A Residual-Based Multivariate Constant Correlation Test

Fang Duan *† Dominik Wied ‡

May 8, 2018

Abstract

We propose a new multivariate constant correlation test based on residuals. This test takes into account the whole correlation matrix instead of the considering merely marginal correlations between bivariate data series. In financial markets, it is unrealistic to assume that the marginal variances are constant. This motivates us to develop a constant correlation test which allows for non-constant marginal variances in multivariate time series. However, when the assumption of constant marginal variances is relaxed, it can be shown that the residual effect leads to nonstandard limit distributions of the test statistics based on residual terms. The critical values of the test statistics are not directly available and we use a bootstrap approximation to obtain the corresponding critical values for the test. We also derive the limit distribution of the test statistics based on residuals under the null hypothesis. Monte Carlo simulations show that the test has appealing size and power properties in finite samples. We also apply our test to the stock returns in Euro Stoxx 50 and integrate the test into a binary segmentation algorithm to detect multiple break points.

Keywords Structural breaks; Hypothesis testing; Correlation; Residual effect

*Fakultät Statistik, Technische Universität Dortmund, Vogelpothsweg 87, 44227 Dortmund, Germany, Email: duan@statistik.tu-dortmund.de

†Ruhr Graduate School in Economics, Hohenzollernstr. 1-3, 45128 Essen, Germany

‡Institut für Ökonometrie und Statistik, Universität zu Köln, Meister-Ekkehart-Str. 9, 50923 Köln, Germany, Email: dwied@uni-koeln.de

JEL Classification C12; C32; C58

Acknowledgements / Conflict of interest: F. Duan gratefully acknowledges funding by Ruhr Graduate School in Economics (RGS Econ). On behalf of all authors, the corresponding author (D. Wied) states that there is no conflict of interest concerning this paper.

1 Introduction

The test of the stability of parameters or coefficients is always of interest in statistical modeling, Brown et al. (1975) introduced the CUSUM-type test on the constancy of coefficients in a linear regression model. It is also prevalent to construct such a test based on residuals, Ploberger and Krämer (1992) proposed a test of the stability of regression coefficients based on OLS residuals. Furthermore, in order to reveal the property of a data generating process, it is crucial to construct change point tests for relevant quantities of data series, such as tests for volatility or dependence measures. For example, Wied, Arnold, Bissantz and Ziggel (2012) presented a fluctuation test for univariate variance constancy. Aue et al. (2009) dealt with the covariance structure stability tests in multivariate time series model. Regarding tests for constant correlation, Wied, Krämer and Dehling (2012) tested the change of correlation between bivariate random vectors using a new functional delta method, while Galeano and Wied (2014) used a binary segmentation algorithm to locate the break points in the correlation structure of bivariate random variables.

Constant correlation tests attract considerable attention as the correlation coefficient is widely accepted as an easily interpretable way of dependence measures between random variables. From a practical point of view, the correlations between financial returns on assets, for example, are rarely assumed to be constant over time in financial models. Given the importance of the impact of structural changes on the dynamic modeling of asset correlations, Berens et al. (2015) empirically investigated whether the performance of the constant conditional correlation (CCC) model and dynamic conditional correlation (DCC) model are improved with the combination of certain structural break tests.

The need for multivariate modeling has also motivated researchers to derive a multivariate constant correlation test instead of merely bivariate tests. Recently, Wied (2017) considered a nonparametric test based on a vector of successively computed pairwise correlations to test the constancy of correlation matrix, and Galeano and Wied (2017) extended their previous work on bivariate random variables to the whole correlation matrix. A potential drawback of these approaches is that the unconditional variance of the random variables under consideration is assumed to be constant. This is a restrictive assumption for financial returns (see e.g. Pape et al., 2016). To circumvent this problem, it would be possible to use residuals, in which these variance changes are calculated out, but this solution potentially leads to a residual effect.

As far as we know, the literature about such a residual effect is rather limited. Recently, Demetrescu and Wied (2017) identified the residual effect in correlation constancy tests. They discussed whether the effect is asymptotically negligible in various scenarios. Their approach falls into the category of inference on the moment hypothesis, which relates to some well-known formal tests on moments, as exemplified by the work of Bai and Ng (2005). Nevertheless, they restricted their work to bivariate data series. In response, we sought to fill the gap with a multivariate test on pairwise correlations with the help of an analytic derivation of the residual effect on the limit distribution given the assumption of non-constant marginal variances. The cost of the flexibility of that assumption is that the critical values of test statistic can not be obtained directly, we use a bootstrap approximation to ascertain the related critical values in the test. The residual-based approach remains useful for detecting structural breaks in the model and turns out to be convenient for handling the problem of inconstant marginal variances.

This paper is organized as follows. We introduce the test and derive its asymptotic properties in Section 2. In Section 3, we discuss the approximation methods of the limit distribution of the test statistics, after which we identify the residual effect and provide a corresponding example in Section 4. In Section 5 we present finite sample simulation studies, followed by an application of the test with real-world data in Section 6. Lastly, we offer our conclusions in Section 7.

2 Testing procedure

We are interested in the pairwise correlations based on a sequence of d-dimensional unobservable variables $\boldsymbol{Z}_t = (Z_{t,1}, Z_{t,2}, \dots, Z_{t,d}), t = 1, \dots, n$. Furthermore, we assume that there is a sample of d-dimensional observable variables $\boldsymbol{X}_t = (X_{t,1}, X_{t,2}, \dots, X_{t,d}), t = 1, \dots, n$ and a filter to connect the unobservable quantities and observable sample. All random variables in our setting have finite $|4+\delta|$-th moments with arbitrary $\delta > 0$. We assume the following simple parametric relationship:

$$X_{t,i} = (\mu_{i,1} + \sigma_{i,1} Z_{t,i})\bar{D}_{t,\lambda_0} + (\mu_{i,2} + \sigma_{i,2} Z_{t,i}) D_{t,\lambda_0}, i = 1, \dots, d, t = 1, \dots, n. \quad (1)$$

Here, we use the period indicator functions $D_{t,\lambda_0} = \mathbb{I}(t/n > \lambda_0)$ for the post-break period, and $\bar{D}_{t,\lambda_0} = 1 - D_{t,\lambda_0}$ for the pre-break period. With $\lambda_0 \in (0,1)$, we denote the true change point in the parameter vector $\boldsymbol{\theta}_{\lambda_0} = (\boldsymbol{\theta}'_{0,1}, \boldsymbol{\theta}'_{0,2})'$. This means that we have parameters in the pre-break subsample $\boldsymbol{\theta}_{0,1} = (\boldsymbol{\mu}'_1, (\boldsymbol{\sigma}^2_1)')'$ and in the post-break subsample $\boldsymbol{\theta}_{0,2} = (\boldsymbol{\mu}'_2, (\boldsymbol{\sigma}^2_2)')'$. The model implies that all parameters share an identical change point. The relation in equation (1) can be considered as a filter $\boldsymbol{f} : \mathbb{R}^{d\times 4d} \to \mathbb{R}^d$, which allows for discontinuities at the break point $[\lambda_0 n]$. The expression $[\lambda_0 n]$ denotes the greatest integer less than or equal to $\lambda_0 n$. The filter is not smooth in the whole sample with respect to the change point λ_0, but smooth in the $\boldsymbol{\theta_{0,1}}$ and $\boldsymbol{\theta_{0,2}}$ in the two subsamples separated by the change point.

Since $\boldsymbol{Z}_t$, which serves as the essential part in the test statistics, is unobservable, one has to estimate the unknown parameters $\boldsymbol{\theta}_{\lambda_0}$ based on sample $\boldsymbol{X}_t$ and uniquely backward induce $\boldsymbol{Z}_t$ assuming the above filter $\boldsymbol{f}$ is invertible. In the full sample estimation of the parameter vector $\boldsymbol{\theta}_{\lambda_0}$, the Generalized Method of Moments (GMM)-type estimators are available for two subsamples (see Hansen, 1982), we choose the simple sample averages here:

$$\hat{\mu}_{i,1} = \frac{1}{[\lambda n]} \sum_{t=1}^{[\lambda n]} X_{t,i} \quad \text{and} \quad \hat{\mu}_{i,2} = \frac{1}{n - [\lambda n]} \sum_{t=[\lambda n]+1}^{n} X_{t,i}, i = 1, \dots, d.$$

$$\hat{\sigma}^2_{i,1} = \frac{1}{[\lambda n]} \sum_{t=1}^{[\lambda n]} (X_{t,i} - \bar{X}_i)^2 \quad \text{and} \quad \hat{\sigma}^2_{i,2} = \frac{1}{n - [\lambda n]} \sum_{t=[\lambda n]+1}^{n} (X_{t,i} - \bar{X}_i)^2, i = 1, \dots, d.$$

Then the vector of estimators is written as

$$\hat{\boldsymbol{\theta}}_{\lambda_0} = (\hat{\mu}_{1,1}, \dots, \hat{\mu}_{d,1}, \hat{\sigma}^2_{1,1}, \dots, \hat{\sigma}^2_{d,1}, \hat{\mu}_{1,2}, \dots, \hat{\mu}_{d,2}, \hat{\sigma}^2_{1,2}, \dots, \hat{\sigma}^2_{d,2})'$$

where $\hat{\mu}_{i,1}, \hat{\sigma}^2_{i,1}, i = 1, \ldots, d$ denote the estimators based on the first subsample and $\hat{\mu}_{i,2}, \hat{\sigma}^2_{i,2}, i = 1, \ldots, d$ denote the estimators using the second subsample. The change point λ is either known or needs to be estimated. If the break point is known, one simply inserts $\lambda = \lambda_0$, whereas, if the break point is unknown, the estimated break point $\lambda = \hat{\lambda}$ is required. One natural estimator of such a break point in the marginal mean or variance is the corresponding time point associated with the maximum of test statistics based on the cumulative first or second moments, which is commonly used in the framework of structural break dating. This decision could also be based on the fluctuation test in Wied, Arnold, Bissantz and Ziggel (2012). As we assume that means and variances change at the same time, in practice one would have to decide for either a mean or variance constancy test or to merge the results appropriately.

Apart from the full sample estimation, the recursive estimation of $\boldsymbol{\theta}_{\lambda_0}$ based on the sample up to $t \leq n$ serves as an alternative, but we will not proceed with this method in this paper.

Based on the parameter estimators, we obtain the residual term

$$\hat{Z}_{t,i} = \frac{X_{t,i} - \hat{\mu}_{i,1}}{\hat{\sigma}_{i,1}} \bar{D}_{t,\lambda} + \frac{X_{t,i} - \hat{\mu}_{i,2}}{\hat{\sigma}_{i,2}} D_{t,\lambda}, \quad i = 1, \ldots, d, \tag{2}$$

which will be used for the test later on.

Define a vector of pairwise correlations as $vech(\boldsymbol{\rho})$ where $\boldsymbol{\rho}$ is the pairwise correlation matrix of d-dimensional variables $\boldsymbol{X}_t$ whose i, j-element is

$$\rho_{i,j} = E(Z_i Z_j), \quad \text{for} \quad 1 \leq i, j \leq d,$$

and $vech(\cdot)$ is the operator that stacks the upper off-diagonal elements in the $d \times d$ correlation matrix $\boldsymbol{\rho}$ as a vector with $d(d-1)/2$ components. The null hypothesis of the test for constant correlation is given as

$$H_0 : vech(\boldsymbol{\rho}_1) = \ldots = vech(\boldsymbol{\rho}_n).$$

The alternative hypothesis that there exists one change point in the correlation at an arbitrary time point k,

$$H_A : vech(\boldsymbol{\rho}_1) = \ldots = vech(\boldsymbol{\rho}_k) \neq vech(\boldsymbol{\rho}_{k+1}) = \ldots = vech(\boldsymbol{\rho}_n).$$

The test statistics is in a multivariate cumulative sum version such that

$$Q_n = \max_{1\le j\le n} \frac{j}{\sqrt{n}} \sqrt{(\boldsymbol{S}_j - \boldsymbol{S}_n)'\Omega^{-1}(\boldsymbol{S}_j - \boldsymbol{S}_n)} \tag{3}$$

where $\boldsymbol{S}_j = \frac{1}{j}\sum_{t=1}^{j} vech(\boldsymbol{Z}_t\boldsymbol{Z}_t')$, and the corresponding feasible test statistics follows

$$\hat{Q}_n = \max_{1\le j\le n} \frac{j}{\sqrt{n}} \sqrt{(\hat{\boldsymbol{S}}_j - \hat{\boldsymbol{S}}_n)'\hat{\Omega}^{-1}(\hat{\boldsymbol{S}}_j - \hat{\boldsymbol{S}}_n)} \tag{4}$$

where the partial sums based on residuals are defined as $\hat{\boldsymbol{S}}_j = \frac{1}{j}\sum_{t=1}^{j} vech(\hat{\boldsymbol{Z}}_t\hat{\boldsymbol{Z}}_t')$. Moreover, $\hat{\Omega}$ is the estimator of the covariance matrix of $vech(\hat{\boldsymbol{Z}}_t\hat{\boldsymbol{Z}}_t')$. The null hypothesis is rejected whenever the test statistics $\hat{Q}_n$ becomes too large, that is, whenever the cumulative sum of at least one pair of the cross products of residuals fluctuates too much over time. We denote convergence in probability by $\to_p$, convergence in distribution by $\to_d$ and weak convergence in the space of càdlàg functions on the interval $[0,1]$ by $\Rightarrow$. In order to derive the limit distribution of the test statistics, we first impose some assumptions.

Assumption 1. *Denote*

$$U_{n,t,1} := \begin{pmatrix} Z_{t,1}Z_{t,2} - \mathbb{E}(Z_{t,1}Z_{t,2}) \\ Z_{t,1}Z_{t,3} - \mathbb{E}(Z_{t,1}Z_{t,3}) \\ \vdots \\ Z_{t,d-1}Z_{t,d} - \mathbb{E}(Z_{t,d-1}Z_{t,d}) \end{pmatrix}, \quad U_{n,t,2} := \begin{pmatrix} \sigma_{1,1,0}Z_{t,1}\bar{D}_{t,\lambda_0} \\ \vdots \\ \sigma_{d,1,0}Z_{t,d}\bar{D}_{t,\lambda_0} \end{pmatrix},$$

$$U_{n,t,3} := \begin{pmatrix} \sigma_{1,1,0}^2(Z_{t,1}^2 - 1)\bar{D}_{t,\lambda_0} \\ \vdots \\ \sigma_{d,1,0}^2(Z_{t,d}^2 - 1)\bar{D}_{t,\lambda_0} \end{pmatrix}, \quad U_{n,t,4} := \begin{pmatrix} \sigma_{1,2,0}Z_{t,1}D_{t,\lambda_0} \\ \vdots \\ \sigma_{d,2,0}Z_{t,d}D_{t,\lambda_0} \end{pmatrix},$$

$$U_{n,t,5} := \begin{pmatrix} \sigma_{1,2,0}^2(Z_{t,1}^2 - 1)D_{t,\lambda_0} \\ \vdots \\ \sigma_{d,2,0}^2(Z_{t,d}^2 - 1)D_{t,\lambda_0} \end{pmatrix}.$$

Moreover, let $\boldsymbol{\Psi}_{\lambda_0}(s)$ *be a* $\left(\frac{d(d-1)}{2} + 4d\right)$*-dimensional Gaussian process with* $\boldsymbol{\Psi}_{\lambda_0}(0) = 0$ *and the* $\left(4d + \frac{d(d-1)}{2}\right) \times \left(4d + \frac{d(d-1)}{2}\right)$ *covariance matrix* $Cov(\boldsymbol{\Psi}_{\lambda_0}(1)) := \Xi_{\lambda_0} = \begin{pmatrix} \Omega & \Lambda_{\lambda_0}' \\ \Lambda_{\lambda_0} & \Sigma_{\lambda_0}. \end{pmatrix}$. *(This implies that* $\boldsymbol{\Psi}_{\lambda_0}^{1:\frac{d(d-1)}{2}}(s)$ *is a Gaussian process with* $\boldsymbol{\Psi}_{\lambda_0}^{1:\frac{d(d-1)}{2}}(1) \sim \mathcal{N}(0,\Omega)$ *whereas* $\boldsymbol{\Psi}_{\lambda_0}^{(\frac{d(d-1)}{2}+1):(\frac{d(d-1)}{2}+4d)}(s)$ *is a Gaussian process with* $\boldsymbol{\Psi}_{\lambda_0}^{(\frac{d(d-1)}{2}+1):(\frac{d(d-1)}{2}+4d)}(1) \sim \mathcal{N}(0,\Sigma_{\lambda_0})$.*)*

Then,

$$U_n(s) := \frac{1}{\sqrt{n}} \sum_{t=1}^{[sn]} \begin{pmatrix} U_{n,t,1} \\ U_{n,t,2} \\ U_{n,t,3} \\ U_{n,t,4} \\ U_{n,t,5} \end{pmatrix} \Rightarrow \boldsymbol{\Psi}_{\lambda_0}(s) := \begin{pmatrix} \Omega^{1/2}\boldsymbol{\Gamma}(s) \\ \Sigma_{\lambda_0}^{1/2}\boldsymbol{\Theta}_{\lambda_0}(s) \end{pmatrix}.$$

Assumption 2. *Define a neighborhood* $\Phi_n = \{\boldsymbol{\theta}^*_{\lambda_0} : ||\boldsymbol{\theta}^*_{\lambda_0} - \boldsymbol{\theta}_{\lambda_0}|| < Cn^{-1/2+\epsilon}, 0 < \epsilon < 1/2, C > 0\}$ *of* $\boldsymbol{\theta}_{\lambda_0}$*, such that*

$$\sup_{\boldsymbol{\theta}^*_{\boldsymbol{\lambda_0}} \in \Phi_n; t=1,\dots,n} || \frac{\partial vech(\boldsymbol{z}\boldsymbol{z}')}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}^*_t} \frac{\partial \boldsymbol{Z}^*_t}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*_{\lambda_0}} - \frac{\partial vech(\boldsymbol{z}\boldsymbol{z}')}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}_t} \frac{\partial \boldsymbol{Z}_t}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}} || \to_p 0$$

where

$$Z^*_{t,i} = \frac{X_{t,i} - \mu^*_{i,1}}{\sigma^*_{i,1}} \bar{D}_{t,\lambda_0} + \frac{X_{t,i} - \mu^*_{i,2}}{\sigma^*_{i,2}} D_{t,\lambda_0}, i = 1, \dots, d$$

and

$$\boldsymbol{\theta}^*_{\lambda_0} = (\mu^*_{1,1}, \dots, \mu^*_{d,1}, (\sigma^*_{1,1})^2, \dots, (\sigma^*_{d,1})^2, \mu^*_{1,2}, \dots, \mu^*_{d,2}, (\sigma^*_{1,2})^2, \dots, (\sigma^*_{d,2})^2).$$

Assumption 3. $\hat{\Xi}_{\lambda_0} \to_p \Xi_{\lambda_0}$ *where* $\hat{\Xi}_{\lambda_0}$ *is the estimator of covariance matrix.*

Assumption 4. *The block length in block bootstrap method* $l \sim n^\alpha, \alpha \in (0,1)$ *such that* $l \to \infty$ *and* $n \to \infty$ *but* $l/n \to 0$.

The first assumption, i.e., the limit process of (cross-)moment conditions, is necessary to derive the asymptotic behaviors of estimated correlation sequence. This means that we impose stationary conditions on the sequence of the Z_t, but not on the observed X_t. Specifically, we do not assume that $X_{t,i}, i = 1, \dots, d$ has constant expectation and variance, whereas Wied (2017) assumed that moments of the observed random variables as being constant. Later on, it will be shown that the time-varying variances play a crucial role in the appearance of residual effect. The setup in Assumption 2 controls the approximation error in the expansion of partial sums term, it ensures that the negligible parts of partial sums are able to be dropped asymptotically. In order to derive the asymptotic distribution of the proposed test, it is necessary to assume the existence of a consistent estimator of covariance matrix. One popular choice of such estimator is HAC estimator by Andrews (1991). Assumption 4 imposes a general restriction that the block length should not become too large compared to the sample size when the block length increases. Intuitively, if the process is more persistent, one needs a larger block length to better capture the dependence structure in the block bootstrapping. Since it is important to choose the block length in overlapping or non-overlapping block bootstrap procedure, there are several seminal works in this topic. Lahiri (1999) theoretically compared various block bootstrap methods for dependent data and determined the MSE-optimal block length, whereas Hall and Horowitz (1996) studied the block bootstrap in test of overidentifying restrictions and t test based on GMM estimation with dependent data. According to Inoue and Shintani (2006), one could set the block length in block bootstrap procedure equal to the HAC truncation parameter in the HAC covariance matrix estimator. Then one might consider the choice of bandwidth of Bartlett kernel in Andrews (1991) to select the order of block length as $O(n^{1/3})$, and this length is close to the average block length in Inoue and Shintani (2006).

If we simply make the test based on $\boldsymbol{Z}_t$, then we have the asymptotic behavior for centered partial sums of correlations as

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[ns]}(vech(\boldsymbol{Z}_t\boldsymbol{Z}_t') - \mathbb{E}(vech(\boldsymbol{Z}_t\boldsymbol{Z}_t')) \Rightarrow \Omega^{1/2}\boldsymbol{\Gamma}(s).$$

One can see that only $\Omega^{1/2}\boldsymbol{\Gamma}(s)$ term is involved in the limit. However, our test relies on $\hat{\boldsymbol{Z}}_t$ instead of $\boldsymbol{Z}_t$ because $\boldsymbol{Z}_t$ is unobservable quantity. Then the effect of the estimation of parameter vector matters, and as a consequence $\boldsymbol{\tau}_\lambda(s)$, $\Sigma_\lambda^{1/2}$ and $\boldsymbol{\Theta}_\lambda(1)$ are relevant terms. Then we propose the asymptotic behavior for the partial sums with $\hat{\boldsymbol{Z}}_t$ which includes the residual effect.

Proposition 1. *(Convergence of partial sums)*
Under Assumptions 1 and 2, if the true break point is known as $\lambda_0 \in (0,1)$, it holds for $n \to \infty$ and $s \in [0,1]$,

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[ns]}(vech(\hat{\boldsymbol{Z}}_t\hat{\boldsymbol{Z}}_t') - \mathbb{E}(vech(\boldsymbol{Z}_t\boldsymbol{Z}_t')) \Rightarrow \Omega^{1/2}\boldsymbol{\Gamma}(s) + \boldsymbol{\tau}_{\lambda_0}(s)\Sigma_{\lambda_0}^{1/2}\boldsymbol{\Theta}_{\lambda_0}(1), \qquad (5)$$

where the asymptotic residual effect term $\boldsymbol{\tau}_{\lambda_0}(s)$ is a deterministic matrix of differential functions which can be obtained via $\frac{1}{n}\sum_{t=1}^{[ns]}\frac{\partial vech(\boldsymbol{z}\boldsymbol{z}')}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}_t}\frac{\partial \boldsymbol{Z}_t}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}} \Rightarrow \boldsymbol{\tau}_{\lambda_0}(s)$.

The proof can be found in appendix. The residual effect term in the limit $\boldsymbol{\tau}_{\lambda_0}(s)$ can also be rewritten as

$$\boldsymbol{\tau}_{\lambda_0}(s) = \begin{pmatrix} \boldsymbol{\tau}_{\boldsymbol{\theta}_{0,1}}(s)\bar{D}_{t,\lambda_0} + \boldsymbol{\tau}_{\boldsymbol{\theta}_{0,1}}(\lambda_0)D_{t,\lambda_0} & (\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,2}}(s) - \boldsymbol{\tau}_{\boldsymbol{\theta}_{0,2}}(\lambda_0))D_{t,\lambda_0} \end{pmatrix}$$

and $\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,1}}(s)$ and $\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,2}}(s)$ denote the matrices of differential functions based on the parameter vectors in pre-break regime and post-break regime, respectively. The explicit expressions of terms $\boldsymbol{\tau}_{\lambda_0}(s)$, $\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,1}}(s)$ and $\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,2}}(s)$ will be given in Section 4. The residual effect exists when $\boldsymbol{\tau}_{\lambda_0}(s) \neq \mathbf{0}$. In contrast, the residual effect does not appear when $\boldsymbol{\tau}_{\lambda_0}(s) = \mathbf{0}, \forall s \in [0, 1]$. Demetrescu and Wied (2017) also named some exceptions that the residual-based tests of moment hypothesis are not affected by the residual effect even when $\boldsymbol{\tau}_{\lambda_0}(s) \neq \mathbf{0}$.

Proposition 2. *(Asymptotic under H_0)*
Under Assumptions 1, 2 and 3, if the true break point is known as $\lambda_0 \in (0, 1)$, under H_0, it holds as $n \to \infty$ that

$$\max_{1\leq j\leq n} \frac{j}{\sqrt{n}}\sqrt{(\hat{\boldsymbol{S}}_j - \hat{\boldsymbol{S}}_n)'\hat{\Omega}^{-1}(\hat{\boldsymbol{S}}_j - \hat{\boldsymbol{S}}_n)} \Rightarrow \sup_{s\in[0,1]} \sqrt{(\hat{\boldsymbol{\Gamma}}(s) - s\hat{\boldsymbol{\Gamma}}(1))'(\hat{\boldsymbol{\Gamma}}(s) - s\hat{\boldsymbol{\Gamma}}(1))}, \quad (6)$$

where $\hat{\boldsymbol{S}}_j = \frac{1}{j}\sum_{t=1}^{j} vech(\hat{\boldsymbol{Z}}_t\hat{\boldsymbol{Z}}_t')$ and $\hat{\boldsymbol{\Gamma}}(s) = \boldsymbol{\Gamma}(s) + \Omega^{-1/2}\boldsymbol{\tau}_{\lambda_0}(s)\Sigma_{\lambda_0}^{1/2}\boldsymbol{\Theta}_{\lambda_0}(1)$.

Note that if we use the unobservable term $\boldsymbol{Z}_t$ to compute the test statistics, we have the form of limit distribution of Q_n as $\sup_{s\in[0,1]} \sqrt{(\boldsymbol{\Gamma}(s) - s\boldsymbol{\Gamma}(1))'(\boldsymbol{\Gamma}(s) - s\boldsymbol{\Gamma}(1))}$, where the residual term does not play a role in the limit distribution of the test statistics, the limit distribution collapses to some functional forms based on multi-dimensional Brownian bridge. We have full information of this limit distribution and the corresponding critical values for the inference are directly available. As corollary 2 in Demetrescu and Wied (2017) pointed out, there are possibilities that Q_n and $\hat{Q}_n$ are asymptotically equivalent if $\boldsymbol{\tau}_{\lambda_0}(s) = \mathbf{0}, \forall s \in [0, 1]$. This occurs, for example, when structural breaks only exist in marginal means in a residual-based constancy test for the second moments. In addition, even if the residual term appears, the asymptotic equivalence of Q_n and $\hat{Q}_n$ also holds when $\boldsymbol{\tau}_{\lambda_0}(s)$ term is linear in s, i.e., $\boldsymbol{\tau}_{\lambda_0}(s) = s\boldsymbol{\tau}_{\lambda_0}(1), \forall s \in [0, 1]$. The leading example would be the case of constant correlation test only with breaks in marginal means, but not in marginal variances. For more details, please see the example in Section 4. If we do not have full information on the true break time point, one needs to plug in the estimated change point $\hat{\lambda}$, and we are able to establish an asymptotic equivalence of convergence with estimated change point and convergence with true change point. Proposition 2 in Demetrescu and Wied (2017) determines the limit distribution of the test statistics with the usage of estimated break point, which can be applied to our case once the moment function is set to be the vector of cross correlations.

3 Approximation of limit distribution

If the residual effect does exist, the asymptotic distribution of the test statistics depends on unknown parameters under the null hypothesis. If the limit random variable can be expressed by a matrix that includes the unknown parameters times a parameter-free random variable, it is possible to make the test statistics pivotal. This is the case, e.g., in the context of testing simple hypotheses, e.g., tests for skewness and kurtosis. However, this task becomes more difficult on the test for constant correlation, as the test statistic in Proposition 2 depends on the whole path of $\mathbf{\Gamma}$ process and $\boldsymbol{\tau}_\lambda$ process instead of $\mathbf{\Gamma}$ and $\boldsymbol{\tau}_\lambda$ only at $s = 1$, then one should consecutively make the correction of covariance matrices over s. In this case, in order to sidestep the problem of appropriate correction of $\hat{\Omega}$, it is easier to resort to bootstrap approach to approximate the asymptotic distribution of test statistics, for example, some block bootstrap strategies for non-IID data generating process, see Lahiri (2003), or bootstrap procedure for piecewise locally stationary time series by Zhou (2013), as the bootstrap is easy to implement, even in high dimensions.

In our bootstrap procedure, we use the demeaned random variables. If we have the IID data sample $\mathcal{X}_n = \boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$, where $\boldsymbol{X}_i, i = 1, \ldots, n$ is d-dimensional vector, it is very straightforward to obtain a bootstrap sample with IID bootstrap: draw a random sample $\mathcal{X}_b^* = \boldsymbol{X}_1^*, \ldots, \boldsymbol{X}_n^*$ with replacement from $\mathcal{X}_n$, repeat this random draw B times, where B is a large number, then obtain the B bootstrap samples. However, if $\mathcal{X}_n$ is m-dependent random variables, i.e., $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k\}$ and $\{\boldsymbol{X}_{k+m+1}, \ldots\}$ are independent for all $k \geq 1$, for some integer $m \geq 0$, the IID bootstrap method is not adequate, since the approximation error of bootstrap limit distribution does not vanish in the limit.

There are several works of bootstrap on dependent case in the literature, we consider the nonoverlapping block bootstrap based on the work of Carlstein (1986). According to Lahiri (2003), we define the length of block as $l \in [1, n]$ and divide the whole time series into b nonoverlapping blocks

$$\mathcal{B}_i = (\boldsymbol{X}_{(i-1)l+1}, \ldots, \boldsymbol{X}_{il})', i = 1, \ldots, b$$

where b is the largest integer such that $lb \leq n$. In each bootstrap repetition, we sample $k = \frac{n}{l}$ times with replacement one of original blocks $\{\mathcal{B}_1, \ldots, \mathcal{B}_b\}$ and concatenate the elements of sampled $\mathcal{B}_1^*, \ldots, \mathcal{B}_k^*$ into a sequence, then one bootstrap sample is obtained:

$$\mathcal{X}_b^* = (\boldsymbol{X}_1^*, \ldots, \boldsymbol{X}_l^*, \ldots, \boldsymbol{X}_{(b-1)l+1}^*, \ldots, \boldsymbol{X}_n^*).$$

We repeat this random draw B times to get B bootstrap samples. Since we allow a change point in each of marginal variances in the original sample, one additional transformation is necessary before we compute the bootstrap test statistics and corresponding bootstrap p-value in both IID bootstrap procedure and nonoverlapping block bootstrap procedure. In order to ensure that each bootstrap sample has the same empirical variance as the original series, after one bootstrap sample $\mathcal{X}_b^*$ is drawn, we transform it in this way: firstly each variate of $\mathcal{X}_b^*$ is split into two parts based on the estimated variance break time points in the original sample and then we variance-standardize both parts to get $\mathcal{X}_{\lambda,b}^*$ which has the same empirical marginal variance as the original sample. Then B bootstrap samples based on $\mathcal{X}_{\lambda,b}^*$ are used in the computation of the bootstrap test statistics $Q_j^*, j = 1, \ldots, B$. One calculates the test statistics $\hat{Q}$ based on the original data sample $\mathcal{X}_n$. The bootstrap p-value is given by

$$p(\hat{Q}) = \frac{1}{B} \sum_{j=1}^{B} \mathbf{1}_{\hat{Q} \leq Q_j^*}.$$

If one use appropriate bootstrap method to capture the data structure such as serial correlations and nonstationarities of the true data generating process if exists, then one should have the following convergence corresponding to the proposition above:

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{[ns]} (vech(\hat{\boldsymbol{Z}}_{t,\lambda,b}^*(\hat{\boldsymbol{Z}}_{t,\lambda,b}^*)') - \mathbb{E}^*(vech(\boldsymbol{Z}_{t,\lambda,b}^*(\boldsymbol{Z}_{t,\lambda,b}^*)') \Rightarrow \Omega^{1/2}\boldsymbol{\Gamma}(s) + \boldsymbol{\tau}_{\lambda_0}(s)\Sigma_{\lambda_0}^{1/2}\boldsymbol{\Theta}_{\lambda_0}(1).$$

The simulation result in Section 5 validates that the empirical size and empirical power of test stay reasonable with bootstrap strategy.

4 Characterization of residual effect

There are several scenarios in which the limit distribution of test statistics is not affected by residual effects, see Demetrescu and Wied (2017). To enumerate few of examples, we have the test of constant correlation when no change point exists in marginal means and in marginal variances, or the test on simple hypothesis of constant variance when a break only appears in the marginal mean. In this section we focus on test for the multivariate constant correlations when the residual effect exists, more explicitly, the case when there is one known break point $\lambda_0 \in (0,1)$ in both marginal variances and marginal means, the residual effect is obtained through deriving the analytic solution of the sequence of essential term $\boldsymbol{\tau}_{\lambda_0}^n$ as well as its asymptotic term $\boldsymbol{\tau}_{\lambda_0}$. Assume that

$$\boldsymbol{X}_t = \boldsymbol{\mu}_1 \bar{D}_{t,\lambda_0} + \boldsymbol{\mu}_2 D_{t,\lambda_0} + \boldsymbol{V}\boldsymbol{Z}_t, t = 1, \dots, n$$

where

$$\boldsymbol{V} = \begin{pmatrix} \sqrt{\sigma_{1,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{1,2}^2 D_{t,\lambda_0}} & 0 & \dots & 0 \\ 0 & \sqrt{\sigma_{2,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{2,2}^2 D_{t,\lambda_0}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \sqrt{\sigma_{d,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{d,2}^2 D_{t,\lambda_0}} \end{pmatrix}$$

and $\boldsymbol{X}_t$, $\boldsymbol{Z}_t$, $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are d-dimensional vectors defined in Section 2. Demetrescu and Wied (2017) showed that the break in marginal means does not have asymptotic effect in bivariate case, which can be consistently extended to the multivariate setting. However, we would like to derive a general result for the multivariate case, so we do not set $(\boldsymbol{\mu}_1', \boldsymbol{\mu}_2')' = \boldsymbol{0}$ at this moment. The residual term is given by

$$\hat{Z}_{t,i} = \frac{X_{t,i} - \hat{\mu}_{i,1}\bar{D}_{t,\lambda_0} - \hat{\mu}_{i,2}D_{t,\lambda_0}}{\sqrt{\hat{\sigma}_{i,1}^2 \bar{D}_{t,\lambda_0} + \hat{\sigma}_{i,2}^2 D_{t,\lambda_0}}}, \quad \text{for} \quad i = 1, \dots, d, \quad t = 1, \dots, n.$$

Staying in line with the notation in previous section, we consider a vectorized version of cross product term $z_i z_j, i = 1, \dots, d-1, j = i+1, \dots, d$:

$$vech(\boldsymbol{z}\boldsymbol{z}') = (z_1 z_2, z_1 z_3, \dots, z_1 z_d, z_2 z_3, \dots, z_2 z_d, \dots, z_{d-1} z_d)'.$$

The $\boldsymbol{\tau}_{\lambda_0}^n$ sequence consists of two parts, the first part is the Jacobian matrix of $vech(\boldsymbol{z}\boldsymbol{z}')$ evaluated at $\boldsymbol{z} = \boldsymbol{Z}_t$, we write the Jacobian matrix of $vech(\boldsymbol{z}\boldsymbol{z}')$ with respect to d-

dimensional vector $\boldsymbol{z}$ in partitioned form:

$$\frac{\partial vech(\boldsymbol{z}\boldsymbol{z}')}{\partial \boldsymbol{z}} = \begin{pmatrix} \boldsymbol{J}_1(\boldsymbol{z}) \\ \boldsymbol{J}_2(\boldsymbol{z}) \\ \vdots \\ \boldsymbol{J}_{d-1}(\boldsymbol{z}) \end{pmatrix}_{\frac{d(d-1)}{2}\times d}$$

where $\boldsymbol{J}_i(\boldsymbol{z}) = \frac{\partial \boldsymbol{z}_i^*}{\partial \boldsymbol{z}}, i = 1, \ldots, d-1$ is the $(d-i)\times d$ Jacobian matrix for the i-th sub-vector of $vech(\boldsymbol{z}\boldsymbol{z}')$:

$$\boldsymbol{z}_i^* = (z_i z_{i+1}, z_i z_{i+2}, \ldots, z_i z_d)', i = 1, \ldots, d-1.$$

More explicitly,

$$\boldsymbol{J}_i(\boldsymbol{z}) = \begin{cases} \begin{pmatrix} \boldsymbol{z}_{-i} & \vdots & z_i \cdot \boldsymbol{I}_{(d-i)\times(d-i)} \end{pmatrix} & \text{for } i = 1 \\ \begin{pmatrix} \boldsymbol{0}_{(d-i)\times(i-1)} & \vdots & \boldsymbol{z}_{-i} & \vdots & z_i \cdot \boldsymbol{I}_{(d-i)\times(d-i)} \end{pmatrix} & \text{for } i = 2, \ldots, d-1. \end{cases}$$

Note that the sub-block matrix $\boldsymbol{0}_{(d-i)\times(i-1)}, i = 2, \ldots, d-1$ in $\boldsymbol{J}_i(\boldsymbol{z})$ is a $(d-i)\times(i-1)$ matrix where all elements are 0, this sub-block matrix is dropped when $i = 1$. The sub-block $\boldsymbol{z}_{-i}, i = 1, \ldots, d-1$ is a column vector with length $d-i$ consisting of $z_{i+1}, \ldots, z_d, \quad i = 1, \ldots, d-1$. In other words, it excludes the first i elements in the vector $\boldsymbol{z}$. The last sub-block matrix $z_i \cdot \boldsymbol{I}_{(d-i)\times(d-i)}, i = 1, \ldots, d-1$ is the diagonal matrix with z_i on the diagonal and 0 at all off-diagonal positions, where $\boldsymbol{I}_{(d-i)\times(d-i)}$ is $(d-i)\times(d-i)$ identity matrix. The second essential part in $\boldsymbol{\tau}_{\lambda_0}^n$ is the gradient of $\boldsymbol{Z}_t$ evaluated at $\boldsymbol{\theta}_{\lambda_0}$:

$$\frac{\partial \boldsymbol{Z}_t}{\partial \boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}} = \begin{pmatrix} \boldsymbol{A}_1 & \vdots & \boldsymbol{A}_2 & \vdots & \boldsymbol{A}_3 & \vdots & \boldsymbol{A}_4 \end{pmatrix}_{d\times 4d}$$

where $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, $\boldsymbol{A}_3$, $\boldsymbol{A}_4$ are $d\times d$ diagonal matrices such that

$$\begin{aligned}
\boldsymbol{A}_1 &= \text{diag}\left((-\frac{\bar{D}_{t,\lambda_0}}{\sqrt{\sigma_{1,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{1,2}^2 D_{t,\lambda_0}}}, -\frac{\bar{D}_{t,\lambda_0}}{\sqrt{\sigma_{2,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{2,2}^2 D_{t,\lambda_0}}}, \ldots, -\frac{\bar{D}_{t,\lambda_0}}{\sqrt{\sigma_{d,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{d,2}^2 D_{t,\lambda_0}}})' \right) \\
\boldsymbol{A}_2 &= \text{diag}\left((-\frac{1}{2}\frac{Z_{t,1}\bar{D}_{t,\lambda_0}}{\sigma_{1,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{1,2}^2 D_{t,\lambda_0}}, -\frac{1}{2}\frac{Z_{t,2}\bar{D}_{t,\lambda_0}}{\sigma_{2,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{2,2}^2 D_{t,\lambda_0}}, \ldots, -\frac{1}{2}\frac{Z_{t,d}\bar{D}_{t,\lambda_0}}{\sigma_{d,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{d,2}^2 D_{t,\lambda_0}})' \right) \\
\boldsymbol{A}_3 &= \text{diag}\left((-\frac{D_{t,\lambda_0}}{\sqrt{\sigma_{1,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{1,2}^2 D_{t,\lambda_0}}}, -\frac{D_{t,\lambda_0}}{\sqrt{\sigma_{2,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{2,2}^2 D_{t,\lambda_0}}}, \ldots, -\frac{D_{t,\lambda_0}}{\sqrt{\sigma_{d,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{d,2}^2 D_{t,\lambda_0}}})' \right) \\
\boldsymbol{A}_4 &= \text{diag}\left((-\frac{1}{2}\frac{Z_{t,1}D_{t,\lambda_0}}{\sigma_{1,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{1,2}^2 D_{t,\lambda_0}}, -\frac{1}{2}\frac{Z_{t,2}D_{t,\lambda_0}}{\sigma_{2,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{2,2}^2 D_{t,\lambda_0}}, \ldots, -\frac{1}{2}\frac{Z_{t,d}D_{t,\lambda_0}}{\sigma_{d,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{d,2}^2 D_{t,\lambda_0}})' \right)
\end{aligned}$$

where the diag$(\cdot)$ operator transforms a column vector into its corresponding diagonal matrix. The sample (cross-) moments of $\boldsymbol{Z}_t$ converges to its theoretical counterparts before and after the change point, respectively. After combine two differentials above, we obtain the asymptotic term of $\boldsymbol{\tau}^n_{\lambda_0}(s)$ as

Proposition 3. *(Characterization of the residual effect in the limit)*
Under Assumption 1, if the true break point is known as $\lambda_0 \in (0,1)$, as $n \to \infty$, we have

$$\boldsymbol{\tau}^n_{\lambda_0}(s) = \frac{1}{n}\sum_{t=1}^{[ns]} \frac{\partial vech(\boldsymbol{z}\boldsymbol{z}')}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z_t}} \frac{\partial \boldsymbol{Z}_t}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}} \Rightarrow \left(\begin{array}{c:c:c:c} \boldsymbol{B}_1 & \boldsymbol{B}_2 & \boldsymbol{B}_3 & \boldsymbol{B}_4 \end{array} \right) \coloneqq \boldsymbol{\tau}_{\lambda_0}(s)$$

where the sub-matrices $\boldsymbol{B}_1, \boldsymbol{B}_2, \boldsymbol{B}_3, \boldsymbol{B}_4$ are of $\frac{d(d-1)}{2} \times d$ dimensions. They can be further expressed in an explicit way:

$$\boldsymbol{B}_1 = \boldsymbol{0}_{\frac{d(d-1)}{2}\times d} \quad and \quad \boldsymbol{B}_3 = \boldsymbol{0}_{\frac{d(d-1)}{2}\times d}$$

$$\boldsymbol{B}_2 = \begin{pmatrix} \boldsymbol{B}^*_{2,1} \\ \boldsymbol{B}^*_{2,2} \\ \vdots \\ \boldsymbol{B}^*_{2,d-1} \end{pmatrix}_{\frac{d(d-1)}{2}\times d} \quad and \quad \boldsymbol{B}_4 = \begin{pmatrix} \boldsymbol{B}^*_{4,1} \\ \boldsymbol{B}^*_{4,2} \\ \vdots \\ \boldsymbol{B}^*_{4,d-1} \end{pmatrix}_{\frac{d(d-1)}{2}\times d}$$

where

$$\boldsymbol{B}^*_{2,i} = \begin{cases} \left(\begin{array}{c:c} -\frac{1}{2}(s\bar{D}_{t,\lambda_0} + \lambda_0 D_{t,\lambda_0})\boldsymbol{\rho}^*_{i,1} & -\frac{1}{2}(s\bar{D}_{t,\lambda_0} + \lambda_0 D_{t,\lambda_0})\boldsymbol{\rho}^{**}_{i,1} \end{array} \right) & i = 1 \\ \left(\begin{array}{c:c:c} \boldsymbol{0}_{(d-i)\times(i-1)} & -\frac{1}{2}(s\bar{D}_{t,\lambda_0} + \lambda_0 D_{t,\lambda_0})\boldsymbol{\rho}^*_{i,1} & -\frac{1}{2}(s\bar{D}_{t,\lambda_0} + \lambda_0 D_{t,\lambda_0})\boldsymbol{\rho}^{**}_{i,1} \end{array} \right) & i = 2,\ldots,d-1 \end{cases}$$

$$\boldsymbol{B}^*_{4,i} = \begin{cases} \left(\begin{array}{c:c} -\frac{1}{2}(s-\lambda_0)D_{t,\lambda_0}\boldsymbol{\rho}^*_{i,2} & -\frac{1}{2}(s-\lambda_0)D_{t,\lambda_0}\boldsymbol{\rho}^{**}_{i,2} \end{array} \right) & i = 1 \\ \left(\begin{array}{c:c:c} \boldsymbol{0}_{(d-i)\times(i-1)} & -\frac{1}{2}(s-\lambda_0)D_{t,\lambda_0}\boldsymbol{\rho}^*_{i,2} & -\frac{1}{2}(s-\lambda_0)D_{t,\lambda_0}\boldsymbol{\rho}^{**}_{i,2} \end{array} \right) & i = 2,\ldots,d-1. \end{cases}$$

The vector $\boldsymbol{\rho}^*_{i,j}$ and the diagonal matrix $\boldsymbol{\rho}^{**}_{i,j}$ follow

$$\boldsymbol{\rho}^*_{i,j} = (\rho_{i,i+1}/\sigma^2_{i,j}, \rho_{i,i+2}/\sigma^2_{i,j}, \ldots, \rho_{i,d}/\sigma^2_{i,j})'$$

$$\boldsymbol{\rho}^{**}_{i,j} = diag\left((\rho_{i,i+1}/\sigma^2_{i+1,j}, \rho_{i,i+2}/\sigma^2_{i+2,j}, \ldots, \rho_{i,d}/\sigma^2_{d,j})'\right).$$

The derivation of this proposition is trivial and not presented here. Both $\boldsymbol{B}_1$ and $\boldsymbol{B}_3$ are $\frac{d(d-1)}{2} \times d$ matrices consisting of nothing but zeros. $\boldsymbol{B}_2$ and $\boldsymbol{B}_4$ are rewritten in partitioned form. In $\boldsymbol{B}^*_{2,i}$ and $\boldsymbol{B}^*_{4,i}$, $i = 1,\ldots,d-1$, the sub-block matrix $\boldsymbol{0}_{(d-i)\times(i-1)}$

remains as a $(d-i)\times(i-1)$ matrix consisting of zeros $\forall i>1$, whereas it disappears when $i=1$. The next sub-block in $\boldsymbol{B}^*_{2,i}$(or $\boldsymbol{B}^*_{4,i}$) which consists of $\boldsymbol{\rho}^*_{i,1}$(or $\boldsymbol{\rho}^*_{i,2}$) is a $(d-i)$ length vector where $(\rho_{i,i+1},\rho_{i,i+2},\ldots,\rho_{i,d})'$, i.e., the i-th sub-vector of vectorized pairwise correlations $vech(\boldsymbol{\rho})$ plays a role. The last sub-block matrix is a $(d-i)\times(d-i)$ diagonal matrix, where the aforementioned i-th sub-vector of $vech(\boldsymbol{\rho})$ emerges once again. Alternatively, one can rewrite $\boldsymbol{\tau}_{\lambda_0}(s)$ in terms of $\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,1}}(s)$ and $\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,2}}(s)$ which capture the residual effect with respect to the parameter vectors before and after the break point:

$$\boldsymbol{\tau}_{\lambda_0}(s)=\Big(\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,1}}(s)\bar{D}_{t,\lambda_0}+\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,1}}(\lambda_0)D_{t,\lambda_0}\quad(\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,2}}(s)-\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,2}}(\lambda_0))D_{t,\lambda_0}\Big)$$

where

$$\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,1}}(s):=\left(\begin{array}{c:c}\boldsymbol{C}_1 & \boldsymbol{C}_2\end{array}\right)\quad\text{and}\quad\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,2}}(s):=\left(\begin{array}{c:c}\boldsymbol{C}_3 & \boldsymbol{C}_4\end{array}\right).$$

Write out the expressions of $\boldsymbol{C}_j, j=1,\ldots,4$ as

$$\boldsymbol{C}_1=\mathbf{0}_{\frac{d(d-1)}{2}\times d}\quad\text{and}\quad\boldsymbol{C}_3=\mathbf{0}_{\frac{d(d-1)}{2}\times d}$$

$$\boldsymbol{C}_2=\begin{pmatrix}\boldsymbol{C}^*_{2,1}\\ \boldsymbol{C}^*_{2,2}\\ \vdots\\ \boldsymbol{C}^*_{2,d-1}\end{pmatrix}_{\frac{d(d-1)}{2}\times d}\quad\text{and}\quad\boldsymbol{C}_4=\begin{pmatrix}\boldsymbol{C}^*_{4,1}\\ \boldsymbol{C}^*_{4,2}\\ \vdots\\ \boldsymbol{C}^*_{4,d-1}\end{pmatrix}_{\frac{d(d-1)}{2}\times d}$$

where

$$\boldsymbol{C}^*_{2,i}=\begin{cases}\left(\begin{array}{c:c}-\frac{1}{2}s\boldsymbol{\rho}^*_{i,1} & -\frac{1}{2}s\boldsymbol{\rho}^{**}_{i,1}\end{array}\right) & i=1\\ \left(\begin{array}{c:c:c}\mathbf{0}_{(d-i)\times(i-1)} & -\frac{1}{2}s\boldsymbol{\rho}^*_{i,1} & -\frac{1}{2}s\boldsymbol{\rho}^{**}_{i,1}\end{array}\right) & i=2,\ldots,d-1\end{cases}$$

$$\boldsymbol{C}^*_{4,i}=\begin{cases}\left(\begin{array}{c:c}-\frac{1}{2}s\boldsymbol{\rho}^*_{i,2} & -\frac{1}{2}s\boldsymbol{\rho}^{**}_{i,2}\end{array}\right) & i=1\\ \left(\begin{array}{c:c:c}\mathbf{0}_{(d-i)\times(i-1)} & -\frac{1}{2}s\boldsymbol{\rho}^*_{i,2} & -\frac{1}{2}s\boldsymbol{\rho}^{**}_{i,2}\end{array}\right) & i=2,\ldots,d-1.\end{cases}$$

The series of sub-matrices $\boldsymbol{C}^*_{2,i}$(or $\boldsymbol{C}^*_{4,i}$), $i=1,\ldots,d-1$ can be clarified in the similar manner as series $\boldsymbol{B}^*_{2,i}$(or $\boldsymbol{B}^*_{4,i}$) $i=1,\ldots,d-1$, although the indicator functions are taken into account only in the later series. The above derivations imply that the $\boldsymbol{\tau}_{\lambda_0}(s)$ term is merely piecewise linear in s as a result of discontinuity of the indicator function. $\boldsymbol{B}_1$ and $\boldsymbol{B}_3$ matrices indicate that the piecewise demeaning does not lead any residual effect whereas this effect is not asymptotically ignorable when the breaks in marginal variances occur, which is implied by $\boldsymbol{B}_2$ and $\boldsymbol{B}_4$. With the help of the analytic results derived above, one could adopt a simulation-based approximation of the limit distribution of our test statistics with consecutive corrections on covariance matrices along the time path, this serves as an alternative to bootstrap method. As discussed in Demetrescu and Wied (2017), one has to check additionally if the linear combination of $\boldsymbol{\Gamma}$ and $\boldsymbol{\Theta}_{\lambda_0}$ has the same properties as $\boldsymbol{\Gamma}$ itself. We leave it to further research and stick to the off-the-shelf bootstrap procedures at this moment. To illustrate how exactly $\boldsymbol{\tau}_{\lambda_0}(s)$ term behaves, we present an example for 3-dimensional variables with breaks both in marginal means and in marginal variances.

Example. *We assume 3-dimensional random variables $\boldsymbol{X}_t$ and $\boldsymbol{Z}_t \sim_{i.i.d.} (\boldsymbol{0}, \boldsymbol{I}_3)$. The breaks in marginal means and in marginal variances occur at the same time fraction λ_0 for simplicity, note that there is only one break point in each variate. We have*

$$\boldsymbol{X}_t = \boldsymbol{\mu}_1 \bar{D}_{t,\lambda_0} + \boldsymbol{\mu}_2 D_{t,\lambda_0} + \boldsymbol{V}\boldsymbol{Z}_t$$

where

$$\boldsymbol{V} = \begin{pmatrix} \sqrt{\sigma_{1,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{1,2}^2 D_{t,\lambda_0}} & 0 & 0 \\ 0 & \sqrt{\sigma_{2,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{2,2}^2 D_{t,\lambda_0}} & 0 \\ 0 & 0 & \sqrt{\sigma_{3,1}^2 \bar{D}_{t,\lambda_0} + \sigma_{3,2}^2 D_{t,\lambda_0}} \end{pmatrix}$$

and the residual is given as

$$\hat{Z}_{t,i} = \frac{X_{t,i} - \hat{\mu}_{i,1}\bar{D}_{t,\lambda_0} - \hat{\mu}_{i,2} D_{t,\lambda_0}}{\sqrt{\hat{\sigma}_{i,1}^2 \bar{D}_{t,\lambda_0} + \hat{\sigma}_{i,2}^2 D_{t,\lambda_0}}}, \quad for \quad i = 1, 2, 3.$$

The true parameters vector is

$$\boldsymbol{\theta}_{\lambda_0} = (\mu_{1,1}, \mu_{2,1}, \mu_{3,1}, \sigma_{1,1}^2, \sigma_{2,1}^2, \sigma_{3,1}^2, \mu_{1,2}, \mu_{2,2}, \mu_{3,2}, \sigma_{1,2}^2, \sigma_{2,2}^2, \sigma_{3,2}^2)'.$$

Let the cross correlation pairs be

$$vech(\boldsymbol{z}\boldsymbol{z}') = \begin{pmatrix} z_1 z_2 & z_1 z_3 & z_2 z_3 \end{pmatrix}'.$$

The corresponding derivative of $vech(\boldsymbol{z}\boldsymbol{z}')$ evaluated at $\boldsymbol{Z}_t$:

$$\frac{\partial vech(\boldsymbol{z}\boldsymbol{z}')}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z_t}} = \begin{pmatrix} Z_{t,2} & Z_{t,1} & 0 \\ Z_{t,3} & 0 & Z_{t,1} \\ 0 & Z_{t,3} & Z_{t,2} \end{pmatrix}.$$

The corresponding derivative of $\boldsymbol{Z}_t$ evaluated at $\boldsymbol{\theta}_{\lambda_0}$:

$$\frac{\partial \boldsymbol{Z}_t}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}} = \left(\boldsymbol{A}_1 \,\vdots\, \boldsymbol{A}_2 \,\vdots\, \boldsymbol{A}_3 \,\vdots\, \boldsymbol{A}_4 \right)$$

where

$$\boldsymbol{A}_1 = \begin{pmatrix} -\frac{\bar{D}_{t,\lambda_0}}{\sqrt{\sigma_{1,1}^2\bar{D}_{t,\lambda_0}+\sigma_{1,2}^2 D_{t,\lambda_0}}} & 0 & 0 \\ 0 & -\frac{\bar{D}_{t,\lambda_0}}{\sqrt{\sigma_{2,1}^2\bar{D}_{t,\lambda_0}+\sigma_{2,2}^2 D_{t,\lambda_0}}} & 0 \\ 0 & 0 & -\frac{\bar{D}_{t,\lambda_0}}{\sqrt{\sigma_{3,1}^2\bar{D}_{t,\lambda_0}+\sigma_{3,2}^2 D_{t,\lambda_0}}} \end{pmatrix}$$

$$\boldsymbol{A}_2 = -\frac{1}{2}\begin{pmatrix} \frac{Z_{t,1}\bar{D}_{t,\lambda_0}}{\sigma_{1,1}^2\bar{D}_{t,\lambda_0}+\sigma_{1,2}^2 D_{t,\lambda_0}} & 0 & 0 \\ 0 & \frac{Z_{t,2}\bar{D}_{t,\lambda_0}}{\sigma_{2,1}^2\bar{D}_{t,\lambda_0}+\sigma_{2,2}^2 D_{t,\lambda_0}} & 0 \\ 0 & 0 & \frac{Z_{t,3}\bar{D}_{t,\lambda_0}}{\sigma_{3,1}^2\bar{D}_{t,\lambda_0}+\sigma_{3,2}^2 D_{t,\lambda_0}} \end{pmatrix}$$

$$\boldsymbol{A}_3 = \begin{pmatrix} -\frac{D_{t,\lambda_0}}{\sqrt{\sigma_{1,1}^2\bar{D}_{t,\lambda_0}+\sigma_{1,2}^2 D_{t,\lambda_0}}} & 0 & 0 \\ 0 & -\frac{D_{t,\lambda_0}}{\sqrt{\sigma_{2,1}^2\bar{D}_{t,\lambda_0}+\sigma_{2,2}^2 D_{t,\lambda_0}}} & 0 \\ 0 & 0 & -\frac{D_{t,\lambda_0}}{\sqrt{\sigma_{3,1}^2\bar{D}_{t,\lambda_0}+\sigma_{3,2}^2 D_{t,\lambda_0}}} \end{pmatrix}$$

$$\boldsymbol{A}_4 = -\frac{1}{2}\begin{pmatrix} \frac{Z_{t,1}D_{t,\lambda_0}}{\sigma_{1,1}^2\bar{D}_{t,\lambda_0}+\sigma_{1,2}^2 D_{t,\lambda_0}} & 0 & 0 \\ 0 & \frac{Z_{t,2}D_{t,\lambda_0}}{\sigma_{2,1}^2\bar{D}_{t,\lambda_0}+\sigma_{2,2}^2 D_{t,\lambda_0}} & 0 \\ 0 & 0 & \frac{Z_{t,3}D_{t,\lambda_0}}{\sigma_{3,1}^2\bar{D}_{t,\lambda_0}+\sigma_{3,2}^2 D_{t,\lambda_0}} \end{pmatrix}.$$

Finally, we have

$$\boldsymbol{\tau}_{\lambda_0}^n(s) = \frac{1}{n}\sum_{t=1}^{[ns]} \frac{\partial vech(\boldsymbol{z}\boldsymbol{z}')}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z_t}} \frac{\partial \boldsymbol{Z}_t}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}} \Rightarrow \left(\mathbf{0}_{(3\times 3)} \,\vdots\, \boldsymbol{B}_2 \,\vdots\, \mathbf{0}_{(3\times 3)} \,\vdots\, \boldsymbol{B}_4 \right) := \boldsymbol{\tau}_{\lambda_0}(s)$$

where

$$\boldsymbol{B}_2 = -\frac{1}{2}\begin{pmatrix} \frac{\rho_{1,2}\mathbb{I}(s<\lambda_0)}{\sigma_{1,1}^2}s + \frac{\rho_{1,2}\mathbb{I}(s\geq\lambda_0)}{\sigma_{1,1}^2}\lambda_0 & \frac{\rho_{1,2}\mathbb{I}(s<\lambda_0)}{\sigma_{2,1}^2}s + \frac{\rho_{1,2}\mathbb{I}(s\geq\lambda_0)}{\sigma_{2,1}^2}\lambda_0 & 0 \\ \frac{\rho_{1,3}\mathbb{I}(s<\lambda_0)}{\sigma_{1,1}^2}s + \frac{\rho_{1,3}\mathbb{I}(s\geq\lambda_0)}{\sigma_{1,1}^2}\lambda_0 & 0 & \frac{\rho_{1,3}\mathbb{I}(s<\lambda_0)}{\sigma_{3,1}^2}s + \frac{\rho_{1,3}\mathbb{I}(s\geq\lambda_0)}{\sigma_{3,1}^2}\lambda_0 \\ 0 & \frac{\rho_{2,3}\mathbb{I}(s<\lambda_0)}{\sigma_{2,1}^2}s + \frac{\rho_{2,3}\mathbb{I}(s\geq\lambda_0)}{\sigma_{2,1}^2}\lambda_0 & \frac{\rho_{2,3}\mathbb{I}(s<\lambda_0)}{\sigma_{3,1}^2}s + \frac{\rho_{2,3}\mathbb{I}(s\geq\lambda_0)}{\sigma_{3,1}^2}\lambda_0 \end{pmatrix}$$

$$\boldsymbol{B}_4 = -\frac{1}{2}\begin{pmatrix} \frac{\rho_{12}\mathbb{I}(s\geq\lambda_0)}{\sigma_{1,2}^2}(s-\lambda_0) & \frac{\rho_{12}\mathbb{I}(s\geq\lambda_0)}{\sigma_{2,2}^2}(s-\lambda_0) & 0 \\ \frac{\rho_{13}\mathbb{I}(s\geq\lambda_0)}{\sigma_{1,2}^2}(s-\lambda_0) & 0 & \frac{\rho_{13}\mathbb{I}(s\geq\lambda_0)}{\sigma_{3,2}^2}(s-\lambda_0) \\ 0 & \frac{\rho_{23}\mathbb{I}(s\geq\lambda_0)}{\sigma_{2,2}^2}(s-\lambda_0) & \frac{\rho_{23}\mathbb{I}(s\geq\lambda_0)}{\sigma_{3,2}^2}(s-\lambda_0) \end{pmatrix}.$$

In this case, $\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,1}}(s)$ *and* $\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,2}}(s)$ *terms follow*

$$\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,1}}(s) = \begin{pmatrix} 0 & 0 & 0 & -\frac{1}{2}s\frac{\rho_{1,2}}{\sigma^2_{1,1}} & -\frac{1}{2}s\frac{\rho_{1,2}}{\sigma^2_{2,1}} & 0 \\ 0 & 0 & 0 & -\frac{1}{2}s\frac{\rho_{1,3}}{\sigma^2_{1,1}} & 0 & -\frac{1}{2}s\frac{\rho_{1,3}}{\sigma^2_{3,1}} \\ 0 & 0 & 0 & 0 & -\frac{1}{2}s\frac{\rho_{2,3}}{\sigma^2_{2,1}} & -\frac{1}{2}s\frac{\rho_{2,3}}{\sigma^2_{3,1}} \end{pmatrix}$$

$$\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,2}}(s) = \begin{pmatrix} 0 & 0 & 0 & -\frac{1}{2}s\frac{\rho_{1,2}}{\sigma^2_{1,2}} & -\frac{1}{2}s\frac{\rho_{1,2}}{\sigma^2_{2,2}} & 0 \\ 0 & 0 & 0 & -\frac{1}{2}s\frac{\rho_{1,3}}{\sigma^2_{1,2}} & 0 & -\frac{1}{2}s\frac{\rho_{1,3}}{\sigma^2_{3,2}} \\ 0 & 0 & 0 & 0 & -\frac{1}{2}s\frac{\rho_{2,3}}{\sigma^2_{2,2}} & -\frac{1}{2}s\frac{\rho_{2,3}}{\sigma^2_{3,2}} \end{pmatrix}.$$

The $\boldsymbol{B}_1$ *and* $\boldsymbol{B}_3$ *matrix is* $\mathbf{0}_{(3\times 3)}$*, so demeaning does not have an asymptotic effect, only the variance estimation affects the limit distribution, in addition, the* $\boldsymbol{\tau}_{\lambda_0}(s)$ *term is only piecewise linear in* s*, the estimation effect does not cancel out, so there is a residual effect in constant correlation test.*

5 Finite Sample Simulation

We analyze the finite sample behavior of the multivariate constant correlation test given time-varying marginal variances. In finite samples, the size and power of the test are appealing. The case of time-varying marginal means is less attractive because the limit distribution is not affected by the residual effect when a break appears in the marginal means, as assured by the analytic derivation in Section 4. In this case, the bootstrap approximation of the asymptotic distribution of test statistics is not necessary, one could obtain the critical values directly via expressions in Kiefer (1959).

Demetrescu and Wied (2017) present simulations which show that the Wied, Krämer and Dehling (2012)-test does not work well in the case of a break in marginal variances. If the marginal variances decrease, the empirical size is lower than the nominal size, if the marginal variances increase, the Wied, Krämer and Dehling (2012)-test is over-sized. On the other hand, it turns out that the residual-based test keeps its size.

To see if this also holds in the multi-dimensional case, in the Monte Carlo simulation design, we assume a series of d-dimensional random variables, $d = 3, 5, 10$. The variables possess two types of dependence structure: serially independent process and

MA(1) process. We summarize the underlying data generating process as

$$\boldsymbol{X}_t = \boldsymbol{u}_t + \boldsymbol{\phi}\boldsymbol{u}_{t-1}, t = 1, \ldots, n$$

where $\boldsymbol{X}_t = (X_{t,1}, \ldots, X_{t,d})'$ and $\boldsymbol{u}_t = (u_{t,1}, \ldots, u_{t,d})' \sim N(\mathbf{0}, \boldsymbol{\Sigma}_d)$. The $d \times d$ matrix $\boldsymbol{\phi}$ is set to $\mathbf{0}_{d\times d}$ and diag$(0.5)_{d\times d}$ for serially independent and dependent cases, respectively.

In the size analysis, two scenarios are further considered: under the null hypothesis, on the one hand, the cross correlations of $\boldsymbol{u}_t$ are restricted to be equal, on the other hand, the cross correlations are allowed to be unequal, since the assumption of equal cross correlations across all dimensions is too restrictive in the applications. In the former scenario, we fix all off-diagonal elements in the correlation matrix to 0.4, whereas, in the latter scenario, we generate data samples from $d = 3, 5, 10$ dimensional normal distribution with the following vectorized pairwise correlation matrices, respectively:

$$vech(\boldsymbol{\rho}_3) = (0.4, 0.7, 0.5)',$$

$$vech(\boldsymbol{\rho}_5) = (0.4, 0.7, 0.1, 0.2, 0.5, 0.6, 0.2, 0.1, 0.3, 0.1)',$$

$$\begin{aligned} vech(\boldsymbol{\rho}_{10}) =&(0.4, 0.4, 0.5, 0.5, 0.5, 0.6, 0.6, 0.7, 0.7, 0.5, 0.5, 0.5, 0.6, 0.6, 0.6, 0.5, 0.5, 0.4, \\ &0.4, 0.4, 0.5, 0.5, 0.6, 0.6, 0.5, 0.5, 0.4, 0.5, 0.4, 0.4, 0.6, 0.5, 0.6, 0.5, 0.4, 0.4, \\ &0.5, 0.6, 0.5, 0.4, 0.5, 0.4, 0.5, 0.6, 0.5)'. \end{aligned}$$

All other settings of the simulation are identical to the first scenario except the cross correlations, and the expression of the vector of cross correlations aligns with the definition in previous section. The marginal variances in the first half of the sample are 1, whereas they take the values $\{0.1, 0.2, \ldots, 1.9, 2.0\}$ in the second half of the sample as a means to check the test performance given different magnitudes of marginal variances. The sample size is $T = 500$ and we run 1000 times Monte Carlo simulations. In order to obtain the critical values for the test, we use the IID bootstrap method and nonoverlapping block bootstrap method for serially independent and dependent cases, respectively, i.e., random draws of single elements or blocks of elements from joint empirical distributions of demeaned $\boldsymbol{X}_t$ with replacement, the number of bootstrap replications is $B = 199$ and the block length is $T^{1/3}$.

In the power analysis, under the alternative hypothesis, the marginal variances are set to 1 in the first half and 2 in the second half of the sample. All pairs of cross correlations in the first half of the sample is set to 0.4, and we have a change in cross correlations in the middle of the sample such that the pairwise correlations take the values $\{0, 0.1, \ldots, 0.6, 0.7, 0.8\}$ in the second half of the sample. This implies that all pairs of cross correlations still remain equal to each other in the second half of the sample, after the shift in the magnitude of pairwise correlations occurs. The cases with the true and estimated break points are both considered.

Figures 1, 2 and 3 respectively correspond to the information stored in Tables 3,4 and 5 in appendix. Figure 1 shows the empirical size of testing procedure for the serially independent data series, whereas Figure 2 shows the empirical size of the test for the serially dependent data sample. The test generally keeps the size in both the case of equal and of unequal cross correlations. The test with true and the test with estimated break points do not yield significant differences in test size. However, when simulating 10-dimensional variables, the test size is slightly smaller than the nominal level of 0.05 given that the true data generating process is MA(1) process. We increase the sample size T from 500 to 1000 to see if a short sample length caused an under-size problem. In Figure 2, the blue curves represent the empirical rejection rates given multiple choices of marginal variances in the second half of the sample after the sample size is increased. We achieved generally better empirical size with the larger sample, i.e., the empirical rejection frequencies are closer to the level of nominal size. Alternative solutions would be trails with higher number of repetitions of simulations or higher numbers of repetitions of bootstraps. At the same time, the standard block bootstrap methods might fail in the case of high dimension, Zhang and Cheng (2014) recently studied on bootstrap inference in high dimensionality with special interest of the interplay between dependence structure and dimensionality. Nevertheless, it is necessary to adopt some modified block bootstrap procedures when the dimension is comparable or even greater than the sample size, our case clearly does not belong to that category, then it is unnecessary to do so.

Concerning the empirical power of the test, please see Figure 3. Again, one can conclude that using the estimated break point does not significantly affect the test power.

Moreover, when $N = 3, 5$, we achieve considerably larger power for larger shifts of correlation coefficients in the second part of the sample, whereas, in 10-dimensional case, we achieve lower empirical rejection rates given large jumps in pairwise correlations in the second subsample. This occurs, for example, when the correlation coefficient changes from 0.4 to 0.8. We double the sample size and replicate the power analysis for 10 variates, the blue curves in Figure 3 indicate that the empirical power is improved for both serially independent and serially dependent data generating processes. In addition, one can observe an asymmetric empirical power level, especially for serially dependent data series. This means that the empirical test rejection rates when the correlation coefficients change to higher values are not as high as those when the correlation coefficients shift downward. This problem is possibly caused by the simulation setting in our approach: an identical block length $T^{1/3}$ in nonoverlapping block bootstrap is selected for all cases regardless of the persistence in data generating process. It makes more sense, for example, to choose a larger block length in the case of stronger correlation in the second subsample. To improve this, one could refer to Politis and White (2004), they proposed a selection procedure of the optimal block length for circular bootstrap and stationary bootstrap in an adaptive way. Since it is not our main concern, we will not proceed with such method given that our result of empirical power is considerable.

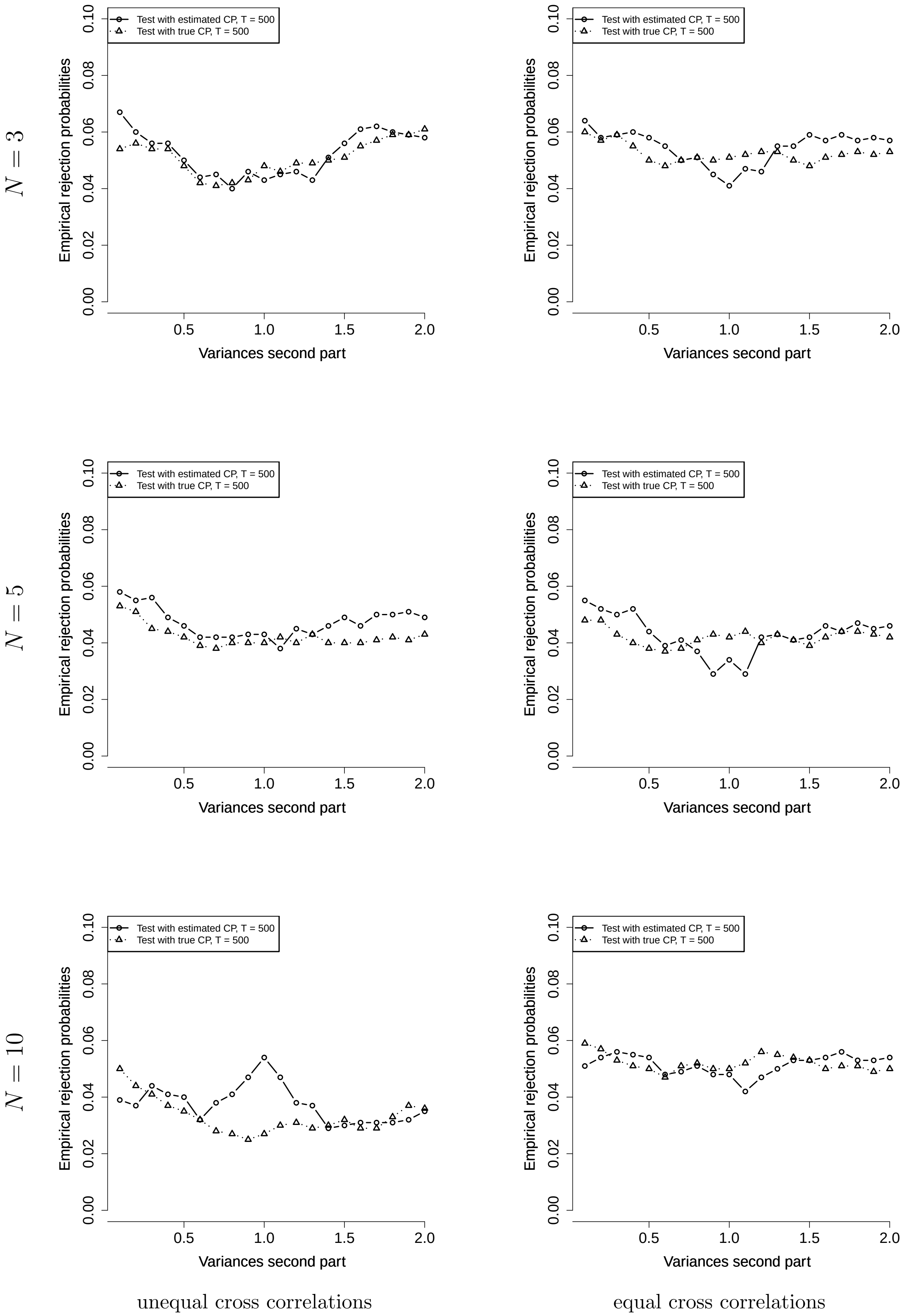

Figure 1: Empirical size for serially independent data generating process: rejection rates given constant cross correlations and nonconstant marginal variances, $N = 3, 5, 10$. CP indicates the change point in the marginal variances.

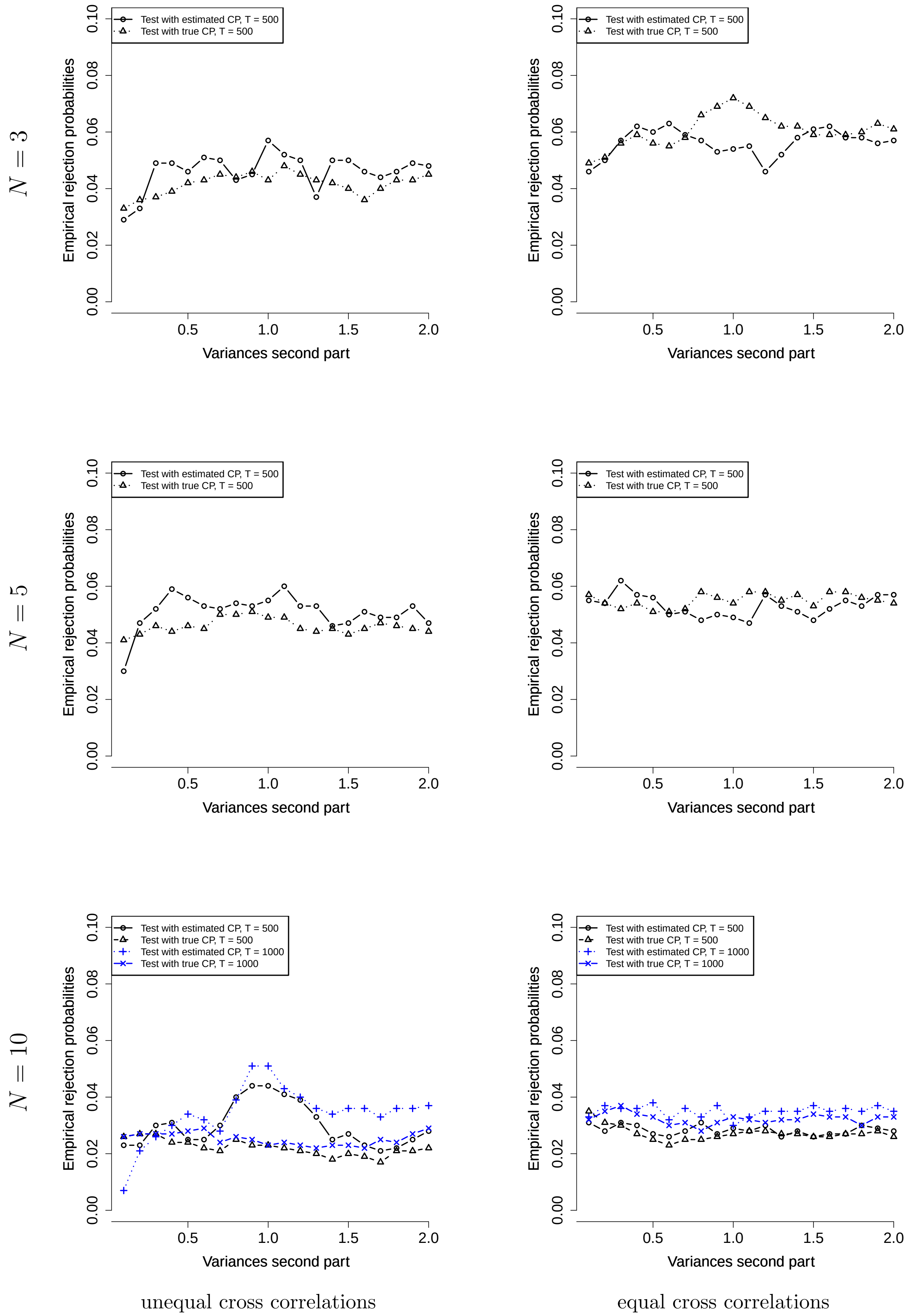

Figure 2: Empirical size for serially dependent data generating process: rejection rates given constant cross correlations and nonconstant marginal variances, $N = 3, 5, 10$. CP indicates the change point in the marginal variances.

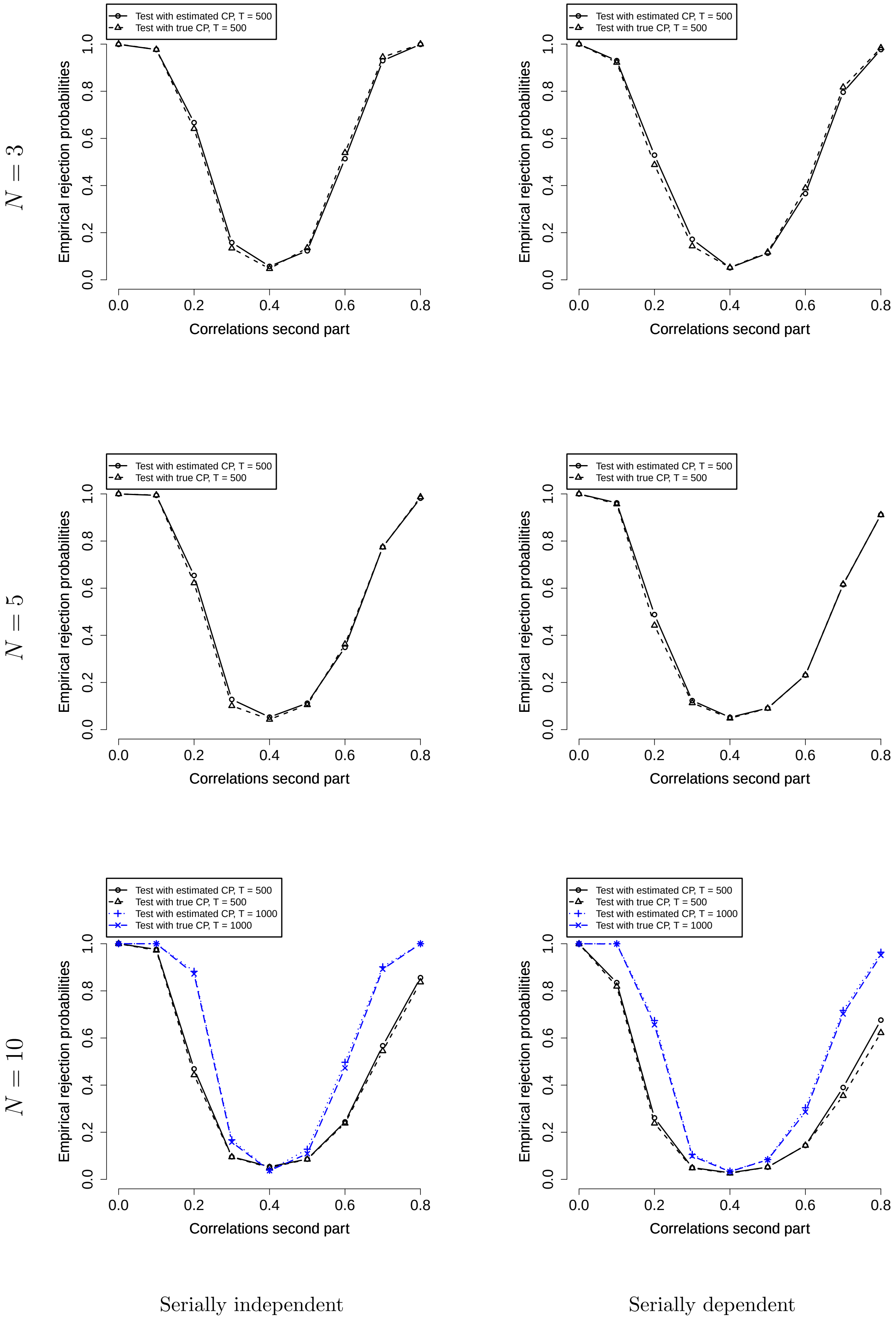

Figure 3: Empirical power: rejection rates given nonconstant cross correlations and nonconstant marginal variances, $N = 3, 5, 10$. CP indicates the change point in the marginal variances.

6 Application

In our empirical study we investigate the log returns of eight stocks from Euro Stoxx 50 from January 1 2005 to January 1 2010, for a sample length of $T = 1304$, which encompasses the period of the recent global financial crisis. The stocks belong to various industry sectors in France, Germany, Italy, Luxembourg and the Netherlands, please see Table 1 for complete information. Data were obtained from database Datastream.

Stock symbol	Variance break date	Country	Sector
ARCELOR	August 29, 2008	LUX	Steel
BASF	September 11, 2008	GER	Chemistry
DAIMLER	July 9, 2008	GER	Automobile
ENEL	September 2, 2008	ITA	Energy
INGGROEP	September 11, 2008	NL	Finance
INTESA	August 26, 2008	ITA	Finance
LVMH	January 7, 2008	FRA	Luxury
SANOFI	January 3, 2008	FRA	Pharmacy

Table 1: Information of eight European stocks including stock symbols, the change points in marginal variances, countries and industry sectors

Figure 4 shows the rolling correlations for all pairs of eight log stock returns. The window of the calculation of the rolling correlations is set to be 120 trading days, which is approximately the length of trading days in half a year. Time-varying correlations are identified. For example, the correlation between DAIMLER and ARCELOR was close to 0 in the beginning of September 2008 and had a significant increase after this.

The correlation matrix based on the full sample of log returns is given as

$$\boldsymbol{R} = \begin{pmatrix} 1.00 & 0.62 & 0.60 & 0.47 & 0.52 & 0.50 & 0.57 & 0.30 \\ 0.62 & 1.00 & 0.69 & 0.56 & 0.59 & 0.60 & 0.60 & 0.45 \\ 0.60 & 0.69 & 1.00 & 0.52 & 0.56 & 0.61 & 0.62 & 0.42 \\ 0.47 & 0.56 & 0.52 & 1.00 & 0.53 & 0.55 & 0.54 & 0.42 \\ 0.52 & 0.59 & 0.56 & 0.53 & 1.00 & 0.65 & 0.57 & 0.33 \\ 0.50 & 0.60 & 0.61 & 0.55 & 0.65 & 1.00 & 0.59 & 0.37 \\ 0.57 & 0.60 & 0.62 & 0.54 & 0.57 & 0.59 & 1.00 & 0.42 \\ 0.30 & 0.45 & 0.42 & 0.42 & 0.33 & 0.37 & 0.42 & 1.00 \end{pmatrix}.$$

Figure 5 shows the rolling marginal variances calculated in rolling windows of 120 days for all series of stock returns, none of which can be assumed to be constant throughout the period. By applying the variance constancy test from Wied, Arnold, Bissantz and Ziggel (2012) combined with a binary segmentation algorithm searching for potential multiple change points, exactly one change point is detected in each of marginal variances. The change points in marginal variances appear in Table 1. Most of the stocks exhibited a change point in summer/autumn 2008, whereas the change points of LVMH and SANOFI occurred in January 2008. Since no additional change points were detected in the refinement procedure, we continue to use those change points in our analysis.

Figure 6 is a representative ACF plot which reveals autocorrelations in the product of residuals of BASF and DAIMLER, see Figure 8 for ACF for all pairs of cross products. Since the cross products of residuals $\hat{Z}_{ti}\hat{Z}_{tj}, \forall i \neq j$ have autocorrelations as Figure 8 shows, we need to use the block bootstrap strategy instead of the IID bootstrap method to approximate the limit distribution of our test statistics, for which the block length is set to $T^{1/3}$ and the number of bootstrap replications $B = 10999$. The test statistics Q_n equals to 4.675097, Figure 9 shows the histogram of asymptotic distribution of test statistics approximated by block bootstrap procedure. The null hypothesis is rejected at the significant level $\alpha = 0.05$, and the approximate p-value is smaller than 0.001 (0.0005455041). The maximum of our test statistics identifies the break point on August 3, 2007, which was very near the beginning of the financial event: liquidity crisis dated from August 9, 2007, when investment bank BNP Paribas was unable to withdraw from two of three hedge funds, and European Central Bank

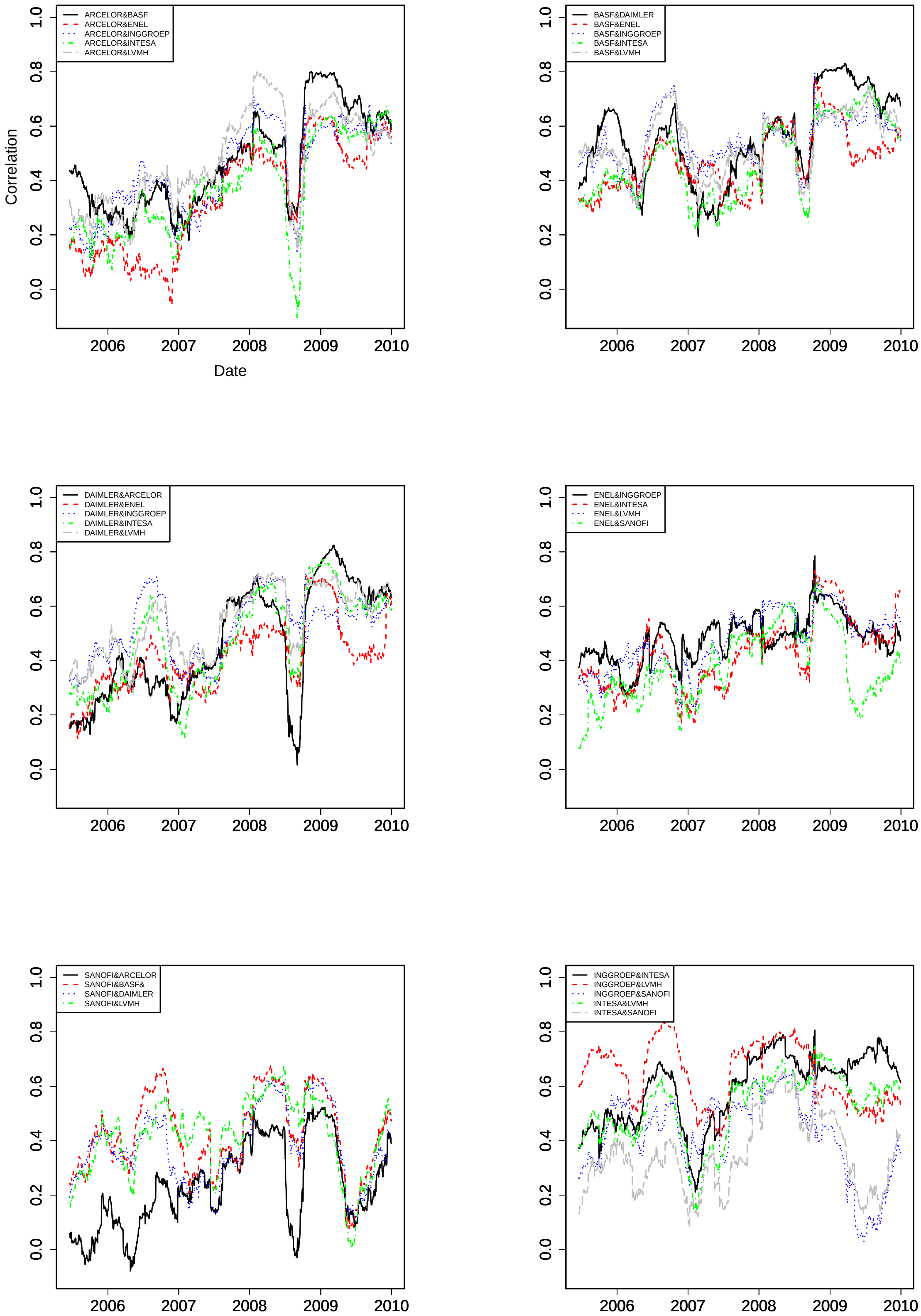

Figure 4: The rolling pairwise correlations between stock log returns

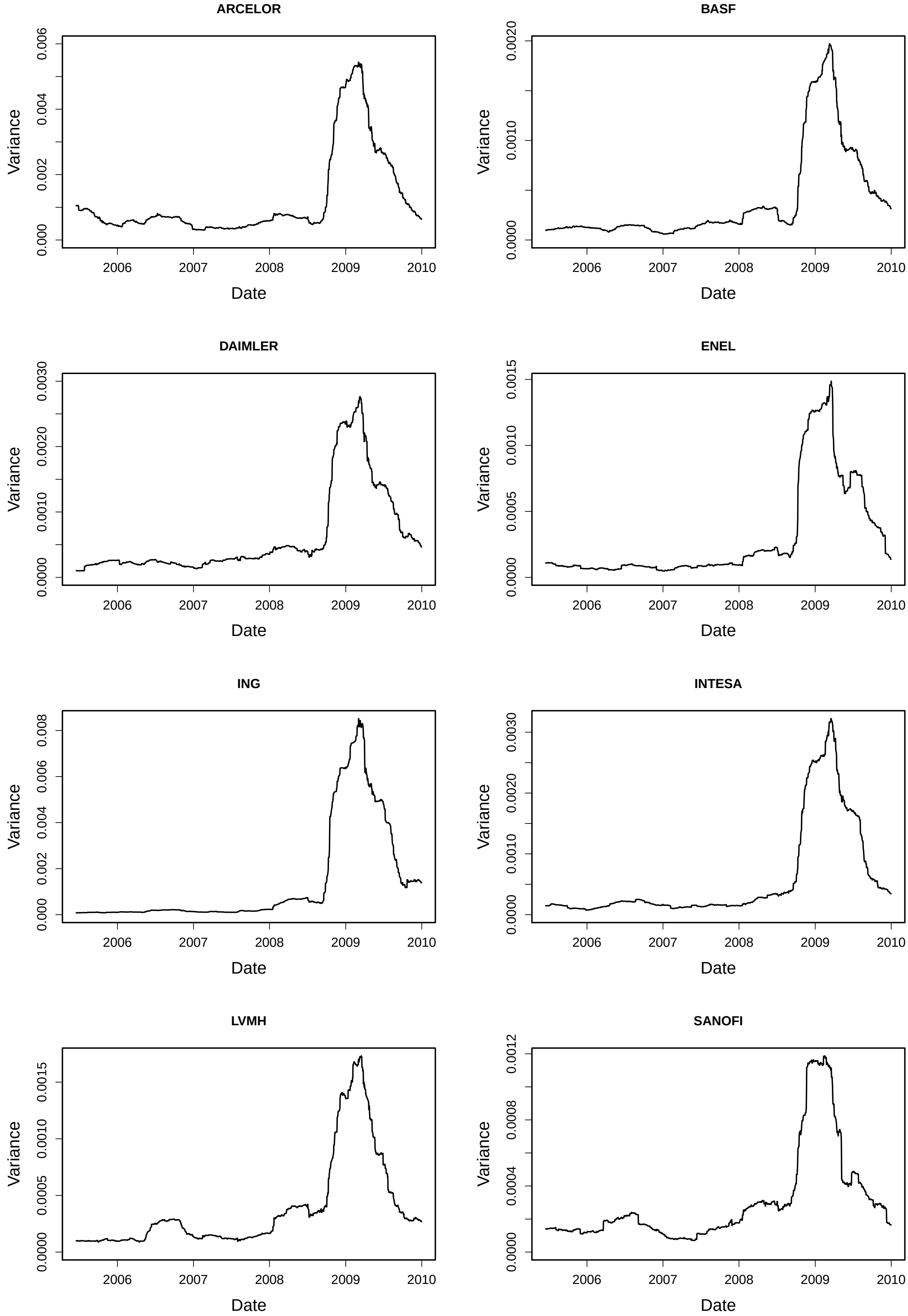

Figure 5: Rolling marginal variances of stock log returns

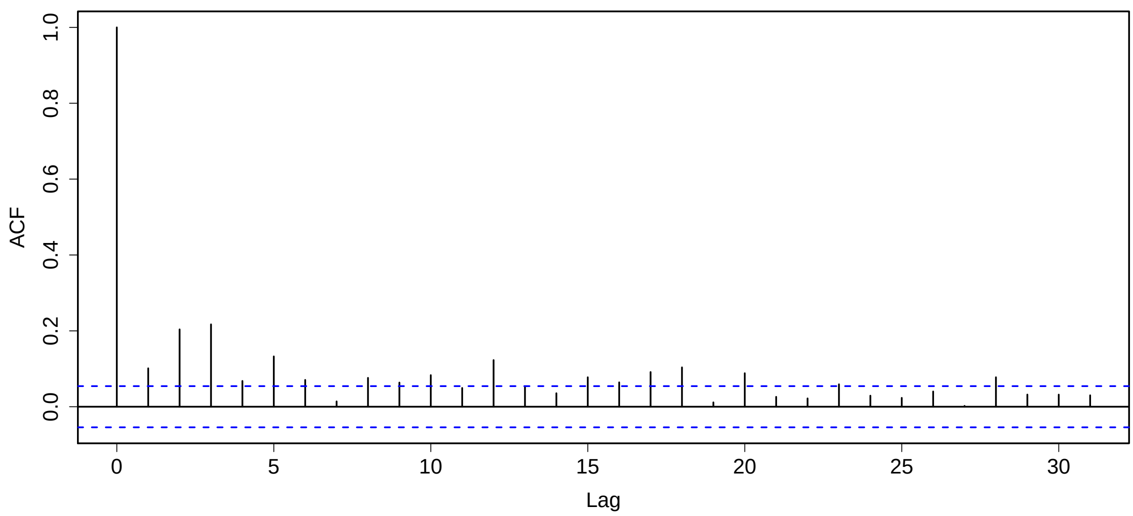

Figure 6: Autocorrelation Function (ACF) plot of cross product of residuals of BASF and DAIMLER

started to intervene in the banking market in order to improve the liquidity, as the timeline listed by Guillén (2015) corroborates.

Figure 7 presents the process of our test statistics against the time axis, which motivates us to investigate other potential break points in the sub-period between 2008 and 2009. We begin with the multivariate constant correlation test proposed by Wied (2017) combined with binary segmentation algorithm, the test statistics is given as

$$A_T = \max_{2 \leq k \leq T} \frac{k}{\sqrt{T}} ||\hat{E}^{-1/2} P_{k,T}||_1 \quad \text{with} \quad ||P_{k,T}||_1 = (\hat{\rho}_k^{ij} - \hat{\rho}_T^{ij})_{1 \leq i < j \leq p}$$

where $(\hat{\rho}_T^{ij})_{1 \leq i < j \leq p}$ is recursively estimated pairwise correlation coefficients, $\hat{E}$ is the bootstrap estimator of the asymptotic covariance of $(\hat{\rho}_k^{ij})_{1 \leq i < j \leq p}$. It pinpoints two break dates: July 6, 2007, and September 2, 2008. Details of the procedure of the binary segmentation algorithm appear in Galeano and Wied (2017).

Results in Table 2 are provided by the new test in this paper combined with binary segmentation algorithm, which show the details of each iteration in the algorithm. In the first iteration, time point $t = 674$ (August 3, 2007) is recognized as the break point in the correlation matrix with test statistics of 4.675 at the significant level 0.05. We split the entire sample into two subintervals $[1, 674]$ and $[675, 1304]$, then try to detect additional change points if available, respectively. In the first subinterval, the

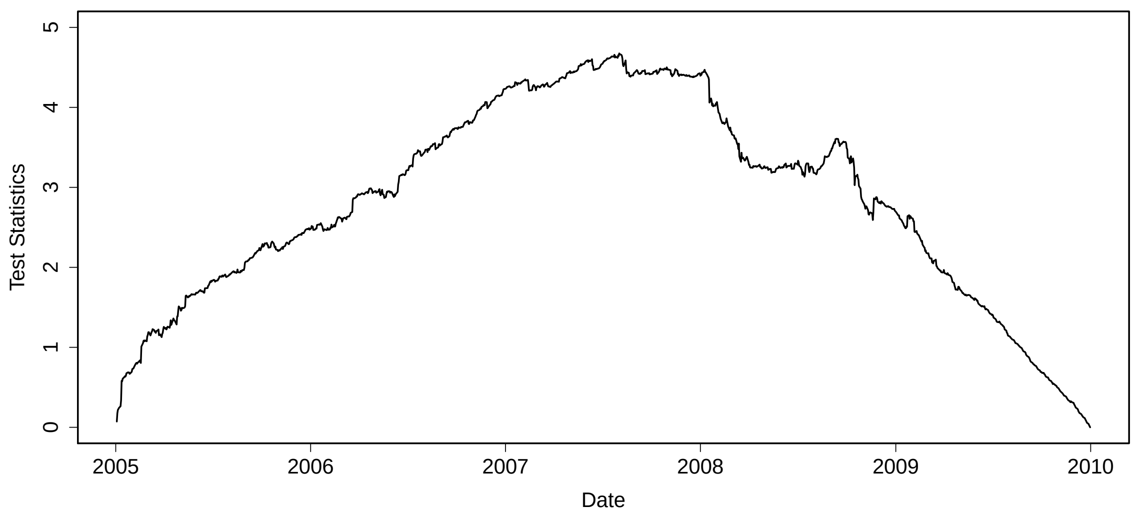

Figure 7: Process of test statistics Q_n

algorithm detects a change point at time point $t = 353$ (May 11, 2006) which is not statistically significant, whereas, in the second subinterval, the statistically significant time point $t = 1016$ (November 25, 2008) is labeled as a change point. Then we split the second subinterval into two subintervals $[675, 1016]$ and $[1017, 1304]$ and search for additional change points. A third statistically significant change point $t = 1137$ (May 13, 2009) appears in subinterval $[1017, 1304]$, then this subinterval needs to be separated into two subintervals $[1017, 1137]$ and $[1138, 1304]$, after which the searching procedure continues in subintervals. No further change points are detected in four subintervals, namely $[1, 674]$, $[675, 1016]$, $[1017, 1137]$ and $[1138, 1304]$, leaving only $t = 674$, $t = 1016$ and $t = 1137$ as change points. In the refinement step, the first change point shifts to July 25, 2007, the second break point November 25, 2008 remains, and the third change point is dropped. The results from test in Wied (2017) and our test can be tied to financial facts: although our test seems to more precisely identify the date related to liquidity crisis, Wied (2017) test provides better results of the change date corresponding to the insolvency of Lehman Brothers.

Time Interval	Change point	Q	Date
	Step 1		
$[1, 1304]$	$t = 674$	4.675^*	August 3, 2007
	Step 2		
$[1, 674]$	$t = 353$	3.308	May 11, 2006
$[675, 1304]$	$t = 1016$	3.991^*	November 25, 2008
$[1, 674]$	$t = 353$	3.308	May 11, 2006
$[675, 1016]$	$t = 811$	3.577	February 12, 2008
$[1017, 1304]$	$t = 1137$	4.056^*	May 13, 2009
$[1, 674]$	$t = 353$	3.308	May 11, 2006
$[675, 1016]$	$t = 811$	3.577	February 12, 2008
$[1017, 1137]$	$t = 1079$	2.716	February 20, 2009
$[1138, 1304]$	$t = 1211$	2.860	August 25, 2009
	Step 3 (Refinement)		
$[1, 1016]$	$t = 667$	4.995^*	July 25, 2007
$[675, 1137]$	$t = 976$	3.296	September 30, 2008
$[1017, 1304]$	$t = 1137$	4.056^*	May 13, 2009
$[1, 1137]$	$t = 667$	4.855^*	July 25, 2007
$[675, 1304]$	$t = 1016$	3.991^*	November 25, 2008

Table 2: Iterations in binary segmentation algorithm, $*$ denotes the statistically significant change points, the initial significance level $\alpha_0 = 0.05$, followed by $\alpha_1 \approx 0.025$, $\alpha_2 \approx 0.017$ in each iterations, respectively.

$$\boldsymbol{R}_1 = \begin{pmatrix} 1.00 & 0.34 & 0.28 & 0.16 & 0.29 & 0.23 & 0.34 & 0.12 \\ 0.34 & 1.00 & 0.47 & 0.40 & 0.52 & 0.37 & 0.51 & 0.38 \\ 0.28 & 0.47 & 1.00 & 0.30 & 0.44 & 0.34 & 0.44 & 0.31 \\ 0.16 & 0.40 & 0.30 & 1.00 & 0.41 & 0.33 & 0.37 & 0.25 \\ 0.29 & 0.52 & 0.44 & 0.41 & 1.00 & 0.49 & 0.65 & 0.41 \\ 0.23 & 0.37 & 0.34 & 0.33 & 0.49 & 1.00 & 0.43 & 0.25 \\ 0.34 & 0.51 & 0.44 & 0.37 & 0.65 & 0.43 & 1.00 & 0.37 \\ 0.12 & 0.38 & 0.31 & 0.25 & 0.41 & 0.25 & 0.37 & 1.00 \end{pmatrix}$$

$$\boldsymbol{R}_2 = \begin{pmatrix} 1.00 & 0.72 & 0.68 & 0.58 & 0.56 & 0.53 & 0.67 & 0.48 \\ 0.72 & 1.00 & 0.73 & 0.65 & 0.62 & 0.61 & 0.62 & 0.62 \\ 0.68 & 0.73 & 1.00 & 0.65 & 0.59 & 0.70 & 0.69 & 0.60 \\ 0.58 & 0.65 & 0.65 & 1.00 & 0.62 & 0.63 & 0.62 & 0.60 \\ 0.56 & 0.62 & 0.59 & 0.62 & 1.00 & 0.68 & 0.61 & 0.48 \\ 0.53 & 0.61 & 0.70 & 0.63 & 0.68 & 1.00 & 0.69 & 0.52 \\ 0.67 & 0.62 & 0.69 & 0.62 & 0.61 & 0.69 & 1.00 & 0.57 \\ 0.48 & 0.62 & 0.60 & 0.60 & 0.48 & 0.52 & 0.57 & 1.00 \end{pmatrix}$$

$$\boldsymbol{R}_3 = \begin{pmatrix} 1.00 & 0.67 & 0.70 & 0.52 & 0.60 & 0.60 & 0.60 & 0.19 \\ 0.67 & 1.00 & 0.74 & 0.51 & 0.61 & 0.67 & 0.63 & 0.24 \\ 0.70 & 0.74 & 1.00 & 0.46 & 0.58 & 0.62 & 0.61 & 0.23 \\ 0.52 & 0.51 & 0.46 & 1.00 & 0.49 & 0.52 & 0.52 & 0.27 \\ 0.60 & 0.61 & 0.58 & 0.49 & 1.00 & 0.68 & 0.54 & 0.16 \\ 0.60 & 0.67 & 0.62 & 0.52 & 0.68 & 1.00 & 0.54 & 0.24 \\ 0.60 & 0.63 & 0.61 & 0.52 & 0.54 & 0.54 & 1.00 & 0.22 \\ 0.19 & 0.24 & 0.23 & 0.27 & 0.16 & 0.24 & 0.22 & 1.00 \end{pmatrix}$$

We also calculate the pairwise correlations based on the three subsamples separated by the change points. The correlation matrices $\boldsymbol{R}_1$, $\boldsymbol{R}_2$ and $\boldsymbol{R}_3$ show that, for example, the correlation between ARCELOR and DAIMLER increases from 0.28 to 0.68 at the first break point, and to 0.70 at the second, which coincides with the usual observation that stocks in financial markets exhibit the synchronized movement during the crisis period.

7 Conclusions

In this paper, we propose a new residual-based constancy test on a correlation matrix in the light of a recently proposed general framework of residual-based inference on the moment hypothesis. Our test is robust to the existence of the residual effect when the marginal variances are not constant over time. The limit distribution of test statistics and its approximation are provided, and the good finite sample behavior has been validated by Monte Carlo simulations under various scenarios. We also examine the performance of the test on the time series of log returns of eight European stocks. In future research, one might develop a general theoretical foundation of the validity of bootstrap approximation in this constant correlation test. It would also be interesting to study whether the correlation models, such as DCC or CCC models can be improved in terms of, for example, Value-at-Risk forecasting accuracy with the combination of the proposed test in detecting and dating structural breaks in correlations.

A Proof

A.1 Proof of Proposition 1

To prove the convergence of the random vector $\frac{1}{\sqrt{n}}\sum_{t=1}^{[ns]} vech(\hat{\boldsymbol{Z}}_t\hat{\boldsymbol{Z}}_t')$, we need to prove each element in this $\frac{d(d-1)}{2}$ dimensional vector has such convergence, for $1 \leq i < j \leq d$:

$$\begin{aligned}
\tfrac{1}{\sqrt{n}}\textstyle\sum_{t=1}^{[ns]}\hat{Z}_{t,i}\hat{Z}_{t,j} &= \tfrac{1}{\sqrt{n}}\textstyle\sum_{t=1}^{[ns]} Z_{t,i}Z_{t,j} + \tfrac{1}{\sqrt{n}}\textstyle\sum_{t=1}^{[ns]}\tfrac{\partial z_i z_j}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}_t^*}\tfrac{\partial \boldsymbol{Z}_t^*}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}^*}(\hat{\boldsymbol{\theta}}_{\lambda_0}-\boldsymbol{\theta}_{\lambda_0})\\
&= \tfrac{1}{\sqrt{n}}\textstyle\sum_{t=1}^{[ns]} Z_{t,i}Z_{t,j} + \tfrac{1}{\sqrt{n}}\textstyle\sum_{t=1}^{[ns]}\tfrac{\partial z_i z_j}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}_t}\tfrac{\partial \boldsymbol{Z}_t}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}}(\hat{\boldsymbol{\theta}}_{\lambda_0}-\boldsymbol{\theta}_{\lambda_0})\\
&+\tfrac{1}{\sqrt{n}}\textstyle\sum_{t=1}^{[ns]}\left(\tfrac{\partial z_i z_j}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}_t^*}\tfrac{\partial \boldsymbol{Z}_t^*}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}^*} - \tfrac{\partial z_i z_j}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}_t}\tfrac{\partial \boldsymbol{Z}_t}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}}\right)(\hat{\boldsymbol{\theta}}_{\lambda_0}-\boldsymbol{\theta}_{\lambda_0})
\end{aligned}$$

Note that $\boldsymbol{\theta}_{\lambda_0}$ is true parameter vector with length $4d$, the estimator $\boldsymbol{\theta}_{\lambda_0}^*$ is the convex combination of $\boldsymbol{\theta}_{\lambda_0}$ and $\hat{\boldsymbol{\theta}}_{\lambda_0}$ such that it lies in the neighborhood Φ_n defined in Assumption 2 (as $\hat{\boldsymbol{\theta}}_{\lambda_0} - \boldsymbol{\theta}_{\lambda_0} = O_p(1/\sqrt{n})$, so $\boldsymbol{\theta}_{\lambda_0}^*$ lies in $\sqrt{n}$-neighborhood of $\boldsymbol{\theta}_{\lambda_0}$ hence in Φ_n). The first equal sign is given by expansion with mean value theorem around $\boldsymbol{\theta}_{\lambda_0}$, and the third term in the second line vanishes asymptotically as

$$\begin{aligned}
&||\tfrac{1}{\sqrt{n}}\textstyle\sum_{t=1}^{[ns]}\left(\tfrac{\partial z_i z_j}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}_t^*}\tfrac{\partial \boldsymbol{Z}_t^*}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}^*} - \tfrac{\partial z_i z_j}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}_t}\tfrac{\partial \boldsymbol{Z}_t}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}}\right)(\hat{\boldsymbol{\theta}}_{\lambda_0}-\boldsymbol{\theta}_{\lambda_0})||\\
\leq\ &||\sqrt{n}(\hat{\boldsymbol{\theta}}_{\lambda_0}-\boldsymbol{\theta}_{\lambda_0})||\sup_{\boldsymbol{\theta}_{\lambda_0}^*, t=1,\ldots,n}||\tfrac{\partial z_i z_j}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}_t^*}\tfrac{\partial \boldsymbol{Z}_t^*}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}^*} - \tfrac{\partial z_i z_j}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}_t}\tfrac{\partial \boldsymbol{Z}_t}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}}||\\
\to_p\ &0
\end{aligned}$$

The last line is guaranteed by Assumption 2. By applying Lemma 18.7 in Davidson (1994), p. 285, for $1 \leq i < j \leq d$, we have

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[ns]}\Big(\frac{\partial z_i z_j}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}_t^*}\frac{\partial \boldsymbol{Z}_t^*}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}^*} - \frac{\partial z_i z_j}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}_t}\frac{\partial \boldsymbol{Z}_t}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}}\Big)(\hat{\boldsymbol{\theta}}_{\lambda_0}-\boldsymbol{\theta}_{\lambda_0}) \to_p 0$$

Note that, from Assumption 1, we have

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_{\lambda_0}-\boldsymbol{\theta}_{\lambda_0}) \to_d \Sigma_{\lambda_0}^{1/2}\boldsymbol{\Theta}_{\lambda_0}(1)$$

In addition, once Assumption 1 is valid, it holds that the sample (cross-) moments of $\boldsymbol{Z}_t$ converges to its theoretical counterparts, respectively, in the subsamples separated by the change point λ_0. As a consequence, the asymptotic property of the essential part of residual effect follows

$$\frac{1}{n}\sum_{t=1}^{[ns]}\frac{\partial vech(\boldsymbol{z}\boldsymbol{z}')}{\partial \boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}_t}\frac{\partial \boldsymbol{Z}_t}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}} \Rightarrow \boldsymbol{\tau}_{\lambda_0}(s)$$

where the asymptotic residual effect part is

$$\boldsymbol{\tau}_{\lambda_0}(s)=\Big(\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,1}}(s)\bar{D}_{t,\lambda_0}+\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,1}}(\lambda_0)D_{t,\lambda_0}\quad(\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,2}}(s)-\boldsymbol{\tau}_{\boldsymbol{\theta}_{0,2}}(\lambda_0))D_{t,\lambda_0}\Big)$$

Following the elementwise convergence derived above, together with continuous mapping theorem and Assumptions 1-2, we have

$$\begin{aligned}
&\tfrac{1}{\sqrt{n}}\textstyle\sum_{t=1}^{[ns]}(vech(\hat{\boldsymbol{Z}}_t\hat{\boldsymbol{Z}}_t')-\mathbb{E}(vech(\boldsymbol{Z}_t\boldsymbol{Z}_t')))\\
=&\tfrac{1}{\sqrt{n}}\textstyle\sum_{t=1}^{[ns]}(vech(\boldsymbol{Z}_t\boldsymbol{Z}_t')-\mathbb{E}(vech(\boldsymbol{Z}_t\boldsymbol{Z}_t')))+\tfrac{1}{\sqrt{n}}\sum_{t=1}^{[ns]}\frac{\partial vech(\boldsymbol{z}\boldsymbol{z}')}{\partial\boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}_t}\frac{\partial\boldsymbol{Z}_t}{\partial\boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}}(\hat{\boldsymbol{\theta}}_{\lambda_0}-\boldsymbol{\theta}_{\lambda_0})\\
&+\tfrac{1}{\sqrt{n}}\textstyle\sum_{t=1}^{[ns]}\left(\frac{\partial vech(\boldsymbol{z}\boldsymbol{z}')}{\partial\boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}_t^*}\frac{\partial\boldsymbol{Z}_t^*}{\partial\boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}^*}-\frac{\partial vech(\boldsymbol{z}\boldsymbol{z}')}{\partial\boldsymbol{z}}\Big|_{\boldsymbol{z}=\boldsymbol{Z}_t}\frac{\partial\boldsymbol{Z}_t}{\partial\boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\lambda_0}}\right)(\hat{\boldsymbol{\theta}}_{\lambda_0}-\boldsymbol{\theta}_{\lambda_0})\\
\Rightarrow\;&\Omega^{1/2}\boldsymbol{\Gamma}(s)+\boldsymbol{\tau}_{\lambda_0}(s)\Sigma_{\lambda_0}^{1/2}\boldsymbol{\Theta}_{\lambda_0}(1)
\end{aligned}$$

A.2 Proof of Proposition 2

Following proposition 1, we have

$$\begin{aligned}
&\tfrac{j}{\sqrt{n}}(\hat{\boldsymbol{S}}_j-\hat{\boldsymbol{S}}_n)\\
=\;&\tfrac{j}{\sqrt{n}}(\tfrac{1}{j}\textstyle\sum_{t=1}^{j}vech(\hat{\boldsymbol{Z}}_t\hat{\boldsymbol{Z}}_t')-\tfrac{1}{n}\sum_{t=1}^{n}vech(\hat{\boldsymbol{Z}}_t\hat{\boldsymbol{Z}}_t'))\\
=\;&\tfrac{1}{\sqrt{n}}\textstyle\sum_{t=1}^{j}[vech(\hat{\boldsymbol{Z}}_t\hat{\boldsymbol{Z}}_t')-\mathbb{E}(vech(\boldsymbol{Z}_t\boldsymbol{Z}_t'))]-\tfrac{j}{n}\tfrac{1}{\sqrt{n}}\sum_{t=1}^{n}[vech(\hat{\boldsymbol{Z}}_t\hat{\boldsymbol{Z}}_t')-\\
&\mathbb{E}(vech(\boldsymbol{Z}_t\boldsymbol{Z}_t'))]\\
\Rightarrow\;&\Omega^{1/2}\boldsymbol{\Gamma}(s)+\boldsymbol{\tau}_{\lambda_0}(s)\Sigma_{\lambda_0}^{1/2}\boldsymbol{\Theta}_{\lambda_0}(1)-\tfrac{j}{n}[\Omega^{1/2}\boldsymbol{\Gamma}(1)+\boldsymbol{\tau}_{\lambda_0}(1)\Sigma_{\lambda_0}^{1/2}\boldsymbol{\Theta}_{\lambda_0}(1)]\\
:=\;&\Omega^{1/2}(\hat{\boldsymbol{\Gamma}}(s)-s\hat{\boldsymbol{\Gamma}}(1))
\end{aligned}$$

where $\hat{\boldsymbol{\Gamma}}(s)=\boldsymbol{\Gamma}(s)+\Omega^{-1/2}\boldsymbol{\tau}_{\lambda_0}(s)\Sigma_{\lambda_0}^{1/2}\boldsymbol{\Theta}_{\lambda_0}(1)$ and $j=[ns]$. The conclusion of this proposition follows with the continuous mapping theorem together with Assumption 3.

B Tables

	unequal cross correlations						equal cross correlations					
	$N=3, T=500$		$N=5, T=500$		$N=10, T=500$		$N=3, T=500$		$N=5, T=500$		$N=10, T=500$	
σ^2	unknown cp	true cp	unknown cp	true cp	unknown cp	true cp	unknown cp	true cp	unknown cp	true cp	unknown cp	true cp
0.1	0.067	0.054	0.058	0.053	0.039	0.050	0.064	0.060	0.055	0.048	0.051	0.059
0.2	0.060	0.056	0.055	0.051	0.037	0.044	0.058	0.057	0.052	0.048	0.054	0.057
0.3	0.056	0.054	0.056	0.045	0.044	0.041	0.059	0.059	0.050	0.043	0.056	0.053
0.4	0.056	0.054	0.049	0.044	0.041	0.037	0.060	0.055	0.052	0.040	0.055	0.051
0.5	0.050	0.048	0.046	0.042	0.040	0.035	0.058	0.050	0.044	0.038	0.054	0.050
0.6	0.044	0.042	0.042	0.039	0.032	0.032	0.055	0.048	0.039	0.037	0.048	0.047
0.7	0.045	0.041	0.042	0.038	0.038	0.028	0.050	0.050	0.041	0.038	0.049	0.051
0.8	0.040	0.042	0.042	0.040	0.041	0.027	0.051	0.051	0.037	0.041	0.051	0.052
0.9	0.046	0.043	0.043	0.040	0.047	0.025	0.045	0.050	0.029	0.043	0.048	0.050
1.0	0.043	0.048	0.043	0.040	0.054	0.027	0.041	0.051	0.034	0.042	0.048	0.050
1.1	0.045	0.046	0.038	0.042	0.047	0.030	0.047	0.052	0.029	0.044	0.042	0.052
1.2	0.046	0.049	0.045	0.040	0.038	0.031	0.046	0.053	0.042	0.040	0.047	0.056
1.3	0.043	0.049	0.043	0.043	0.037	0.029	0.055	0.053	0.043	0.043	0.050	0.055
1.4	0.051	0.050	0.046	0.040	0.029	0.030	0.055	0.050	0.041	0.041	0.053	0.054
1.5	0.056	0.051	0.049	0.040	0.030	0.032	0.059	0.048	0.042	0.039	0.053	0.053
1.6	0.061	0.055	0.046	0.040	0.031	0.029	0.057	0.051	0.046	0.042	0.054	0.050
1.7	0.062	0.057	0.050	0.041	0.031	0.029	0.059	0.052	0.044	0.044	0.056	0.051
1.8	0.060	0.059	0.050	0.042	0.031	0.033	0.057	0.053	0.047	0.044	0.053	0.051
1.9	0.059	0.059	0.051	0.041	0.032	0.037	0.058	0.052	0.045	0.043	0.053	0.049
2.0	0.058	0.061	0.049	0.043	0.035	0.036	0.057	0.053	0.046	0.042	0.054	0.050

Table 3: Empirical size of the multivariate constant correlation test for serially independent random variables with 1000 Monte Carlo simulations. σ^2 denotes the marginal variances in the second half of the sample. *unknown cp* and *true cp* indicate the test with estimated and the test with true change points in the marginal variances, respectively.

	unequal cross correlations								equal cross correlations							
	$N=3,T=500$		$N=5,T=500$		$N=10,T=500$		$N=10,T=1000$		$N=3,T=500$		$N=5,T=500$		$N=10,T=500$		$N=10,T=1000$	
σ^2	unknown cp	true cp	unknown cp	true cp	unknown cp	true cp	unknown cp	true cp	unknown cp	true cp	unknown cp	true cp	unknown cp	true cp	unknown cp	true cp
0.1	0.029	0.033	0.030	0.041	0.023	0.026	0.007	0.026	0.046	0.049	0.055	0.057	0.031	0.035	0.033	0.032
0.2	0.033	0.036	0.047	0.043	0.023	0.027	0.021	0.027	0.050	0.051	0.054	0.054	0.028	0.031	0.037	0.035
0.3	0.049	0.037	0.052	0.046	0.030	0.027	0.026	0.027	0.057	0.056	0.062	0.052	0.031	0.030	0.036	0.037
0.4	0.049	0.039	0.059	0.044	0.031	0.024	0.030	0.027	0.062	0.059	0.057	0.054	0.030	0.027	0.036	0.034
0.5	0.046	0.042	0.056	0.046	0.025	0.024	0.034	0.028	0.060	0.056	0.056	0.051	0.027	0.025	0.038	0.033
0.6	0.051	0.043	0.053	0.045	0.025	0.022	0.032	0.029	0.063	0.055	0.050	0.051	0.026	0.023	0.032	0.030
0.7	0.050	0.045	0.052	0.050	0.030	0.021	0.028	0.024	0.059	0.058	0.051	0.052	0.028	0.025	0.036	0.031
0.8	0.043	0.044	0.054	0.050	0.040	0.025	0.039	0.026	0.057	0.066	0.048	0.058	0.031	0.025	0.033	0.028
0.9	0.045	0.046	0.053	0.051	0.044	0.023	0.051	0.025	0.053	0.069	0.050	0.056	0.027	0.026	0.037	0.031
1.0	0.057	0.043	0.055	0.049	0.044	0.023	0.051	0.023	0.054	0.072	0.049	0.054	0.029	0.027	0.030	0.033
1.1	0.052	0.048	0.060	0.049	0.041	0.022	0.043	0.024	0.055	0.069	0.047	0.058	0.028	0.028	0.033	0.032
1.2	0.050	0.045	0.053	0.045	0.039	0.021	0.040	0.023	0.046	0.065	0.057	0.058	0.030	0.028	0.035	0.031
1.3	0.037	0.043	0.053	0.044	0.033	0.020	0.036	0.022	0.052	0.062	0.053	0.055	0.026	0.027	0.035	0.032
1.4	0.050	0.042	0.046	0.045	0.025	0.018	0.034	0.023	0.058	0.062	0.051	0.057	0.028	0.027	0.035	0.032
1.5	0.050	0.040	0.047	0.043	0.027	0.020	0.036	0.023	0.061	0.059	0.048	0.053	0.026	0.026	0.037	0.034
1.6	0.046	0.036	0.051	0.045	0.023	0.019	0.036	0.022	0.062	0.059	0.052	0.058	0.027	0.026	0.035	0.033
1.7	0.044	0.040	0.049	0.047	0.021	0.017	0.033	0.025	0.058	0.059	0.055	0.058	0.027	0.027	0.036	0.033
1.8	0.046	0.043	0.049	0.046	0.022	0.021	0.036	0.024	0.058	0.060	0.053	0.056	0.030	0.027	0.035	0.030
1.9	0.049	0.043	0.053	0.045	0.025	0.021	0.036	0.027	0.056	0.063	0.057	0.055	0.029	0.028	0.037	0.033
2.0	0.048	0.045	0.047	0.044	0.028	0.022	0.037	0.029	0.057	0.061	0.057	0.054	0.028	0.026	0.035	0.033

Table 4: Empirical size of the multivariate constant correlation test for serially dependent random variables with 1000 Monte Carlo simulations. σ^2 denotes the marginal variances in the second half of the sample. *unknown cp* and *true cp* indicate the test with estimated and the test with true change points in the marginal variances, respectively.

	Serially independent random variables								Serially dependent random variables							
	$N = 3, T = 500$		$N = 5, T = 500$		$N = 10, T = 500$		$N = 10, T = 1000$		$N = 3, T = 500$		$N = 5, T = 500$		$N = 10, T = 500$		$N = 10, T = 1000$	
ρ	unknown cp	true cp	unknown cp	true cp	unknown cp	true cp	unknown cp	true cp	unknown cp	true cp	unknown cp	true cp	unknown cp	true cp	unknown cp	true cp
0	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	1.000	1.000
0.1	0.977	0.977	0.994	0.995	0.976	0.972	1.000	1.000	0.930	0.922	0.962	0.957	0.835	0.819	1.000	0.999
0.2	0.667	0.641	0.654	0.622	0.469	0.443	0.883	0.872	0.529	0.488	0.488	0.442	0.261	0.238	0.674	0.656
0.3	0.158	0.134	0.128	0.101	0.096	0.095	0.168	0.158	0.172	0.143	0.123	0.113	0.050	0.048	0.106	0.099
0.4	0.057	0.047	0.053	0.043	0.054	0.050	0.039	0.037	0.051	0.052	0.052	0.048	0.028	0.026	0.035	0.033
0.5	0.123	0.135	0.112	0.105	0.086	0.085	0.128	0.109	0.113	0.117	0.091	0.090	0.052	0.052	0.084	0.083
0.6	0.514	0.539	0.349	0.362	0.244	0.238	0.497	0.473	0.366	0.389	0.231	0.231	0.144	0.144	0.304	0.286
0.7	0.930	0.945	0.775	0.774	0.567	0.545	0.902	0.892	0.796	0.817	0.615	0.617	0.390	0.355	0.717	0.702
0.8	0.999	1.000	0.982	0.987	0.856	0.837	1.000	1.000	0.977	0.984	0.912	0.911	0.676	0.622	0.964	0.951

Table 5: Empirical power of the multivariate constant correlation test for both serially independent and serially dependent random variables with 1000 Monte Carlo simulations. ρ denotes the cross correlations in the second half of the sample. *unknown cp* and *true cp* indicate the test with estimated and the test with true change points in the marginal variances, respectively.

C Figures

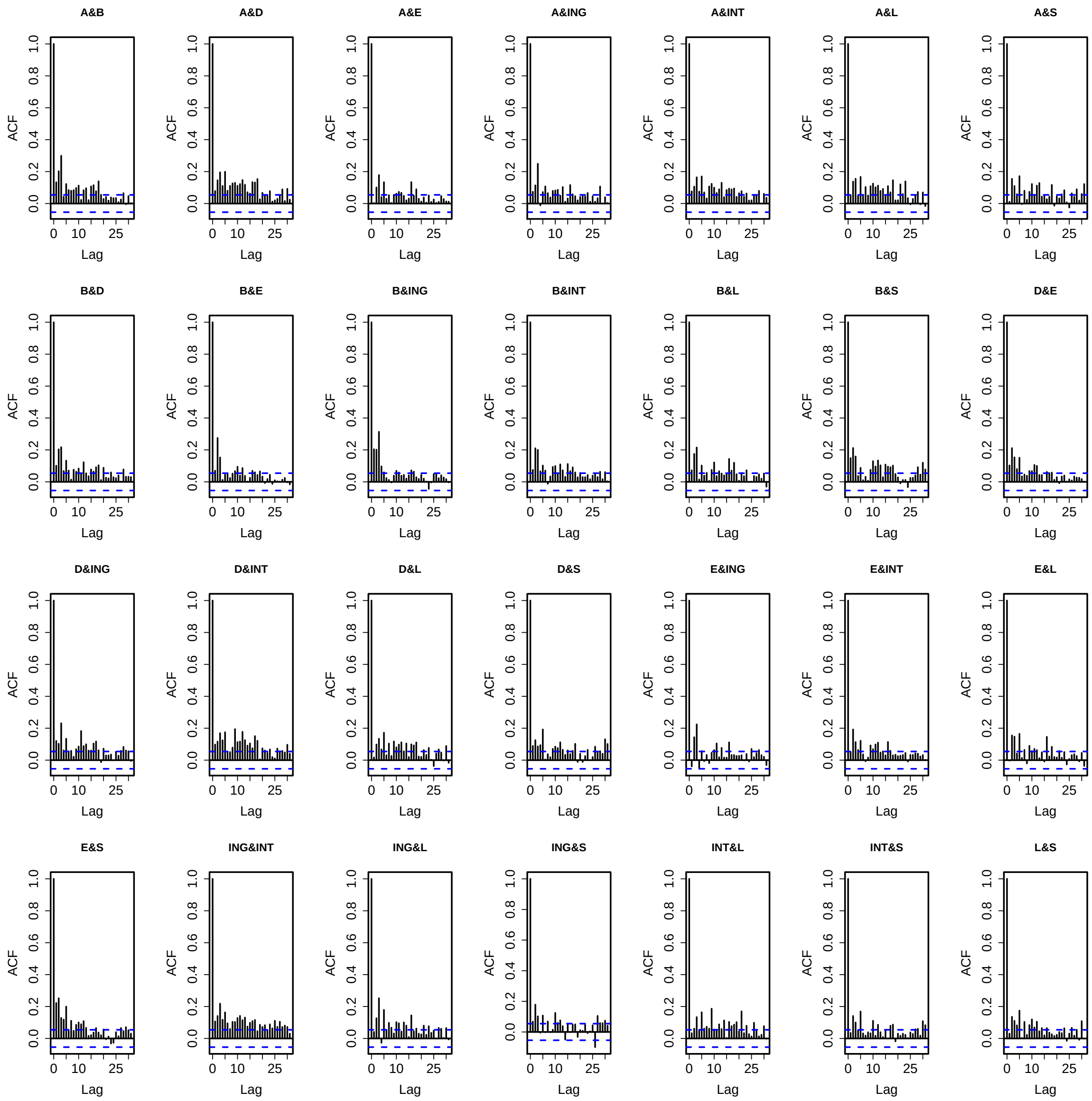

Figure 8: ACF plots of cross products of residuals $\hat{Z}_{t,i}\hat{Z}_{t,j}, \forall i \neq j$ for stocks, and A, B, D, E, ING, INT, L, S are abbreviations for *ARCELOR*, *BASF*, *DAIMLER*, *ENEL*, *INGGROEP*, *INTESA*, *LVMH*, *SANOFI*, respectively

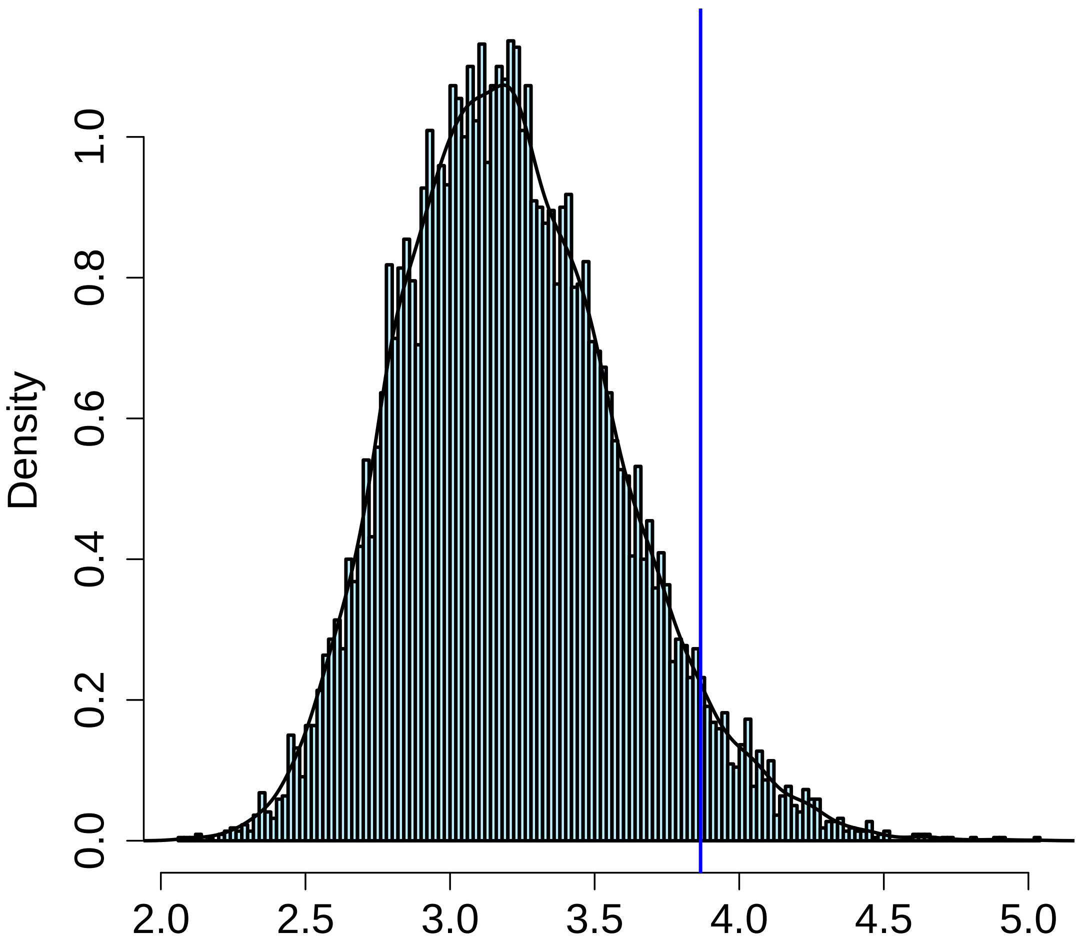

Figure 9: Histogram of asymptotic distribution of test statistics Q_n approximated by block bootstrap procedure with $B = 10999$ bootstrap replications and block length $T^{1/3}$ (the blue vertical line indicates the 0.95 quantile)

References

Andrews, D. (1991), 'Heteroskedasticity and autocorrelation consistent covariance matrix estimation', *Econometrica* **59**(3), 817–858.

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

Bai, J. and Ng, S. (2005), 'Tests for skewness, kurtosis, and normality for time series data', *Journal of Business & Economic Statistics* **23**(1), 49–60.

Berens, T., Weiß, G. N. and Wied, D. (2015), 'Testing for structural breaks in correlations: Does it improve Value-at-Risk forecasting?', *Journal of Empirical Finance* **32**(C), 135 – 152.

Brown, R. L., Durbin, J. and Evans, J. M. (1975), 'Techniques for testing the constancy of regression relationships over time', *Journal of the Royal Statistical Society. Series B (Methodological)* **37**(2), 149–192.

Carlstein, E. (1986), 'The use of subseries values for estimating the variance of a general statistic from a stationary sequence', *The Annals of Statistics* **14**(3), 1171–1179.

Davidson, J. (1994), *Stochastic Limit Theory*, Advanced Texts in Econometrics, Oxford University Press, UK.

Demetrescu, M. and Wied, D. (2017), 'Residual-based inference on moment hypotheses, with an application to testing for constant correlation', *SSRN Research Paper Nr. 3119290* .

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

Galeano, P. and Wied, D. (2017), 'Dating multiple change points in the correlation matrix', *TEST: An Official Journal of the Spanish Society of Statistics and Operations Research* **26**(2), 331–352.

Guillén, M. F. (2015), 'The global economic & financial crisis: A timeline', *Unpublished manuscript* .

Hall, P. and Horowitz, J. (1996), 'Bootstrap critical values for tests based on generalized-method-of-moments estimators', *Econometrica* **64**(4), 891–916.

Hansen, L. (1982), 'Large sample properties of generalized method of moments estimators', *Econometrica* **50**(4), 1029–54.

Inoue, A. and Shintani, M. (2006), 'Bootstrapping GMM estimators for time series', *Journal of Econometrics* **133**(2), 531 – 555.

Kiefer, J. (1959), 'K-sample analogues of the Kolmogorov-Smirnov and Cramér-V. Mises tests', *The Annals of Mathematical Statistics* **30**(2), 420–447.

Lahiri, S. N. (1999), 'Theoretical comparisons of block bootstrap methods', *The Annals of Statistics* **27**(1), 386–404.

Lahiri, S. N. (2003), *Resampling Methods for Dependent Data*, Springer Series in Statistics, 1 edn, Springer-Verlag New York.

Pape, K., Wied, D. and Galeano, P. (2016), 'Monitoring multivariate variance changes', *Journal of Empirical Finance* **39**(PA), 54–68.

Ploberger, W. and Krämer, W. (1992), 'The CUSUM test with OLS residuals', *Econometrica* **60**(2), 271–285.

Politis, D. N. and White, H. (2004), 'Automatic block-length selection for the dependent bootstrap', *Econometric Reviews* **23**(1), 53–70.

Wied, D. (2017), 'A nonparametric test for a constant correlation matrix', *Econometric Reviews* **36(10)**, 1157–1172.

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

Wied, D., Krämer, W. and Dehling, H. (2012), 'Testing for a change in correlation at an unknown point in time using an extended functional delta method', *Econometric Theory* **28**(03), 570–589.

Zhang, X. and Cheng, G. (2014), 'Bootstrapping high dimensional time series', *ArXiv e-prints* .

Zhou, Z. (2013), 'Heteroscedasticity and autocorrelation robust structural change detection', *Journal of the American Statistical Association* **108**(502), 726–740.