,t}, 0.06 - 0.01s, s, 8) && \text{for } t = l_1, ..., n \\
&X_{2,t} = F_t^{-1}(U_{2,t}, 0.05, 1, 8) && \text{for } t = 1, ..., l_2 \\
&X_{2,t} = F_t^{-1}(U_{2,t}, 0.06 - 0.01s, s, 8) && \text{for } t = l_2, ..., n
\end{aligned}
$$

with

$$
P_1 = \begin{pmatrix} 1 & 0.4 \\ 0.4 & 1 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 1 & \rho_2 \\ \rho_2 & 1 \end{pmatrix}
$$

Generalizations to higher-dimensional cases are obtained by extending the correlation matrix and subsequently transform data with the quantile function accordingly. We set $l_2 = l_3 = ... = l_m = 300$ and 1050, respectively. Under this DGP we compare the non-parametric benchmark-test based on the empirical copula from Bücher et al. [2014], lined out in the appendix, with the sup-LR test under correct specification (section 3.4) and under misspecification as Gaussian copula (section 3.2).

Figure 4.14: t-Copula, Scenario 1: Rejection Rates under H_0

s	$n = 500$				$n = 1500$			
	Fluctuation		sup-LR		Fluctuation		sup-LR	
margins	X_1	X_2	X_1	X_2	X_1	X_2	X_1	X_2
0.2	0.999	0.999	1	1	1	1	1	1
1/3	0.999	0.995	0.999	0.999	1	1	1	1
0.5	0.917	0.859	0.920	0.896	1	1	1	1
0.75	0.242	0.192	0.208	0.179	0.670	0.598	0.517	0.485
1	0.043	0.039	0.048	0.031	0.049	0.038	0.017	0.010
4/3	0.240	0.233	0.209	0.219	0.650	0.641	0.513	0.484
2	0.906	0.920	0.918	0.910	1	1	1	1
3	0.999	0.999	1	0.999	1	1	1	1
5	1	1	1	1	1	1	1	1

$m = 2$	EC-test	t-Cop	Gauss	Gauss	EC-test	t-Cop	Gauss	Gauss
l_D				boot				boot
0.2	0.053	0.048	0.143	0.065	0.059	0.052	0.171	0.050
1/3	0.051	0.056	0.151	0.066	0.063	0.055	0.177	0.051
0.5	0.048	0.073	0.192	0.079	0.062	0.057	0.187	0.054
0.75	0.051	0.084	0.204	0.080	0.062	0.100	0.219	0.092
1	0.052	0.048	0.115	0.057	0.061	0.047	0.181	0.049
4/3	0.049	0.076	0.198	0.083	0.058	0.101	0.241	0.077
2	0.053	0.070	0.187	0.069	0.052	0.062	0.182	0.054
3	0.053	0.051	0.156	0.060	0.062	0.052	0.173	0.052
5	0.052	0.049	0.142	0.064	0.063	0.050	0.162	0.054

$m = 3$	EC-test	t-Cop	Gauss	Gauss	EC-test	t-Cop	Gauss	Gauss
l_D				boot				boot
0.2	0.054	0.043	0.234	0.055	0.060	0.036	0.213	0.052
1/3	0.054	0.044	0.251	0.048	0.055	0.043	0.250	0.048
0.5	0.048	0.062	0.295	0.057	0.056	0.046	0.270	0.048
0.75	0.052	0.073	0.301	0.068	0.055	0.096	0.320	0.094
1	0.051	0.042	0.265	0.045	0.053	0.029	0.265	0.043
4/3	0.052	0.059	0.302	0.068	0.056	0.102	0.352	0.090
2	0.058	0.060	0.286	0.065	0.060	0.046	0.265	0.048
3	0.055	0.043	0.242	0.053	0.051	0.044	0.249	0.049
5	0.054	0.038	0.220	0.051	0.060	0.042	0.230	0.056

$m = 5$	EC-test	t-Cop	Gauss	Gauss	EC-test	t-Cop	Gauss	Gauss
l_D				boot				boot
0.2	0.051	0.026	0.378	0.044	0.060	0.026	0.379	0.046
1/3	0.048	0.026	0.414	0.043	0.058	0.028	0.409	0.048
0.5	0.053	0.045	0.465	0.060	0.058	0.032	0.422	0.042
0.75	0.050	0.040	0.482	0.070	0.054	0.070	0.516	0.110
1	0.048	0.025	0.397	0.033	0.056	0.022	0.414	0.034
4/3	0.050	0.035	0.476	0.062	0.060	0.076	0.524	0.086
2	0.046	0.040	0.455	0.061	0.060	0.026	0.428	0.042
3	0.048	0.029	0.407	0.042	0.054	0.028	0.403	0.046
5	0.051	0.026	0.372	0.044	0.054	0.028	0.369	0.048

The empirical copula test keeps its size when testing for a constant copula, indicating that piecewise standardization appropriately accounts changes in the margins and the i.i.d. multiplier process can be applied. Except for the case of moderate changes at the margins, where Hypothesis Pair 2 is rejected too frequently, this also holds for the correctly specified sup-LR test. The result is in accordance with the Gaussian and t_5-case. Unsurprisingly, the misspecified sup-LR test does not keep its size given the nominal level of 5 %, this is however appropriately corrected by the residual bootstrap. Using the scheme lined out in section 3, empirical rejection rates are similar to the sup-LR test under correct specification.

Figure 4.15: t-Copula, n=500, Scenario 2: Empirical Power

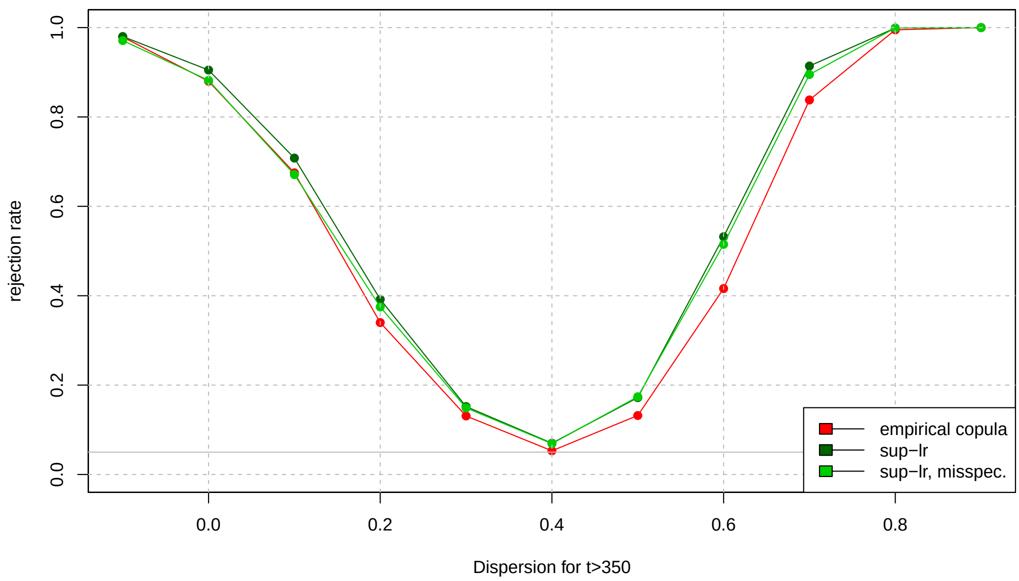


Figure 4.16: t-Copula, n=1500, Scenario 2: Empirical Power

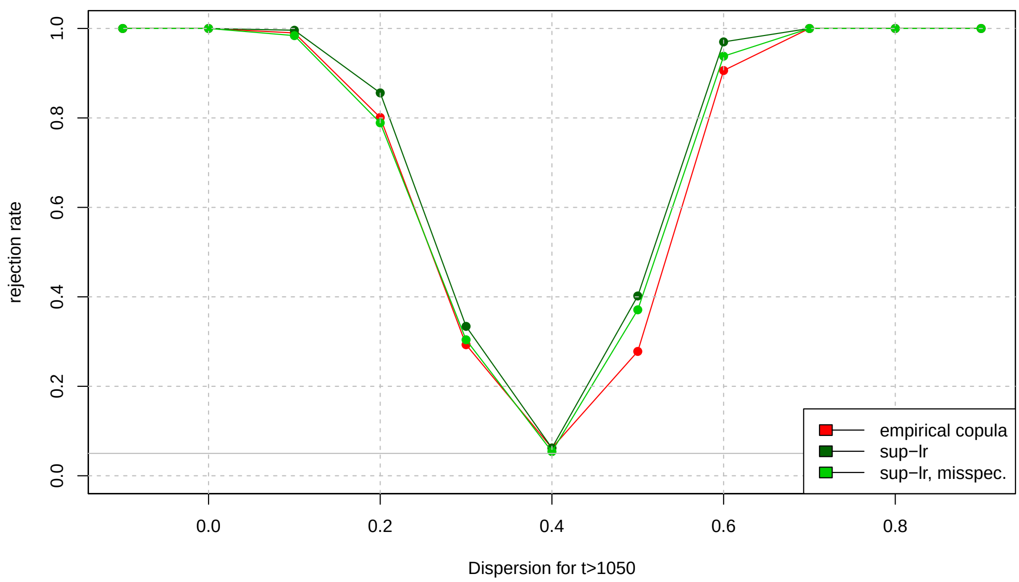


Figure 4.17: 3-variate t-Copula, n=500, Scenario 2: Empirical Power

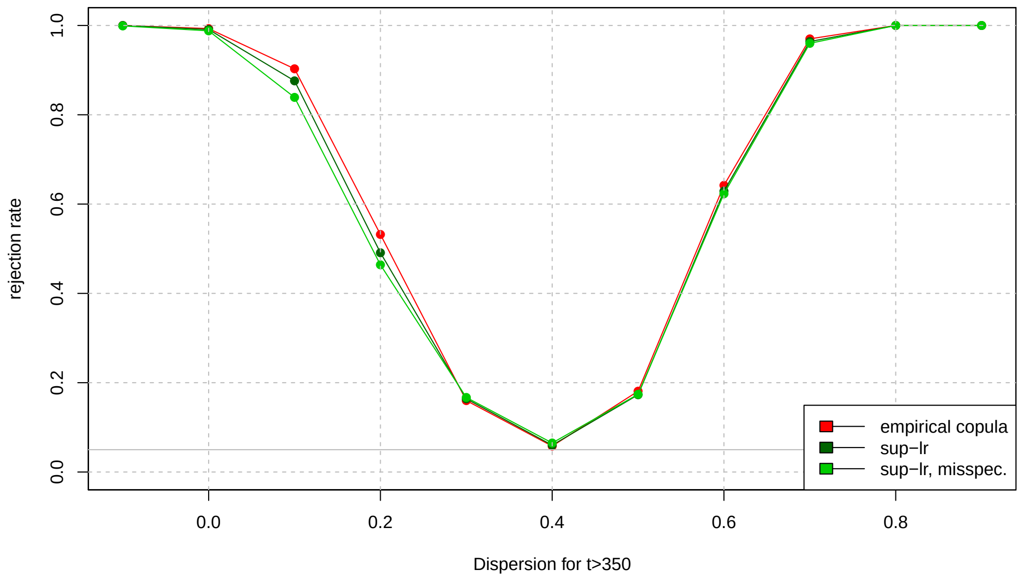


Figure 4.18: 3-variate t-Copula, n=1500, Scenario 2: Empirical Power

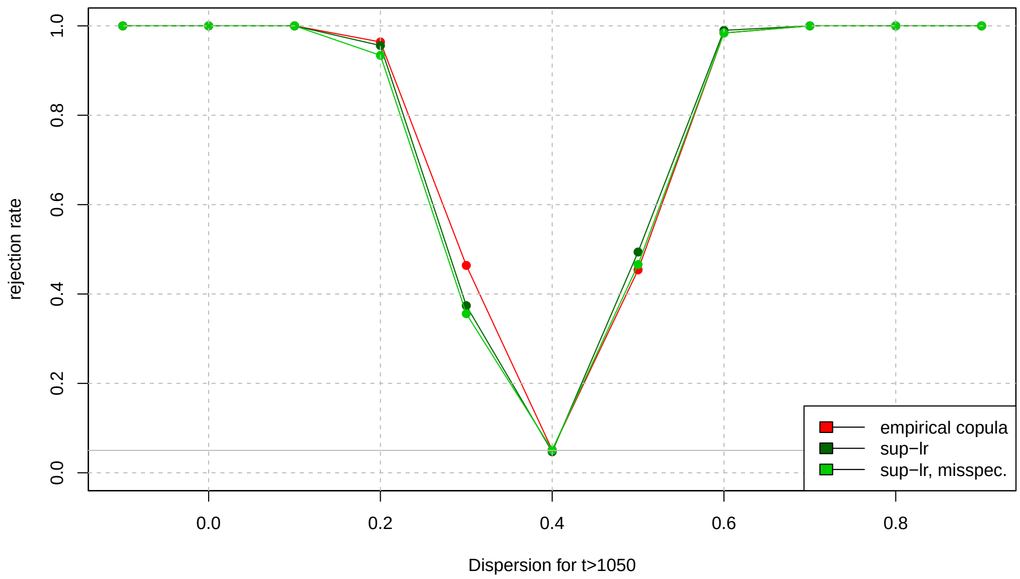


Figure 4.19: 5-variate t-Copula, n=500, Scenario 2: Empirical Power

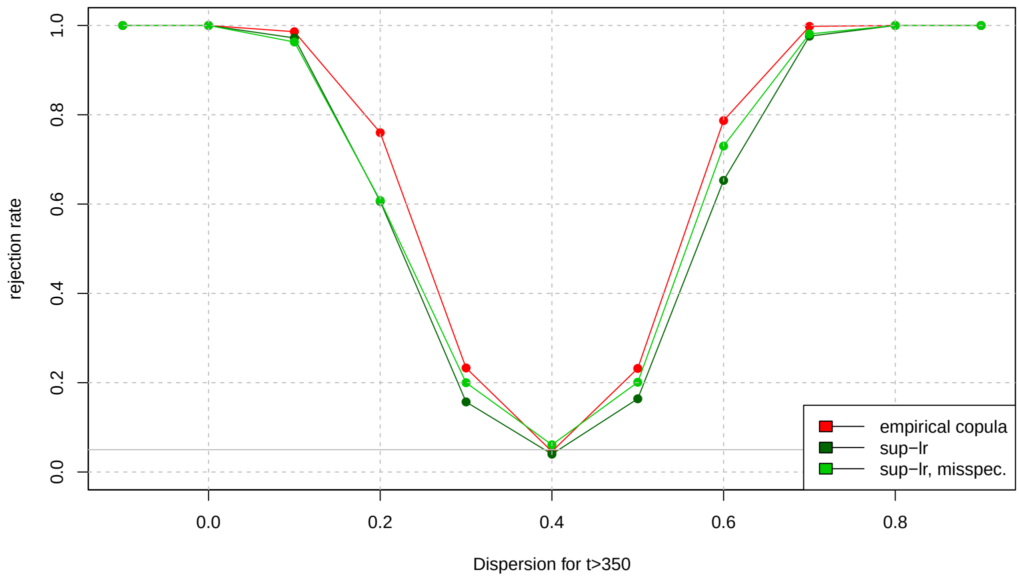
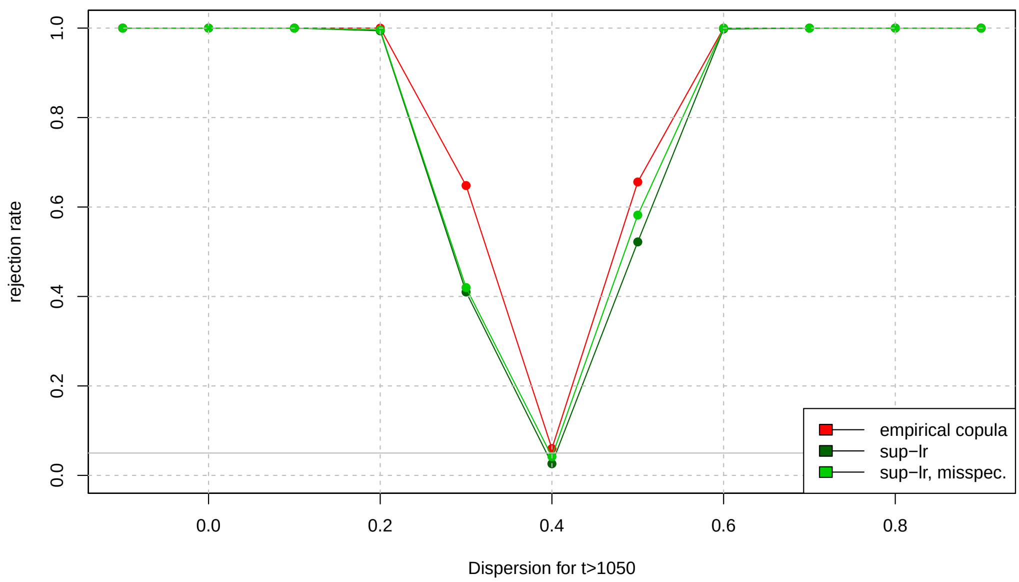


Figure 4.20: 5-variate t-Copula, n=1500, Scenario 2: Empirical Power



Results on empirical power in figure 4.15 - 4.20 suggest using the correctly specified sup-LR test in larger samples, where the t-copula can be reliably estimated. What counts as a large sample depends on the number of parameters to be estimated and thus on dimensionality: in the bivariate case, $n = 1500$ is already sufficient for the parametric framework under correct specification to outperform the non-parametric and misspecified test frameworks. Some care has to be taken for the case $n = 500$, here the parametric test frameworks reject H_0 slightly too often (see second panel of figure 4.14).
In the three-dimensional case, where results tend to favor the empirical copula test for

$n = 500$ and results start to shift for $n = 1500$. Moving to the five-dimensional case, neither the correctly nor misspecified tests reach the power of the non-parametric framework and in addition, testing under the Gaussian-copula assumptions yields higher power than testing under the correct specification. Since the sup-LR test requires estimates of 10 correlation parameters, ν_D in case of correct specification and (ν_i, μ_i, ξ_i) for every margin, a total 25 or 26 parameters, respectively, have to be estimated in the five-dimensional case and it comes not at much surprise, that the test lacks power for $n = 1500$.
However, the sup-LR test yields considerably better results in estimating the change-point, irrespective of whether the model is correctly specified or not, as can be seen in figure 4.21. Even for $n = 1500$ and $\rho_2 = 0.9$, l_D is visibly biased in the fluctuation test framework, while l_D is already precisely estimated for $\rho_2 = 0.7$. We also observe similar or even better results of the misspecified sup-LR test compared to its correctly specified counterpart for samples, that might be too small for efficiently estimating a higher-dimensional t-copula. Therefore we suggest using the sup-LR test whenever precise estimation of the change-point is required. Within the sup-LR framework usage of any advanced model is only advantageous if one has high confidence on the model's appropriateness and the sample size (relative to dimension) permits reliable estimation. If the empirical researcher is merely interested in detecting regime-shifts and the sample size is insufficiently small, better results can be achieved in higher dimensions with the empirical copula test.

Figure 4.21: Multivariate t-Copula, n=500, Scenario 2: Copula-Break Point Estimation

	$n = 500$						$n = 1500$					
ρ_2	EC-test		sup-LR		sup-LR, mis.		EC-test		sup-LR		sup-LR, mis.	
$m = 2$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$
-0.1	-16.64	36.37	-4.52	37.44	-5.32	40.42	-19.88	44.44	0.18	30.01	0.12	35.28
0.1	-33.00	65.76	-18.86	70.46	-21.62	77.34	-42.39	95.71	-5.51	82.97	-7.09	89.21
0.3	-79.44	115.54	-72.52	132.35	-76.53	133.93	-200.34	310.01	-148.00	332.92	-163.65	350.30
0.5	-85.97	119.52	-73.75	133.72	-75.54	135.24	-195.07	298.51	-125.00	302.52	-141.19	323.58
0.7	-34.18	60.76	-9.07	47.99	-8.48	48.72	-42.96	81.92	-2.20	31.67	-1.34	38.23
0.9	-13.04	23.80	0.01	5.68	1.10	5.54	-14.12	26.39	0.16	4.41	1.50	6.28
ρ_2	EC-test		sup-LR		sup-LR, mis.		EC-test		sup-LR		sup-LR, mis.	
$m = 3$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$
-0.1	-11.26	22.25	-0.41	13.22	0.69	14.53	-12.69	26.37	1.28	9.69	2.21	12.44
0.1	-24.48	48.26	-8.53	47.23	-12.16	55.71	-21.83	50.78	1.57	44.64	0.63	44.13
0.3	-78.39	113.25	-76.20	131.47	-75.50	133.44	-145.39	250.86	-122.72	304.22	-136.03	319.16
0.5	-78.98	111.31	-73.45	130.87	-71.60	128.93	-145.28	248.97	-100.69	270.65	-103.99	278.64
0.7	-22.88	43.32	-1.41	28.15	-0.07	30.89	-24.19	54.54	0.09	17.34	2.12	22.80
0.9	-8.21	16.33	0.74	3.37	2.17	5.16	-7.62	15.11	0.74	2.60	1.93	4.38
ρ_2	EC-test		sup-LR		sup-LR, mis.		EC-test		sup-LR		sup-LR, mis.	
$m = 5$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$	$bias(l_D)$	$rmse(l_D)$
-0.1	-9.94	18.54	0.25	4.83	1.13	5.43	-10.78	20.05	0.47	4.11	1.42	4.98
0.1	-21.23	38.37	-4.89	27.32	-5.76	31.59	-20.24	40.48	0.62	13.32	1.49	17.62
0.3	-70.23	102.33	-60.10	117.79	-64.71	122.83	-111.41	199.28	-109.42	280.00	-119.52	294.02
0.5	-61.31	95.48	-65.59	124.34	-59.41	120.61	-97.78	187.68	-78.76	237.63	-50.87	212.54
0.7	-13.36	28.75	-0.40	19.43	1.70	19.37	-16.91	37.61	0.56	10.58	2.03	15.25
0.9	-4.76	9.77	0.75	2.45	2.23	3.55	-5.01	10.45	0.96	2.07	2.34	3.92

5 Application to Commodity and Equity Index Data

A first empirical application uses the methods subject to the simulation studies section 4.1 and section 4.2: Daily log-returns of real estate and equity indices are sequentially tested for constancy of correlation once under the Gaussian (section 3.1) and once under the assumption of a bivariate t-distribution (section 3.3). Both tests are benchmarked against the non-parametric fluctuation test outlined in Appendix A. We test 3 for Crude Oil spot market returns [1] and the European equity sector, which we proxy by the EUROSTOXX50[2] over the time-period 1991-04-17 to 2003-03-05. Foreign involvment in petrol-exporting countries has been fairly low following the early 1990s until 2003. Additionally events in the late 1980s and later technological changes in oil production, the financial crises and relaxed monetary policy probably did not influence the fundamental market environment over the sample. However markets experienced a period of increased volatility around 2000, associated with events such as the burst of the dotcom-bubble among others. This can be observed in figure 5.3 and 5.4. Rolling correlations in figure 5.2 however indicate a rather stable correlation pattern over the test period and thus making the sample a plausbile candidate to test for Hypothesis Pair 3. Reported numbers are annualized (business-)daily log-returns and their volatilities (annualized standard-deviations) in percent.

Figure 5.1: Estimation of European Crude Oil and Equity Data

	Fluctuation Test		sup-LR test, Gauss		sup-LR test, t	
	Crude Oil	Equity	Crude Oil	Equity	Crude Oil	Equity
$\hat{l}_i$	1998-01-26	1997-07-16	1996-03-15	1997-06-27	1996-03-18	1997-06-30
test statistic	4.78	7.76	378.32	906.98	222.12	526.74
$\hat{\mu}_1$	-4.32	15.00	-0.57	14.03	7.77	18.13
$\hat{\mu}_2$	16.27	-3.94	8.12	-2.75	14.87	5.39
$\hat{\sigma}_1$	26.22	12.36	23.59	12.28	23.66	11.38
$\hat{\sigma}_2$	42.81	27.56	40.28	27.53	40.31	26.41
test statistic	0.771		4.9		7.82	
p-value (boot)	0.595		0.495		0.745	
$\hat{l}_D$	1995-07-24		1995-07-24		2001-05-25	
ρ_0	0.0141		0.008		0.032	
ρ_1	-0.0433		0.011		0.027	
ρ_2	0.0438		-0.002		0.022	
$\hat{\nu}$					4.99	

All procedures strongly reject the hypothesis of constant margins, the critical values at 99 % for the fluctuation test being 1.84 and for the sup-LR tests 15.51. Break-point estimates lie together very closely for both specifications of the sup-LR tests; based on results in figure 4.13 we favor the estimates based on the sup-LR test with t-distribtional assumption. When it comes to testing constant correlation, our empirical findings from section 4.2 directly carry over to this particular example: Following figure 4.9, it is crucial to apply a suitable bootstrap here. Using the bootstrapped p-values around or larger than 0.5, neither fluctuation test and not the sup-LR tests reject Hypothesis Pair 3. Had one used the incorrect asymptotic value for the sup-LR test, which is 7.17 at 90 % confidence level, one might incorrectly reject H_0 using the empirically plausible t-distributional assumption.
It has been previously established that incorrectly assuming constant variances when testing for constant correlation - implicitly by considering covariances as Aue et al. [2009] or explicitly by directly using the procedure of Wied et al. [2012b] - leads to flawed inference. But, as the preceding application points out, even if changes at the marginal distributions are

[1]Europe Brent, Data is taken from the U.S. Energy Information Administration: https://www.eia.gov/dnav/pet/hist/

[2]ISIN: EU0009658145, returns are calculated from the closing price of the last trading day each month.

taken into account correctly, applying invalid standard asymptotics may lead to incorrectly rejecting constant cross-sectional dependence.

Figure 5.2: Rolling Correlations, Equity and Crude Oil

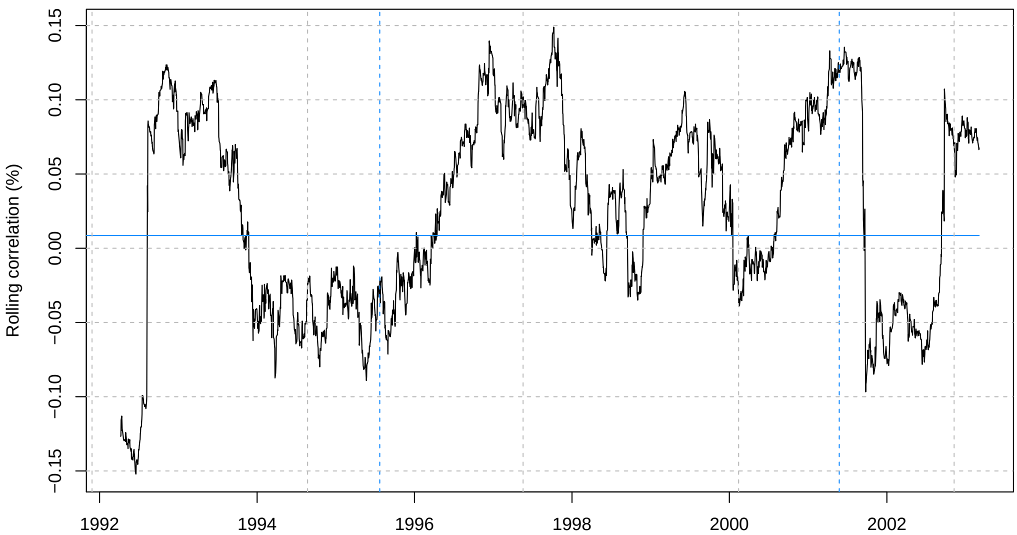


Figure 5.3: Crude Oil, Rolling Volatility

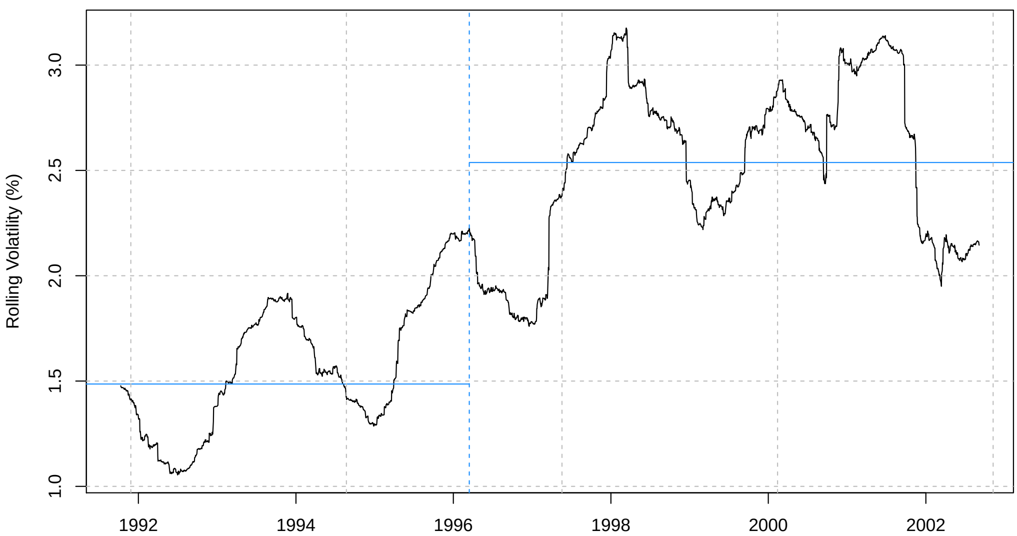
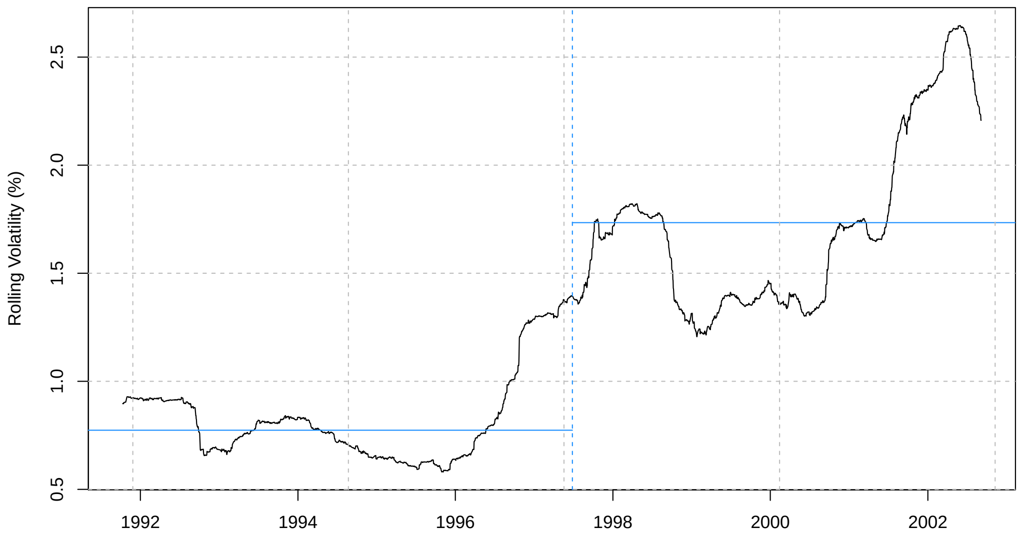


Figure 5.4: Equity, Rolling Volatility



6 Concluding Remarks

We have proposed and analyzed parametric two-step procedures for assessing the stability of cross-sectional dependency measures in the presence of potential breaks in the marginal distributions. We have focused on sup-LR tests and it could be interesting or further research to also look at sup-Wald, sup-LM or exponentially weighted test statistics in the spirit of Andrews and Ploberger [1994]. Moreover, while we have tackled the case of serial dependence when discussing volatility filtering, it might be interesting to also investigate e.g. changes in VAR-filtering in the first step of the procedure. Also changes in GARCH or AR parameters could be investigated.

References

D. Andrews and W. Ploberger. Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica*, 62:1383–1414, 1994.

D. Andrews and W. Ploberger. Admissibility of the likelihood ratio test when a nuisance parameter is present only under the alternative. *Annals of Statistics*, 23:1609–1629, 1995.

D. W. K. Andrews. Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61(4):821–856, 1993.

A. Aue, S. Hörmann, L. Horvath, and M. Reimherr. Break detection in the covariance structure of multivariate time series models. *Annals of Statistics*, 37(6B):4046–4087, 2009.

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

D. Blatt, B. Candelon, and H. Manner. Detecting contagion in a multivariate time series system: An application to sovereign bond markets in europe. *Journal of Banking and Finance*, 59(C):1–13, 2015.

N. Chan, J. Chen, X. Chen, Y. Fan, and L. Peng. Statistical inference for multivariate resiudal copula of garch models. *Statistica Sinica*, 19:53–70, 2009.

R. Cont. Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1:223–236, 2001.

S. Demarta and A. McNeil. The t copula and related copulas. *International Statistical Review*, 73(1):111–129, 2005.

M. Demetrescu and D. Wied. Residual-based inference on moment hypotheses, with an application to testing for constant correlation. *Econometrics Journal, forthcoming, doi: 10.1111/ectj.12116*, 2018+.

A. Dias and P. Embrechts. Change-point analysis for dependence structures in finance and insurance. In C. Szegoe, editor, *Risk Measures for the 21st Century*. Wiley Finance Series, Chichester, 2004.

F. Duan and D. Wied. A residual-based multivariate constant correlation test. *Metrika, forthcoming, doi: 10.1007/s00184-018-0675-y*, 2018.

K. Forbes and R. Rigobon. No contagion, only interdependence: Measuring stock market comovements. *Journal of Finance*, 57(5):2223–2261, 2002.

P. Galeano and D. Peña. Covariance changes detection in multivariate time series. *Journal of Statistical Planning and Inference*, 137:194–211, 2007.

J. Kiefer. K-sample analogues of the kolmogorov-smirnov and cramér-v.mises tests. *Annals of Mathematical Statistics*, 30(2):420–447, 1959.

M. King and S. Wadhwani. Transmission of volatility between stock markets. *Review of Financial Studies*, 3(1):281–307, 1990.

C. Liu and D. Rubin. Ml-estimation of the t-distribution using em and its extensions, ecm and ecme. *Statistica Sinica*, 5:19–39, 1995.

R. Mashal and A. Zeevi. Beyond correlation: extreme co-movements between financial assets. *Unpublished*, 2002.

Z. Qu and P. Perron. Estimating and testing multipile structural changes in multivariate regressions. *Econometrica*, 75(2):459–502, 2007.

B. Rémillard. Goodness-of-fit tests for copulas of multivariate time series. *Technical Report, HEC Montréal*, 2010.

D. Wied, M. Arnold, N. Bissantz, and D. Ziggel. A new fluctuation test for constant variances with applications to finance. *Metrika*, 75(8):1111–1127, 2012a.

D. Wied, W. Krämer, and H. Dehling. Testing for a change in correlation at unknown point in time using an extended functional delta method. *Econometric Theory*, 28(3):570–589, 2012b.